Forthcoming in: *Economics Letters*

Tractable Likelihood-Based Estimation of Non-Linear DSGE Models

Robert Kollmann [(*)]
ECARES, Université Libre de Bruxelles & CEPR

August 29, 2017

This paper presents a simple and fast maximum likelihood estimation method for non-linear DSGE models that are solved using a second- (or higher-) order accurate approximation. The method requires that the number of observables equals the number of exogenous shocks. Exogenous innovations are extracted recursively by inverting the observation equation, which allows easy computation of the likelihood function.

Keywords: Estimation of non-linear DSGE models, observation equation inversion.
JEL codes: C51, C63, C68, E37

[(*)] I thank Wouter den Haan, Tom Holden, Matteo Iacoviello, Jinill Kim, Johannes Pfeifer, Stephanie Schmitt-Grohé, Frank Schorfheide, Chris Sims, Martin Uribe and Raf Wouters for useful discussions. This research was funded by EU-FP7 grant 612796 (MACFINROBODS).
Address: European Centre for Advanced Research in Economics and Statistics (ECARES), Université Libre de Bruxelles, 50 Av. Roosevelt, 1050-Brussels, Belgium.
robert_kollmann@yahoo.com http://www.robertkollmann.com

1. Introduction

Dynamic Stochastic General Equilibrium (DSGE) models are the workhorses of modern macroeconomics. A large literature has empirically estimated *linearized* DSGE models using likelihood-based methods (e.g., Kim (2000), Otrok (2001), Ireland (2004)). Linearized models cannot capture the effect of big shocks, or the role of risk for economic behavior—*non-linear* model solutions are needed for studying these phenomena. This paper presents a simple and fast maximum likelihood estimation method for non-linear DSGE models that are solved using second- (or higher-) order Taylor approximations (e.g., Sims (2000), Kollmann (2002), Schmitt-Grohé and Uribe (2004), Lombardo and Sutherland (2007)). Those approximations provide the most tractable and widely used non-linear solution technique for medium- and large-scale DSGE models (Kollmann et al. (2011)). Thus, it is important to develop efficient methods for taking higher-order approximated models to data.

The *estimation* method discussed here requires that the number of observed variables (used for estimation) equals the number of exogenous shocks in the DSGE model. Exogenous innovations are extracted recursively by inverting the observation equation, which allows easy computation of the sample likelihood. A challenge for this approach is that, in higher-order approximated models, the decision rule (solution) for endogenous variables depends on powers of exogenous innovations--multiple exogenous innovations are thus consistent with the data. To overcome this problem, I posit a *modified* higher-order decision rule in which *powers* of exogenous innovations are replaced by their unconditional expected value. This allows straightforward observation equation inversion. A numerical example suggests that the estimation method here provides accurate parameter estimates, even for models with strong curvature and big shocks. [1]

Other likelihood-based estimation methods for non-linear DSGE models use particle filters or deterministic filters to infer exogenous shocks. The estimation method here is much simpler and faster; it can thus be applied to larger models. [2]

[1] Observation equation inversion is an intuitive and popular statistical technique (e.g., Guerrieri and Iacoviello (2014), Deák et al. (2015)), but has so far not been used to estimate higher-order approximated DSGE models. The paper here shows how this can be accomplished.

[2] Particle filters (PFs) use Monte Carlo methods to infer latent states (An and Schorfheide (2007)), and are thus computationally slow. Deterministic filters (DFs) are much faster than PFs, as DFs do not use Monte Carlos; instead updating rules akin to the standard Kalman filter are employed; this requires computation of conditional moments of the state vector (e.g., Kollmann (2015)). The method here is fastest as it does not involve computing moments of states. In contrast to the method here, PFs and DFs can be used when there are less observables than shocks.

2. Model and estimation method

Standard DSGE models can be expressed as:

$$E_t G(\Omega_{t+1}, \Omega_t, \varepsilon_{t+1}) = 0, \qquad (1)$$

where $G: R^{2n+m} \to R^n$ is a function, and Ω_t is an nx1 vector of endogenous and exogenous variables known at date t; $\varepsilon_{t+1} \sim N(0, \xi^2 \Sigma_\varepsilon)$ is an mx1 vector of innovations to exogenous variables; ξ is a scalar indexing shock size. The model solution is a 'decision rule' $\Omega_{t+1} = F(X_t, \varepsilon_{t+1}, \xi)$, where X_t is a vector of state variables (predetermined endogenous variables and exogenous variables), i.e. $X_t = \Lambda\Omega_t$, where Λ is a matrix that picks the state variables among the elements of Ω_t. The decision rule has to satisfy $E_t G(F(\Lambda\Omega_t, \varepsilon_{t+1}, \xi), \Omega_t, \varepsilon_{t+1}) = 0 \;\; \forall \Omega_t$. Following Sims (2000) and Schmitt-Grohé and Uribe (2004), this paper focuses on second-order accurate model solutions, namely on second-order Taylor approximations of the decision rule around a deterministic steady state, i.e. around $\xi = 0$ and a vector Ω such that $0 = G(\Omega, \Omega, 0)$.

The paper presents an estimation method for *second*-order approximated models. It is straightforward to extend the estimation method to models that are approximated to a *higher* order--see Appendix (Not-for-Publication).

Second-order accurate solution and pruning

Let $\omega_t \equiv \Omega_t - \Omega, x_t \equiv X_t - X$ (with $X = \Lambda\Omega$) denote deviations of Ω_t, X_t from steady state. The second-order accurate model solution has the form

$$\omega_{t+1} = F_0 \xi^2 + F_1 x_t + F_2 \varepsilon_{t+1} + F_{11} x_t \otimes x_t + F_{12} x_t \otimes \varepsilon_{t+1} + F_{22} \varepsilon_{t+1} \otimes \varepsilon_{t+1}, \;\; x_t = \Lambda\omega_t, \qquad (2)$$

where $\otimes$ denotes the Kronecker product. $F_0, F_1, F_2, F_{11}, F_{12}, F_{22}$ are vectors/matrices that are functions of structural parameters. The first-order accurate (linearized) model solution is:

$$\omega_{t+1}^{(1)} = F_1 x_t^{(1)} + F_2 \varepsilon_{t+1}, \text{ with } x_t^{(1)} = \Lambda\omega_t^{(1)}. \qquad (3)$$

The superscript $^{(1)}$ denotes variables solved to first-order accuracy. I assume that the eigenvalues of F_1 are strictly inside the unit circle, i.e. that the linearized model is stable.

I use the 'pruning' scheme of Kim et al. (2008), under which $x_t \otimes x_t$ and $x_t \otimes \varepsilon_{t+1}$ are replaced by $x_t^{(1)} \otimes x_t^{(1)}$ and $x_t^{(1)} \otimes \varepsilon_{t+1}$, respectively, in (2):

$$\omega_{t+1} = F_0 \xi^2 + F_1 x_t + F_2 \varepsilon_{t+1} + F_{11} x_t^{(1)} \otimes x_t^{(1)} + F_{12} x_t^{(1)} \otimes \varepsilon_{t+1} + F_{22} \varepsilon_{t+1} \otimes \varepsilon_{t+1}. \qquad (4)$$

Note that $x_t \otimes x_t = x_t^{(1)} \otimes x_t^{(1)}$ and $x_t \otimes \varepsilon_{t+1} = x_t^{(1)} \otimes \varepsilon_{t+1}$ hold, up to second-order accuracy. Thus, (4) is a valid second-order accurate solution. The justification for pruning is that (2) has spurious steady states (not present in the original model); some of those steady states mark transitions to unstable behavior. Large shocks can thus trigger explosive trajectories. Pruning eliminates this problem. Stability of the first-order solution (3) ensures that the pruned second-order solution is stable. Pruning is thus essential for applied work based on second-order approximated models.

The estimation method below uses data on ω_{t+1} to extract the exogenous innovation ε_{t+1}. As ω_{t+1} depends on squares of ε_{t+1}, multiple innovations are consistent with given data. To allow observation equation inversion, I replace the term $\varepsilon_{t+1} \otimes \varepsilon_{t+1}$ in (4) by its expected value $E(\varepsilon_{t+1} \otimes \varepsilon_{t+1})$.[3] This produces the '*modified'* decision rule

$$\omega_{t+1} = F_0 \xi^2 + F_1 x_t + F_2 \varepsilon_{t+1} + F_{11} x_t^{(1)} \otimes x_t^{(1)} + F_{12} x_t^{(1)} \otimes \varepsilon_{t+1} + F_{22} E(\varepsilon_{t+1} \otimes \varepsilon_{t+1}) \qquad (5)$$

that is *linear* in ε_{t+1}, but non-linear in lagged state variables. The subsequent discussion assumes that (5) is the **true** data generating process (DGP).[4]

Observation equation inversion

The estimation method here requires that the number of observables equals the number of exogenous innovations, *m*. Assume that the econometrician observes a vector z_{t+1} comprising *m* elements of the vector ω_{t+1}. Thus, the observation equation is $z_{t+1} = Q\omega_{t+1}$, where Q is an *m*x*n* matrix. Substituting (5) into the observation equation gives $z_{t+1} = \gamma_t + \lambda_t \varepsilon_{t+1}$, where $\gamma_t \equiv Q \cdot (F_0 \xi^2 + F_1 x_t + F_{11} x_t^{(1)} \otimes x_t^{(1)} + F_{22} E(\varepsilon_{t+1} \otimes \varepsilon_{t+1}))$ and λ_t is an mxm matrix such that $\lambda_t \varepsilon_{t+1} \equiv Q \cdot (F_2 \varepsilon_{t+1} + F_{12} x_t^{(1)} \otimes \varepsilon_{t+1})$. Provided λ_t is non-singular, we thus have:

$$\varepsilon_{t+1} = \lambda_t^{-1} (z_{t+1} - \gamma_t) \,. \qquad (6)$$

---

[3]I thank Chris Sims for suggesting this approach. Dropping the term $F_{22} \varepsilon_{t+1} \otimes \varepsilon_{t+1}$ from (4) also permits observation equation inversion and produces very similar estimation results.

[4]For the illustrative DSGE model below, (4) and (5) are virtually indistinguishable: feeding the same sequence of innovations $\{\varepsilon\}$ into (4) and (5) produces almost identical time series $\{\omega\}$. Using (5) to extract the exogenous innovations (see below) from the time series generated by (4) and (5) yields very similar parameter estimates. Thus, even if the true DGP is (4), one obtains reliable estimates by *positing* (5) for observation equation inversion.

Sample likelihood

Given the initial states $x_0^{(1)}, x_0$ and data $\{z_t\}_{t=1}^T$ one can recursively extract the exogenous innovations $\{\varepsilon_t\}_{t=1}^T$ using (3),(5),(6). The log likelihood of the data (conditional on $x_0^{(1)}, x_0$) is:

$$\ln L(\{z_t\}_{t=1}^T | x_0^{(1)}, x_0) = -\tfrac{mT}{2}\ln(2\pi) - \tfrac{T}{2}\ln|\xi^2\Sigma_\varepsilon| - \tfrac{1}{2}\sum_{t=1}^T \{\varepsilon_t' (\xi^2\Sigma_\varepsilon)^{-1}\varepsilon_t - \ln|\lambda_{t-1}|. \qquad (7)$$

Structural model parameters (and initial states) can be estimated by maximizing this function.

3. Illustration

The method is tested for a basic Real Business Cycle (RBC) model. Assume a representative household who maximizes date *t* lifetime utility V_t given by $V_t = \frac{1}{1-\sigma}C_t^{1-\sigma} - \frac{1}{1+\eta}\psi_t N_t^{1+\eta} + \lambda_t \beta E_t V_{t+1}$, where C_t and N_t are consumption and hours worked. $\sigma,\eta>0$ are risk aversion and the inverse of labor supply elasticity. $\lambda_t\beta$ is the subjective discount factor between *t* and *t*+1. $\psi_t,\lambda_t>0$ are preference shocks. The resource/technology constraints are $C_t+I_t+G_t=Y_t$, $Y_t=\theta_t K_t^\alpha N_t^{1-\alpha}$, $K_{t+1}=(1-\delta)K_t+I_t$, $0<\alpha,\delta<1$. $Y_t, K_t, I_t, G_t, \theta_t$ denote GDP, capital, investment, and exogenous government purchases and productivity. The exogenous variables follow $\ln(\theta_t)=\rho_\theta\ln(\theta_{t-1})+\varepsilon_{\theta,t}$, $\ln(G_t/G)=\rho_G\ln(G_{t-1}/G)+\varepsilon_{G,t}$, $\ln(\psi_t)=\rho_\psi\ln(\psi_{t-1})+\varepsilon_{\psi,t}$, $\ln(\lambda_t)=\rho_\lambda\ln(\lambda_{t-1})+\varepsilon_{\lambda,t}$, where $\varepsilon_{\theta,t}, \varepsilon_{G,t}, \varepsilon_{\psi,t}, \varepsilon_{\lambda,t}$ are normal white noises with standard deviations $s_\theta, s_G, s_\psi, s_\lambda$. As conventional in (quarterly) models, I set $\beta=0.99, \eta=0.25, \alpha=0.3, \delta=0.025, \rho_\theta=\rho_G=\rho_\psi=\rho_\lambda=0.99$. The steady state government purchases/GDP ratio is $G/Y=0.2$. Risk aversion is set at a high value, $\sigma=10$, to generate strong curvature and permit non-negligible differences between first- and second-order model approximations.

I normalize $\xi=1$. One model variant assumes standard deviations $s_\theta=s_G=s_\psi=1\%, s_\lambda=0.025\%$ ('small shocks' variant) as in typical RBC models. The relative size of the shocks ensures that each shock accounts for a non-negligible share of the variance of GDP (see Appendix). I also consider a 'big shocks' variant in which shocks are 5 times larger: $s_\theta=s_G=s_\psi=5\%, s_\lambda=0.125\%$. I solve the model using DYNARE.

For each model variant, I generated 30 simulation runs of 100 periods.[5] For each run, I estimated 10 model parameters by maximizing the likelihood function (7): the risk aversion coefficient (σ), labor supply parameter (η) and autocorrelations and standard deviations of the exogenous variables. As the model has four exogenous shocks, four observables are needed for estimation. GDP, consumption, investment and hours worked are used as observables.

In computing the sample likelihood, I *assume* that initial states $x_0^{(1)}, x_0$ equal their unconditional mean. Although true initial states (in a given sample) differ from the assumed initial states, the recursively extracted exogenous innovations converge to the true innovations after a few periods. I thus use the first 10 periods of each simulation run as a training sample (the first 10 periods are dropped from the likelihood function).

One evaluation of the likelihood takes merely 0.014 seconds on a personal computer (Intel i7-7700K processor). This allows rapid maximization of the likelihood.

The Table reports the median, mean and standard deviation of the estimated model parameters across the 30 simulation runs, for the 'small shocks' model variant (Columns (1)-(3)) and for the 'big shocks' variant (Cols. (4)-(6)). Most model parameters are tightly estimated: the median and mean parameter estimates (across runs) are close to true parameter values, and the standard deviations of the parameter estimates are generally small.

[5]To eliminate the influence of initial conditions, the model was simulated over 5100 periods; the first 5000 periods were discarded.

References

An, S. and F. Schorfheide, 2007. Bayesian Analysis of DSGE models. Econometric Reviews 26, 113–172.

Deák, S., T. Holden and A. Mele, 2015. Advanced Course on DSGE Modelling. University of Surrey.

Guerrieri, L. and M. Iacoviello, 2014. Collateral Constraints and Macroeconomic Asymmetries. WP, Federal Reserve Board.

Ireland, P., 2004. A Method for Taking Models to the Data. Journal of Economic Dynamics and Control 28, 1205 – 1226.

Kim, J., 2000. Constructing and Estimating a Realistic Optimizing Model of Monetary Policy. Journal of Monetary Economics 45, 329-359.

Kim, J., Kim, S., Schaumburg, E. and C. Sims, 2008. Calculating and Using Second-Order Accurate Solutions of Discrete-Time Dynamic Equilibrium Models, Journal of Economic Dynamics and Control 32, 3397-3414.

Kollmann, R., 2002. Monetary Policy Rules in the Open Economy: Effects on Welfare and Business Cycles. Journal of Monetary Economics 49, 989-1015.

Kollmann, R., S. Maliar, B. Malin and P. Pichler, 2011. Comparison of Numerical Solutions to a Suite of Multi-Country Models. Journal of Economic Dynamics and Control 35, pp.186-202

Kollmann, R., 2015. Tractable Latent State Filtering for Non-Linear DSGE Models Using a Second-Order Approximation and Pruning, Computational Economics 45, 239-260.

Lombardo, G. and A. Sutherland, 2007. Computing Second-Order Accurate Solutions for Rational Expectation Models Using Linear Solution Methods. Journal of Economic Dynamics and Control 31, 515-530.

Otrok, C., 2001. On Measuring the Welfare Cost of Business Cycles. Journal of Monetary Economics 47, 61-92.

Schmitt-Grohé, S. and M. Uribe, 2004. Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function. Journal of Economic Dynamics and Control 28, 755 – 775.

Sims, C., 2000. Second Order Accurate Solution of Discrete Time Dynamic Equilibrium Models. WP, Princeton University.

Monte Carlo: parameter estimates for second-order approximated RBC model

Parameter	Model variant with 'small shocks'			Model variant with 'big shocks'		
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
	Median	**Mean**	**Std**	**Median**	**Mean**	**Std**
σ	10.18	12.01	5.11	10.08	10.29	1.65
η	0.27	0.59	0.79	0.22	0.32	0.26
ρ_θ	0.99	0.99	0.002	0.99	0.99	0.002
ρ_G	0.99	0.98	0.02	0.98	0.98	0.01
ρ_ψ	0.99	0.99	0.01	0.99	0.99	0.01
ρ_λ	0.97	0.93	0.09	0.99	0.96	0.07
s_θ (%)	1.00	1.00	0.07	4.94	4.99	0.45
s_G (%)	0.99	1.00	0.08	5.04	5.09	0.48
s_ψ (%)	1.02	1.28	0.73	4.99	5.16	1.17
s_λ (%)	0.027	0.059	0.069	0.12	0.16	0.13

Note: The Table summarizes parameters estimates across 30 simulation runs (T=100). Cols. labelled 'Median', 'Mean' and 'Std' report the median, mean and standard deviation of estimated parameters (see left-most column) across the 30 runs. Cols. labelled (1)-(3) [(4)-(6)]: 'small shocks' ['big shocks'] model variant. The ***true*** parameter values are: $\sigma=10, \eta=0.25, \rho_\theta=\rho_G=\rho_\psi=\rho_\lambda=0.99$. True standard deviations of exogenous innovations in 'small shocks' variant: $s_\theta=s_G=s_\psi=1\%, s_\lambda=0.025\%$. 'Big shocks' variant: $s_\theta=s_G=s_\psi=5\%, \sigma_\lambda=0.125\%$.

Not-for-publication Appendix for

‘Tractable Likelihood-Based Estimation of Non-Linear DSGE Models’

Part A. of this Appendix provides supplementary information for the second-order approximated RBC model discussed in the main text. Part B. shows how *third*-order approximated DSGE models can be estimated using observation equation inversion.

A. Supplementary information for the second-order approximated RBC model

● Comparison between decision rule (4) and modified decision rule (5)

Table a1 documents that the decision rule (4) and the modified decision rule (5) are (essentially) indistinguishable. An identical sequence of random exogenous innovations of length T=500,000 was fed into (4) and into (5). Table a1 shows that the resulting time series of endogenous variables are almost perfectly correlated across (4) and (5), and that they have (essentially) the same standard deviation. This holds both for levels and for first differences of logged simulated endogenous variables.

● Standard deviations of first- and second-order approximated models

Table a2 reports predicted standard deviations of first- and second-order approximated variables (log levels and log first differences). The Table documents that each of the four types of exogenous shocks accounts for a sizable share of the variance of GDP (see Panel (a), Col. (1)). In the ‘small shocks’ model variant, the first- and second-order approximated models produce almost identical standard deviations of endogenous variables (see Panel (a)). In the ‘big shocks’ model variant, by contrast, the second-order approximated variables are more volatile than the first-order approximated variables; this is, especially, the case for GDP, investment and hours worked (see Panel (b)).

Table a1. Second-order approximated RBC model: correlations across time series generated by decision rule (4) [ω] and time series generated by the 'modified' decision rule (5) [ω^{mod}]

	Y	*C*	*I*	*N*	*K*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
(a) Model variant with 'small shocks' ($\sigma_\theta=\sigma_G=\sigma_\psi=1\%, \sigma_\lambda=0.025\%$)					
Correlations between ω and ω^{mod}					
Levels	1.0000	1.0000	1.0000	1.0000	1.0000
First differences	1.0000	1.0000	0.9999	1.0000	1.0000
Relative standard deviations: std(ω)/std(ω^{mod})					
Levels	1.0000	1.0000	1.0000	1.0000	1.0000
First differences	1.0000	1.0000	1.0000	1.0000	1.0000
Relative standard deviation of difference between decision rules: std($\omega - \omega^{mod}$)/std(ω^{mod})					
Levels	0.0006	0.0001	0.0030	0.0003	0.0001
First differences	0.0039	0.0012	0.0168	0.0033	0.0020
(b) Model variant with 'big shocks' ($\sigma_\theta=\sigma_G=\sigma_\psi=5\%, \sigma_\lambda=0.125\%$)					
Correlations between ω and ω^{mod}					
Levels	1.0000	1.0000	0.9999	1.0000	1.0000
First differences	0.9998	1.0000	0.9967	0.9999	0.9999
Relative standard deviations: std(ω)/std(ω^{mod})					
Levels	1.0000	1.0000	0.9999	1.0000	1.0000
First differences	1.0000	1.0000	0.9994	1.0000	1.0000
Relative standard deviation of difference between decision rules: std($\omega - \omega^{mod}$)/std(ω^{mod})					
Levels	0.0027	0.0005	0.0147	0.0014	0.0007
First differences	0.0192	0.0061	0.0818	0.0164	0.0101

Note: Correlations of simulated time series (of variables listed above Cols. (1)-(5)) generated by the decision rule (4) and by the 'modified' decision rule (5) are reported, as well as the relative standard deviation of these two sets of time series. These statistics are reported for variables in log levels, and for variables in log first differences. *Y*: GDP; *C*: consumption; *I*: gross investment; *N*: hours worked; *K*: capital stock. Correlations greater than 0.99995 are reported as 1.0000. Reported statistics are based on one sequence of T=500,000 random exogenous innovations that was fed into (4) and (5).

Table a2. RBC model: predicted standard deviations (in %). Comparison between 1st order and 2nd order accurate model solutions

	Y	*C*	*I*	*N*	*K*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
(a) Model variant with 'small shocks' $(\sigma_{\theta}=\sigma_{G}=\sigma_{\psi}=1\%, \sigma_{\lambda}=0.025\%)$					
Variables in levels					
1st order, all shocks	3.34	1.57	10.43	9.68	7.59
1st order, just θ shock	2.07	1.36	6.20	9.32	4.52
1st order, just G shock	1.66	0.08	1.50	1.97	1.08
1st order, just ψ shock	1.14	0.75	3.43	0.96	2.49
1st order, just λ shock	1.66	0.21	7.51	1.61	5.48
2nd order, all shocks	3.34	1.57	10.43	9.68	7.59
First-differenced variables					
1st order, all shocks	0.67	0.17	2.60	1.13	0.18
2nd order, all shocks	0.67	0.17	2.60	0.13	0.18
(b) Model variant with 'big shocks' $(\sigma_{\theta}=\sigma_{G}=\sigma_{\psi}=5\%, \sigma_{\lambda}=0.125\%)$					
Variables in levels					
1st order, all shocks	16.72	7.83	52.14	48.39	37.95
2nd order, all shocks	17.11	7.83	52.97	48.67	38.21
First-differenced variables					
1st order, all shocks	3.33	0.86	12.98	5.66	0.91
2nd order, all shocks	3.41	0.87	13.37	5.77	0.92

Note: Standard deviations (in %) of simulated variables (listed above Cols. (1)-(5)) are shown for the RBC model. Rows labeled '1st order' and '2nd order' show standard deviations predicted by the first- and second-order accurate model solutions, respectively. The statistics are reported for variables in log levels, and for variables in log first differences. *Y*: GDP; *C*: consumption; *I*: gross investment; *N*: hours worked; *K*: capital stock. All statistics are computed using one simulation run of 500,000 periods.

B. Tractable Likelihood-Based Estimation of Third-Order Approximated DSGE Models

The technique described in the main text can also be used for likelihood estimation of DSGE models that are approximated to an order that is higher than the second order. This is illustrated here for **third-order** approximated models.

The third-order accurate model solution of the DSGE model (1) is given by:

$$\omega_{t+1}=F_0\xi^2+(F_1+F_{1\xi}\xi^2)x_t+(F_2+F_{2\xi}\xi^2)\varepsilon_{t+1}+F_{11}x_t\otimes x_t+F_{12}x_t\otimes\varepsilon_{t+1}+F_{22}\varepsilon_{t+1}\otimes\varepsilon_{t+1}+...$$

$$F_{111}x_t\otimes x_t\otimes x_t+F_{112}x_t\otimes x_t\otimes\varepsilon_{t+1}+F_{122}x_t\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1}+F_{222}\varepsilon_{t+1}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1},\ \text{with}\ x_t=\Lambda\omega_t. \qquad \text{(B.1)}$$

$F_{1\xi},F_{2\xi},F_{111},F_{112},F_{122},F_{222}$ are matrices that are functions of the structural model parameters ($F_0,F_1,F_2,F_{11},F_{12},F_{22}$ are identical to the corresponding coefficients in the second-order accurate model solution; see (2)).

'Pruning' is also essential for applied work based on third-order approximated models--the 'un-pruned' system (B.1) can exhibit explosive dynamics, in response to big shocks (see discussion in main text). To apply the logic of pruning to equation (B.1), note that the following conditions hold up to third-order accuracy:

$$\xi^2x_t=\xi^2x_t^{(1)},\quad x_t\otimes x_t=x_t^{(2)}\otimes x_t^{(1)}+x_t^{(2)}\otimes(x_t^{(2)}-x_t^{(1)}),\quad x_t\otimes\varepsilon_{t+1}=x_t^{(2)}\otimes\varepsilon_{t+1},$$

$$x_t\otimes x_t\otimes x_t=x_t^{(1)}\otimes x_t^{(1)}\otimes x_t^{(1)},\ x_t\otimes x_t\otimes\varepsilon_{t+1}=x_t^{(1)}\otimes x_t^{(1)}\otimes\varepsilon_{t+1},\ x_t\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1}=x_t^{(1)}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1},\ ^{6} \qquad \text{(B.2)}$$

where the superscript $^{(i)}$ denotes variables solved to i$^{\text{th}}$ accuracy and $x_t^{(i)}=\Lambda\omega_t^{(i)}$. The Dynare toolbox (Adjemian et al. (2014)) implements a pruned version of the third-order solution in which product terms in equation (B.1) are replaced by their third-order accurate equivalents stated in (B.2):

$$\omega_{t+1}=F_0\xi^2+F_1x_t+F_{1\xi}\xi^2x_t^{(1)}+(F_2+F_{2\xi}\xi^2)\varepsilon_{t+1}+F_{11}\{x_t^{(2)}\otimes x_t^{(1)}+x_t^{(1)}\otimes(x_t^{(2)}-x_t^{(1)})\}+F_{12}x_t^{(2)}\otimes\varepsilon_{t+1}+F_{22}\varepsilon_{t+1}\otimes\varepsilon_{t+1}+...$$

$$F_{111}x_t^{(1)}\otimes x_t^{(1)}\otimes x_t^{(1)}+F_{112}x_t^{(1)}\otimes x_t^{(1)}\otimes\varepsilon_{t+1}+F_{122}x_t^{(1)}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1}+F_{222}\varepsilon_{t+1}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1}. \qquad \text{(B.3)}$$

(This pruned third-order solution was also proposed by Kollmann (2004).) The dynamics of the first- and second-order approximated quantities is governed by (3) and (4) in the main text, restated here for convenience:

---

[6]For variable a_t we can write $a_t=a_t^{(1)}+R^{(2)}$ and $a_t=a_t^{(2)}+R^{(3)}$, where $R^{(n)}$ contains terms of order n or higher in deviations from the steady state. The product a_tb_t can thus be expressed as $a_tb_t=(a_t^{(1)}+a_t^{(2)}-a_t^{(1)}+R^{(3)})(b_t^{(1)}+b_t^{(2)}-b_t^{(1)}+R^{(3)})=a_t^{(1)}b_t^{(2)}+(a_t^{(2)}-a_t^{(1)})b_t^{(1)}+R^{(4)}$; hence, $(a_tb_t)^{(3)}=a_t^{(1)}b_t^{(2)}+(a_t^{(2)}-a_t^{(1)})b_t^{(1)}$. (Note that $a_t^{(2)}-a_t^{(1)}=R^{(2)}$, and hence $(a_t^{(2)}-a_t^{(1)})(b_t^{(2)}-b_t^{(1)})=R^{(4)}$.) The same logic shows that $(a_tb_tc_t)^{(3)}=a_t^{(1)}b_t^{(1)}c_t^{(1)}$.

$$\omega^{(1)}_{t+1}=F_1x^{(1)}_t+F_2\varepsilon_{t+1},\;\; \omega^{(2)}_{t+1}=F_0\xi^2+F_1x^{(2)}_t+F_2\varepsilon_{t+1}+F_{11}x^{(1)}_t\otimes x^{(1)}_t+F_{12}x^{(1)}_t\otimes\varepsilon_{t+1}+F_{22}\varepsilon_{t+1}\otimes\varepsilon_{t+1}. \qquad \text{(B.4)}$$

The moving average representation of the third-order pruned solution (B.3) depends on first-, second and third-order terms in exogenous innovations (ε), but not on higher-order terms. The third-order pruned system (B.3) is stationary if the first-order system is stationary.

To allow observation equation inversion, I replace squares and cubes of ε_{t+1} in (B.3) by their expected values. This gives the '*modified*' decision rule

$$\omega_{t+1}=F_0\xi^2+F_1x_t+F_{1\xi}\xi^2x^{(1)}_t+(F_2+F_{2\xi}\xi^2)\varepsilon_{t+1}+F_{11}\{x^{(2)}_t\otimes x^{(1)}_t+x^{(1)}_t\otimes(x^{(2)}_t-x^{(1)}_t)\}+F_{12}x^{(2)}_t\otimes\varepsilon_{t+1}+F_{22}E(\varepsilon_{t+1}\otimes\varepsilon_{t+1})+...$$
$$F_{111}x^{(1)}_t\otimes x^{(1)}_t\otimes x^{(1)}_t+F_{112}x^{(1)}_t\otimes x^{(1)}_t\otimes\varepsilon_{t+1}+F_{122}x^{(1)}_t\otimes E(\varepsilon_{t+1}\otimes\varepsilon_{t+1}). \qquad \text{(B.5)}$$

Note that $E(\varepsilon_{t+1}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1})=0$, because ε_{t+1} is normally distributed. The subsequent discussion assumes that (B.5) is the **true** data generating process.

Assume that the econometrician observes a vector z_{t+1} comprising m elements of the vector ω_{t+1} (recall that m is the number of exogenous innovations). Thus, the observation equation is $z_{t+1}=Q\cdot\omega_{t+1}$, where Q is an mxn selection matrix. Substitution of equation (B.5) into the observation equation gives $z_{t+1}=\gamma_t+\lambda_t\varepsilon_{t+1}$, where

$$\gamma_t\equiv Q\cdot[F_0\xi^2+F_1x_t+F_{1\xi}\xi^2x^{(1)}_t+F_{11}\{x^{(2)}_t\otimes x^{(1)}_t+x^{(1)}_t\otimes(x^{(2)}_t-x^{(1)}_t)\}+F_{22}E(\varepsilon_{t+1}\otimes\varepsilon_{t+1})+F_{111}x^{(1)}_t\otimes x^{(1)}_t\otimes x^{(1)}_t+F_{122}x^{(1)}_t\otimes E(\varepsilon_{t+1}\otimes\varepsilon_{t+1})]$$

and λ_t is an mxm matrix such that $\lambda_t\varepsilon_{t+1}\equiv Q\cdot[(F_2+F_{2\xi}\xi^2)\varepsilon_{t+1}+F_{12}x^{(2)}_t\otimes\varepsilon_{t+1}+F_{112}x^{(1)}_t\otimes x^{(1)}_t\otimes\varepsilon_{t+1}]$. Provided λ_t is non-singular, we thus have:

$$\varepsilon_{t+1}=\lambda_t^{-1}(z_{t+1}-\gamma_t)\,. \qquad \text{(B.6)}$$

Given the initial states $x^{(1)}_0, x^{(2)}_0, x_0$ and data $\{z_t\}^T_{t=1}$ one can recursively extract the innovations $\{\varepsilon_t\}^T_{t=1}$ using (B.4),(B.5) and (B.6). The log likelihood of the data (conditional on $x^{(1)}_0, x^{(2)}_0, x_0$) is:

$$\ln L(\{z_t\}^T_{t=1}|\, x^{(1)}_0, x^{(2)}_0, x_0)=-\tfrac{mT}{2}\ln(2\pi)-\tfrac{T}{2}\ln|\xi^2\Sigma_\varepsilon|-\tfrac{1}{2}\sum\nolimits_{t=1}^{T}\{\varepsilon_t'\,(\xi^2\Sigma_\varepsilon)^{-1}\varepsilon_t-\ln|\lambda_{t-1}|\,. \qquad \text{(B.7)}$$

Structural model parameters (and the initial states) can be estimated by maximizing this function.

Illustration: RBC model, approximated to third-order

I compute a third-order approximation of the RBC model described in the main text. Both the 'small shocks' variant of that model, and the 'big shocks' variant are considered. Table b1 documents that decision rule (B.3) and the modified decision rule (B.5) are (essentially) indistinguishable. An identical sequence of random exogenous innovations of length T=500,000 was fed into (B.3) and into (B.5). Table b1 shows that the resulting time series of endogenous variables are almost perfectly correlated across (B.3) and (B.5), and that they have (essentially) the same standard deviations. This holds both for levels and for first differences of logged simulated endogenous variables.

Table b2 reports predicted standard deviations of first-, second- and third-order approximated variables (log levels and log first differences). In the 'big shocks' RBC model variant, GDP, investment and capital are noticeably more volatile under a third-order approximation than under first- or second-order approximations (see Panel (b)).

Finally, I estimate the model parameters using simulated time series, by maximizing the likelihood function (B.7). As for the Monte Carlo described in the main text, I generated 30 simulation runs of 100 periods each.[7] In computing the sample likelihood, I *assume* that the initial states $x_0^{(1)}, x_0^{(2)}, x_0$ equal their unconditional mean. The first 10 periods in each simulation run are used as a training sample. Table b3 reports the median, mean and standard deviation of the estimated model parameters across the 30 simulation runs, for the 'small shocks' model variant (Columns (1)-(3)) and for the 'big shocks' variant (Cols. (4)-(6)). As for the second-order accurate model discussed in the main text, most model parameters are tightly estimated.

References

Adjemian, S., H. Bastani, M. Juillard, F. Mihoubi, G. Perendia, J. Pfeifer, M. Ratto, S. Villemot, 2014. Dynare: reference manual, Version 4.4.3., Working Paper, CEPREMAP.

Kollmann, R., 2004. Solving Non-Linear Rational Expectations Models: Approximations Based on Taylor Expansions, WP, University of Paris XII.

[7]To eliminate the influence of initial conditions, the model was simulated over 5100 periods; the first 5000 periods were discarded.

Table b1. Third-order approximated RBC model: correlations across time series generated by decision rule (B.3) [ω] and time series generated by 'modified' decision rule (B.5) [ω^{mod}]

	Y	*C*	*I*	*N*	*K*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
(a) Model variant with 'small shocks' ($\sigma_\theta=\sigma_G=\sigma_\psi=1\%, \sigma_\lambda=0.025\%$)					
Correlations between ω and $\omega^{non-\text{mod}}$					
Levels	1.0000	1.0000	1.0000	1.0000	1.0000
First differences	1.0000	1.0000	0.9999	1.0000	1.0000
Relative standard deviations: std(ω)/std(ω^{mod})					
Levels	1.0000	1.0000	1.0000	1.0000	1.0000
First differences	1.0000	1.0000	0.9996	1.0000	1.0000
Relative standard deviation of difference between decision rules: std($\omega - \omega^{\text{mod}}$)/std($\omega^{\text{mod}}$)					
Levels	0.0006	0.0001	0.0030	0.0003	0.0001
First differences	0.0039	0.0012	0.0169	0.0033	0.0020
(b) Model variant with 'big shocks' ($\sigma_\theta=\sigma_G=\sigma_\psi=5\%, \sigma_\lambda=0.125\%$)					
Correlations between ω and $\omega^{non-\text{mod}}$					
Levels	1.0000	1.0000	0.9999	1.0000	1.0000
First differences	0.9998	1.0000	0.9961	0.9999	1.0000
Relative standard deviations: std(ω)/std(ω^{mod})					
Levels	1.0000	1.0000	0.9994	1.0000	1.0000
First differences	0.9998	1.0000	0.9910	1.0000	1.0000
Relative standard deviation of difference between decision rules: std($\omega - \omega^{\text{mod}}$)/std($\omega^{\text{mod}}$)					
Levels	0.0024	0.0005	0.0155	0.0015	0.0007
First differences	0.0179	0.0065	0.0896	0.0168	0.0101

Note: Correlations of simulated time series (of variables listed above Cols. (1)-(5)) generated by decision rule (B.3) and by the 'modified' decision rule (B.5) are reported, as well as the relative standard deviation of these two sets of time series. The statistics are reported for log levels and for log first differences of endogenous variables. *Y*: GDP; *C*: consumption; *I*: gross investment; *N*: hours worked; *K*: capital stock. Correlations greater than 0.99995 are reported as 1.0000. Reported statistics are based on one simulation run of 500,000 periods.

Table b2. RBC model: predicted standard deviations (in %). Comparison between 1st order, 2nd order and 3rd order accurate model solutions

	Y	*C*	*I*	*N*	*K*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
(a) Model variant with 'small shocks' (σ_θ=σ_G=σ_ψ=1%, σ_λ=0.025%)					
Variables in levels					
1st order	3.34	1.57	10.43	9.68	7.59
2nd order	3.34	1.57	10.43	9.68	7.59
3rd order	3.36	1.57	10.48	9.67	7.62
First-differenced variables					
1st order	0.67	0.17	2.60	1.13	0.18
2nd order	0.67	0.17	2.60	1.13	0.18
3rd order	0.67	0.17	2.61	1.13	0.18
(b) Model variant with 'big shocks' (σ_θ=σ_G=σ_ψ=5%, σ_λ=0.125%)					
Variables in levels					
1st order	16.72	7.83	52.14	48.39	37.95
2nd order	17.11	7.83	52.97	48.67	38.21
3rd order	19.70	7.83	59.36	47.90	42.86
First-differenced variables					
1st order	3.33	0.86	12.98	5.66	0.91
2nd order	3.41	0.87	13.37	5.77	0.92
3rd order	3.80	0.85	14.46	5.87	0.96

Note: Standard deviations (in %) of simulated variables (listed above Cols. (1)-(5)) are shown for the RBC model. Rows labeled '1st order', '2nd order' and '3rd order' show standard deviations predicted by the first-, second- and third- order accurate model solutions, respectively. The statistics are reported for log levels and for log first differences of endogenous variables. *Y*: GDP; *C*: consumption; *I*: gross investment; *N*: hours worked; *K*: capital stock. All statistics are computed using one simulation run of 500,000 periods.

Table b3. Monte Carlo: parameter estimates for <u>third-order</u> approximated RBC model

Parameter	Model variant with 'small shocks'			Model variant with 'big shocks'		
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
	Median	**Mean**	**Std**	**Median**	**Mean**	**Std**
σ	10.67	11.17	2.74	10.95	12.42	3.89
η	0.31	0.43	0.47	0.45	0.69	0.69
ρ_θ	0.99	0.99	0.003	0.99	0.99	0.003
ρ_G	0.98	0.98	0.01	0.98	0.98	0.03
ρ_ψ	0.99	0.99	0.01	0.99	0.98	0.01
ρ_λ	0.96	0.95	0.05	0.98	0.96	0.05
s_θ (%)	0.99	1.00	0.07	4.97	4.97	0.34
s_G (%)	0.98	0.98	0.09	4.77	4.88	0.59
s_ψ (%)	0.97	1.13	0.40	5.34	6.86	3.24
s_λ (%)	0.035	0.042	0.025	0.18	0.19	0.09

Note: The Table summarizes parameters estimates across 30 simulation runs of 100 periods. Cols. labelled 'Median', 'Mean' and 'Std' report the median, mean and standard deviation of estimated parameters (listed in left-most column) across the 30 runs. Cols. labelled (1)-(3): 'small shocks' model variant. Cols. (4)-(6): 'big shocks' model variant.

The ***true*** parameter values are: σ=10, η=0.25, $\rho_\theta=\rho_G=\rho_\psi=\rho_\lambda=0.99$. True standard deviations of exogenous innovations in 'small shocks' model variant: $s_\theta=s_G=s_\psi=1\%$, $s_\lambda=0.025\%$. 'Big shocks' variant: $s_\theta=s_G=s_\psi=5\%$, $\sigma_\lambda=0.125\%$.