Dynare Working Papers Series
http://www.dynare.org/wp/

Estimating nonlinear DSGE models with moments based methods

Ivashchenko Sergey

Working Paper no. 32

January 2014

CEPREMAP
CENTRE POUR LA RECHERCHE ECONOMIQUE ET SES APPLICATIONS
142, rue du Chevaleret — 75013 Paris — France
http://www.cepremap.fr

Estimating nonlinear DSGE models with moments based methods

By Ivashchenko Sergey[1]

Abstract

This article suggests the new approach to an approximation of nonlinear DSGE models moments. This approach is fast and accurate enough to use it for an estimation of nonlinear DSGE models. The small financial DSGE model is repeatedly estimated by several modifications of suggested approach. Approximations of moments are close to the results of large sample Monte Carlo estimation. Quality of parameters estimation with suggested approach is close to the Central Difference Kalman Filter (the CDKF) based. At the same time suggested approach is much faster.

Keywords: DSGE; DSGE-VAR, GMM, nonlinear estimation

Jel-codes: C13; C32; E32

[1]St. Petersburg Institute for Economics and Mathematics (Russian Academy of Sciences); 1 Tchaikovsky Str., St. Petersburg, 191187 RUSSIA
e-mail: glucke_ru@pisem.net; tel: +7-921-746-32-12

1. Introduction

Modern macroeconomics seeks to explain the aggregate economy using theories based on strong microeconomic foundations. The advantage of such an approach is a description of models in terms of “deep structural” parameters that are not influenced by economic policy [Wickens (2008)]. Nevertheless, these parameters should be estimated for usage of DSGE models. There are different econometric techniques for models estimation but empirical studies have concentrated their attention on an estimation of first-order linearized DSGE models [Tovar (2008)].

Non-linear approximations of DSGE models have several important advantages, in particular, they allow uncertainty to influence economic choices [Ruge-Murcia (2012)]. Parameters drifting and stochastic volatility are examples of important elements that are essentially worthless with linear approximations [Fernandez-Villaverde, Guerron et al. (2010)]. Linear approximation of DSGE models behavior may significantly differ from higher order (more accurate) approximations [Collard and Juillard (2001)]. Second order approximation makes difference between the models and the approximation behavior much smaller. There are a few methods for non-linear approximations of DSGE models, but the perturbation method is the most widely used [Schmitt-Grohe and Uribe (2004)]. Thus advantages of non-linear approximations are known and techniques for approximations are developed. However, these advantages can not be fully implemented due to lack of estimation techniques for non-linear approximations (existing techniques are too computationally expensive for usage of them with medium-scale DSGE models).

There are two main approaches for an estimation of DSGE models: moments based and likelihood based [DeJong and Dave (2007), Canova (2007)]. The likelihood based approaches use non-linear filters for the construction of the likelihood function. The first of them is the particle filter. This tool could produce all advantages of nonlinear approximation, including sharper the likelihood function and smaller variance of parameters estimator [An and

Schorfheide (2006), Fernandez-Villaverde, Guerron et al. (2010)]. However, the particle filters have some disadvantages: the likelihood evaluation with the particle filter is a random variable (each evaluation of the likelihood for the same model and observations produces different value). Thus usual maximization algorithms can't be used and Markov Chain Monte Carlo inefficiency greatly increases - Pitt, Silva et al. (2012) show that number of draws should be 10 (from 5 to 400 depending on the number of particles) times higher for the same accuracy of MCMC.

There are alternative filters that can be used for the likelihood calculation. For example, Andreasen (2008) has shown an advantage of the Central Difference Kalman Filter over a few versions of particle filter (in terms of the quality and computing time (100 times faster)). The Quadratic Kalman Filter has an advantage in the quality with some loss in computing time over the Unscented Kalman filter (Julier and Uhlmann (1997)) and the Central Difference Kalman Filter (Ivashchenko (2013)).

Moments based approaches for estimation of DSGE models are more robust [Ruge-Murcia (2007), Creel and Kristensen (2011)]. The first of them is the instrumental variables approach which is a special case of the generalized method of moments [Canova (2007)]. All variables of the DSGE model should be observed for usage of this approach. It can be true only for small-scale DSGE models. Another version of the GMM is used for a linearized model because a wide range of empirical targets can be calculated analytically [DeJong and Dave (2007)]. The simulated method of moments has almost the same statistical efficient as the GMM, but it is more computational demanding [Ruge-Murcia (2007)]. The SMM can be implemented for non-linear DSGE models [Ruge-Murcia (2012), Kim and Ruge-Murcia (2009)]. But it's too slow for an estimation of medium-scale or large-scale models that are used for policymaking. Another moments based approach is the indirect inference. Theoretically it's more efficient than the GMM [Creel and Kristensen (2011)]. The usage of the indirect inference is more complicated than the GMM, because it requires knowledge of the moments distribution function. It can be calculated easily only for a narrow range of moments (usually it is parameters estimation of an econometric model, example is DSGE-VAR model). So, usually this approach is described differently [DeJong and Dave (2007)].

This article suggests an approach for a fast calculation of non-linear DSGE model's moments. This moments calculations are compared with alternative approaches. A small non-linear DSGE model is estimated with moments based approaches. The quality of such estimations is compared with linear maximum likelihood estimation and the CDKF based quasi-maximum likelihood estimation.

2. The approach for moments calculation

The equation (1) describes the data generating process for state variables (X_t) which is approximation of a rational expectation model with perturbation method. Exogenous shocks (ε_t) have normal distribution with covariance matrix Ω_ε and mean equal zero. The measurement equation (2) describes the dependence of observed variables (Y_t) on state variables and measurement errors (u_t) that have normal distribution with zero mean and covariance matrix Ω_u.

$$X_t = \begin{bmatrix} B_X & B_\varepsilon \end{bmatrix} \begin{bmatrix} X_{t-1} \\ \varepsilon_t \end{bmatrix} + C + \begin{bmatrix} A_{xx} & A_{x\varepsilon} & 0 & A_{\varepsilon\varepsilon} \end{bmatrix} \begin{bmatrix} X_{t-1} \otimes X_{t-1} \\ X_{t-1} \otimes \varepsilon_t \\ \varepsilon_t \otimes X_{t-1} \\ \varepsilon_t \otimes \varepsilon_t \end{bmatrix} \qquad (1)$$

$$Y_t = S + DX_t + u_t \qquad (2)$$

The measurement equation (2) is linear, therefore dependence between moments of observed variables (Y_t) and state variables (X_t) is standard. Equations for first and for second moments of state variables (3)-(5) are more complicated.

$$E(X_t) = B_X E(X_t) + C + A_{xx} E(X_t \otimes X_t) + A_{\varepsilon\varepsilon}\Omega_\varepsilon \qquad (3)$$

$$\begin{aligned} E(X_t \otimes X_t) = {} & (\mathrm{B_X} \otimes \mathrm{B_X}) E(X_t \otimes X_t) + (\mathrm{B_X} \otimes \mathrm{C}) E(X_t) + (\mathrm{B_X} \otimes \mathrm{A_{XX}}) E(X_t \otimes X_t \otimes X_t) + \\ & (\mathrm{B_X} \otimes \mathrm{A}_{\varepsilon\varepsilon})(E(X_t) \otimes \Omega_\varepsilon) + (\mathrm{B}_\varepsilon \otimes \mathrm{B}_\varepsilon)\Omega_\varepsilon + (\mathrm{B}_\varepsilon \otimes \mathrm{A}_{X\varepsilon})\mathrm{U_{EXE\_XEE}}(E(X_t) \otimes \Omega_\varepsilon) + \\ & (\mathrm{C} \otimes \mathrm{B_X}) E(X_t) + \mathrm{C} \otimes \mathrm{C} + (\mathrm{C} \otimes \mathrm{A_{XX}}) E(X_t \otimes X_t) + (\mathrm{C} \otimes \mathrm{A}_{\varepsilon\varepsilon})\Omega_\varepsilon + \\ & (\mathrm{A_{XX}} \otimes \mathrm{B_X}) E(X_t \otimes X_t \otimes X_t) + (\mathrm{A_{XX}} \otimes \mathrm{C}) E(X_t \otimes X_t) + (\mathrm{A_{XX}} \otimes \mathrm{A}_{\varepsilon\varepsilon})(E(X_t \otimes X_t) \otimes \Omega_\varepsilon) + \\ & (\mathrm{A}_{X\varepsilon} \otimes \mathrm{B}_\varepsilon)(\mathrm{X} \otimes \Omega_\varepsilon) + (\mathrm{A}_{X\varepsilon} \otimes \mathrm{A}_{X\varepsilon})\mathrm{U_{XEXE\_EEXX}}(\Omega_\varepsilon \otimes E(X_t \otimes X_t)) + \\ & (\mathrm{A}_{\varepsilon\varepsilon} \otimes \mathrm{B_X})(\Omega_\varepsilon \otimes E(X_t)) + (\mathrm{A}_{\varepsilon\varepsilon} \otimes \mathrm{C})\Omega_\varepsilon + (\mathrm{A}_{\varepsilon\varepsilon} \otimes \mathrm{A_{XX}})(\Omega_\varepsilon \otimes E(X_t \otimes X_t)) + \\ & (\mathrm{A}_{\varepsilon\varepsilon} \otimes \mathrm{A}_{\varepsilon\varepsilon})\Omega_{\varepsilon\varepsilon} + (\mathrm{A_{XX}} \otimes \mathrm{A_{XX}}) E(X_t \otimes X_t \otimes X_t \otimes X_t) \end{aligned} \qquad (4)$$

$$E(X_t \otimes X_{t-s}) = (B_X \otimes I_X)E(X_t \otimes X_{t-s+1}) + (C \otimes I_X)E(X_t) + (A_{xx} \otimes I_X)E(X_t \otimes X_t \otimes X_{t-s+1}) + (A_{\varepsilon\varepsilon} \otimes I_X)(\Omega_\varepsilon \otimes E(X_t)) \quad (5)$$

where $\Omega_{\varepsilon\varepsilon}$ is the fourth moment of vector ε_t, I_X is identity matrix of the same size as state variables vector (X_t), U_{EXE_XEE} and U_{XEXE_EEXX} are permutation matrixes that describe dependence between vectors indicated in the subscript (equation (6) is an example):

$$\varepsilon_t \otimes X_t \otimes \varepsilon_t = U_{EXE\_XEE}(X_t \otimes \varepsilon_t \otimes \varepsilon_t) \quad (6)$$

The equation (7) shows matrix formula for calculation of the fourth moment (for ε_t):

$$vec(\Omega_{\varepsilon\varepsilon}; n_\varepsilon^4; 1) = vec(vec(\Omega_\varepsilon; n_\varepsilon; n_\varepsilon) \otimes vec(\Omega_\varepsilon; n_\varepsilon^2; 1) + vec(\Omega_\varepsilon; n_\varepsilon; n_\varepsilon) \otimes vec(\Omega_\varepsilon; n_\varepsilon; n_\varepsilon) + vec(\Omega_\varepsilon; n_\varepsilon; n_\varepsilon) \otimes vec(\Omega_\varepsilon; 1; n_\varepsilon^2); n_\varepsilon^4; 1) \quad (7)$$

where n_ε is the number of elements in the vector ε_t, *vec(M;n;m)* is the vectorization function that transforms matrix M to matrix with n rows and m columns.

The solution of the equation (3) requires knowledge of vector's X_t second moment. The equation (4) for the second moment requires knowledge of the third and the fourth moments. So, it is impossible to solve these equations directly. There are some approaches for approximation of equation (3)-(5) solution. The key question is what to do with higher order moments. The first way is to approximate the higher order moments (normal distribution would be used for this approximation). The second way is to assume that the higher order moments are equal to zero.

Equations (8)-(9) are the formulas for the third and the fourth moments of normal distribution:

$$E(X_t \otimes X_t \otimes X_t) = \mu_X \otimes \mu_X \otimes \mu_X + \Omega_X \otimes \mu_X + \mu_X \otimes \Omega_X + U_{X\mu X\_\mu XX}(\mu_X \otimes \Omega_X) \quad (8)$$

$$E(X_t \otimes X_t \otimes X_t \otimes X_t) = \mu_X \otimes \mu_X \otimes \mu_X \otimes \mu_X + \Omega_X \otimes \mu_X \otimes \mu_X + U_{X\mu X\mu\_\mu\mu XX}(\mu_X \otimes \mu_X \otimes \Omega_X) + U_{X\mu\mu X\_\mu\mu XX}(\mu_X \otimes \mu_X \otimes \Omega_X) + U_{\mu XX\mu\_\mu\mu XX}(\mu_X \otimes \mu_X \otimes \Omega_X) + U_{\mu X\mu X\_\mu\mu XX}(\mu_X \otimes \mu_X \otimes \Omega_X) + \mu_X \otimes \mu_X \otimes \Omega_X + \Omega_{XX} \quad (9)$$

where μ_X is the first moment of a vector X_t, Ω_X is the second central moment of a vector X_t, Ω_{XX} is the fourth central moment of a vector X_t.

Equations (3)-(4) with normal approximation of the higher order moments are nonlinear (in moments). Analytical solution of a large non-linear system is problematic. Thus numeric solution is required. The simple numeric method is used:

$$\mu_{X,k+1} = B_X \mu_{X,k} + C + A_{xx}\mu_{XX,k} + A_{\varepsilon\varepsilon}\Omega_\varepsilon \quad (10)$$

$$
\begin{aligned}
\mu_{XX,k+1} = {} & (\mathrm{B_X} \otimes \mathrm{B_X})\mu_{XX,k} + (\mathrm{B_X} \otimes \mathrm{C})\mu_{X,k} + (\mathrm{B_X} \otimes \mathrm{A_{XX}})\mu_{XXX,k} + \\
& (\mathrm{B_X} \otimes \mathrm{A}_{\varepsilon\varepsilon})(\mu_{X,k} \otimes \Omega_\varepsilon) + (\mathrm{B}_\varepsilon \otimes \mathrm{B}_\varepsilon)\Omega_\varepsilon + (\mathrm{B}_\varepsilon \otimes \mathrm{A}_{X\varepsilon})\mathrm{U_{EXE\_XEE}}(\mu_{X,k} \otimes \Omega_\varepsilon) + \\
& (\mathrm{C} \otimes \mathrm{B_X})\mu_{X,k} + \mathrm{C} \otimes \mathrm{C} + (\mathrm{C} \otimes \mathrm{A_{XX}})\mu_{XX,k} + (\mathrm{C} \otimes \mathrm{A}_{\varepsilon\varepsilon})\Omega_\varepsilon + \\
& (\mathrm{A_{XX}} \otimes \mathrm{B_X})\mu_{XXX,k} + (\mathrm{A_{XX}} \otimes \mathrm{C})\mu_{XX,k} + (\mathrm{A_{XX}} \otimes \mathrm{A}_{\varepsilon\varepsilon})(\mu_{XX,k} \otimes \Omega_\varepsilon) + \\
& (\mathrm{A}_{X\varepsilon} \otimes \mathrm{B}_\varepsilon)(\mu_{X,k} \otimes \Omega_\varepsilon) + (\mathrm{A}_{X\varepsilon} \otimes \mathrm{A}_{X\varepsilon})\mathrm{U_{XEXE\_EEXX}}(\Omega_\varepsilon \otimes \mu_{XX,k}) + \\
& (\mathrm{A}_{\varepsilon\varepsilon} \otimes \mathrm{B_X})(\Omega_\varepsilon \otimes \mu_{X,k}) + (\mathrm{A}_{\varepsilon\varepsilon} \otimes \mathrm{C})\Omega_\varepsilon + (\mathrm{A}_{\varepsilon\varepsilon} \otimes \mathrm{A_{XX}})(\Omega_\varepsilon \otimes \mu_{XX,k}) + \\
& (\mathrm{A}_{\varepsilon\varepsilon} \otimes \mathrm{A}_{\varepsilon\varepsilon})\Omega_{\varepsilon\varepsilon} + (\mathrm{A_{XX}} \otimes \mathrm{A_{XX}})\mu_{XXXX,k}
\end{aligned}
\qquad (11)
$$

where $\mu_{X,k}$, $\mu_{XX,k}$, $\mu_{XXX,k}$, $\mu_{XXXX,k}$ are the first, the second, the third and the fourth moments of a vector X_t after iteration k.

It should be noted that only a few iterations can be calculated due to computational costs. It means that general properties of the convergence are less important. The numerical comparisons of moments and estimation based on them are made instead of it.

The equation (5) can be solved without numerical approximation (just with normal approximation of third moment of vector $[X_t;X_{t-s}]$). This approach would be called the normal approximation of the higher than 2nd moments (the NAHM2).

Alternative approach (where the higher order moments are zeros) is also used in this article. Due to lower computational costs the third moment of X_t can be calculated (large formula (12)).

$$
\begin{aligned}
&\mu_{XXX,k+1} = (B_X \otimes B_X \otimes B_X)\mu_{XXX,k} + (B_X \otimes B_X \otimes C)\mu_{XX,k} + (B_X \otimes B_X \otimes A_{\varepsilon\varepsilon})(\mu_{XX,k} \otimes \Omega_\varepsilon) + \\
&(B_X \otimes B_\varepsilon \otimes B_\varepsilon)(\mu_{X,k} \otimes \Omega_\varepsilon) + (B_X \otimes B_\varepsilon \otimes A_{X\varepsilon})U_{XEXE\_EEXX}(\Omega_\varepsilon \otimes \mu_{XX,k}) + (B_X \otimes C \otimes B_X)\mu_{XX,k} + \\
&(B_X \otimes C \otimes C)\mu_{X,k} + (B_X \otimes C \otimes A_{XX})\mu_{XXX,k} + (B_X \otimes C \otimes A_{\varepsilon\varepsilon})(\mu_{X,k} \otimes \Omega_\varepsilon) + (B_X \otimes A_{XX} \otimes C)\mu_{XXX,k} + \\
&(B_X \otimes A_{XX} \otimes A_{\varepsilon\varepsilon})(\mu_{XXX,k} \otimes \Omega_\varepsilon) + (B_X \otimes A_{X\varepsilon} \otimes B_\varepsilon)(\mu_{XX,k} \otimes \Omega_\varepsilon) + \\
&(B_X \otimes A_{X\varepsilon} \otimes A_{X\varepsilon})U_{XXEXE\_EEXXX}(\Omega_\varepsilon \otimes \mu_{XXX,k}) + (B_X \otimes A_{\varepsilon\varepsilon} \otimes B_X)U_{XEEX\_EEXX}(\Omega_\varepsilon \otimes \mu_{XX,k}) + \\
&(B_X \otimes A_{\varepsilon\varepsilon} \otimes C)(\mu_{X,k} \otimes \Omega_\varepsilon) + (B_X \otimes A_{\varepsilon\varepsilon} \otimes A_{XX})U_{XEEXX\_EEXXX}(\Omega_\varepsilon \otimes \mu_{XXX,k}) + \\
&(B_X \otimes A_{\varepsilon\varepsilon} \otimes A_{\varepsilon\varepsilon})(\mu_{X,k} \otimes \Omega_{\varepsilon\varepsilon}) + (B_\varepsilon \otimes B_X \otimes B_\varepsilon)U_{EXE\_XEE}(\mu_{X,k} \otimes \Omega_\varepsilon) + \\
&(B_\varepsilon \otimes B_X \otimes A_{X\varepsilon})U_{EXXE\_EEXX}(\Omega_\varepsilon \otimes \mu_{XX,k}) + (B_\varepsilon \otimes B_\varepsilon \otimes B_X)(\Omega_\varepsilon \otimes \mu_{X,k}) + \\
&(B_\varepsilon \otimes B_\varepsilon \otimes C)\Omega_\varepsilon + (B_\varepsilon \otimes B_\varepsilon \otimes A_{XX})(\Omega_\varepsilon \otimes \mu_{XX,k}) + (B_\varepsilon \otimes B_\varepsilon \otimes A_{\varepsilon\varepsilon})\Omega_{\varepsilon\varepsilon} + (B_\varepsilon \otimes C \otimes B_\varepsilon)\Omega_\varepsilon + \\
&(B_\varepsilon \otimes C \otimes A_{X\varepsilon})U_{EXE\_XEE}(\mu_{X,k} \otimes \Omega_\varepsilon) + (B_\varepsilon \otimes A_{XX} \otimes B_\varepsilon)U_{EXXE\_EEXX}(\Omega_\varepsilon \otimes \mu_{XX,k}) + \\
&(B_\varepsilon \otimes A_{XX} \otimes A_{X\varepsilon})U_{EXXXE\_EEXXX}(\Omega_\varepsilon \otimes \mu_{XXX,k}) + (B_\varepsilon \otimes A_{X\varepsilon} \otimes B_X)U_{EXEX\_EEXX}(\Omega_\varepsilon \otimes \mu_{XX,k}) + \\
&(B_\varepsilon \otimes A_{X\varepsilon} \otimes C)U_{EXE\_XEE}(\mu_{X,k} \otimes \Omega_\varepsilon) + (B_\varepsilon \otimes A_{X\varepsilon} \otimes A_{XX})U_{EXEXX\_EEXXX}(\Omega_\varepsilon \otimes \mu_{XXX,k}) + \\
&(B_\varepsilon \otimes A_{X\varepsilon} \otimes A_{\varepsilon\varepsilon})U_{EXEEE\_XEEEE}(\mu_{X,k} \otimes \Omega_{\varepsilon\varepsilon}) + (B_\varepsilon \otimes A_{\varepsilon\varepsilon} \otimes B_\varepsilon)\Omega_{\varepsilon\varepsilon} + \\
&(B_\varepsilon \otimes A_{\varepsilon\varepsilon} \otimes A_{X\varepsilon})U_{EEEXE\_XEEEE}(\mu_{X,k} \otimes \Omega_{\varepsilon\varepsilon}) + (C \otimes B_X \otimes B_X)\mu_{XX,k} + (C \otimes B_X \otimes C)\mu_{X,k} + \\
&(C \otimes B_X \otimes A_{XX})\mu_{XXX,k} + (C \otimes B_X \otimes A_{\varepsilon\varepsilon})(\mu_{X,k} \otimes \Omega_\varepsilon) + (C \otimes B_\varepsilon \otimes B_\varepsilon)\Omega_\varepsilon + \\
&(C \otimes B_\varepsilon \otimes A_{X\varepsilon})U_{EXE\_XEE}(\mu_{X,k} \otimes \Omega_\varepsilon) + (C \otimes C \otimes B_X)\mu_{X,k} + C \otimes C \otimes C + \\
&(C \otimes C \otimes A_{XX})\mu_{XX,k} + (C \otimes C \otimes A_{\varepsilon\varepsilon})\Omega_\varepsilon + (C \otimes A_{XX} \otimes B_X)\mu_{XXX,k} + (C \otimes A_{XX} \otimes C)\mu_{XX,k} + \\
&(C \otimes A_{XX} \otimes A_{\varepsilon\varepsilon})(\mu_{XX,k} \otimes \Omega_\varepsilon) + (C \otimes A_{X\varepsilon} \otimes B_\varepsilon)(\mu_{X,k} \otimes \Omega_\varepsilon) + \\
&(C \otimes A_{X\varepsilon} \otimes A_{X\varepsilon})U_{XEXE\_EEXX}(\Omega_\varepsilon \otimes \mu_{XX,k}) + (C \otimes A_{\varepsilon\varepsilon} \otimes B_X)(\Omega_\varepsilon \otimes \mu_{X,k}) + (C \otimes A_{\varepsilon\varepsilon} \otimes C)\Omega_\varepsilon + \\
&(C \otimes A_{\varepsilon\varepsilon} \otimes A_{XX})(\Omega_\varepsilon \otimes \mu_{XX,k}) + (C \otimes A_{\varepsilon\varepsilon} \otimes A_{\varepsilon\varepsilon})\Omega_{\varepsilon\varepsilon} + (A_{XX} \otimes B_X \otimes C)\mu_{XXX,k} +
\end{aligned}
\tag{12}
$$

$$
\begin{aligned}
&+(A_{XX}\otimes B_X\otimes A_{\varepsilon\varepsilon})(\mu_{XXX,k}\otimes\Omega_\varepsilon)+(A_{XX}\otimes B_\varepsilon\otimes B_\varepsilon)(\mu_{XX,k}\otimes\Omega_\varepsilon)+\\
&(A_{XX}\otimes B_\varepsilon\otimes A_{X\varepsilon})U_{XXEXE\_EEXXX}(\Omega_\varepsilon\otimes\mu_{XXX,k})+(A_{XX}\otimes C\otimes B_X)\mu_{XXX,k}+\\
&(A_{XX}\otimes C\otimes C)\mu_{XX,k}+(A_{XX}\otimes C\otimes A_{\varepsilon\varepsilon})(\mu_{XX,k}\otimes\Omega_\varepsilon)+(A_{XX}\otimes A_{X\varepsilon}\otimes B_\varepsilon)(\mu_{XXX,k}\otimes\Omega_\varepsilon)+\\
&(A_{XX}\otimes A_{\varepsilon\varepsilon}\otimes B_X)U_{XXEEX\_EEXXX}(\Omega_\varepsilon\otimes\mu_{XXX,k})+(A_{XX}\otimes A_{\varepsilon\varepsilon}\otimes C)(\mu_{XX,k}\otimes\Omega_\varepsilon)+\\
&(A_{XX}\otimes A_{\varepsilon\varepsilon}\otimes A_{\varepsilon\varepsilon})(\mu_{XX,k}\otimes\Omega_{\varepsilon\varepsilon})+(A_{X\varepsilon}\otimes B_X\otimes B_\varepsilon)U_{XEXE\_EEXX}(\Omega_\varepsilon\otimes\mu_{XX,k})+\\
&(A_{X\varepsilon}\otimes B_X\otimes A_{X\varepsilon})U_{XEXXE\_EEXXX}(\Omega_\varepsilon\otimes\mu_{XXX,k})+(A_{X\varepsilon}\otimes B_\varepsilon\otimes B_X)U_{XEEX\_EEXX}(\Omega_\varepsilon\otimes\mu_{XX,k})+\\
&(A_{X\varepsilon}\otimes B_\varepsilon\otimes C)(\mu_{X,k}\otimes\Omega_\varepsilon)+(A_{X\varepsilon}\otimes B_\varepsilon\otimes A_{XX})U_{XEEXX\_EEXXX}(\Omega_\varepsilon\otimes\mu_{XXX,k})+\\
&(A_{X\varepsilon}\otimes B_\varepsilon\otimes A_{\varepsilon\varepsilon})(\mu_{X,k}\otimes\Omega_{\varepsilon\varepsilon})+(A_{X\varepsilon}\otimes C\otimes B_\varepsilon)(\mu_{X,k}\otimes\Omega_\varepsilon)+\\
&(A_{X\varepsilon}\otimes C\otimes A_{X\varepsilon})U_{XEXE\_EEXX}(\Omega_\varepsilon\otimes\mu_{XX,k})+(A_{X\varepsilon}\otimes A_{XX}\otimes B_\varepsilon)U_{XEXXE\_EEXXX}(\Omega_\varepsilon\otimes\mu_{XXX,k})+\\
&(A_{X\varepsilon}\otimes A_{X\varepsilon}\otimes B_X)U_{XEXEX\_EEXXX}(\Omega_\varepsilon\otimes\mu_{XXX,k})+(A_{X\varepsilon}\otimes A_{X\varepsilon}\otimes C)U_{XEXE\_EEXX}(\Omega_\varepsilon\otimes\mu_{XX,k})+\\
&(A_{X\varepsilon}\otimes A_{X\varepsilon}\otimes A_{\varepsilon\varepsilon})U_{XEXEEE\_EEEEXX}(\Omega_{\varepsilon\varepsilon}\otimes\mu_{XX,k})+(A_{X\varepsilon}\otimes A_{\varepsilon\varepsilon}\otimes B_\varepsilon)(\mu_{X,k}\otimes\Omega_{\varepsilon\varepsilon})+\\
&(A_{X\varepsilon}\otimes A_{\varepsilon\varepsilon}\otimes A_{X\varepsilon})U_{XEEEXE\_EEEEXX}(\Omega_{\varepsilon\varepsilon}\otimes\mu_{XX,k})+(A_{\varepsilon\varepsilon}\otimes B_X\otimes B_X)(\Omega_\varepsilon\otimes\mu_{XX,k})+\\
&(A_{\varepsilon\varepsilon}\otimes B_X\otimes C)(\Omega_\varepsilon\otimes\mu_{X,k})+(A_{\varepsilon\varepsilon}\otimes B_X\otimes A_{XX})(\Omega_\varepsilon\otimes\mu_{XXX,k})+\\
&(A_{\varepsilon\varepsilon}\otimes B_X\otimes A_{\varepsilon\varepsilon})U_{EEXEE\_XEEEE}(\mu_{X,k}\otimes\Omega_{\varepsilon\varepsilon})+(A_{\varepsilon\varepsilon}\otimes B_\varepsilon\otimes B_\varepsilon)\Omega_{\varepsilon\varepsilon}+\\
&(A_{\varepsilon\varepsilon}\otimes B_\varepsilon\otimes A_{X\varepsilon})U_{EEEXE\_XEEEE}(\mu_{X,k}\otimes\Omega_{\varepsilon\varepsilon})+(A_{\varepsilon\varepsilon}\otimes C\otimes B_X)(\Omega_\varepsilon\otimes\mu_{X,k})+(A_{\varepsilon\varepsilon}\otimes C\otimes C)\Omega_\varepsilon+\\
&(A_{\varepsilon\varepsilon}\otimes C\otimes A_{XX})(\Omega_\varepsilon\otimes\mu_{XX,k})+(A_{\varepsilon\varepsilon}\otimes C\otimes A_{\varepsilon\varepsilon})\Omega_{\varepsilon\varepsilon}+(A_{\varepsilon\varepsilon}\otimes A_{XX}\otimes B_X)(\Omega_\varepsilon\otimes\mu_{XXX,k})+\\
&(A_{\varepsilon\varepsilon}\otimes A_{XX}\otimes C)(\Omega_\varepsilon\otimes\mu_{XX,k})+(A_{\varepsilon\varepsilon}\otimes A_{XX}\otimes A_{\varepsilon\varepsilon})U_{EEXXEE\_EEEEXX}(\Omega_{\varepsilon\varepsilon}\otimes\mu_{XX,k})+\\
&(A_{\varepsilon\varepsilon}\otimes A_{X\varepsilon}\otimes B_\varepsilon)U_{EEXEE\_XEEEE}(\mu_{X,k}\otimes\Omega_{\varepsilon\varepsilon})+(A_{\varepsilon\varepsilon}\otimes A_{X\varepsilon}\otimes A_{X\varepsilon})U_{EEXEXE\_EEEEXX}(\Omega_{\varepsilon\varepsilon}\otimes\mu_{XX,k})+\\
&(A_{\varepsilon\varepsilon}\otimes A_{\varepsilon\varepsilon}\otimes B_X)(\Omega_{\varepsilon\varepsilon}\otimes\mu_{X,k})+(A_{\varepsilon\varepsilon}\otimes A_{\varepsilon\varepsilon}\otimes C)\Omega_{\varepsilon\varepsilon}+\\
&(A_{\varepsilon\varepsilon}\otimes A_{\varepsilon\varepsilon}\otimes A_{XX})(\Omega_{\varepsilon\varepsilon}\otimes\mu_{XX,k})+(A_{\varepsilon\varepsilon}\otimes A_{\varepsilon\varepsilon}\otimes A_{\varepsilon\varepsilon})\Omega_{\varepsilon\varepsilon\varepsilon}
\end{aligned}
$$
(12)

The equation for the third moment (with zero higher moments) is linear in the moments of state variables X_t, but it is a large scale equation. Equation for the second moments is Lyapunov equation and there are special algorithms for solving it. The equation for the third moments is large and there are no special algorithms (which use specific structure of matrixes) for solving it. This is why the same iteration algorithm is used for moments calculation. This approach would be called the zero approximation of the higher than 3 moments (the ZAHM3). The case with zero approximation of the higher than the second moment is calculated too (the ZAHM2). It should be noted that difference between the NAHM2 and the ZAHM2 is 3 of 20 sum components in the dispersion formula (11).

The suggested approaches use properties of exogenous shocks (ε_t) normal distribution for calculation of exogenous shocks higher moments. It means that these approaches can be easily modified for other distribution of exogenous shocks. However, an approximation of a rational expectation model with the perturbation method (1) can have infinite unconditional moments if shocks have heavier tails such as t-distribution. The reason is following: t-distribution has infinite higher moments (depending on degrees of freedom); formula (1) means that X_t depend on $(\varepsilon_t)^2$, $(\varepsilon_{t-1})^4$, $(\varepsilon_{t-2})^8$ and so on. Thus, the suggested approaches can be modified for non-normal distribution of exogenous shocks, if all moments of this distribution are finite.

3. Comparison techniques

Finance is one of those areas where linear approximation of DSGE models is unsuitable. Therefore a finance model is used here for comparing different estimation approaches. The same model as [Ivashchenko (2013)] is used (but with additional observed variables). Households maximize the expected utility function (13) with budget constraint (14). There are 3 types of an expenditure: consumption (Ct) with exogenous price ($Z_{P,t}$), one period bonds (B_t), and stocks (X_t), the price of which is S_t. There are 3 sources of income: exogenous income ($S_t Z_{I,t}$), bonds with interest that were bought one period ago ($R_{t-1}\ B_{t-1}$), and stocks with dividend that were bought one period ago ($X_{t-1}(S_t+D_t)$).

$$E_0 \sum_{t=0}^{\infty} \beta^t \frac{(C_t)^\gamma}{\gamma} \to \max_{C;B;X} \tag{13}$$

$$Z_{P,t} C_t + B_t + X_t S_t = R_{t-1} B_{t-1} + X_{t-1}(S_t + D_t) + S_t Z_{I,t} \tag{14}$$

This model suggests that dividend growth is exogenous (15), the bond amount is set by government (16), and amount of stocks is equal to 1 (17).

$$\frac{D_t}{D_{t-1}} = Z_{D,t} \tag{15}$$

$$B_t = Z_{B,t} S_t \tag{16}$$

$$X_t = 1 \tag{17}$$

The model (13)-(17) is transformed into (18)-(22) where stationarized variables are used. Table 1 shows the relation between the original and stationarized variables.

TABLE 1. The DSGE model variables

Variable	Description	Stationary variable
B_t	Value of bonds bought by households at period t	$b_t = B_t / S_t$
C_t	Consumption at time t	$c_t = \ln\left(Z_{P,t} C_t / S_t\right)$
D_t	Dividends at time t	$d_t = \ln\left(D_t / S_t\right)$
R_t	Interest rate at time t	$r_t = \ln\left(R_t\right)$
S_t	Price of stocks at time t	$s_t = \ln\left(S_t / S_{t-1}\right)$
X_t	Amount of stocks bought by households at period t	$x_t = X_t$
Λ_t	Lagrange multiplier corresponding to budget restriction of households at period t	$\lambda_t = \Lambda_t$
$Z_{A,B,t}$	Exogenous process corresponding to near-rationality of households with their bond position	$z_{A,B,t} = Z_{A,B,t}$
$Z_{A,C,t}$	Exogenous process corresponding to near-rationality of households with their consumption	$z_{A,C,t} = Z_{A,C,t}$
$Z_{A,S,t}$	Exogenous process corresponding to near-rationality of households with their stocks position	$z_{A,C,t} = Z_{A,C,t}$
$Z_{B,t}$	Exogenous process corresponding to bond amount sold by the government	$z_{B,t} = Z_{B,t}$
$Z_{D,t}$	Exogenous process corresponding to growth of dividends	$z_{D,t} = Z_{D,t}$
$Z_{I,t}$	Exogenous process corresponding to income of households	$z_{T,t} = Z_{T,t}$
$Z_{P,t}$	Exogenous process corresponding to price level	$z_{P,t} = \ln\left(Z_{P,t} / Z_{P,t-1}\right)$

$$E_0 \sum_{t=0}^{\infty} \beta^t e^{\gamma \sum_{i=0}^{t} (s_i - z_{P,i})} \frac{\left(e^{c_t}\right)^{\gamma}}{\gamma} \to \max_{c;b;x} \tag{18}$$

$$e^{c_t} + b_t + x_t = e^{r_{t-1} - s_t} b_{t-1} + x_{t-1}\left(1 + e^{d_t}\right) + z_{I,t} \tag{19}$$

$$d_t - d_{t-1} + s_t = z_{D,t} \tag{20}$$

$$b_t = z_{B,t} \tag{21}$$

$$x_t = 1 \tag{22}$$

The optimal conditions of (18)-(19) problems with additional exogenous processes ($z_{A,S,t}$, $z_{A,B,t}$, $z_{A,C,t}$) are the following:

$$e^{\lambda_t + z_{A,S,t}} = E_t e^{\lambda_{t+1} + \ln(\beta) + \gamma\left(s_{t+1} - z_{P,t+1}\right)}\left(1 + e^{d_{t+1}}\right) \tag{23}$$

$$e^{\lambda_t + z_{A,B,t}} = E_t e^{\lambda_{t+1} + r_t - s_{t+1} + \ln(\beta) + \gamma\left(s_{t+1} - z_{P,t+1}\right)} \tag{24}$$

$$\gamma c_t = \lambda_t + c_t + z_{A,C,t} \tag{25}$$

An additional exogenous process ($z_{A,S,t}$, $z_{A,B,t}$, $z_{A,C,t}$) can be interpreted as near-rational households (these processes have zero mean). Another interpretation is a compensation of approximation errors (these processes allows the use of a linear approximation for parameter estimation). All the exogenous processes are AR(1) with the following parameterization:

$$z_{*,t} = \eta_{0,*,t}(1-\eta_{1,*,t}) + \eta_{1,*,t} z_{*,t-1} + \varepsilon_{*,t} \quad (26)$$

The model parameters are estimated with indirect inference approach (DSGE-VAR with 4 lags and 5 iterations with zero higher than second moments). Monthly data (Average rate on 1-month certificates of deposit; MSCI USA price return; MSCI USA gross return; personal consumption expenditures; compensation of employees) from January 1975 till December 2012 is used. Estimated values are used for generating observations.

The following two approaches are used for comparison of the moments estimation techniques. The first of them is a comparison of moments approximations with a usual simulation based approach for moment's calculation (it should be more accurate in case of a large sample simulation). The second approach is a comparison of estimation error (generation of the observed variables by the DSGE model and the estimation of models with different moments based approaches).

The DSGE model is simulated for 100 000 observations. The moments are calculated. This procedure is repeated 10 times. Mean and standard deviation of the moments are reported. Deviations from this moment's mean (errors) are reported for each approach.

The following procedure is used for comparison of the estimation results:

1. Generation of 400 observations (600 and drop of the first 200 observations) from the second-order approximation of model
2. Parameters estimation by different approaches (linear maximum likelihood; the CDKF based maximum likelihood; the indirect inference maximum likelihood and the GMM with different moment's calculation approaches). The true values of parameters are used as the initial values.
3. Steps 1-2 are repeated 100 times.

The indirect inference maximum likelihood is DSGE-VAR (with infinite weight of prior parameter) maximum likelihood estimation (Del Negro and Schorfheide (2004)). The Newey-West estimator (with window equal $7 \approx 4(T/100)^{2/9}$)) is used for calculation of moments variance for the GMM [Hamilton (1994)].

The CDKF is used as a benchmark nonlinear estimation technique because it has advantage over other approaches (in terms of speed with appropriate quality) that are well known in DSGE literature [Andreasen (2008), Ivashchenko (2013)].

4. The results

Table 2 shows the results of moment's estimation. All approaches have very high errors after the first iteration, but these errors greatly decrease after the second iteration (to the same level for each approach). The errors after 5 and 10 iterations are very close that indicates fast convergence. It should be noted that the errors of all approaches are close to errors of mean over 100 000 sample (for the most of the moments errors are smaller than standard deviation). This indicates a very high quality of such simple moment approximations. Thus suggested approaches produce more accurate estimation of the moments than existing approach (simulation) and require less computer time (14 sec. for simulation and less than 1 second for the NAHM2 and the ZAHM2). The quality of the NAHM2 and the ZAHM3 is almost the same. The quality of the ZAHM2 is 3-4 times worse after 5-10 iterations.

It should be noted that the RMSE of all moments for the ZAHM2 with 10 iterations is a little bit worse than ones with 5 iterations. A more detailed view shows that RMSE of the first moments are much smaller for 10 iterations case, but RMSE of second moments is a little bit higher for 10 iterations case. There are only 5 first moments and 115 second moments (including all 4 lags); therefore influence of the second moments is higher for the RMSE of all moments. This difference can be more important for the GMM approach due to smaller variance of the second moments (which means higher weight of second moments). This effect can be important for the NAHM2 and the ZAHM3 because there are a few moments that have higher errors in 10 iterations case than in 5 iterations case (it can make score function of the GMM with 5 iterations better than with 10 iterations for some weight matrixes).

TABLE 2. The DSGE model moments estimation

		Mean	Std	NAHM2				ZAHM2	
iteration number				1	2	5	10	1	2
First	obs_pp	9.32E-03	1.76E-04	-1.92E-01	8.18E-04	1.99E-05	1.89E-05	-1.92E-01	8.18E-04
	obs_pg	1.57E-02	1.70E-04	-1.92E-01	8.71E-04	7.53E-05	7.42E-05	-1.92E-01	8.71E-04
	obs_r	1.69E-02	6.56E-05	7.51E-04	-3.99E-05	8.72E-06	4.99E-06	7.51E-04	-3.99E-05
	obs_c	9.32E-03	1.76E-04	-5.88E-02	-1.46E-01	8.04E-04	1.66E-05	-5.88E-02	-1.46E-01
	obs_i	9.32E-03	1.76E-04	-1.92E-01	8.19E-04	2.02E-05	1.92E-05	-1.92E-01	8.18E-04
Second	obs_pp	3.09E-03	1.98E-05	-4.05E-02	1.55E-05	5.96E-06	5.98E-06	-4.05E-02	4.76E-05
	obs_pg	3.22E-03	2.11E-05	-4.30E-02	3.66E-05	7.74E-06	7.75E-06	-4.30E-02	6.77E-05
	obs_r	2.90E-04	2.21E-06	2.48E-05	-1.51E-06	2.69E-07	1.42E-07	2.48E-05	-1.36E-06
	obs_c	2.15E-03	1.45E-05	-5.30E-03	-2.85E-02	-8.35E-04	-8.07E-04	-4.84E-03	-2.41E-02
	obs_i	6.11E-03	2.39E-05	-4.05E-02	-9.33E-06	1.25E-05	1.25E-05	-4.05E-02	2.29E-05
Second lagged	obs_pp	8.87E-05	7.13E-06	-3.42E-03	1.90E-05	4.11E-08	3.29E-08	-3.44E-03	1.23E-06
	obs_pg	2.35E-04	8.35E-06	-6.15E-03	3.20E-05	1.92E-06	1.89E-06	-6.17E-03	1.43E-05
	obs_r	2.90E-04	2.21E-06	2.56E-05	-1.37E-06	2.80E-07	1.53E-07	2.55E-05	-1.39E-06
	obs_c	5.89E-04	1.00E-05	-9.48E-03	4.50E-03	4.15E-04	3.86E-04	-8.57E-03	-1.35E-03
	obs_i	6.41E-05	1.06E-05	-3.42E-03	3.16E-05	-5.22E-06	-5.20E-06	-3.44E-03	1.14E-05
RMSE of first moments				1.51E-01	6.54E-02	3.62E-04	3.61E-05	1.51E-01	6.54E-02
RMSE of second moments				2.58E-02	5.80E-03	1.79E-04	1.75E-04	2.58E-02	5.18E-03
RMSE of second lagged moments (1 lag)				1.13E-02	1.77E-03	9.15E-05	8.46E-05	1.03E-02	1.04E-03
RMSE of all second moments (4 lags)				1.29E-02	2.87E-03	9.02E-05	8.71E-05	1.27E-02	2.55E-03
RMSE of all moments				3.22E-02	1.31E-02	1.13E-04	8.57E-05	3.22E-02	1.31E-02

TABLE 2 (continued). The DSGE model moments estimation

		Mean	Std	ZAHM2		ZAHM3			
iteration number				5	10	1	2	5	10
First	obs_pp	9.32E-03	1.76E-04	2.01E-05	1.89E-05	-1.92E-01	8.18E-04	2.02E-05	1.89E-05
	obs_pg	1.57E-02	1.70E-04	7.52E-05	7.40E-05	-1.92E-01	8.71E-04	7.56E-05	7.42E-05
	obs_r	1.69E-02	6.56E-05	8.60E-07	5.88E-07	7.51E-04	-3.99E-05	8.87E-06	6.73E-06
	obs_c	9.32E-03	1.76E-04	-1.97E-04	1.90E-05	-5.88E-02	-1.46E-01	-1.41E-03	1.84E-05
	obs_i	9.32E-03	1.76E-04	2.04E-05	1.92E-05	-1.92E-01	8.18E-04	2.05E-05	1.92E-05
Second	obs_pp	3.09E-03	1.98E-05	1.82E-05	1.83E-05	-4.05E-02	1.64E-05	6.02E-06	6.00E-06
	obs_pg	3.22E-03	2.11E-05	1.98E-05	1.98E-05	-4.30E-02	3.75E-05	7.81E-06	7.77E-06
	obs_r	2.90E-04	2.21E-06	6.01E-08	5.14E-08	2.48E-05	-1.46E-06	2.79E-07	2.07E-07
	obs_c	2.15E-03	1.45E-05	3.06E-04	3.80E-04	-4.84E-03	-2.67E-02	-5.15E-04	-6.29E-04
	obs_i	6.11E-03	2.39E-05	2.47E-05	2.48E-05	-4.05E-02	-8.41E-06	1.25E-05	1.25E-05
Second lagged	obs_pp	8.87E-05	7.13E-06	-6.11E-06	-6.18E-06	-3.42E-03	1.61E-05	-7.80E-07	-8.86E-07
	obs_pg	2.35E-04	8.35E-06	-4.18E-06	-4.27E-06	-6.15E-03	2.94E-05	1.13E-06	1.01E-06
	obs_r	2.90E-04	2.21E-06	5.83E-09	-3.43E-09	2.56E-05	-1.46E-06	2.84E-07	2.12E-07
	obs_c	5.89E-04	1.00E-05	-1.95E-04	-1.96E-04	-9.77E-03	5.80E-04	3.75E-04	4.62E-04
	obs_i	6.41E-05	1.06E-05	-1.17E-05	-1.18E-05	-3.42E-03	2.87E-05	-5.98E-06	-6.12E-06
RMSE of first moments				9.50E-05	3.63E-05	1.51E-01	6.54E-02	6.31E-04	3.64E-05
RMSE of second moments				6.70E-04	6.75E-04	2.58E-02	5.45E-03	1.27E-04	1.43E-04
RMSE of second lagged moments (1 lag)				4.64E-04	4.68E-04	1.13E-02	1.11E-03	1.13E-04	1.30E-04
RMSE of all second moments (4 lags)				3.64E-04	3.67E-04	1.29E-02	2.70E-03	8.02E-05	8.82E-05
RMSE of all moments				3.58E-04	3.60E-04	3.22E-02	1.31E-02	1.47E-04	8.68E-05

RMSE of parameters estimation by the indirect inference (DSGE-VAR with 4 lags) with different moment's calculation techniques presented at the table 3. RMSE for the GMM approach presented at the table 4. It should be noted that the indirect inference with moments calculated by the NAHM2 or the ZAHM2 with 2 iterations produce errors covariance matrix which is not positive-definite. Thus, RMSE of the indirect inference presented only for 5 and 10 iterations. The results for the ZAHM3 are not presented due to computational expense of this approach (the ZAHM3 with 5 iterations requires about 21 second; it requires about 11 seconds for 2 iterations what is much higher than the CDKF).

TABLE 3. The RMSE of parameters estimation

	the linear likelihood	the CDKF	DSGE-VAR(4) the NAHM2, 10 iter.	DSGE-VAR(4) the NAHM2, 5 iter.	DSGE-VAR(4) the ZAHM2, 10 iter.	DSGE-VAR(4) the ZAHM2, 5 iter.
Std of $\varepsilon_{A,B}$	1.18E-04	8.93E-05	3.83E-05	5.22E-05	1.13E-02	1.06E-03
Std of $\varepsilon_{A,C}$	6.66E-02	1.51E-02	8.01E-03	1.21E-02	1.60E-02	1.43E-02
Std of $\varepsilon_{A,S}$	1.94E-03	5.41E-03	2.76E-03	1.98E-03	9.51E-03	6.58E-03
Std of ε_B	7.13E-02	4.41E-02	6.96E-03	6.98E-03	1.38E-01	6.64E-02
Std of ε_D	1.87E-02	4.50E-03	7.15E-03	8.93E-03	1.09E-02	1.03E-02
Std of ε_I	1.31E-02	3.96E-02	1.08E-02	4.51E-03	4.11E-03	3.49E-03
Std of ε_P	1.86E+00	3.73E-03	7.16E-03	5.30E-03	1.84E-02	2.08E-02
$ln(\beta)$	3.01E-02	5.45E-03	3.16E-03	2.58E-03	3.65E-03	4.48E-03
γ	8.13E-02	1.58E-03	1.50E-03	1.54E-03	2.12E-02	3.47E-02
$\eta_{0,B}$	8.21E-02	3.50E-01	1.26E-01	3.10E-01	3.85E-01	3.73E-01
$\eta_{0,D}$	2.15E-03	1.66E-03	1.80E-03	2.35E-03	3.42E-03	3.17E-03
$\eta_{0,I}$	2.14E-01	1.35E+00	2.62E-01	4.78E-01	9.31E-01	3.12E-01
$\eta_{0,P}$	2.11E-03	3.99E-03	3.85E-03	2.39E-03	1.75E-03	9.07E-04
$\eta_{1,AB}$	1.74E-01	1.98E-01	9.39E-03	8.72E-03	9.92E-02	2.16E-01
$\eta_{1,AC}$	1.15E-01	1.05E-01	6.36E-04	2.03E-01	1.19E-01	2.03E-01
$\eta_{1,AS}$	2.75E-01	1.71E-01	9.06E-02	6.92E-02	3.92E-01	3.75E-01
$\eta_{1,B}$	2.91E-02	2.78E-01	1.33E+00	1.47E+00	5.92E-01	4.68E-01
$\eta_{1,D}$	1.90E-01	1.35E-01	1.55E-01	1.28E-01	1.61E-01	1.68E-01
$\eta_{1,I}$	1.15E-02	1.91E-02	9.80E-03	6.46E-03	1.34E-01	8.07E-02
$\eta_{1,P}$	1.50E-01	5.07E-01	7.16E-01	8.11E-01	7.34E-01	5.10E-01
Sum of RMSE	3.39E+00	3.23E+00	2.75E+00	3.53E+00	3.79E+00	2.87E+00
Sum of MSE	3.72E+00	2.37E+00	2.39E+00	3.20E+00	2.15E+00	9.86E-01
Time for likelihood calculation (sec)*	0.03	3.95	0.44	0.25	0.125	0.078

*PC used: Intel core 2 Duo E8400 3 GHz, 1 Gb RAM, Windows XP.

It should be noted that a few parameters have much higher RMSE for each approach which mean that its influence is critical for such measure as sum of MSE. The indirect inference with the ZAHM2 with 5 iterations produces extremely high quality of estimation. The reason is a high amount of local extremums (many parameters values produce errors covariance matrix which is not positive-definite). About half of cases produce low value of log-likelihood function (local extremums which are close to the initial values (true values), produce low RMSE).

The NAHM2 with 10 iterations produces the best quality according to sum of RMSE. However, the ZAHM2 with 10 iterations produces the best quality according to sum of MSE. The CDKF is the second best for the both measures of quality. It should be noted that the NAHM2 with 10 iterations has almost the same sum of MSE as the CDKF.

Unexpected result is that the ZAHM2 is better than the NAHM2 for estimation purpose (GMM with 2 and 5 iterations) despite worse quality of moments calculation. The NAHM2 is better than the ZAHM2 (GMM with 10 iterations) according to the one of quality measures. An advantage of the indirect inference over the GMM is expectable [Creel and Kristensen (2011)].

TABLE 4. The RMSE of parameters estimation (GMM)

	GMM the NAHM2 10 iter	GMM the NAHM2 5 iter	GMM the NAHM2 2 iter	GMM the ZAHM2 10 iter	GMM the ZAHM2 5 iter	GMM the ZAHM2 2 iter
Std of $\varepsilon_{A,B}$	8.94E-03	1.06E-02	2.00E-02	2.25E-02	1.52E-02	1.54E-02
Std of $\varepsilon_{A,C}$	1.18E-02	2.02E-02	1.83E-02	2.10E-02	1.32E-02	1.72E-02
Std of $\varepsilon_{A,S}$	1.76E-03	2.66E-03	4.12E-03	2.79E-03	6.04E-04	1.29E-03
Std of ε_B	1.50E-02	4.11E-02	7.58E-02	2.65E-02	4.38E-02	6.21E-02
Std of ε_D	3.68E-02	3.44E-02	4.22E-02	4.31E-02	3.61E-02	4.70E-02
Std of ε_I	2.34E-02	2.84E-02	2.31E-02	1.96E-02	2.28E-02	2.32E-02
Std of ε_P	2.23E-02	1.58E-02	2.56E-03	1.39E-02	5.93E-03	5.55E-03
$ln(\beta)$	2.59E-03	2.85E-03	3.47E-03	2.84E-03	1.96E-03	2.74E-03
γ	1.04E-01	1.66E-01	2.88E-01	1.96E-01	1.92E-01	4.72E-01
$\eta_{0,B}$	4.18E-01	5.85E-01	7.67E-01	6.08E-01	1.07E+00	9.36E-01
$\eta_{0,D}$	2.02E-03	1.49E-03	2.52E-03	2.58E-03	2.15E-03	2.40E-03
$\eta_{0,I}$	2.32E+00	3.61E+00	4.78E+00	1.82E+00	2.56E+00	4.54E+00
$\eta_{0,P}$	6.93E-04	6.43E-04	1.28E-03	1.64E-03	1.27E-03	5.58E-03
$\eta_{1,AB}$	1.04E+00	1.12E+00	1.09E+00	9.43E-01	1.03E+00	9.00E-01
$\eta_{1,AC}$	5.51E-01	6.31E-01	7.81E-01	6.83E-01	7.55E-01	5.00E-01
$\eta_{1,AS}$	5.78E-02	3.41E-01	5.21E-01	5.00E-01	3.73E-02	6.12E-01
$\eta_{1,B}$	1.38E+00	1.28E+00	1.06E+00	1.16E+00	1.24E+00	1.06E+00
$\eta_{1,D}$	8.48E-01	8.56E-01	9.18E-01	9.77E-01	9.78E-01	1.04E+00
$\eta_{1,I}$	6.66E-01	8.81E-01	5.79E-01	6.53E-01	8.58E-01	7.45E-01
$\eta_{1,P}$	4.80E-01	5.92E-01	7.81E-01	6.21E-01	6.79E-01	7.18E-01
Sum of RMSE	7.99E+00	1.02E+01	1.18E+01	8.32E+00	9.54E+00	1.17E+01
Sum of MSE	1.02E+01	1.87E+01	2.85E+01	8.45E+00	1.31E+01	2.65E+01
Time for likelihood calculation(sec)*	0.44	0.25	0.125	0.125	0.078	0.047

*PC used: Intel core 2 Duo E8400 3 GHz, 1 Gb RAM, Windows XP.

The speed of likelihood functions calculations for normal approximations of higher moments (the NAHM2) is between linear-likelihood and the CDKF-likelihood. For the ZAHM2 the speed is almost the same as for linear-likelihood.

5. Conclusions

This article suggests the new approach to approximation of nonlinear DSGE models moments. These approximations are fast and accurate enough to use them for an estimation of parameters of nonlinear DSGE models (it produces a more accurate estimation of moments than simulation of 100 000 sample). Existing method of approximation of nonlinear DSGE models moments (Monte-Carlo simulations) is 32 to 112 times slower than the suggested approaches (depending on version).

The suggested approaches are 9 (0.44 sec.) or 31 (0.125 sec.) times faster (depending on version) than the CDKF (3.95 sec.). Combination of the suggested approaches with the GMM doesn't produce high quality estimation, but combination of the suggested approaches with indirect inference has almost the same quality as the CDKF. One of quality measure (sum of RMSE) is 17.5% (3.79) worse or 15.0% (2.75) better (depending on version) than the CDKF (3.23). Another measure (sum of MSE) is 0.7% (2.39) worse or 9.4% (2.15) better (depending on version) than the CDKF (2.37). Thus, the suggested approaches are close in the terms of estimation quality but much faster than one of the best existing nonlinear estimation approaches (the CDKF).

Literature

An S. and Schorfheide F., (2006). Bayesian analysis of DSGE models. Working Papers from Federal Reserve Bank of Philadelphia. No. 06-5.

Andreasen M.M., (2008). Non-linear DSGE models, the Central Difference Kalman Filter, and the Mean Shifted Particle Filter. CREATES Research Paper 2008-33. Available at SSRN: http://ssrn.com/abstract=1148079.

Canova F., (2007). Methods for Applied Macroeconomic Research. Princeton University Press.

Collard F. and Juillard M., (2001). Accuracy of stochastic perturbation methods: The case of asset pricing models. Journal of Economic Dynamics and Control, 25 (6-7): 979-999.

Creel M. and Kristensen D. (2011). Indirect Likelihood Inference. Dynare Working Papers from CEPREMAP, No 8,

DeJong D.N. with C. Dave (2007). Structural Macroeconometrics. Princeton University Press.

Fernandez-Villaverde J., Guerron P.A. and Rubio-Ramirez J.F. (2010). Reading the recent monetary history of the United States, 1959-2007 // Review, 2010, issue May, pages 311-338

Hamilton J.D. (1994) Time Series Analysis. Princeton University Press.

Ivashchenko S., (2013). DSGE Model Estimation on the Basis of Second-Order Approximation. Computational Economics, 2013, DOI 10.1007/s10614-013-9363-1

Jinill K. and Ruge-Murcia F.J. (2009). How much inflation is necessary to grease the wheels? Journal of Monetary Economics, 56 (3): 365-377.

Julier S. J. and Uhlmann J. K. (1997). A new extension of the Kalman filter to nonlinear systems. in Proc. AeroSense: 11th Int. Symp. Aerospace/Defense Sensing, Simulation and Controls, pp. 182–193.

Del Negro, M., and F. Schorfheide (2004). Priors from General Equilibrium Models for VARs. International Economic Review, 45 (2): 643–673

Pitt M.K., Silva R.S., Giordani P. and Kohn R.J. (2012). On some properties of Markov chain Monte Carlo simulation methods based on the particle filter. Journal of Econometrics, 171 (2): 134-151

Ruge-Murcia F.J. (2012). Estimating nonlinear DSGE models by the simulated method of moments: With an application to business cycles. Journal of Economic Dynamics and Control, 36 (6): 914-938

Ruge-Murcia F.J. (2007). Methods to estimate dynamic stochastic general equilibrium models. Journal of Economic Dynamics and Control, 31 (8): 2599-2636

Schmitt-Grohe S. and Uribe M. (2004). Solving dynamic general equilibrium models using a second-order approximation to the policy function // Journal of Economic Dynamics and Control, 2004, vol. 28, issue 4, pages 755-775

Tovar C. E. (2008). DSGE models and central banks. BIS Working Papers from Bank for International Settlements. No. 258.

Wickens M.R. (2008). MACROECONOMIC THEORY – A DYNAMIC GENERAL EQUILIBRIUM APPROACH. Princeton University Press, Princeton.