Work in progress

Tractable Estimation of Non-Linear DSGE Models Using Observation Equation Inversion

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

First version: February 2, 2015
This version: July 7, 2016

This paper presents a tractable approach for computing the likelihood function of non-linear Dynamic Stochastic General Equilibrium (DSGE) models that are solved using second- and third order accurate approximations. The method assumes that the number of exogenous shocks equals the number of observables. For given initial states, exogenous innovations are computed recursively by inverting the observation equation. It is hence easy to compute the sample likelihood function. Initial states and model parameters can be estimated by maximizing the likelihood function. Numerical examples suggest that the method provides reliable estimates of model parameters, even for highly non-linear economies with big shocks. By contrast to particle filters, no stochastic simulations are needed to compute the likelihood function. The method here is, hence, much faster and it is thus suitable for the estimation of medium-scale non-linear models.

Keywords: Likelihood-based estimation of non-linear DSGE models, higher-order approximations, pruning, latent state variables.
JEL codes: C63, C68, E37

- -

[(*)] I am very grateful to Tom Holden, Matteo Iacoviello, Johannes Pfeifer, Raf Wouters and Chris Sims for useful suggestions, and to conference participants at the National Bank of Belgium for helpful comments. Special thanks are also due to Johannes Pfeifer for advice about the Dynare software. The research leading to these results has received funding from Académie Universitaire Wallonie-Bruxelles (Action de recherche concertée, grant ARC-AUWB/2010-15/ULB-11) and from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 612796, Project MACFINROBODS ('Integrated Macro-Financial Modelling for Robust Policy Design'). Address: ECARES, CP 114, ULB; 50 Av. F. Roosevelt, B-1050 Brussels, Belgium. robert_kollmann@yahoo.com

1. Introduction

During the last three decades, Dynamic Stochastic General Equilibrium (DSGE) models have become the workhorse of macroeconomic research. These models are also invaluable tools for economic policy analysis and forecasting. Due to their complexity, numerical approximations are required to solve DSGE models. The bulk of DSGE-based analysis uses linear approximations. A fast growing recent literature has taken *linearized* DSGE models to the data, using likelihood-based methods, building on contributions by i.a. Kim (2000), Schorfheide (2000) and Otrok (2001). Linearity (in state variables) greatly facilitates model estimation, as it allows to use the standard Kalman filter to infer latent variables and to compute sample likelihood functions based on prediction error decompositions. However, linear approximations are inadequate for models with big shocks, and they cannot capture the effect of risk on economic decisions and welfare. Non-linear approximations are thus, for example, needed for welfare calculations in stochastic models, or for studying asset pricing and non-linearities due to financial frictions and constraints.

Recent research has begun to estimate *non-linear* DSGE models. That work has mainly used particle filters, i.e. filters that infer latent states using Monte Carlo methods (see Fernández-Villaverde and Rubio-Ramírez (2007) and An and Schorfheide (2007) for early applications). Particle filters are slow computationally, which limits their use to small models.

This paper presents a tractable estimation method for non-linear DSGE models. The method assumes that the number of exogenous shocks equals the number of observables (data variables). For given initial values of the state variables, one then recursively infer the exogenous innovations by inverting the observation equation. This method also produces estimates of the trajectories of all latent state variables. It thus allows to compute the sample likelihood function without using a simulations-based filter. The paper uses this idea to estimate DSGE models that are solved by second- or third- order Taylor expansions of the decision rules around a deterministic steady state.[1] A challenge for observation equation inversion is that approximating decision rules include powers of innovations to exogenous variables: given the state variables realized at date t-1, multiple date t exogenous innovations are consistent with the period t observables. To overcome this problem, I consider truncated date t decision rules from which powers of the exogenous innovations have been suppressed. The truncated decision rules are,

[1] Guerrieri and Iacoviello (2014) and Deák, Holden and Mele (2015) also discuss estimation of DSGE models via observation equation inversion. These authors do not consider second- or third-order approximated models (that are the focus of the present paper).

hence, *linear* in date t exogenous innovations (the coefficients of those innovations are, however, functions of lagged state variables), but non-linear in lagged state variables. Hence, it is straightforward to invert the observation equation. I present examples of DSGE models for which the truncated decision rules are observationally indistinguishable from decision rules that include higher-order powers of contemporaneous exogenous innovations. Numerical examples show that the estimation method here is fast and accurate, even for models with strong curvature and big shocks. The method requires a guess about the initial state vector. In the numerical example provided here, I use a training sample to reduce the effect of the initial state vector on the inferred exogenous innovations. It is also possible to estimate the initial states, by maximizing the likelihood function with respect to those states.

The numerical DSGE-model solution technique (second- or third-order approximations) considered here is the most tractable non-linear solution method for medium-scale models, and it has thus widely been used in macroeconomics (see Kollmann (2002) and Kollmann et al. (2011) for detailed references).[2] When simulating higher-order approximated models, it is common to use the 'pruning' scheme of Kim et al. (2008), under which second-order terms are replaced by the products of the linearized solution (third-order terms are replaced by products of linearized and second-order approximated variables). Unless the pruning algorithm is used, second-order (and third-order) approximated models often generate exploding simulated paths. Pruning is therefore crucial for applied work based on second-order (and third-order) approximated models. The paper here also uses the pruning scheme.

The present paper is complementary to Kollmann (2015a,b) who developed a tractable *deterministic* filter for pruned second-order and third-order approximated DSGE models. That filter exploits the fact that a pruned second-order accurate system is *linear* in an extended state vector that consists of variables solved to second- and first-order accuracy, and of products of first-order accurate variables.[3] Linearity in that extended state vector allows closed-form determination of the state vector's (conditional) mean and variance. Kollmann (2015a,b) applies the linear updating rule of the standard Kalman filter to the pruned state equation.[4] The filter based on the linear updating rule is more accurate than particle filters (unless a very large

[2]Computer code that solves higher-order approximated models is freely available; see, e.g., Sims (2000), Schmitt-Grohé and Uribe (2004), and Adjemian et al. (2014).

[3]The pruned *third*-order accurate system is linear in a state vector that consists of variables solved to third-, second- and first-order, of products of first-order accurate variables and of second-order accurate variables (see Appendix).

[4]Ivashchenko (2014) also develops a Kalman filter for second-order approximated models, but his method does not use the pruning scheme.

number of particles is used), and it too is much faster than particle filters. A practical issue for the Kollmann (2015a,b) filter is that the dimension of the pruned state space increases rapidly when model size increases. As there is no need to compute the moments of the pruned state vector, the observation equation inversion method developed in the present paper is faster than the Kollmann (2015a,b) filter. However, the method here requires that the number of exogenous shocks equals the number of observables, while the Kollmann (2015a,b) method allows to handle situations in which the number of shocks exceeds the number of observables.

2. Model format

2.1. Model and second-order accurate solution

Standard DSGE models can be expressed as:

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

where E_t is the mathematical expectation conditional on date *t* information; $G: R^{2n+m} \to R^n$ is a function, and Ω_t is an *n*x1 vector of endogenous and exogenous variables known at *t*; ε_{t+1} is an *m*x1 vector of serially independent innovations to exogenous variables. In what follows, ε_t is Gaussian: $\varepsilon_t \sim N(0, \xi^2 \Sigma_\varepsilon)$, where ξ is a scalar that indexes the size of shocks. The model solution is a "policy function" ("decision rule") $\Omega_{t+1} = F(X_t, \varepsilon_{t+1}, \xi)$, where X_t is a vector consisting of the 'state' variables of the economy, i.e. of the predetermined endogenous variables (e.g., the physical capital stock) and the exogenous variables included in the vector Ω_t, i.e. $X_t = \Lambda \Omega_t$, where Λ is a matrix of zeros and ones that selects the 'state' variables among the elements of Ω_t. The policy function has to be such that such that $E_t G(F(\Lambda \Omega_t, \varepsilon_{t+1}, \xi), \Omega_t, \varepsilon_{t+1}) = 0 \ \forall \Omega_t$. See, e.g., Sims (2000) and Schmitt-Grohé and Uribe (2004). Following these authors, this paper focuses on Taylor series expansions of the policy function around a deterministic stead state, i.e. around $\xi = 0$ and a vector Ω such that $\Omega = F(\Omega, \Omega, 0)$. Let $\omega_t \equiv \Omega_t - \Omega$ and $x_t \equiv X_t - X$ denote deviations from steady state. The second-order accurate model solution has the form

$$\omega_{t+1} = F_0 \xi^2 + F_1 x_t + F_2 \varepsilon_{t+1} + F_{11} x_t \otimes x_t + F_{12} x_t \otimes \varepsilon_{t+1} + F_{22} \varepsilon_{t+1} \otimes \varepsilon_{t+1}, \text{ with } x_t = \Lambda \omega_t. \tag{2}$$

Here $\otimes$ denotes the Kronecker product. $F_0, F_1, F_2, F_{11}, F_{12}, F_{22}$ are matrices that are functions of the structural model parameters, but do not depend on the scale of shocks (ξ); see, Schmitt-Grohé and Uribe (2004). The first-order accurate (linearized) model solution is:

$$\omega_{t+1}^{(1)}=F_1 x_t^{(1)}+F_2\varepsilon_{t+1}. \quad (3)$$

The superscript $^{(1)}$ denotes a variable solved to first-order accuracy.

2.2. The pruned second-order accurate system

When simulating second-order accurate models, it is common to use the 'pruning' scheme of Kim, Kim, Schaumburg and Sims (2008), under which products of state variables are replaced by products of variables approximated to lower order, i.e. $x_t \otimes x_t$ 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. With pruning, the solution (2) is thus replaced by:

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

Note that $x_t \otimes x_t = x_t^{(1)} \otimes x_t^{(1)}$ and $x_t \otimes \varepsilon_{t+1} = x_t^{(1)} \otimes \varepsilon_{t+1}$ holds, up to second-order accuracy. [5] Thus, (4) is a valid second-order accurate solution. The motivation for pruning is that, in repeated applications of (2), third and higher-order terms of state variables appear; e.g., when ω_{t+1} is quadratic in ω_t, then ω_{t+2} is quartic in ω_t; pruning removes these higher-order terms. The non-pruned system (2) has extraneous steady states (not present in the original model)--some of these steady states mark transitions to unstable behavior. Large shocks can thus move the model into an unstable region. Pruning overcomes this problem. If the first-order solution is stable, then the pruned second-order solution (4) too is stable.

The following exposition focuses on the estimation of the pruned second-order system. Extension of the method to third (or higher) order approximated system is straightforward. See the Appendix.

2.3. Inferring the exogenous innovations from observables

Assume that, at date t+1, the econometrician knows the state vectors $x_t^{(1)}, x_t^{(2)}$ and observes 'm' of the elements of the vector $\omega_{t+1}^{(2)}$ (or 'm' linear combinations of the elements of $\omega_{t+1}^{(2)}$), i.e. the vector of observables is

[5] For any variable a_t we can write $a_t=a_t^{(1)}+R^{(2)}$ where $R^{(n)}$ contains terms of order *n* or higher in deviations from the steady state. Thus, $a_t b_t=(a_t^{(1)}+R^{(2)})(b_t^{(1)}+R^{(2)})=a_t^{(1)}b_t^{(1)}+a_t^{(1)}R^{(2)}+b_t^{(1)}R^{(2)}+R^{(4)}=a_t^{(1)}b_t^{(1)}+R^{(3)}$ and hence $(a_t b_t)^{(2)}=a_t^{(1)}b_t^{(1)}$. This logic implies $(x_t \otimes x_t)^{(2)}=x_t^{(1)}\otimes x_t^{(1)}$ and $(x_t \otimes \varepsilon_{t+1})^{(2)}=x_t^{(1)}\otimes\varepsilon_{t+1}$. See, e.g., Kollmann (2004; 2015a,b), Lombardo and Sutherland (2007) and Lombardo and Uhlig (2015).

$$z_{t+1} \equiv Q\omega_{t+1}^{(2)}, \tag{5}$$

where Q is a known matrix of dimension mxn. (Recall that 'm' and 'n' are the number of exogenous innovations and of endogenous variables, respectively.) (4) implies:

$$z_{t+1} = Q(F_0\xi^2 + F_1 x_t^{(2)} + F_{11} x_t^{(1)} \otimes x_t^{(1)}) + Q(F_2\varepsilon_{t+1} + F_{12} x_t^{(1)} \otimes \varepsilon_{t+1} + F_{22}\varepsilon_{t+1} \otimes \varepsilon_{t+1}). \tag{6}$$

As the right-hand side of (6) includes squares of ε_{t+1}, one cannot uniquely solve (6) for the unknown vector of innovations ε_{t+1}, given the data z_{t+1} and the states $x_t^{(1)}, x_t^{(2)}$. There is no tractable method for computing all vectors ε_{t+1} that solve (6) when m is larger than 3 or 4. [6] In what follows, I replace the term $F_{22}\varepsilon_{t+1} \otimes \varepsilon_{t+1}$ by its expected value $F_{22}E(\varepsilon_{t+1} \otimes \varepsilon_{t+1})$ in (6). [7] Thus, equations (4),(6) are replaced by

$$\omega_{t+1}^{(2)} = F_0\xi^2 + F_1 x_t^{(2)} + F_2\varepsilon_{t+1} + F_{11} x_t^{(1)} \otimes x_t^{(1)} + F_{12} x_t^{(1)} \otimes \varepsilon_{t+1} + F_{22}E(\varepsilon_{t+1} \otimes \varepsilon_{t+1}) \text{ and} \tag{7}$$

$$z_{t+1} = Q(F_0\xi^2 + F_1 x_t^{(2)} + F_{11} x_t^{(1)} \otimes x_t^{(1)} + F_{22}E(\varepsilon_{t+1} \otimes \varepsilon_{t+1})) + Q(F_2\varepsilon_{t+1} + F_{12} x_t^{(1)} \otimes \varepsilon_{t+1}). \tag{8}$$

Numerical experiments, for a range of models, show that this modification only has a minor effect on the model dynamics; see below.[8] Note that the observation equation (8) is *linear* in ε_{t+1}: $z_{t+1} = \gamma_t + \lambda_t\varepsilon_{t+1}$, where $\gamma_t \equiv Q(F_0\xi^2 + F_1 x_t^{(2)} + F_{11} x_t^{(1)} \otimes x_t^{(1)} + F_{22}E(\varepsilon_{t+1} \otimes \varepsilon_{t+1}))$, while λ_t is an (n *x* n) matrix such that $\lambda_t\varepsilon_{t+1} \equiv Q(F_2\varepsilon_{t+1} + F_{12} x_t^{(1)} \otimes \varepsilon_{t+1})$. If λ_t is non-singular, one can solve (7) for the vector of exogenous innovations: $\varepsilon_{t+1} = \lambda_t^{-1}(z_{t+1} - \gamma_t)$.

2.3. Sample likelihood

Assume that (3),(7),(8) is the true data generating process. Given the initial state $x_0^{(1)}, x_0^{(2)}$ and data $\{z_t\}_{t=1}^T$ one can recursively compute the innovations $\{\varepsilon_t\}_{t=1}^T$ and the states $\{\omega_t^{(i)}\}_{t=1}^T$ for *i*=1,2 using (3),(6),(8). The log likelihood of the data, conditional on $x_0^{(1)}, x_0^{(2)}$ is:

---

[6] If it were possible to compute all ε_{t+1} that solve (6) then one could pick the most likely of these vectors (i.e. the one for which $\varepsilon_{t+1}'(\Sigma_\varepsilon)^{-1}\varepsilon_{t+1}$ is smallest) as an estimate of the true ε_{t+1}.

[7] I thank Chris Sims for suggesting this approach. In earlier versions of the algorithm, I dropped the term $F_{22}\varepsilon_{t+1} \otimes \varepsilon_{t+1}$ from (6), which likewise produces an observation equation that is linear in ε_{t+1}. The two approaches produce very similar results. An advantage of replacing $F_{22}\varepsilon_{t+1} \otimes \varepsilon_{t+1}$ by its expected value $F_{22}E(\varepsilon_{t+1} \otimes \varepsilon_{t+1})$ is that the unconditional mean of the 'modified' process equals that of the original process.

[8] Feeding a given sequence of exogenous innovations $\{\varepsilon_t\}$ into the system (3),(4) and into the system (3),(7) generates paths for the endogenous variables that are almost perfectly correlated across the two systems.

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

One can estimate structural model parameters (and the initial state), by maximizing the likelihood function with respect to the parameters.

3. Application: basic RBC model

I illustrate the method for a standard RBC model. Assume a closed economy with a representative infinitely-lived household whose date t expected lifetime utility V_t is 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, at t, respectively. $\sigma>0$ and $\eta>0$ are the risk aversion coefficient and the inverse of the (Frisch) labor supply elasticity. $\lambda_t\beta$ is the subjective discount factor between periods t and $t+1$. $\psi_t>0$ and $\lambda_t>0$ are exogenous preference shocks: ψ_t is a labor supply shock, while λ_t is a shock to the subjective discount factor. ψ_t and λ_t equal unity in the deterministic steady state. The household maximizes expected lifetime utility subject to the period t resource constraint

$$C_t+I_t+G_t=Y_t,$$

where Y_t and I_t are output, gross investment and exogenous government consumption, respectively. The production function is

$$Y_t=\theta_t K_t^\alpha N_t^{1-\alpha}$$

where K_t is the beginning-of-period t capital stock, and $\theta_t>0$ is exogenous total factor productivity (TFP). The law of motion of the capital stock is

$$K_{t+1}=(1-\delta)K_t+I_t.$$

$0<\alpha,\delta<1$ are the capital share and the capital depreciation rate, respectively. The household's first-order conditions are:

$$\lambda_t E_t\beta(C_{t+1}/C_t)^{-\sigma}(\theta_{t+1}\alpha K_{t+1}^{\alpha-1}N_{t+1}^{1-\alpha}+1-\delta)=1, \quad C_t^{-\sigma}(1-\alpha)\theta_t K_t^\alpha N_t^{-\alpha}=\psi_t N_t^{1/\eta}.$$

The forcing variables follow independent autoregressive processes:

$\ln(\theta_t/\theta)=\rho_\theta\ln(\theta_{t-1}/\theta)+\varepsilon_{\theta,t}$, $\ln(G_t/G)=\rho_G\ln(G_{t-1}/G)+\varepsilon_{G,t}$, $\ln(\psi_t)=\rho_\psi\ln(\psi_{t-1})+\varepsilon_{\psi,t}$, $\ln(\lambda_t)=\rho_\lambda\ln(\lambda_{t-1})+\varepsilon_{\lambda,t}$,

with $0<\rho_\theta,\rho_G,\rho_\Psi,\rho_\lambda<1$, where θ and G are steady state TFP and steady state government purchases. $\varepsilon_{\theta,t},\varepsilon_{G,t},\varepsilon_{\psi,t}$ and $\varepsilon_{\lambda,t}$ are normal i.i.d. white noises with standard deviations $\sigma_\theta,\sigma_G,\sigma_\psi$ and σ_λ.

The numerical simulations discussed below assume $\beta=0.99,\eta=0.25,\alpha=0.3,\delta=0.025$; the steady state ratio of government purchases to GDP (G/Y) is set at 0.2. The autocorrelations of all forcing variables is set at $\rho_\theta=\rho_G=\rho_\psi=\rho_\lambda=0.98$, i.e. the exogenous variables undergo persistent fluctuations. Parameter values in that range are standard in (quarterly) macro models. The risk aversion coefficient is set at a high value, $\sigma=10$, so that the model has enough curvature to produce non-negligible differences between the second- and third-order model approximations and the linearized model. In all model variants, I set the scalar ξ (that indexes the size of shocks; see Section 2.1) at $\xi=1$. One model variant, referred to as the 'small shocks' variant, assumes $\sigma_\theta=\sigma_G=\sigma_\psi=1\%$ and $\sigma_\lambda=0.025\%$. Those shock sizes (i.e. rate of time preference shocks 40-times smaller than the other shocks) ensure that each shock accounts for a non-negligible share of the variance of the endogenous variables (see Table 1). That 'small shocks' calibration is standard in the RBC literature, and it implies that the volatility of the endogenous variables in the model is roughly consistent with the empirical volatility. In the 'small shocks' variant, the behavior of endogenous variables predicted by the second- and third-order approximated model is broadly similar to that predicted by the linearized model (see the Appendix for a presentation of the third-order accurate model solution). I thus also consider model variants with much bigger shocks—in those variants, the higher-order approximated model generates predicted behavior that differs noticeably from behavior in the first-order approximated model. In one model variant, the standard deviations or shocks are 5 times greater than in the 'small shocks' variant ($\sigma_\theta=\sigma_G=\sigma_\psi=5\%,\sigma_\lambda=0.125\%$); I also consider a variant in which the standard deviation of exogenous innovations is 10 time greater ($\sigma_\theta=\sigma_G=\sigma_\psi=10\%,\sigma_\lambda=0.250\%$). I refer to these model variants as the 'big shocks' variant and the 'very big shocks' variant, respectively.

I solve the model using the Dynare toolbox (Adjemian et al. (2014)). The Taylor expansions of the model equations are taken with respect to logs of all variables.

3.1. Predicted moments

Table 1 reports predicted standard deviations of GDP, consumption, investment, hours worked and the capital stock. The predicted moments are shown for variables in logged levels, as well as for first-differenced logged variables. In the 'small shocks' variant, the order of approximation does not matter much for predicted behavior. For example, the predicted standard deviation of GDP is 3.00% (3.09%) [3.04%] under the first- (second-) [third-] order accurate model approximation.

By contrast, in the model variants with 'big' and with 'very big' shocks, the second- and third-order approximations generate markedly greater volatility of the endogenous variables than the linear approximation. In the 'big shocks' ['very big shocks'] variant the predicted volatility of GDP rises by one quarter [doubles] when the third-order approximation is used, instead of the linear approximation.

Under the linear approximation, the unconditional means of all endogenous variables equals their values in the deterministic steady state. Under the second- and third-order approximations, the unconditional means can differ from the deterministic steady state (unconditional means implied by the second and third-order approximations are identical). In the 'small shocks' variant, the mean of capital stock and mean GDP exceeds steady state values by 0.81% and 0.25%, respectively. This is due to precautionary saving that is captured by the second-order approximation. In the 'big shocks' ['very big shocks'] model variant, the mean capital stock and mean GDP are 20.39% and 6.26% [81.56% and 25.05%] above the deterministic steady state.

3.2. Comparing the 'truncated' versions of the third-order accurate model

Table 2 documents that the 'truncated' version of the (pruned) third-order accurate model is observationally equivalent to the 'non-truncated' version (see Appendix). The correlations of simulated time series across the truncated and non-truncated variants are very close to unity, both in levels and in first differences, and that even when shocks are very big.

3.3. Estimating structural parameters

I now evaluate the ability of the method to estimate structural model parameters, for the case of the third-order accurate model approximation. For each of the three model variants, I generated 30 simulation runs of 5100 periods (each simulation run was initiated at the unconditional means

of the state variables). I use the last 100 periods of each simulation run for estimation. As the model has four exogenous shocks, four observables are needed for estimation. First differences of log GDP, consumption, investment and hours worked are used as observables. I add equations defining the first difference of these four variables to the model equations. Dynare then identifies 9 state variables (the capital stock, and lagged values of the four exogenous variables and of GDP, consumption, investment and hours). I assume that the initial values of the states $x_0^{(1)}, x_0^{(2)}, x_0^{(3)}$ equal the unconditional means of these vectors. Although true initial values differ from the assumed values, the inferred estimates of the exogenous innovations converge fast to the true values, in most of the simulation runs. I thus use the first 10 periods of each sample period (of length T=100) as a training sample (i.e. the first 10 periods are dropped in the construction of the likelihood function). I estimate the following 10 structural parameters: the risk aversion coefficient (σ), labor supply parameter (η), as well as the autocorrelations and standard deviations of the four exogenous variables.

Panel (a) of Table 3 reports the mean, median and standard deviation of the estimated model parameters across the 30 simulation runs, for the 'small shocks' model variant (Columns (1)-(3)), the 'big shocks' variant (Cols. (4)-(6)) and the 'very big shocks' model variant (Cols. (7)-(9)). Table 3 shows that the risk aversion coefficient, the autocorrelations of the exogenous variables and the standard deviations of exogenous innovations are relatively tightly estimated: the mean and median parameter estimates (across runs) are close to the true parameter values, and the standard deviations of the parameter estimates are mostly small. The labor supply parameter η is less precisely estimated.

For the model and sample length considered here, one evaluation of the likelihood function (for a given parameter vector) takes 0.25 second for the third-order accurate model (using a PC with an Intel i7-2600 processor, 3.40Ghz). By contrast, one evaluation of the likelihood takes 0.80 second when the Kollmann (2015a,b) filter is used. Thus, the method here is markedly faster. The gain in speed is even greater when the number of states is bigger. For the two-country DSGE model with 19 state variables discussed in Kollmann (2015b), one evaluation of the likelihood function takes 1.80 second when the method here is used. Estimation of a model of this size is hence feasible. The Kalman filter method of Kollmann (2015a,b) is an order of magnitude slower.

APPENDIX: Estimation method for third-order accurate approximate model solutions

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

$$\omega_{t+1}=F_0\xi^2+(F_1+F_{1\xi}\xi^2)x_t+(F_2+F_{2\xi}\xi^2)\varepsilon_{t+1}+F_{11}x_t\otimes x_t+F_{12}x_t\otimes\varepsilon_{t+1}+F_{22}\varepsilon_{t+1}\otimes\varepsilon_{t+1}+...$$
$$F_{111}x_t\otimes x_t\otimes x_t+F_{112}x_t\otimes x_t\otimes\varepsilon_{t+1}+F_{122}x_t\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1}+F_{222}\varepsilon_{t+1}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1}, \quad (A.1)$$

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

$$(\xi^2x_t)^{(3)}=\xi^2x_t^{(1)},\quad (x_t\otimes x_t)^{(3)}=x_t^{(2)}\otimes x_t^{(1)}+x_t^{(2)}\otimes(x_t^{(2)}-x_t^{(1)}),\quad (x_t\otimes\varepsilon_{t+1})^{(3)}=x_t^{(2)}\otimes\varepsilon_{t+1},$$
$$(x_t\otimes x_t\otimes x_t)^{(3)}=x_t^{(1)}\otimes x_t^{(1)}\otimes x_t^{(1)},\ (x_t\otimes x_t\otimes\varepsilon_{t+1})^{(3)}=x_t^{(1)}\otimes x_t^{(1)}\otimes\varepsilon_{t+1},\ (x_t\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1})^{(3)}=x_t^{(1)}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1}.$$ [9]

The pruned third-order accurate system is thus given by:

$$\omega_{t+1}^{(3)}=F_0\xi^2+F_1x_t^{(3)}+F_{1\xi}\xi^2x_t^{(1)}+(F_2+F_{2\xi}\xi^2)\varepsilon_{t+1}+F_{11}\{x_t^{(2)}\otimes x_t^{(1)}+x_t^{(1)}\otimes(x_t^{(2)}-x_t^{(1)})\}+F_{12}x_t^{(2)}\otimes\varepsilon_{t+1}+F_{22}\varepsilon_{t+1}\otimes\varepsilon_{t+1}+...$$
$$F_{111}x_t^{(1)}\otimes x_t^{(1)}\otimes x_t^{(1)}+F_{112}x_t^{(1)}\otimes x_t^{(1)}\otimes\varepsilon_{t+1}+F_{122}x_t^{(1)}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1}+F_{222}\varepsilon_{t+1}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1}, \quad (A.2)$$

where the first- and second order accurate quantities $x_t^{(1)}$ and $x_t^{(2)}$ obey (3) and (4):

$$\omega_{t+1}^{(1)}=F_1x_t^{(1)}+F_2\varepsilon_{t+1}, \quad (3)$$

$$\omega_{t+1}^{(2)}=F_0\xi^2+F_1x_t^{(2)}+F_2\varepsilon_{t+1}+F_{11}x_t^{(1)}\otimes x_t^{(1)}+F_{12}x_t^{(1)}\otimes\varepsilon_{t+1}+F_{22}\varepsilon_{t+1}\otimes\varepsilon_{t+1}. \quad (4)$$

Unless the pruning algorithm is used, third-order approximated models often generate exploding simulated time paths. Pruning ensures that (A.2) is non-explosive if the first-order system (3) is stationary. To permit inversion of the observation equation, I replace squares and cubes of ε_{t+1} by their expected values, on the right-hand side of (A.2):

$$\omega_{t+1}^{(3)}=F_0\xi^2+F_1x_t^{(3)}+F_{1\xi}\xi^2x_t^{(1)}+(F_2+F_{2\xi}\xi^2)\varepsilon_{t+1}+F_{11}\{x_t^{(2)}\otimes x_t^{(1)}+x_t^{(1)}\otimes(x_t^{(2)}-x_t^{(1)})\}+F_{12}x_t^{(2)}\otimes\varepsilon_{t+1}+F_{22}E(\varepsilon_{t+1}\otimes\varepsilon_{t+1})+...$$
$$F_{111}x_t^{(1)}\otimes x_t^{(1)}\otimes x_t^{(1)}+F_{112}x_t^{(1)}\otimes x_t^{(1)}\otimes\varepsilon_{t+1}+F_{122}x_t^{(1)}\otimes E(\varepsilon_{t+1}\otimes\varepsilon_{t+1}). \quad (A.3)$$

Note that $E(\varepsilon_{t+1}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1})=0$, because ε_{t+1} is normally distributed. Assume that the true data generating process is given by (A.3),(3),(4), and that the econometrician observes 'm' of the

[9] 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 $(a_t^{(2)}-a_t^{(1)})(a_t^{(2)}-a_t^{(1)})=R^{(4)}$.) The same logic shows that $(a_tb_tc_t)^{(3)}=a_t^{(1)}b_t^{(1)}c_t^{(1)}$.

third-order accurate variables, $z_{t+1} \equiv Q\omega_{t+1}^{(3)}$, where Q is a known matrix of dimension mxn. Given initial state $x_0^{(1)}, x_0^{(2)}, x_0^{(3)}$ and data $\{z_t\}_{t=1}^T$ one can recursively compute the innovations $\{\varepsilon_t\}_{t=1}^T$ and the states $\{\omega_t^{(i)}\}_{t=1}^T$ for *i*=1,2,3 using (3),(4) and (A.2).

The log likelihood of the data, conditional on $x_0^{(1)}, x_0^{(2)}, x_0^{(3)}$ is:

$$\ln L(\{z_t\}_{t=1}^T | x_0^{(1)}, x_0^{(2)}, x_0^{(3)}) = -\tfrac{mT}{2}\ln(2\pi) - \tfrac{T}{2}\ln|\xi^2\Sigma_\varepsilon| - \tfrac{1}{2}\sum_{t=1}^T \{\varepsilon_t' (\xi^2\Sigma_\varepsilon)^{-1}\varepsilon_t - \ln|\lambda_{t-1}|, \qquad \text{(A.4)}$$

where λ_t is the (m *x* m) matrix such that $\lambda_t\varepsilon_{t+1} = (F_2 + F_{2\xi}\xi^2)\varepsilon_{t+1} + F_{12}x_t^{(2)} \otimes \varepsilon_{t+1} + F_{112}x_t^{(1)} \otimes x_t^{(1)} \otimes \varepsilon_{t+1}$.

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.

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

Deák, S., T. Holden and A. Mele, 2015. An Advanced Course On The Science and Art of DSGE Modelling. University of Surrey.

Fernández-Villaverde, J. and J. Rubio-Ramírez, 2007. Estimating Macroeconomic Models: a Likelihood Approach. Review of Economic Studies 74, 1059–1087.

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

Ivashchenko, S., 2014. DSGE Model Estimation on the Basis of Second-Order Approximation. Computational Economics 43, 73-81.

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 of Welfare and Business Cycles. Journal of Monetary Economics 49, 989-1015.

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

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

Kollmann, R., 2015b. Tractable Latent State Filtering for Non-Linear DSGE Models Using a Third- Order Approximation, Pruning and a Kalman filter, work in progress.

Lombardo, G. and A. Sutherland, 2007. Computing Second-Order Accurate Solutions for Rational Expectations Models Using Linear Solution Methods. Journal of Economic Dynamics and Control 31, 203-206.

Lombardo, G. and H. Uhlig, 2015. A Theory of Pruning, Working Paper, BIS and University of Chicago.

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.

Schorfheide, F., 2000. Loss Function-Based Evaluation of DSGE Models. Journal of Applied Econometrics 16, 645-670.

Sims, C., 2000. Second Order Accurate Solution of Discrete Time Dynamic Equilibrium Models. Working Paper, Economics Department, Princeton University

Table 1. RBC model: predicted standard deviations (in%)

	Y	*C*	*I*	*N*	*K*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
(a) Model variant with small shocks ($\sigma_\theta=\sigma_G=\sigma_\psi=0.01, \sigma_\lambda=0.00025$)					
(a.1) Variables in levels (logs)					
1st order, all shocks	3.00	1.46	10.35	10.46	7.80
1st order, just θ shock	2.08	1.37	6.21	9.43	4.58
1st order, just *G* shock	1.59	0.08	1.46	1.90	1.03
1st order, just ψ shock	1.08	0.70	3.26	0.93	2.32
1st order, just λ shock	1.68	0.22	7.60	1.55	5.81
2nd order, all shocks	3.09	1.46	10.40	10.45	7.82
3rd order, all shocks	3.04	1.46	10.45	10.44	7.86
(a.2) First-differenced variables (logs)					
1st order, all shocks	0.67	0.17	2.62	1.12	0.17
2nd order, all shocks	0.67	0.17	2.62	0.12	0.17
3rd order, all shocks	0.68	0.17	2.63	0.13	0.17
(b) Model variant with big shocks ($\sigma_\theta=\sigma_G=\sigma_\psi=0.05, \sigma_\lambda=0.00125$)					
(b.1) Variables in levels (logs)					
1st order, all shocks	14.99	7.33	51.76	52.31	39.05
2nd order, all shocks	15.89	7.32	53.87	52.29	39.07
3rd order, all shocks	18.71	7.33	60.18	51.62	44.98
(b.2) First-differenced variables (logs)					
1st order, all shocks	3.35	0.85	13.09	5.63	0.86
2nd order, all shocks	3.56	0.85	13.45	5.77	0.88
3rd order, all shocks	4.00	0.83	14.63	5.94	0.92
(c) Model variant with very big shocks ($\sigma_\theta=\sigma_G=\sigma_\psi=0.10, \sigma_\lambda=0.00250$)					
(c.1) Variables in levels (logs)					
1st order, all shocks	29.99	14.66	103.52	104.62	78.01
2nd order, all shocks	35.41	14.65	115.43	105.46	86.33
3rd order, all shocks	58.77	14.87	166.39	103.71	123.94
(c.2) First-differenced variables (logs)					
1st order, all shocks	6.71	1.70	26.19	11.27	1.72
2nd order, all shocks	7.92	1.71	29.08	12.33	1.86
3rd order, all shocks	11.54	1.63	39.34	14.80	2.21

Note: Standard deviations (std.) of logged variables (listed above Cols. (1)-(5)) are shown for the RBC model. All moments are computed based on one simulation run of 5000 periods (the run is initiated at the unconditional mean of the state variables). Rows labeled '1st order', '2nd order' and '3rd order' show standard deviations predicted by the first-, second- and third-order accurate model variants, respectively. *Y*: GDP; *C*: consumption; *I*: gross investment; *N*: hours worked; *K*: capital stock.

Table 2. RBC model, third-order approximation: correlations between variables predicted by 'truncated' and 'non-truncated' model versions

	Y	*C*	*I*	*N*	*K*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
(a) Model variant with small shocks ($\sigma_\theta=\sigma_G=\sigma_\psi=0.01, \sigma_\lambda=0.00025$)					
(a.1) Variables in levels (logs)					
All shocks	1.000	1.000	1.000	1.000	1.000
Just θ shock	1.000	1.000	1.000	1.000	1.000
Just *G* shock	1.000	1.000	1.000	1.000	1.000
Just ψ shock	1.000	1.000	1.000	1.000	1.000
Just λ shock	1.000	1.000	1.000	1.000	1.000
(a.2) First-differenced variables (logs)					
All shocks	1.000	1.000	1.000	1.000	1.000
Just θ shock	1.000	1.000	1.000	1.000	1.000
Just *G* shock	1.000	1.000	1.000	1.000	1.000
Just ψ shock	1.000	1.000	1.000	1.000	1.000
Just λ shock	1.000	1.000	1.000	1.000	1.000
(b) Model variant with big shocks ($\sigma_\theta=\sigma_G=\sigma_\psi=0.05, \sigma_\lambda=0.00125$)					
(b.1) Variables in levels (logs)					
All shocks	1.000	1.000	1.000	1.000	1.000
Just θ shock	1.000	1.000	1.000	1.000	1.000
Just *G* shock	1.000	1.000	1.000	1.000	1.000
Just ψ shock	1.000	1.000	1.000	1.000	1.000
Just λ shock	1.000	1.000	1.000	1.000	1.000
(b.2) First-differenced variables (logs)					
All shocks	1.000	1.000	0.996	1.000	1.000
Just θ shock	1.000	1.000	0.999	1.000	1.000
Just *G* shock	0.999	0.999	1.000	0.999	1.000
Just ψ shock	1.000	1.000	1.000	1.000	1.000
Just λ shock	1.000	1.000	0.998	1.000	1.000
(c) Model variant with very big shocks ($\sigma_\theta=\sigma_G=\sigma_\psi=0.10, \sigma_\lambda=0.00250$)					
(c.1) Variables in levels (logs)					
All shocks	1.000	1.000	0.999	1.000	1.000
Just θ shock	1.000	1.000	1.000	1.000	1.000
Just *G* shock	1.000	1.000	1.000	1.000	1.000
Just ψ shock	1.000	1.000	1.000	1.000	1.000
Just λ shock	1.000	1.000	1.000	1.000	1.000
(c.2) First-differenced variables (logs)					
All shocks	1.000	1.000	0.984	0.999	1.000
Just θ shock	1.000	1.000	0.995	1.000	1.000
Just *G* shock	0.998	0.998	0.998	0.998	1.000
Just ψ shock	1.000	1.000	0.998	1.000	1.000
Just λ shock	1.000	1.000	0.992	1.000	1.000

Note: Correlations between variables predicted by the 'full' and 'truncated' third-order models are reported. 'All shocks': simulations with all 4 shocks. 'Just θ shocks', 'Just *G* shocks' etc. pertain to simulations in which just one type of shock is fed into the model; the other exogenous variables are set at steady state values (model is solved assuming 4 shocks). Reported statistics are based on one simulation run of 5000 periods (the run is initiated at the unconditional mean of the state variables). *Y*: GDP; *C*: consumption; *I*: gross investment; *N*: hours worked; *K*: capital stock. Correlations greater than 0.9995 are reported as 1.000.

Table 3. RBC model: estimates of structural parameters, 30 simulation runs (100 periods)

	Model variant with 'small shocks'			Model variant with 'big shocks'			Model variant with 'very big shocks'		
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*
(a) Parameter estimates									
	Median	**Mean**	**Std**	**Median**	**Mean**	**Std**	**Median**	**Mean**	**Std**
σ	9.98	12.28	8.54	9.82	11.59	4.91	10.35	11.21	3.17
η	0.21	0.74	1.81	0.12	0.36	0.55	0.04	0.25	0.56
ρ_θ	0.98	0.98	0.003	0.98	0.98	0.014	0.98	0.98	0.016
ρ_G	0.98	0.99	0.019	0.99	0.98	0.019	0.98	0.99	0.023
ρ_ψ	0.98	0.98	0.008	0.98	0.98	0.021	0.99	0.98	0.014
ρ_λ	0.99	0.98	0.05	0.99	0.98	0.037	0.98	0.98	0.027
σ_θ (%)	1.02	1.00	0.07	5.02	5.40	2.20	10.22	11.15	3.21
σ_G (%)	0.97	0.98	0.11	4.08	4.39	1.45	8.43	10.11	5.51
σ_ψ (%)	0.97	1.36	1.34	5.01	6.11	3.33	9.85	11.22	5.18
σ_λ (%)	0.021	0.026	0.011	0.094	0.108	0.037	0.227	0.279	0.225

Note: The Table summarizes estimation results across 30 simulation runs of 100 periods each (the model is simulated over 5100 periods, the last 100 periods are used for estimation). The observables are the first differences of logged GDP, consumption, investment and hours worked.

Panel (a) reports the mean, median and standard deviation of the estimated model parameters across the 30 runs, for the 'small shocks' model variant (Columns (1)-(3)), the 'big shocks' variant (Cols. (4)-(6)) and the 'very big shocks' variant (Cols. (7)-(9)). The ***true*** values of the estimated parameters are: σ=10, η=0.25, $\rho_\theta=\rho_G=\rho_\psi=\rho_\lambda$=0.98. In the 'small shocks' model variant, the true standard deviations of exogenous innovations are: $\sigma_\theta=\sigma_G=\sigma_\psi$=1%, σ_λ=0.025%. 'Big shocks' model variant: $\sigma_\theta=\sigma_G=\sigma_\psi$=5%, σ_λ=0.125%. 'Very big shocks' model variant: $\sigma_\theta=\sigma_G=\sigma_\psi$=10%, σ_λ=0.250%.