

Giesen, Sebastian; Scheufele, Rolf

Working Paper

Effects of Incorrect Specification on the Finite Sample Properties of Full and Limited Information Estimators in DSGE Models

IWH Discussion Papers, No. 8/2013

Provided in Cooperation with:

Halle Institute for Economic Research (IWH) – Member of the Leibniz Association

Suggested Citation: Giesen, Sebastian; Scheufele, Rolf (2013) : Effects of Incorrect Specification on the Finite Sample Properties of Full and Limited Information Estimators in DSGE Models, IWH Discussion Papers, No. 8/2013, Leibniz-Institut für Wirtschaftsforschung Halle (IWH), Halle (Saale), <https://nbn-resolving.de/urn:nbn:de:gbv:3:2-22479>

This Version is available at:

<https://hdl.handle.net/10419/75409>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

**Effects of Incorrect Specification on the Finite Sample
Properties of Full and Limited Information Estimators
in DSGE Models**

*Sebastian Giesen
Rolf Scheufele*

May 2013

No. 8

Authors: *Sebastian Giesen*
Halle Institute for Economic Research
E-mail: Sebastian.Giesen@iwh-halle.de

Rolf Scheufele
Swiss National Bank
E-mail: Rolf.Scheufele@snb.ch

The responsibility for discussion papers lies solely with the individual authors. The views expressed herein do not necessarily represent those of the IWH. The papers represent preliminary work and are circulated to encourage discussion with the authors. Citation of the discussion papers should account for their provisional character; a revised version may be available directly from the authors.

Comments and suggestions on the methods and results presented are welcome.

IWH Discussion Papers are indexed in RePEc-Econpapers and in ECONIS.

Editor:
HALLE INSTITUTE FOR ECONOMIC RESEARCH – IWH
The IWH is a member of the Leibniz Association.

Address: Kleine Maerkerstrasse 8, D-06108 Halle (Saale), Germany
Postal Address: P.O. Box 11 03 61, D-06017 Halle (Saale), Germany
Phone: +49 345 7753 60
Fax: +49 345 7753 820
Internet: <http://www.iwh-halle.de>

ISSN 1860-5303 (Print)
ISSN 2194-2188 (Online)

Effects of Incorrect Specification on the Finite Sample Properties of Full and Limited Information Estimators in DSGE Models¹

Abstract

In this paper we analyze the small sample properties of full information and limited information estimators in a potentially misspecified DSGE model. Therefore, we conduct a simulation study based on a standard New Keynesian model including price and wage rigidities. We then study the effects of omitted variable problems on the structural parameters estimates of the model. We find that FIML performs superior when the model is correctly specified. In cases where some of the model characteristics are omitted, the performance of FIML is highly unreliable, whereas GMM estimates remain approximately unbiased and significance tests are mostly reliable.

Keywords: FIML, GMM, finite sample bias, misspecification, Monte Carlo, DSGE

JEL Classification: C26; C36; C51; E17

¹ We want to thank Gregor Bäurle, Jörg Breitung, Fabio Canova, Jean-Marie Dufour, Andre Kurmann, Jesper Linde, Ludger Linnemann, Roberto Motto, Federico Ravenna, Barbara Rossi and Michael R. Wickens for valuable comments. The paper has been presented at the 12th IWH-CIREQ Conference 2011, the North American Summer Meeting and the Australasian Meeting of the Econometric Society 2012 as well as the 8th Dynare Conference 2012. The views expressed in this paper are those of the authors and not necessarily those of the Swiss National Bank.

Effects of Incorrect Specification on the Finite Sample Properties of Full and Limited Information Estimators in DSGE Models

Zusammenfassung

In diesem Beitrag analysieren wir die Eigenschaften von verschiedenen Schätzern in einem potenziell fehlspezifizierten DSGE-Modell in kleinen Stichproben. Hierfür wird eine Simulationsstudie auf Basis eines neukeynesianischen Standardmodells durchgeführt. Es werden die Effekte von verschiedenen Fehlspezifikationen auf die Schätzer der strukturellen Parameter untersucht. Diese Studie zeigt, dass *Full Information Maximum Likelihood* überlegen ist, wenn das Modell korrekt spezifiziert ist. In Fällen, in denen einige Modelleigenschaften verändert werden, ist diese Methode jedoch höchst unzuverlässig, während Ansätze mit begrenzter Information (hier GMM) unverzerrt bleiben und auch Signifikanztests zuverlässiger sind.

Schlagwörter: FIML, GMM, Kleinstichprobenverzerrung, Fehlspezifikation, Monte Carlo, DSGE

JEL-Klassifikation: C26; C36; C51; E17

Effects of Incorrect Specification on the Finite Sample Properties of Full and Limited Information Estimators in DSGE Models*

Sebastian Giesen[†] Rolf Scheufele[‡]

April 22, 2013

Abstract

In this paper we analyze the small sample properties of full information and limited information estimators in a potentially misspecified DSGE model. Therefore, we conduct a simulation study based on a standard New Keynesian model including price and wage rigidities. We then study the effects of omitted variable problems on the structural parameters estimates of the model. We find that FIML performs superior when the model is correctly specified. In cases where some of the model characteristics are omitted, the performance of FIML is highly unreliable, whereas GMM estimates remain approximately unbiased and significance tests are mostly reliable.

Keywords: FIML, GMM, finite sample bias, misspecification, Monte Carlo, DSGE

JEL Classification: C26; C36; C51; E17

*We want to thank Gregor Bäurle, Jörg Breitung, Fabio Canova, Jean-Marie Dufour, André Kurmann, Jesper Lindé, Ludger Linnemann, Roberto Motto, Federico Ravenna, Barbara Rossi and Michael R. Wickens for valuable comments. The paper has been presented at the 12th IWH-CIREQ conference 2011, the North American Summer Meeting and the Australasian Meeting of the Econometric Society 2012 as well as the 8th Dynare Conference 2012. The views expressed in this paper are those of the authors and not necessarily those of the Swiss National Bank.

[†]Halle Institute for Economic Research, Sebastian.Giesen@iwh-halle.de.

[‡]Swiss National Bank, Rolf.Scheufele@snb.ch.

1 Introduction

The choice of a specific estimator for estimating parameters in structural macroeconomic models has been disputed for several decades (see Theil, 1971, Ch. 10 for a summary). Typically, one can differentiate between two different classes of methods: full information or limited information methods. Full information methods provide the complete range of statistical properties associated with the model under investigation. Normally, this is preferable in terms of efficiency, given a correctly specified model (see e.g. Cragg, 1967, for an early contribution). Limited information methods do not require a fully specified model, instead it is enough to set up certain moment conditions to estimate the parameters of interest. Thus, there is the classical trade-off between efficiency and the sensitivity to model misspecification known from simultaneous equation models (see Cragg, 1968; Hale, Mariano and Ramage, 1980).

Today, New Keynesian DSGE models became the workhorse macroeconomic models used for policy analysis and forecasting (e.g. Smets and Wouters, 2007; Schorfheide, 2011). This model class, as opposed to simultaneous equation systems, typically involves rational expectations as well as non-linearities in the structural model parameters. However, the choice among full and limited information econometric techniques to estimate New Keynesian models has still to be made.

First attempts to estimate DSGE models relied on limited information methods (e.g. Christiano and Eichenbaum, 1992), mainly because limited information methods are less computationally demanding. Another important advantage of limited information techniques is that it can be applied to study certain aspects of a model and leaving apart those of little interest (e.g. it can be applied in a single equations

framework), which has become quite popular (see Galí and Gertler, 1999; Clarida, Galí and Gertler, 2000; Eichenbaum and Fisher, 2007). The standard estimation technique in this setting is Generalized Methods of Moments (GMM) (see Canova, 2007, Ch. 5, for GMM estimation of DSGE models).

In recent years, the availability of computational power and the increasing popularity of Bayesian techniques strongly simplified the estimation of these models by full information methods, which are based on the likelihood of the state-space solution of the log-linearized DSGE model (see e.g. Ireland, 2004; An and Schorfheide, 2007) and thus become standard in applied research and policy institutions (e.g. central banks). Most recently, there is also growing interest in the econometric aspects of these models. One major concern has been potential identification difficulties (Canova and Sala, 2009; Iskrev, 2010; Andrews and Mikusheva, 2011; Dufour, Khalaf and Kichian, 2012; Guerron-Quintana, Inoue and Kilian, 2012). The issue of weak identification is not only relevant for full information techniques, but may be also evident in limited information settings (see e.g. Dufour, Khalaf and Kichian, 2006; Kleibergen and Mavroeidis, 2009).

In this paper, we concentrate on another econometric aspect of New Keynesian DSGE models: the consequences of potential misspecification. We therefore investigate the properties of limited (GMM) and full information (FIML) techniques in a standard (and potentially misspecified) DSGE model with nominal price and wage rigidities (as in Erceg, Henderson and Levin, 2000). Several forms of misspecification are taken into account. First, we look at the consequences of estimating a model with price rigidities but omit nominal wage rigidities. Second, we omit price indexation in the Phillips curve. Finally, we investigate the case of misspecified shocks (missing autocorrelation) in the IS curve. In a simulation study we

document the properties of the different estimation techniques for point estimates and standard significance tests.

The paper is closely related to Ruge-Murcia's (2007) work, who compares the properties of different estimators in a stylized RBC model. Instead, we look at a New Keynesian DSGE model which forms the basis of today's macroeconomic models. Additionally, Lindé (2005) as well as Jondeau and Le Bihan (2008) also take into account model misspecification. While Jondeau and Le Bihan (2008) look at misspecification within one equation of interest, we look at the consequences of misspecification in the whole system. Lindé (2005) only takes into account quite moderate degrees of misspecification (mainly the misspecification of shocks and lagged persistence). Both, Lindé (2005) and Jondeau and Le Bihan (2008) compare ML with GMM, but use the standard two-step GMM estimator in a single equation set-up. Instead of doing so, we propose the continuous-updating GMM (CUGMM) as advocated by Hansen, Heaton and Yaron (1996), which does much better in finite-samples and DSGE type relations. Furthermore, we apply GMM to a multi-equation framework, to estimate all relevant relationships jointly.

Our results suggest that full information procedures should be used with caution. Clearly, FIML is dominant when it comes to estimating the model under the null hypothesis (without misspecification). Parameter estimates are unbiased and quite precise. Moreover, confidence intervals are very reliable in this case, i.e. they provide at least the pre-specified coverage (even in smaller samples). GMM estimates (based on CUGMM) turn out to be less efficient in that case, but remain unbiased. Estimated standard errors are also reliable in most cases and under the proposed model.

In the case of model misspecification the performance of FIML worsens substan-

tially. When nominal wage rigidities are ignored, all parameters are heavily biased and confidence intervals are totally unreliable. More milder forms of misspecifications result in slightly better estimates, but even in the case of misspecified shocks, FIML estimates turn out to be biased and some parameter estimates are far from their true value.

GMM estimates based on CUGMM remain roughly unbiased in all considered cases, although estimated standard errors get slightly less reliable. In general, our results suggest that if one has not very strong beliefs in the usefulness of all aspects of the model, one should stick to limited information methods. In particular, CUGMM applied to multi-equations does a good job in finite samples and is much less sensitive to model misspecification than FIML based on the Kalman filter. Therefore, the classical trade-off between efficiency and potential biases due to misspecification of the different estimation strategies remain valid for standard DSGE models.

The remainder of the paper is organized as follows. The subsequent section presents the model structure. Section 3 discusses the different estimation strategies. Section 4 outlines the simulation setup. Section 5 shows our results and Section 6 concludes.

2 The Model Economy

Our starting point is a stylized New Keynesian Dynamic Stochastic General Equilibrium (DSGE) model, the workhorse structural macroeconomic model for short-term analysis. Those models are well established in the academic literature (see e.g. Christiano, Eichenbaum and Evans, 2005; Smets and Wouters, 2007) and one

of the main tools for policy analysis in central banks and policy institutions.

To make things transparent and as simple as possible, we consider a small scale DSGE inspired by Erceg et al. (2000) that incorporates both nominal price and wage rigidities. In particular, we stick to a model version which has been successfully analyzed empirically by Rabanal and Rubio-Ramirez (2005) by means of Bayesian techniques. The theoretical framework consists of inter-temporally optimizing households and firms. Monopolistically competitive firms face nominal price rigidities (Calvo pricing). This implies an equation linking price inflation to the gap between real wages and the marginal product of labor. Also households have monopolistic power and set nominal wages in staggered contracts. Nominal wage rigidities translate into a relation linking wage inflation to the gap between the real wage and the marginal rate of substitution of consumption for leisure. The remaining model elements comprise a standard Euler equation for aggregate output as well as an interest rate reaction function reflecting monetary policy. In the following we present the model structure that can be derived by a log-linear approximation around the steady-state symmetric equilibrium with zero price and wage inflation rates (see Erceg et al., 2000; Rabanal and Rubio-Ramirez, 2005, for a detailed derivation).

2.1 The Full Model Structure

The model comprises an Euler equation:

$$y_t = E_t y_{t+1} - \sigma(r_t - E_t \Delta p_{t+1} + E_t g_{t+1} - g_t), \quad (1)$$

where y_t describes output, r_t denotes the nominal interest rate, Δp_t is the inflation

rate, g_t describes a preference shock and the parameter σ is the intertemporal elasticity of substitution.

Then we have a production function and real marginal costs which are written as

$$y_t = a_t + (1 - \delta)n_t, \quad (2)$$

$$mc_t = w_t - p_t + n_t - y_t, \quad (3)$$

where a_t is a technology shock, n_t denotes hours worked and δ is the capital share parameter. mc_t denotes marginal costs given by the wage share of national income ($w - p$ reflect real wages).

The marginal rate of substitution and real wage growth is given by

$$mrs_t = g_t + \frac{1}{\sigma}y_t + \gamma n_t, \quad (4)$$

where mrs_t is the marginal rate of substitution and γ is the elasticity of labor supply. Real wage growth results from wage inflation minus price inflation.

The interest rate reaction function is characterized by

$$r_t = \rho_r r_{t-1} + (1 - \rho_r)(\gamma_\pi \Delta p_t + \gamma_y y_t) + ms_t, \quad (5)$$

where ρ_r is the interest rate smoothing parameter, γ_π and γ_y describe the response of the monetary authority to deviations of inflation and output from their steady-state values, and ms_t is a monetary shock.

The Phillips curve and nominal wage growth are described as

$$\Delta p_t = \gamma_b \Delta p_{t-1} + \gamma_f E_t \Delta p_{t+1} + \kappa_p (mc_t + \lambda_t) \quad (6)$$

$$\Delta w_t = \beta E_t \Delta w_{t+1} + \kappa_w (mrs_t - (w_t - p_t)), \quad (7)$$

where λ describes the mark-up shock and γ_b , γ_f , κ_p and κ_w can be written in terms of the structural parameters β , ω , θ_p , θ_w , and $\bar{\varepsilon}$; which denote the discount parameter, the price indexation parameter, the Calvo parameters for prices and wages and the steady state markup, respectively. In formal terms:

$$\gamma_b = \omega / (1 + \omega\beta) \quad (8)$$

$$\gamma_f = \beta / (1 + \omega\beta) \quad (9)$$

$$\kappa_p = \frac{(1 - \delta)(1 - \theta_p\beta)(1 - \theta_p) / (\theta_p(1 + \delta(\bar{\varepsilon} - 1)))}{1 + \omega\beta} \quad (10)$$

$$\kappa_w = \frac{(1 - \theta_w)(1 - \beta\theta_w)}{\theta_w(1 + \phi\gamma)}. \quad (11)$$

The structural shocks are specified as follows:

$$a_t = \rho_a a_{t-1} + \varepsilon_t^a \quad (12)$$

$$g_t = \rho_g g_{t-1} + \varepsilon_t^g \quad (13)$$

$$ms_t = \varepsilon_t^{ms} \quad (14)$$

$$\lambda_t = \varepsilon_t^\lambda. \quad (15)$$

In here, ρ_i describes the autoregressive parameters (for $i = a, g$) and the innovations ε_t^i are assumed to follow Gaussian normal distributions with mean zero and variance σ_i^2 for $i = a, g, ms, \lambda$. The innovations are assumed to be uncorrelated

with each other.

2.2 The Estimated Model under Misspecification

In order to analyze the effects of misspecification, we estimate several models under wrong structural assumptions. Three different forms are taken into account: omitted wage rigidities, omitted price indexation and misspecified shock processes. In all types of misspecification the model is simulated under the true model, but estimated under wrong assumptions.

For the case where we omit nominal wage rigidity from the model (flexible wages case) the corresponding equation (7) is substituted by:

$$w_t - p_t = mrs_t, \quad (16)$$

such that real wages are always equal to the marginal rate of substitution.

Second, we estimate a model without price indexation (a pure forward-looking price equation). In this model the Phillips curve equation (6) reduces to

$$\Delta p_t = \beta E_t \Delta p_{t+1} + \kappa_p^* (mc_t + \lambda_t) \quad (17)$$

with

$$\kappa_p^* = (1 - \delta)(1 - \theta_p \beta)(1 - \theta_p) / (\theta_p(1 + \delta(\bar{\varepsilon} - 1))). \quad (18)$$

Finally, we look at an even lighter form of misspecification where the first order autoregressive coefficient of the preference shifter shock g_t in the IS curve is esti-

mated as an *i.i.d.* process. Hence, we substitute equation (13) by setting $g_t = \varepsilon_t^g$, with ε_t^g being normally distributed with mean zero and variance σ_g^2 .

2.3 Model Solution

The above outlined linearized model equations represent a dynamic system that describe the joint determination of the endogenous variables. This model structure consists of

$$AE_t Y_{t+1} = BY_t + C_0 X_t + C_1 X_{t+1}, \quad (19)$$

where Y_t and X_t denote the endogenous variables and exogenous (i.e. shocks) system variables, respectively.¹

The rational expectations solution to this system can be found using standard procedures (we made use of the methods provided by King and Watson, 2002; King and Kurmann, 2005). The resulting solution can be written in state space form according to

$$\begin{aligned} Z_t &= \Pi S_t \\ S_t &= MS_{t-1} + Ge_t, \end{aligned} \quad (20)$$

where Z_t is a 15×1 dimensional vector containing all relevant variables of the system (contemporaneous endogenous variables, predetermined variables and the shocks). S_t is of dimension 7×1 and includes the predetermined variables (Δp_{t-1} ,

¹ Y_t and X_t are given by $Y_t = (y_t, \Delta p_t, \Delta w_t, r_t, n_t, mc_t, mrs_t, (w-p)_t, \Delta p_{t-1}, r_{t-1}, (w-p)_{t-1})'$ and $X_t = (g_t, \lambda_t, ms_t, a_t)'$. The matrices A , B , C_0 and C_1 are described in Appendix A.1.

$r_{t-1}, (w-p)_{t-1})'$ and the four shocks $(a_t, g_t, ms_t, \lambda_t)'$. e_t contains the underlying structural shocks of the economy $(\epsilon_t^a, \epsilon_t^g, \epsilon_t^{ms}, \epsilon_t^\lambda)'$. The matrices Π , M and G are functions of the structural parameters and are obtained using the solution algorithm.

3 The Estimation Techniques

3.1 Maximum Likelihood

The full information set up is concerned with estimating the described DSGE model by means of Maximum Likelihood (ML). Therefore, equation 20 is slightly modified to setting up the likelihood function and to estimate the structural parameters of the model. In particular, the Kalman filter is used for this purpose (see e.g. Ireland, 2003; Ireland, 2004; Adolfson and Lindé, 2011, for a similar strategy). For the observation equation we take output, the inflation rate, nominal wage growth and interest rate as vector $d_t = (y_t, \Delta p_t, \Delta w_t, r_t)'$ of observables. Note that the model has four structural shocks which implies that at least 4 variables can be used for estimating the model.² Therefore, the observation equation (see eq. 20) can be reformulated for estimation as:

$$d_t = hZ_t = h\Pi S_t = HS_t, \quad (21)$$

²Due to the problem of stochastic singularity, DSGE models using ML cannot be estimated with more observable variables than structural shocks. Therefore, one has either to omit additional observables or to include error terms to the observation equation. We opt for the first alternative and include the four most prominent variables in line with Rabanal and Rubio-Ramirez (2005).

where $h = [I, 0]$ is a selection matrix with I being a 4×4 identity matrix and d_t is of size 4×1 . Next, we can evaluate the likelihood function $L(\{d_t\}_{t=1}^T | \Upsilon)$, whereas the vector Υ contains the structural model parameters, given the sample of observations $\{d_t\}_{t=1}^T$. Following Rabanal and Rubio-Ramirez (2005) and due to potential identification difficulties, we only estimate a subset of the entire model parameters denoted by $\Upsilon' = [\sigma, \rho_r, \gamma_\pi, \gamma_y, \theta_p, \omega, \theta_w, \rho_g, \rho_a, \sigma_g, \sigma_l, \sigma_m, \sigma_a]'$. The remaining model parameters $\beta, \bar{\epsilon}, \phi, \delta$ and γ are calibrated.

The ML estimator of Υ is obtained as:

$$\hat{\Upsilon}_{ml} = \max_{\Upsilon} \sum_{t=1}^T \log L(d_t | \Upsilon). \quad (22)$$

The Kalman filter algorithm with its subsequent updating and forecasting of S_t is conducted along the lines outlined by Hamilton (1994) Chapter 13. Under standard regularity conditions $\hat{\Upsilon}_{ml}$ is consistent and

$$T^{\frac{1}{2}}(\hat{\Upsilon}_{ml} - \Upsilon_0) \xrightarrow{d} N\left(0, -\bar{\mathfrak{S}}^{-1}\right), \quad (23)$$

where $\text{plim}_{T \rightarrow \infty} T^{-1} \mathfrak{S} = \bar{\mathfrak{S}}$ and \mathfrak{S} describes the Hessian matrix associated with $\log L(\{d_t\}_{t=1}^T | \Upsilon)$. Υ_0 are the true parameter values. This implies a covariance matrix

$$\Sigma = E \left[\frac{\partial^2 \log L(\{d_t\}_{t=1}^T | \Upsilon)}{\partial \Upsilon \partial \Upsilon'} \right]^{-1}, \quad (24)$$

which is evaluated at $\hat{\Upsilon}_{ml}$. Therefore, the standard errors of $\hat{\Upsilon}_{ml}$ are given by the square root of the diagonal elements of $\hat{\Sigma}_{ml}$. Note that ML estimation requires an expression for the final solution of the model (a linearized rational expectation

solution) which is not required for limited information techniques.

3.2 GMM

As an alternative to the full information method, we subsequently present the estimation of the model's structural parameters by limited information methods, namely Generalized Methods of Moments (GMM). Estimation by GMM does not require a rational expectations solution to the model (given by eq. 20).³ Instead, certain moment conditions are enough to obtain parameter estimates for GMM. Additionally, we do not have to consider all aspects (i.e. equations) of a model. A (limited) information subset of equilibrium conditions is sufficient to investigate certain features of a model. It is even possible to estimate one single equation by forming appropriate moment conditions and to ignore all other aspects of the model.

System of equations set up: To take into account more aspects of the model, the GMM framework can be augmented by additional equilibrium conditions of the analyzed DSGE model (see e.g. Christiano and Eichenbaum, 1992; Fuhrer, Moore and Schuh, 1995; Burnside and Eichenbaum, 1996; Ruge-Murcia, 2007, for GMM applications in a multiequation framework). Even when interest lies exclusively on some model parameters, e.g. the NKPC parameters θ_p and ω , due to cross equation restrictions and contemporaneous correlations of the error terms, one would expect efficiency gains by augmenting the information set (i.e. by using additional moment conditions). Therefore, it would be natural to use the complete

³In principle, the estimation strategy is closely related to McCallum's (1976) approach for estimating rational expectations models.

set of Euler equations of the model and to estimate the same parameter vector Υ as with ML. We can basically refer to equation 19 (i.e. the system of equations) that represents the collection of expectational difference equations that depend non-linear on the structural parameters. For the GMM system estimation, we slightly modify the original system to form the orthogonality conditions for GMM, which can be expressed as:

$$E_{t-1} \{ \mathbf{h}(\Upsilon, \mathbf{w}_t) \} = E_{t-1} \{ [\Psi(Y_{t+1}, Y_t, Y_{t-1})] \mathbf{z}_{t-1} \} = 0. \quad (25)$$

The vector Y_t contains all necessary observable variables of the model which is $Y_t = (y_t, \Delta p_t, \Delta w_t, r_t, n_t, (w - p)_t)'$ and thus include two additional observable variables more compared to ML estimation. $\Psi(Y_{t+1}, Y_t, Y_{t-1})$ includes all available moment conditions of the model which equals

$$\Psi(Y_{t+1}, Y_t, Y_{t-1}) = PY_t - QY_{t+1} - RY_{t-1}, \quad (26)$$

where the matrices P , Q and R are of dimension 6×6 in our case and contain all structural parameters Υ of the model.⁴ The instrument set consists of predetermined variables of our system up to lag two, which implies 12 instruments given by $\mathbf{z}_{t-1} = [Y_{t-1}, Y_{t-2}]'$. Furthermore, we take the same instrument set for each equation, which implies that $\mathbf{h}(\Upsilon, \mathbf{w}_t)$ consists of 72 moment conditions (6 equations

⁴In this setting we can take into account 6 observable variables (see Appendix A.2 for the model structure). In general it is also possible to estimate any sub-system of these equation. Additionally, one may reduce the number of variables by means of equation wise substitution. However, this may imply a loss of information which would translate into less efficient estimates.

using 12 instruments for each relation). Note also that we do not specify a certain parametric form of the error terms as it is done in the ML case. Instead, we can allow for flexible forms of autocorrelation and heteroscedasticity by employing a HAC estimator.

Single equation set up: Applying GMM within a single equation framework is by far the most frequent application, even for DSGE type relations (see e.g. Galí and Gertler, 1999; Clarida et al., 2000; Galí, Gertler and López-Salido, 2001; Fuhrer and Rudebusch, 2004). The single equation estimation strategy is simply a special case of the equations (25) and (26), where each equation is estimated separately. As instruments (\mathbf{z}_{t-1}) we include the same set of variables as in our system estimation. Irrespectively whether applied to a single equation or to a multiple equation system, the GMM estimator solves the quadratic minimization problem

$$\hat{\Upsilon}_{GMM} = \min_{\Upsilon} \left[\frac{1}{T} \sum_{t=1}^T \mathbf{h}(\Upsilon, \mathbf{w}_t) \right]' \mathbf{W} \left[\frac{1}{T} \sum_{t=1}^T \mathbf{h}(\Upsilon, \mathbf{w}_t) \right], \quad (27)$$

where \mathbf{w}_t contains either all model variables Y_t (including potential leads and lags) or just the variables of interest in the single relation together with the instrument set \mathbf{z}_{t-1} . \mathbf{W} is a positive-definite weighting matrix. The optimal weighting matrix is the inverse of the variance covariance matrix Ω . To allow $\mathbf{h}(\Upsilon, \mathbf{w}_t)$ to be autocorrelated and heteroskedastic, a consistent estimate of Ω is given by

$$\hat{\Omega} = \hat{\Gamma}_0 + \sum_{v=1}^s \left(1 - \frac{v}{s+1} \right) (\hat{\Gamma}_v + \hat{\Gamma}'_v), \quad \text{with} \quad (28)$$

$$\hat{\Gamma}_v = \frac{1}{T} \sum_{t=v+1}^T \mathbf{h}(\hat{\Upsilon}, \mathbf{w}_t) \mathbf{h}(\hat{\Upsilon}, \mathbf{w}_{t-v})', \quad (29)$$

where the estimator for Ω is due to Newey and West (1987).⁵

Hansen (1982) has shown that this estimator

$$\hat{\Upsilon}_{GMM} \xrightarrow{d} N(\Upsilon_0, (D\Omega^{-1}D')^{-1}), \quad (30)$$

where D is a gradient matrix estimated at $\hat{\Upsilon}$, given by $\hat{D} = \frac{\partial \mathbf{h}(\Upsilon, \mathbf{w}_t)}{\partial \Upsilon'}|_{\Upsilon=\hat{\Upsilon}}$. Then standard errors for significance tests can be obtained using the square root of the diagonal elements of $\hat{\Omega}$.

It is important to note that the optimal weighting matrix $\mathbf{W} = \hat{\Omega}^{-1}$ depends on Υ and therefore Hansen (1982) proposed a multi-step procedure (two steps were often found to be sufficient) to arrive at the final GMM estimates. This *two-step GMM* (2GMM) estimator minimizes the objective function (27) using an identity matrix in the first step and afterwards it is re-optimized using the updated weighting matrix $W = \hat{\Omega}(\hat{\Upsilon}^1)^{-1}$ which is based on the parameter estimates of the first step $\hat{\Upsilon}^1$.⁶ Since it is often found that the 2GMM estimator performs poorly in finite samples and exhibit substantial bias (see Hansen et al. (1996)), we employ the *continuous-updating GMM* estimator (CUGMM) as our baseline estimator. This estimator optimizes the objective function entirely along with the corresponding

⁵In this application we set the bandwidth $s = [T^{1/3}]$. In principle one could try to improve the HAC estimation by automatic bandwidth selection and/or prewhitening (see e.g. Den Haan and Levin, 1997). However this requires other choices to make and turned out to be less robust compared to applying the standard procedure of Newey and West given a fixed bandwidth.

⁶Due to problems of scaling we find it more convenient to work with the initial weighting matrix as $W^1 = (I_N \otimes Z'Z)^{-1}$ which reduces the computational time markedly.

weighting matrix. The additional advantage of the CUGMM estimator is that it is unaffected by transformations of the orthogonality conditions.

4 Monte Carlo Simulation

To investigate the performance of the various estimation procedures, we take the DSGE model as outlined above and generate a sequence of observable variables. More precisely, we solve the model and take the matrices Π , M and G (see equation 20) to construct our data generating process. Let $Z_t = [Y_t, X_t]'$ denote the data matrix comprising the endogenous variables Y_t and the exogenous variables X_t , such that our data can be constructed from:

$$Z_t = \Pi M^t S_0 + \Pi \{G e_t + M G e_{t-1} + \dots + M^{t-1} G e_1\}, \quad (31)$$

with S_0 describing the initial state vector and e_t, \dots, e_1 be vectors of normally distributed shock realizations.

From this, we generate artificial samples of length $1000 + T$ ($T = 400$ and $T = 100$). The first 1000 observations are discarded to avoid sensitivity with respect to the initial state vector. To reach convergence of the parameter distributions, we calculate 2500 different data samples. We fix the structural parameters at values that are in line with the literature (we use similar values as Rabanal and Rubio-Ramirez, 2005, obtained in their analysis). Table 1 presents the values used for the model simulations.

For both GMM and FIML, we employ the Chris Sims optimizer CSMINWEL.

Table 1: Calibrated Parameters

parameter	calibrated value	upper bound	lower bound
σ	0.20	0.01	0.99
$\bar{\epsilon}$	11.0		—
ϕ	6.20		—
β	0.99		—
ω	0.25	0.01	0.99
γ	2.20		—
δ	0.36		—
θ_p	0.65	0.01	0.99
γ_π	1.10	1.01	2.00
ρ_r	0.74	0.01	0.99
θ_w	0.30	0.01	0.99
γ_y	0.40	0.01	0.99
ρ_g	0.80	0	0.99
ρ_a	0.75	0	0.99
σ_g	4.30		positive
σ_l	9.00		positive
σ_{ms}	0.35		positive
σ_a	1.30		positive

Notes: The parameters β , $\bar{\epsilon}$, ϕ , δ and γ are not part of the estimated parameter vector Υ' . Thus, there is no need to specify any boundaries.

Additionally, we work with parameter restrictions as displayed in Table 1. We save the point estimates along with the standard errors from each estimation round. In the first step we obtain the results for the model under the null hypothesis (i.e. estimation is based on the true model structure). The second step consists of introducing certain types of misspecification in the estimation step.

5 Results

In the following, we discuss the small-sample properties of limited and full estimation under different settings. First, we present the results under the null, i.e. when the model is correctly specified. Next, we turn to the model estimates under misspecification. In all cases we compare the point estimates with their true values to investigate the consistency properties. Therefore, we present means and medians as well as their corresponding average bias. Mean square errors (MSE) are computed to summarize the precision of the estimators for each parameter. To compare the performance of the whole system, we also provide a multivariate accuracy measure, which is the trace MSE (TMSE).

Besides point estimation, we are also interested in how well the asymptotic standard errors match the true variability of the estimates in finite samples. To decide whether the estimated standard errors are reliable, we compare the average asymptotic standard error (ASE) with the standard error (SD) of the empirical parameter distribution from our Monte Carlo experiment. Whenever SD is larger than ASE this indicates that the asymptotic formula for the standard error understates the true variability of the estimate in small samples.

Finally, the tables report the size of a standard t -test. The size is the proportion of times that the null hypothesis that a parameter takes the true value is rejected by using a t -test with a nominal size of 5 percent. For a test to be reliable, the effective size (as displayed in the table) should be close to (or at least not larger than) the nominal size (in our case: 0.05).

5.1 Estimation of the True Model

FIML estimation: The full information results under the correctly specified model turn out to be quite precise (see Table 2 and Figure 1 in the Appendix). The estimators match closely their true values and the MSEs are extremely small, even for the smaller sample size of $T = 100$. All distributions turn out to be symmetric and look approximately normal. The only exception is γ_π , where in a number of cases the lower bound is hit and therefore the distribution is somehow truncated. The estimated standard errors turn out to be larger than the empirical ones which implies that the effective size is even smaller than the nominal sizes. Therefore, standard tests turn out to be conservative.

The results are line with those of Ruge-Murcia (2007) and Adolfson and Lindé (2011) who also document quite promising results for FIML in their Monte Carlo studies. We find no indication of potential identification problems (as documented by Canova and Sala, 2009; Iskrev, 2010) that might adversely affect the ML estimation. Even in small samples FIML provide reliable point estimates and confidence intervals.⁷

System GMM estimation: The baseline limited information results under the correctly specified model (based on the systems GMM estimator) are clearly inferior compared to full information estimation (Table 3 and Figure 5 in the Appendix for a graphical representation). Obviously, the dispersion of the point estimates is substantially larger compared to FIML (see SD and MSEs). In some cases the distribution is skewed and clearly deviates from normality. However, the CUGMM

⁷Due to the conservativeness of the confidence intervals based on asymptotic standard errors, there might be room for improved (i.e. more efficient) standard errors. Maybe bootstrapping could offer some further improvements.

Table 2: FIML Estimates (correct specification)

	true	mean	median	bias	MSE	ASE	SD	size
<i>T</i> = 400								
σ	0.20	0.2010	0.1999	0.0010	0.0004	0.0594	0.0189	0.0000
ρ_r	0.74	0.7391	0.7400	-0.0009	0.0003	0.0540	0.0172	0.0000
γ_π	1.10	1.1025	1.0982	0.0025	0.0049	0.0920	0.0699	0.0008
γ_y	0.40	0.4036	0.4000	0.0036	0.0093	0.0341	0.0966	0.0000
ω	0.25	0.2483	0.2484	-0.0017	0.0006	0.0775	0.0247	0.0000
θ_p	0.65	0.6494	0.6497	-0.0006	0.0001	0.0246	0.0106	0.0000
θ_w	0.30	0.2997	0.2998	-0.0003	0.0003	0.0411	0.0169	0.0004
TMSE: 0.0159								
<i>T</i> = 100								
σ	0.20	0.2046	0.2000	0.0046	0.0013	0.0949	0.0364	0.0000
ρ_r	0.74	0.7381	0.7401	-0.0019	0.0010	0.0980	0.0309	0.0000
γ_π	1.10	1.1269	1.0997	0.0269	0.0154	0.1319	0.1212	0.0008
γ_y	0.40	0.4070	0.4004	0.0070	0.0227	0.0464	0.1505	0.0000
ω	0.25	0.2415	0.2440	-0.0085	0.0023	0.1344	0.0475	0.0000
θ_p	0.65	0.6462	0.6486	-0.0038	0.0004	0.0424	0.0200	0.0000
θ_w	0.30	0.2976	0.2983	-0.0024	0.0010	0.0650	0.0309	0.0000
TMSE: 0.0441								

Note: Mean is the arithmetic average of the estimated parameter values. ASE is the average asymptotic standard error; Median and SD are the median and standard deviation of the empirical parameter distribution; Size is an estimate of the actual size of the *t*-test with nominal size of 5 per cent of the null hypothesis that the parameter takes its true value; and MSE is the mean squared error of the parameter estimates. TMSE (trace MSE) is a multivariate generalization of the MSE.

Table 3: System GMM Estimates (correct specification)

	true	mean	median	bias	MSE	ASE	SD	size
<i>T</i> = 400								
σ	0.20	0.1827	0.1851	-0.0173	0.0037	0.1241	0.0584	0.0108
ρ_r	0.74	0.7163	0.7351	-0.0237	0.0069	0.1647	0.0797	0.0332
γ_π	1.10	1.1014	1.0999	0.0014	0.0016	0.1344	0.0396	0.0092
γ_y	0.40	0.4001	0.4002	0.0001	0.0060	0.0560	0.0776	0.0156
ω	0.25	0.2389	0.2455	-0.0111	0.0054	0.2547	0.0726	0.0076
θ_p	0.65	0.7249	0.6799	0.0749	0.0154	0.6748	0.0992	0.0076
θ_w	0.30	0.3616	0.3297	0.0616	0.0223	0.1590	0.1361	0.2561
TMSE: 0.0614								
<i>T</i> = 100								
σ	0.20	0.2083	0.1891	0.0083	0.0179	0.1073	0.1335	0.0756
ρ_r	0.74	0.7128	0.7394	-0.0272	0.0192	0.1530	0.1359	0.0736
γ_π	1.10	1.1096	1.1000	0.0096	0.0076	0.1255	0.0869	0.0128
γ_y	0.40	0.4039	0.4001	0.0039	0.0154	0.0699	0.1243	0.0320
ω	0.25	0.2601	0.2488	0.0101	0.0152	0.2366	0.1229	0.0344
θ_p	0.65	0.7078	0.6728	0.0578	0.0222	0.5925	0.1375	0.0408
θ_w	0.30	0.3790	0.3280	0.0790	0.0416	0.2472	0.1881	0.2105
TMSE: 0.1392								

Note: See Table 2

point estimates are approximately unbiased. Although, the median seems to be closer to the true value compared to the mean (which is not surprising, as long as the CUGMM estimator has no moments, see Hansen et al., 1996). Furthermore, standard significance tests are mostly reliable, even under the small sample size of $T = 100$. The only exception is the test for θ_w , where larger overrejections are detected. Therefore, given the model structure and the prefixed values of some parameters, we find no indication that identification problems invalidates standard hypothesis tests.

When we employ the standard two-step GMM (2GMM) estimator instead of CUGMM to estimate the system, this results in considerable parameter biases (see Figure 8). Even for the larger sample size ($T = 400$), the estimates of the structural parameters turn out to be highly unreliable. In particular the Taylor rule parameters are far from their true value. We can therefore confirm the results of Tauchen (1986) and Kocherlakota (1990) who also find considerable parameter bias of 2GMM in small to moderate sample sizes. In our situation, this suggests that the 2GMM estimator is inappropriate for estimating new Keynesian sticky price models. One explanation may be the inclusion of too many instruments, where some of them are redundant. It is well known that the bias of the 2GMM increases with additional irrelevant instruments, whereas CUGMM does not. This might be one explanation why we find CUGMM to be approximately unbiased and can be recommended in this setting.⁸

Single equation GMM estimation: Estimating the individual equations separately by CUGMM, gives satisfactory results (see Table 4). Clearly — compared

⁸The bias of the 2GMM estimator, although at a more moderate degree, is also found in the single equation setup.

Table 4: Single Equation GMM Estimates

	Eq	true	mean	median	bias	MSE	ASE	SD	size
<i>T</i> = 400									
σ	IS	0.20	0.3091	0.3178	0.1091	0.0264	0.1334	0.1206	0.2768
ρ_r	TR	0.74	0.7347	0.7330	-0.0053	0.0005	0.4988	0.0216	0.0000
γ_π	TR	1.10	1.1004	1.1001	0.0004	0.0001	0.3332	0.0113	0.0008
γ_y	TR	0.40	0.4012	0.4008	0.0012	0.0012	0.3480	0.0343	0.0008
ω	PC	0.25	0.2605	0.2596	0.0105	0.0007	1.2233	0.0265	0.0000
θ_p	PC	0.65	0.6482	0.6472	-0.0018	0.0008	1.7194	0.0279	0.0000
θ_w	WC	0.30	0.3410	0.2041	0.0410	0.0803	0.5548	0.2805	0.0616
<i>T</i> = 100									
σ	IS	0.20	0.2673	0.2822	0.0673	0.0414	0.2653	0.1922	0.2684
ρ_r	TR	0.74	0.7212	0.7259	-0.0188	0.0055	0.8397	0.0719	0.0032
γ_π	TR	1.10	1.1013	1.1001	0.0013	0.0012	0.5707	0.0339	0.0008
γ_y	TR	0.40	0.4020	0.4006	0.0020	0.0049	0.6055	0.0702	0.0004
ω	PC	0.25	0.2756	0.2519	0.0256	0.0145	1.9186	0.1178	0.0068
θ_p	PC	0.65	0.6570	0.6530	0.0070	0.0088	3.1762	0.0938	0.0112
θ_w	WC	0.30	0.4700	0.3300	0.1700	0.1441	1.1628	0.3395	0.0928

Note: See Table 2. The parameter estimates are obtained from different single equation models: the dynamic IS curve (IS), the Taylor rule (TR), the Phillips curve (PC) and the wage curve (WC).

to FIML — the estimates are more dispersed. Generally, the parameter estimates are more sensitive in comparison with the system estimation approach. However, the parameters of the Phillips curve and the Taylor rule are estimated with even higher precision (which might explain why single equation estimation of these two relationships has remained quite popular). For the IS curve and the wage equation, single equation estimates are less reliable compared to their system GMM counterparts. This loss of efficiency may come from ignoring the cross equation restriction of σ , which shows up in both equations. Here, single equation estimates turn out to be more biased. Estimated standard errors are close or larger than their true value. Furthermore, it is important to recognize that single equation estimates are per construction less prone to misspecification (since they ignore all other parts of the model).

5.2 Dropping Wage Rigidity

In the first case of misspecification, we omit nominal wage rigidity from the model (flexible wages case). When we look at the results of the FIML procedure our results change radically. This form of misspecification introduces large biases in the parameter estimates (see Table 5 and Figure 2). In this case, the omission of one model relation leads to inconsistent estimates everywhere in the model. All relevant model parameters seem to be affected. In the ML case, σ for instance is highly upward biased and often lays at the upper boundary. The distributions are highly asymmetric, but still consider not much variation (as in the case of correct specification). The large biases together with nonstandard distributions clearly invalidates any hypothesis test. More observations do not help to circumvent any

Table 5: FIML Estimates (ignoring wage rigidity)

	true	mean	median	bias	MSE	ASE	SD	size
$T = 400$								
σ	0.20	0.8764	0.9427	0.6764	0.4832	0.2643	0.1602	0.7364
ρ_r	0.74	0.6681	0.6717	-0.0719	0.0068	0.0234	0.0401	0.7844
γ_π	1.10	1.0010	1.0010	-0.0990	0.0098	0.0120	0.0000	0.9352
γ_y	0.40	0.3872	0.3859	-0.0128	0.0018	0.0053	0.0405	0.7088
ω	0.25	0.3040	0.2785	0.0540	0.0107	0.0123	0.0880	0.7080
θ_p	0.65	0.5245	0.5246	-0.1255	0.0162	0.0260	0.0213	0.9924
TMSE: 0.5284								
$T = 100$								
σ	0.20	0.8789	0.9459	0.6789	0.4860	0.2743	0.1584	0.7252
ρ_r	0.74	0.6678	0.6714	-0.0722	0.0068	0.0242	0.0393	0.7852
γ_π	1.10	1.0010	1.0010	-0.0990	0.0098	0.0128	0.0000	0.9256
γ_y	0.40	0.3880	0.3860	-0.0120	0.0021	0.0054	0.0444	0.6992
ω	0.25	0.3054	0.2792	0.0554	0.0108	0.0128	0.0879	0.7112
θ_p	0.65	0.5248	0.5244	-0.1252	0.0161	0.0260	0.0216	0.9916
TMSE: 0.5316								

Note: See Table 2.

Table 6: System GMM Estimates (ignoring wage rigidity)

	true	mean	median	bias	MSE	ASE	SD	size
$T = 400$								
σ	0.20	0.1604	0.1816	-0.0396	0.0054	0.0774	0.0618	0.0084
ρ_r	0.74	0.7002	0.7048	-0.0398	0.0071	0.1531	0.0744	0.0256
γ_π	1.10	1.1024	1.0998	0.0024	0.0023	0.1327	0.0476	0.0132
γ_y	0.40	0.4020	0.4058	0.0020	0.0062	0.0705	0.0784	0.0184
ω	0.25	0.2677	0.2508	0.0177	0.0072	0.0917	0.0832	0.1096
θ_p	0.65	0.6694	0.6571	0.0194	0.0046	0.1678	0.0651	0.0460
TMSE: 0.0328								
$T = 100$								
σ	0.20	0.1977	0.1897	-0.0023	0.0216	0.0831	0.1470	0.1012
ρ_r	0.74	0.6971	0.7281	-0.0429	0.0337	0.1744	0.1784	0.1273
γ_π	1.10	1.1167	1.0999	0.0167	0.0122	0.1380	0.1092	0.0336
γ_y	0.40	0.4043	0.4005	0.0043	0.0211	0.0954	0.1452	0.0548
ω	0.25	0.2871	0.2515	0.0371	0.0322	0.1013	0.1757	0.1349
θ_p	0.65	0.6651	0.6550	0.0151	0.0215	0.1771	0.1460	0.1208
TMSE: 0.1423								

Note: See Table 2

of these problems. This result suggest that estimating an oversimplified model can be very dangerous and may lead to wrong conclusions about the structural parameters of the model.

Estimates based on GMM are much less affected by this form of misspecification (see Table 6 and Figure 6). Many of the estimates remain approximately unbiased. Interestingly, the precision of some parameter estimates remain more or less the same. Therefore, including the wage curve does not much improve the estimation precision. But, significance tests, at least in case of $T = 100$, suffer slightly from the omitted wage equation. This kind of misspecification has practically only little consequences on GMM estimation. The superiority of GMM under this form of misspecification is not surprising since the specified moment conditions remain valid. By omitting the wage equation, we just exclude some source of information.

5.3 Omitted Price Indexation

In this type of misspecification where price indexation is omitted, in contrast to the case of omitted wage rigidities, the estimated model is a nested version of the true model (where the indexation parameter ω is zero).

Under this form of misspecification FIML remains unreliable and most parameter estimates exhibit substantial biases (see Table 7 and Figure 3). Only the parameters of the Taylor rule seem to be less affected by this type of misspecification. For the Phillips curve and the wage curve, the omitted lagged inflation term leads to a downward bias of the Calvo parameters θ_p and θ_w . While σ in the IS curve is upward biased.

The GMM estimates remain again largely unaffected by this form of misspecifica-

Table 7: FIML Estimates (omitted price indexation)

	true	mean	median	bias	MSE	ASE	SD	size
$T = 400$								
σ	0.20	0.5808	0.5746	0.3808	0.1498	0.2041	0.0697	0.1956
ρ_r	0.74	0.6995	0.6999	-0.0405	0.0019	0.0596	0.0171	0.0016
γ_π	1.10	1.1190	1.1215	0.0190	0.0047	0.1101	0.0662	0.1488
γ_y	0.40	0.3333	0.3202	-0.0667	0.0091	0.0735	0.0683	0.0156
θ_p	0.65	0.5891	0.5892	-0.0609	0.0038	0.0287	0.0081	0.7024
θ_w	0.30	0.2219	0.2217	-0.0781	0.0062	0.0523	0.0112	0.0388
TMSE: 0.1750								
$T = 100$								
σ	0.20	0.5785	0.5734	0.3785	0.1650	0.3996	0.1474	0.0164
ρ_r	0.74	0.7002	0.7019	-0.0398	0.0028	0.1070	0.0351	0.0000
γ_π	1.10	1.1360	1.1178	0.0360	0.0161	0.1211	0.1217	0.0948
γ_y	0.40	0.3416	0.3370	-0.0584	0.0210	0.1350	0.1328	0.0020
θ_p	0.65	0.5882	0.5887	-0.0618	0.0041	0.0544	0.0161	0.0184
θ_w	0.30	0.2257	0.2221	-0.0743	0.0062	0.0987	0.0256	0.0124
TMSE: 0.2151								

Note: See Table 2.

Table 8: System GMM Estimates (omitted price indexation)

	true	mean	median	bias	MSE	ASE	SD	size
$T = 400$								
σ	0.20	0.1792	0.1831	-0.0208	0.0030	0.1208	0.0510	0.0096
ρ_r	0.74	0.7159	0.7346	-0.0241	0.0064	0.1606	0.0766	0.0448
γ_π	1.10	1.1019	1.0998	0.0019	0.0023	0.1192	0.0481	0.0128
γ_y	0.40	0.3986	0.4002	-0.0014	0.0056	0.0624	0.0748	0.0176
θ_p	0.65	0.7017	0.6693	0.0517	0.0126	0.0893	0.0996	0.0620
θ_w	0.30	0.3684	0.3341	0.0684	0.0247	0.1455	0.1417	0.2353
TMSE: 0.0741								
$T = 100$								
σ	0.20	0.2116	0.1891	0.0116	0.0172	0.0936	0.1308	0.0908
ρ_r	0.74	0.7171	0.7379	-0.0229	0.0147	0.1422	0.1192	0.0932
γ_π	1.10	1.1071	1.1000	0.0071	0.0064	0.1074	0.0799	0.0172
γ_y	0.40	0.3990	0.4004	-0.0010	0.0111	0.0726	0.1053	0.0296
θ_p	0.65	0.6902	0.6659	0.0402	0.0199	0.1076	0.1352	0.1437
θ_w	0.30	0.3766	0.3189	0.0766	0.0428	0.2929	0.1923	0.2097
TMSE: 0.1122								

Note: See Table 2

tion (see Table 8 and Figure 7). This result might be surprising since in contrast to the case where we estimated a flexible wage model, this form of model misspecification implies invalid moment conditions (as a lagged inflation term is included in the instrument set, although it should be included in the equation). However, the results are in line with Galí and Gertler (1999) who also estimate a Phillips curve with and without a lagged inflation term, but the results for θ hardly changed. There is even an improvement of the precision of the estimate for θ_p , since ω could not be estimated correctly even under the true model. Therefore, we can conclude that for limited information methods it is sufficient to use a forward looking term in the Phillips curve. The indexation term, although significant, has practically no implications for the model.

5.4 Misspecified Shocks

Finally, we look at an even lighter form of misspecification where the first order autoregressive coefficient of the disturbance term is omitted. When estimating this model, we find that the biases in the FIML estimation remain a problem (see Table 9 and Figure 4). But the parameters are affected to various degrees. The parameters in the price and wage relations remain approximately unbiased (which is in line with Lindé, 2005, who find that misspecified shocks have no quantitative impact on FIML parameter estimates of the Phillips curve), while the IS relation with the estimate for σ is affected the most. Also the inflation response parameter γ_π and the interest smoothing parameter ρ_r in the Taylor rule are biased downward. However, this form of misspecification has by far the smallest implication. But inference seems to be a problem for the parameters in

Table 9: FIML Estimates (misspecified shock)

	true	mean	median	bias	MSE	ASE	SD	size
<i>T</i> = 400								
σ	0.20	0.1205	0.1172	-0.0795	0.0068	0.0504	0.0216	0.2032
ρ_r	0.74	0.6971	0.6982	-0.0429	0.0022	0.0445	0.0178	0.0708
γ_π	1.10	1.0013	1.0010	-0.0987	0.0098	0.0256	0.0054	0.9928
γ_y	0.40	0.4013	0.4167	0.0013	0.0038	0.0141	0.0619	0.2848
ω	0.25	0.1791	0.1760	-0.0709	0.0065	0.1046	0.0388	0.0048
θ_p	0.65	0.6442	0.6449	-0.0058	0.0002	0.0156	0.0110	0.0012
θ_w	0.30	0.2913	0.2901	-0.0087	0.0006	0.0307	0.0238	0.0880
TMSE: 0.0299								
<i>T</i> = 100								
σ	0.20	0.1284	0.1277	-0.0716	0.0058	0.0897	0.0267	0.0000
ρ_r	0.74	0.6937	0.6959	-0.0463	0.0031	0.0812	0.0305	0.0040
γ_π	1.10	1.0067	1.0010	-0.0933	0.0097	0.0517	0.0321	0.4680
γ_y	0.40	0.4025	0.4180	0.0025	0.0074	0.0248	0.0861	0.0196
ω	0.25	0.1769	0.1845	-0.0731	0.0083	0.1703	0.0547	0.0000
θ_p	0.65	0.6410	0.6433	-0.0090	0.0005	0.0317	0.0208	0.0004
θ_w	0.30	0.2824	0.2790	-0.0176	0.0010	0.0559	0.0257	0.0052
TMSE: 0.0359								

Note: See Table 2.

the IS curve and the Taylor rule which turn out to be largely unreliable. Per construction, this form of misspecification leaves the GMM estimates unaffected since we left the error process unspecified in our GMM strategy.

6 Conclusion

In this paper, we systematically investigated GMM and FIML under several forms of model misspecification. We employed a standard New Keynesian model including nominal price and wage rigidities (as in Erceg et al., 2000; Rabanal and Rubio-Ramirez, 2005). By estimating the correct model structure, we can show that FIML provides superior results in small samples. While GMM also provides unbiased estimates (when using a CUGMM estimator), the dispersion of this estimator is larger but standard asymptotic theory for inference works satisfactory in most cases. However, when the misspecified model is estimated, FIML estimates are substantially biased and significance tests highly unreliable. At the same time, CUGMM remains roughly unbiased. Hence, this paper partly rehabilitates the use of limited information methods for evaluating DSGE models. We find that they are mainly unaffected by typical model misspecifications which can be of great advantage in practice where model uncertainty is high. In particular, the use of the CUGMM estimator results in mostly unbiased estimates. Furthermore, it is often advantageous to evaluate several model equations simultaneously by using a system-based CUGMM estimator. In general, this paper suggests that limited information methods should be used at least as supplementary tools in the empirical evaluation of DSGE models, since the blind use of full information techniques might lead to wrong conclusions about the structural parameters of a model.

A Appendix

A.1 Model Matrices

Note that the linearized model can be cast in the form:

$$AE_t Y_{t+1} = BY_t + C_0 X_t + C_1 X_{t+1}, \quad (32)$$

where Y_t and X_t denotes the endogenous variables and exogenous (i.e. shocks) system variables, respectively. The matrices A , B , C_0 and C_1 are defined as

$$A = \begin{pmatrix} 1 & \sigma & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma_f & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, C_0 = \begin{pmatrix} -\sigma & 0 & 0 & 0 \\ 0 & -\kappa_p & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, C_1 = \begin{pmatrix} \sigma & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$B = \begin{pmatrix} 1 & 0 & 0 & \sigma & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1+1 & 0 & 0 & 0 & -\kappa_p & 0 & 0 & -\gamma_b & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -\kappa_w & \kappa_w & 0 & 0 & 0 \\ (1-\rho_r)\gamma_y & (1-\rho_r)\gamma_\pi & 0 & -1+1 & 0 & 0 & 0 & 0 & 0 & \rho_r & 0 \\ -1 & 0 & 0 & 0 & (1-\delta) & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & -1 & 0 & 1 & 0 & 0 & 0 \\ \frac{1}{\sigma} & 0 & 0 & 0 & \gamma & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & -1+1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

A.2 The System for GMM Estimation

For the system GMM estimator, we collect equations 1 to 7 and reformulate the stochastic equations including the relation for real wages to

$$y_t = y_{t+1} - \sigma(r_t - \Delta p_{t+1}) + u_t^y \quad (33)$$

$$\Delta p_t = \gamma_f \Delta p_{t+1} + \gamma_b \Delta p_{t-1} + \kappa_p ((w - p)_t + n_t - y_t) + u_t^p, \quad (34)$$

$$r_t = \rho_r r_{t-1} + (1 - \rho_r) [\gamma_\pi \Delta p_t + \gamma_y y_t] + u_t^r, \quad (35)$$

$$\Delta w_t = \beta \Delta w_{t+1} + \kappa_w \left(\frac{1}{\sigma} y_t + \gamma n_t - (w - p)_t \right) + u_t^w, \quad (36)$$

$$n_t = \frac{1}{1 - \delta} y_t + u_t^n, \quad (37)$$

$$(w - p)_t = \Delta w_t - \Delta p_t + (w - p)_{t-1}. \quad (38)$$

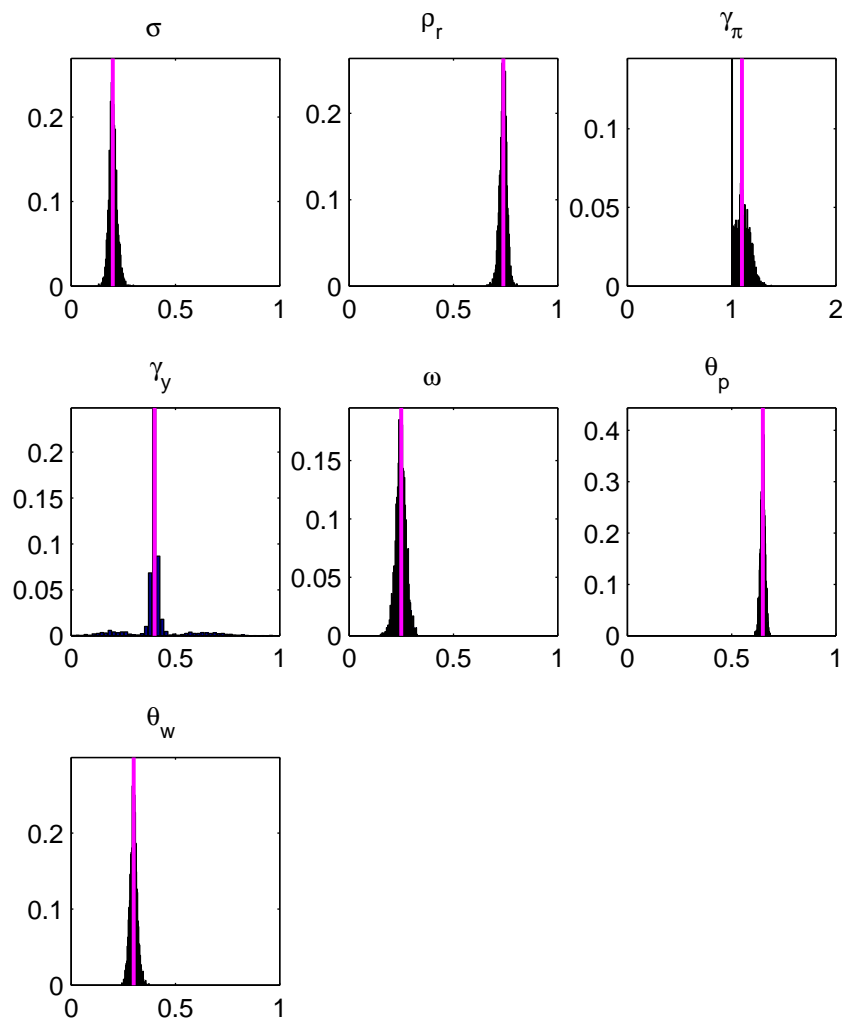
Note that we replaced the expectation term by its future realized value by assuming $E_t x_{t+1} = x_{t+1} + \nu_{t+1}$, which implies that expectations are unbiased predictions. Further it is assumed that the expectation error ν_{t+1} is orthogonal to past information (i.e. instrument validity of the instrument vector \mathbf{z}_t). This implies, that the residuals u_t^y , u_t^p , u_t^r , u_t^n and u_t^w do not only depend on the structural shocks of the model, but also contain the expectational errors.

In matrix notation the model structure is $PY_t = QY_{t+1} + RY_{t-1} + U_t$, which can be represented as

$$\begin{pmatrix} 1 & 0 & \sigma & 0 & 0 & 0 \\ \kappa_p & 1 & 0 & 0 & -\kappa_p & -\kappa_p \\ -(1-\rho_r)\gamma_y & -(1-\rho_r)\gamma_\pi & 1 & 0 & 0 & 0 \\ -\kappa_w \frac{1}{\sigma} & 0 & 0 & 1 & -\kappa_w \gamma & \kappa_w \\ -\frac{1}{1-\delta} & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \end{pmatrix} \begin{bmatrix} y_t \\ \Delta p_t \\ r_t \\ \Delta w_t \\ n_t \\ (w-p)_t \end{bmatrix} = \\
\begin{pmatrix} 1 & \sigma & 0 & 0 & 0 & 0 \\ 0 & \gamma_f & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{bmatrix} y_{t+1} \\ \Delta p_{t+1} \\ r_{t+1} \\ \Delta w_{t+1} \\ n_{t+1} \\ (w-p)_{t+1} \end{bmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma_b & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_r & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} y_{t-1} \\ \Delta p_{t-1} \\ r_{t-1} \\ \Delta w_{t-1} \\ n_{t-1} \\ (w-p)_{t-1} \end{bmatrix} + \begin{bmatrix} u_t^y \\ u_t^p \\ u_t^r \\ u_t^w \\ u_t^n \\ 0 \end{bmatrix}.$$

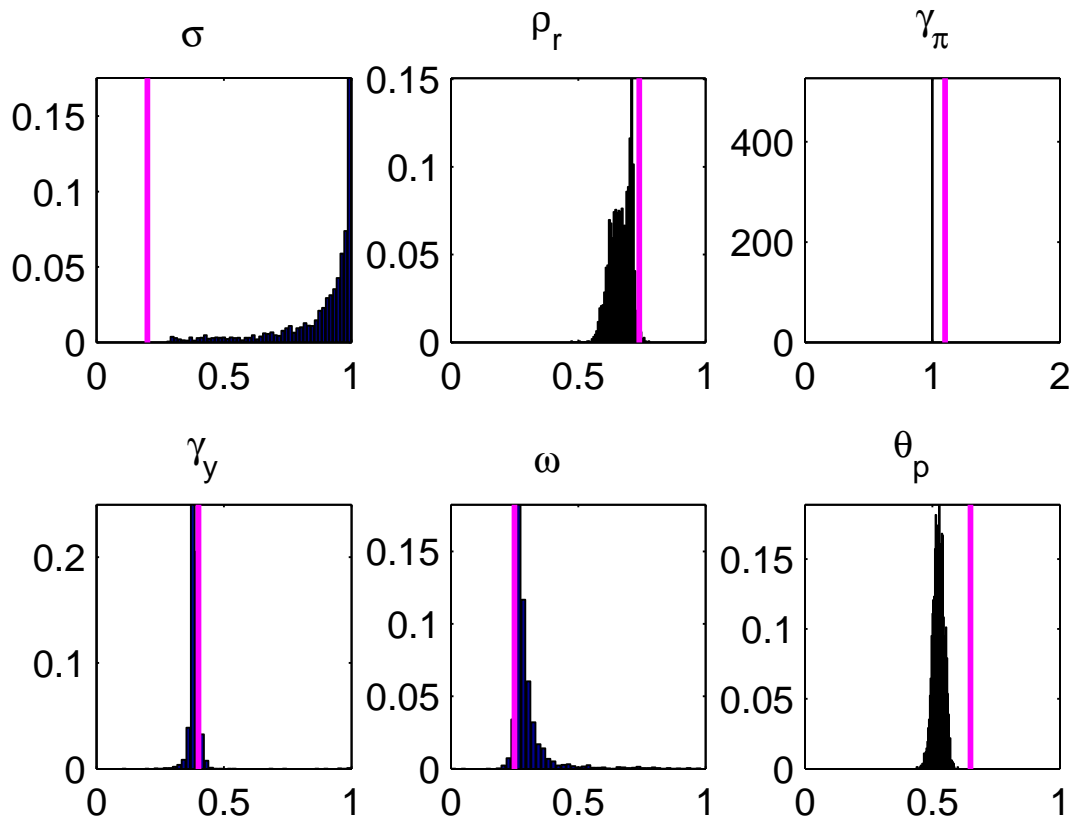
A.3 Figures

Figure 1: FIML Estimates (correct specification)



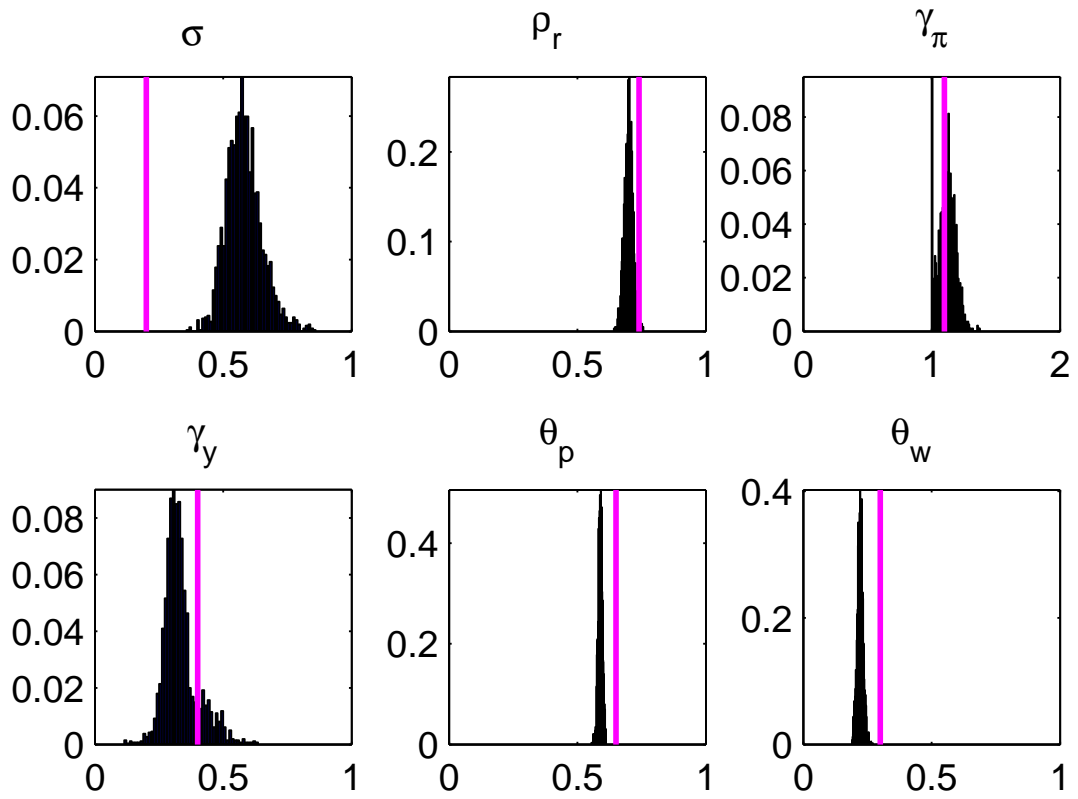
Notes: Small sample distribution of the estimated structural parameters. The true parameters are given by the vertical bars. $T = 400$ observations in each of the $N = 2500$ artificial samples.

Figure 2: FIML Estimates (ignoring wage rigidity)



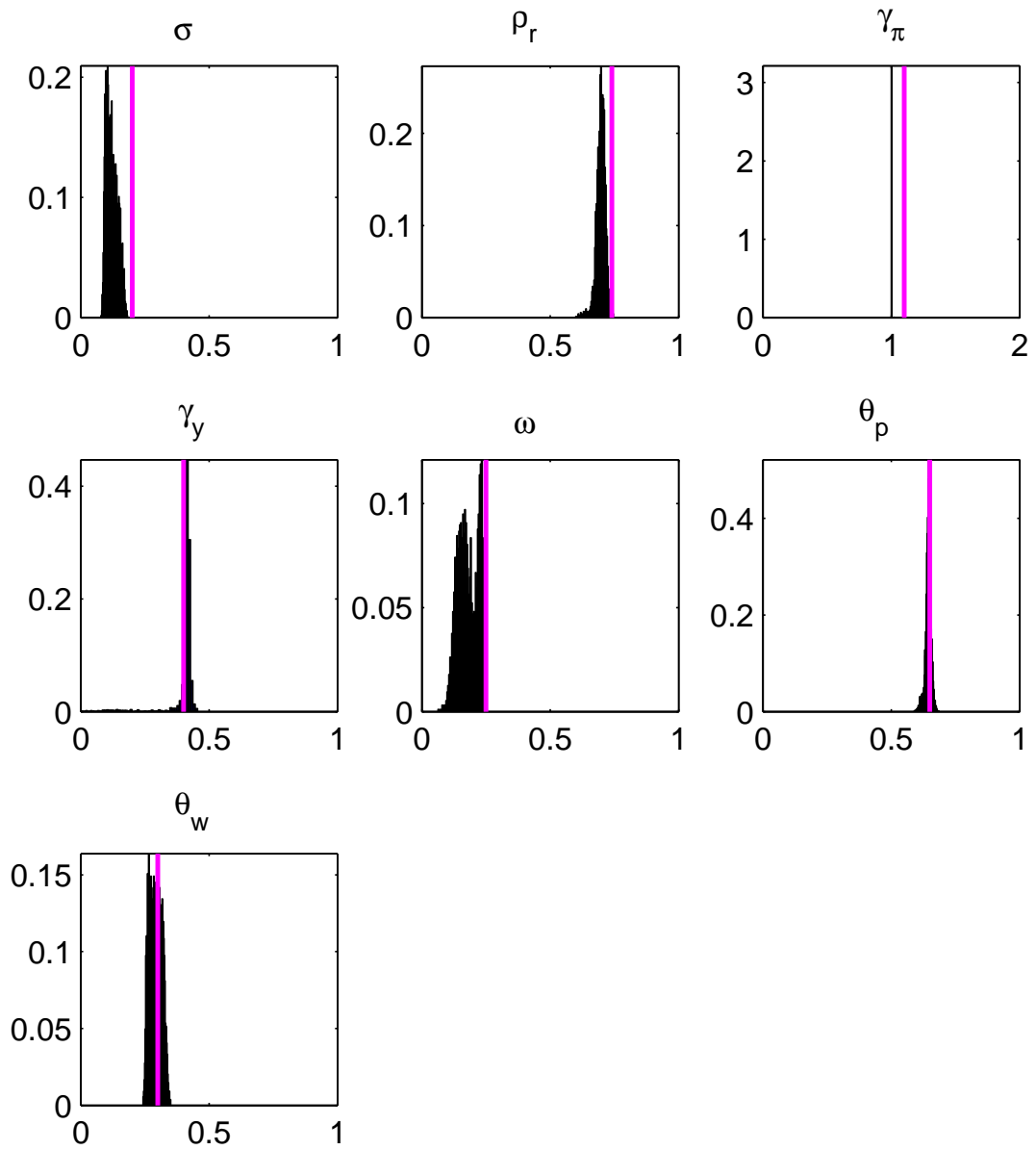
Notes: See Figure 1

Figure 3: FIML Estimates (omitted price indexation)



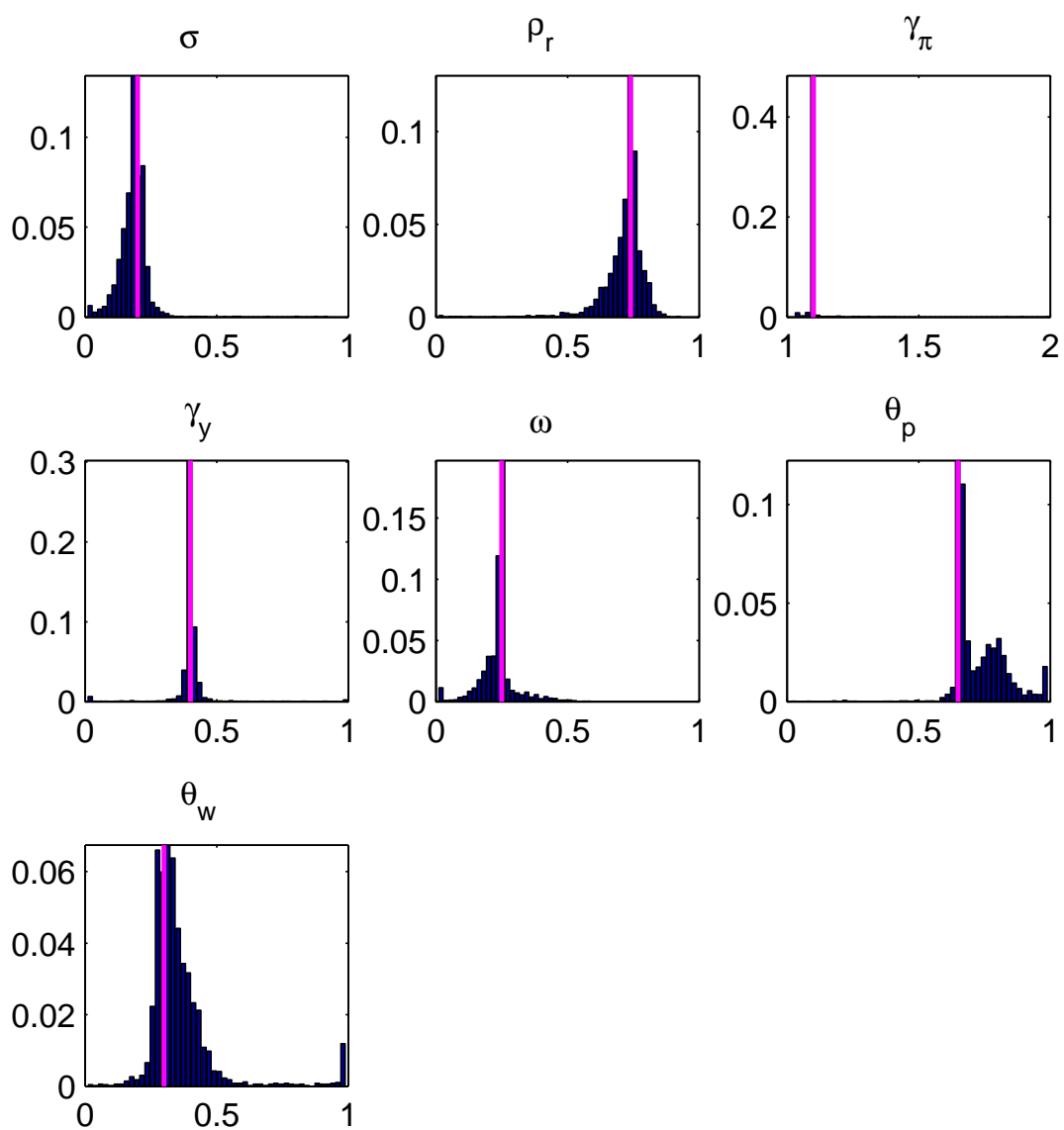
Notes: See Figure 1

Figure 4: FIML Estimates (misspecified shock)



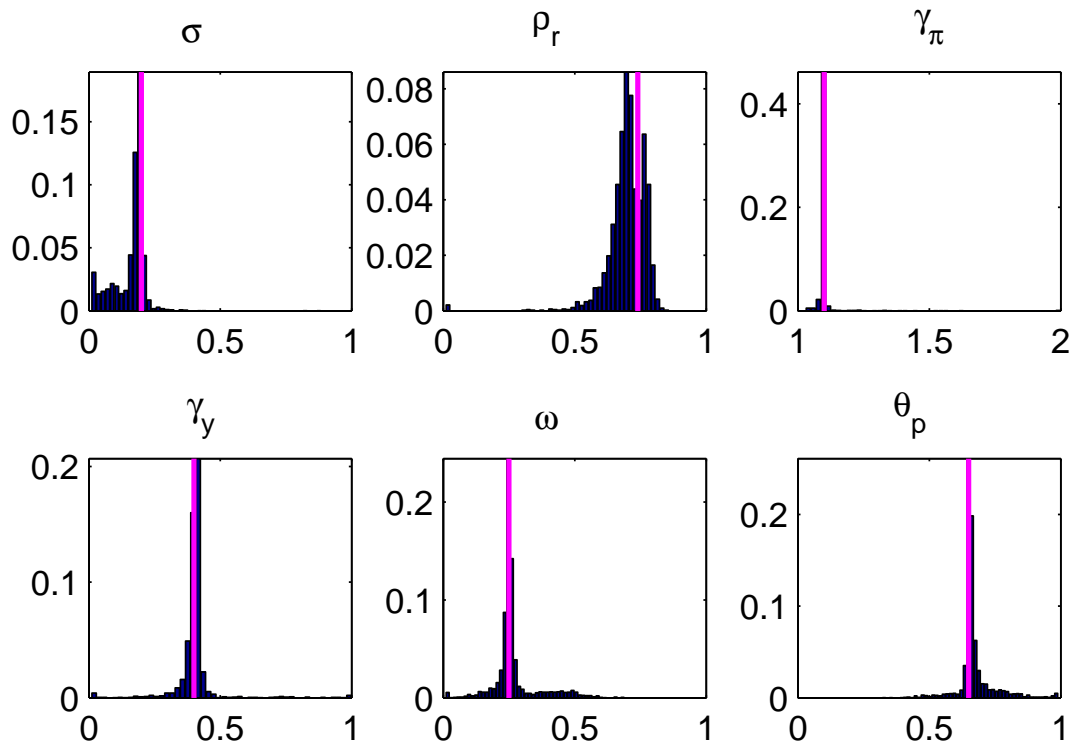
Notes: See Figure 1

Figure 5: System GMM Estimates (correct specification)



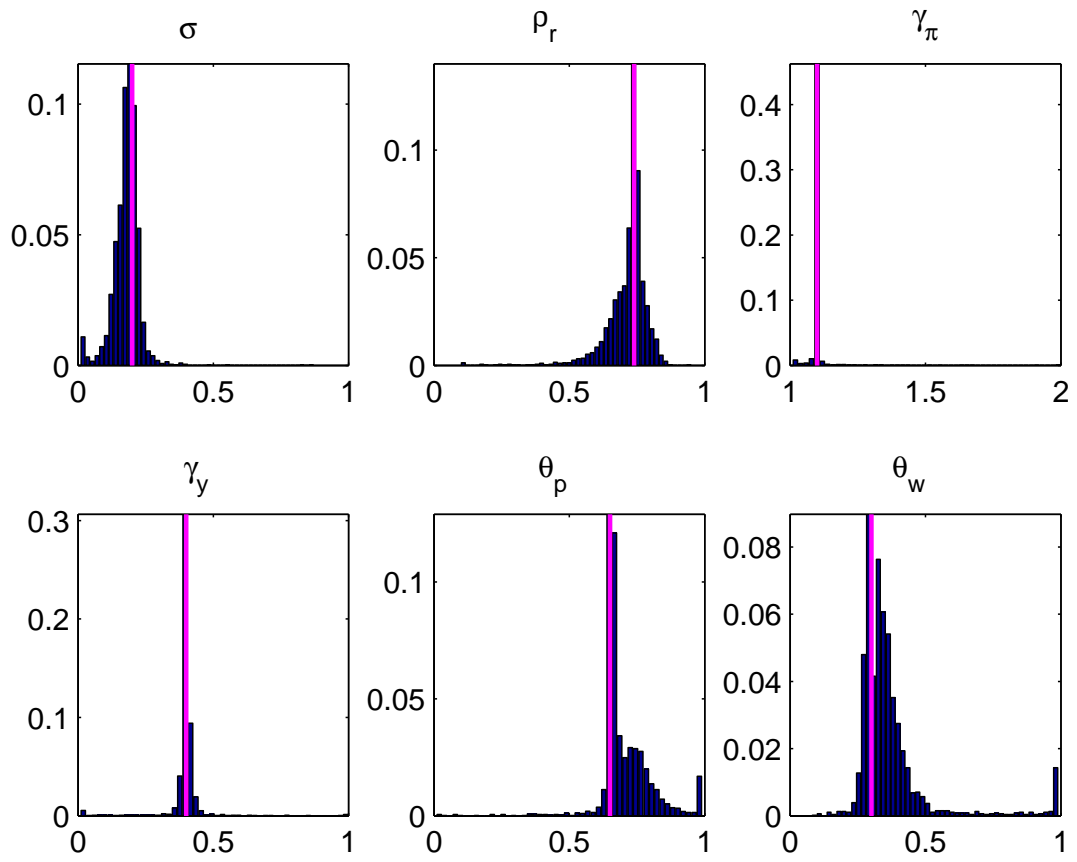
Notes: See Figure 1

Figure 6: System GMM Estimates (ignoring wage rigidity)



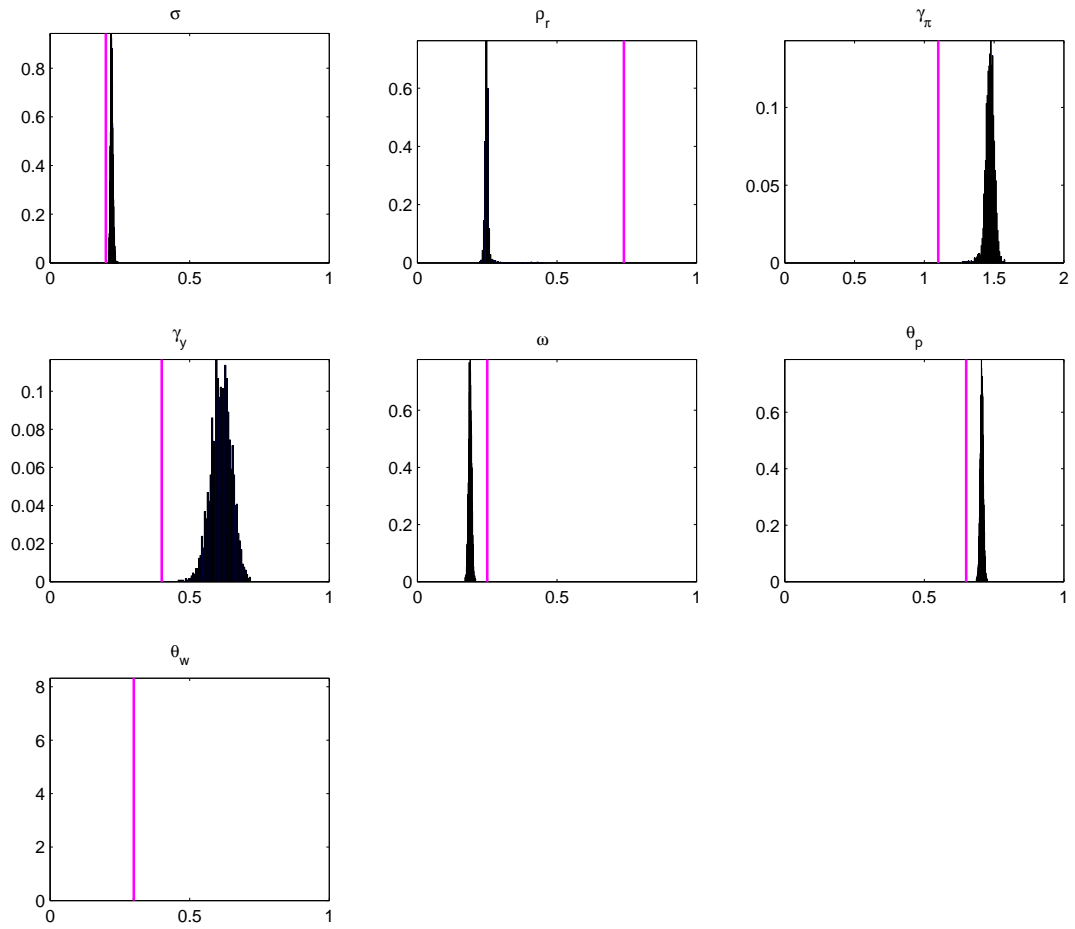
Notes: See Figure 1

Figure 7: System GMM Estimates (omitted price indexation)



Notes: See Figure 1

Figure 8: Two-step GMM Estimates (correct specification)



Notes: See Figure 1

References

- Adolfson, M. and Lindé, J. (2011). Parameter Identification in an Estimated New Keynesian Open Economy Model, *Working Paper Series 251*, Sveriges Riksbank.
- An, S. and Schorfheide, F. (2007). Bayesian Analysis of DSGE Models, *Econometric Reviews* **26**(2-4): 113–172.
- Andrews, I. and Mikusheva, A. (2011). Maximum Likelihood Inference in Weakly Identified DSGE Models, *Working Paper 11-03*, MIT Department of Economics.
- Burnside, C. and Eichenbaum, M. (1996). Small-Sample Properties of GMM-Based Wald Tests, *Journal of Business and Economic Statistics* **14**(3): 294–308.
- Canova, F. (2007). *Methods for Applied Macroeconomic Research*, Princeton University Press.
- Canova, F. and Sala, L. (2009). Back to square one: Identification issues in DSGE models, *Journal of Monetary Economics* **56**(4): 431–449.
- Christiano, L. and Eichenbaum, M. (1992). Current Real-Business-Cycle Theories and Aggregate Labor Market Fluctuations, *American Economic Review* **82**(3): 430–450.
- Christiano, L. J., Eichenbaum, M. and Evans, C. L. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy, *Journal of Political Economy* **113**(1): 1–45.

- Clarida, R., Galí, J. and Gertler, M. (2000). Monetary Policy Rules And Macroeconomic Stability: Evidence And Some Theory, *The Quarterly Journal of Economics* **115**(1): 147–180.
- Cragg, J. G. (1967). On the Relative Small-Sample Properties of Several Structural-Equation Estimators, *Econometrica* **35**(1): 89–110.
- Cragg, J. G. (1968). Some Effects of Incorrect Specification on the Small-Sample Properties of Several Simultaneous-Equation Estimators, *International Economic Review* **9**(1): 63–86.
- Den Haan, W. J. and Levin, A. (1997). A practitioner’s guide to robust covariance matrix estimation, in *Handbook of Statistics: Robust Inference (Vol 15)*, North-Holland, pp. 291–341.
- Dufour, J.-M., Khalaf, L. and Kichian, M. (2006). Inflation dynamics and the New Keynesian Phillips Curve: An identification robust econometric analysis, *Journal of Economic Dynamics and Control* **30**(9-10): 1707–1727.
- Dufour, J.-M., Khalaf, L. and Kichian, M. (2012). Identification-Robust Analysis of DSGE and Structural Macroeconomic Models.
- Eichenbaum, M. and Fisher, J. D. M. (2007). Estimating the Frequency of Price Re-Optimization in Calvo-Style Models, *Journal of Monetary Economics* **54**(7): 2032–2047.
- Erceg, C. J., Henderson, D. W. and Levin, A. T. (2000). Optimal Monetary Policy with Staggered Wage and Price Contracts, *Journal of Monetary Economics* **46**(2): 281–313.

- Fuhrer, J. C., Moore, G. R. and Schuh, S. D. (1995). Estimating the Linear-Quadratic Inventory Model Maximum Likelihood versus Generalized Methods of Moments, *Journal of Monetary Economics* **35**(1): 115–157.
- Fuhrer, J. C. and Rudebusch, G. D. (2004). Estimating the Euler Equation for Output, *Journal of Monetary Economics* **51**(6): 1133–1153.
- Galí, J. and Gertler, M. (1999). Inflation Dynamics: A Structural Econometric Analysis, *Journal of Monetary Economics* **44**(2): 195–222.
- Galí, J., Gertler, M. and López-Salido, J. D. (2001). European Inflation Dynamics, *European Economic Review* **45**(7): 1237–1270.
- Guerron-Quintana, P., Inoue, A. and Kilian, L. (2012). Frequentist Inference in Weakly Identified DSGE Models.
- Hale, C., Mariano, R. S. and Ramage, J. G. (1980). Finite Sample Analysis of Misspecification in Simultaneous Equation Models, *Journal of the American Statistical Association* **75**(370): 418–427.
- Hamilton, J. D. (1994). *Time Series Analysis*, Princeton University Press.
- Hansen, L. P. (1982). Large Sample Properties of Generalized Methods of Moments Estimators, *Econometrica* **50**(4): 1029–1054.
- Hansen, L. P., Heaton, J. and Yaron, A. (1996). Finite Sample Properties of Some Alternative GMM Estimators, *Journal of Business & Economic Statistics* **14**(3): 262–280.
- Ireland, P. N. (2003). Endogenous Money or Sticky Prices?, *Journal of Monetary Economics* **50**(8): 1623–1648.

- Ireland, P. N. (2004). A Method for Taking Models to the Data, *Journal of Economic Dynamics and Control* **28**(6): 1205–1226.
- Iskrev, N. (2010). Local identification in DSGE models, *Journal of Monetary Economics* **57**(2): 189–202.
- Jondeau, E. and Le Bihan, H. (2008). Examining Bias in Estimators of Linear Rational Expectations Models under Misspecification, *Journal of Econometrics* **143**(2): 375 – 395.
- King, R. G. and Kurmann, A. (2005). Solving Linear Rational Expectations Models.
- King, R. G. and Watson, M. W. (2002). System Reduction and Solution Algorithms for Singular Linear Difference Systems under Rational Expectations , *Computational Economics* **20**(1-2): 57–86.
- Kleibergen, F. and Mavroeidis, S. (2009). Weak instrument robust tests in GMM and the New Keynesian Phillips curve, *Journal of Business and Economic Statistics* **27**(3): 293–339.
- Kocherlakota, N. R. (1990). On Tests of Representative Consumer Asset Pricing Models, *Journal of Monetary Economics* **26**(2): 285–304.
- Lindé, J. (2005). Estimating New-Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach, *Journal of Monetary Economics* **52**(6): 1135–1149.
- McCallum, B. T. (1976). Rational Expectations and the Natural Rate Hypothesis: Some Consistent Estimates, *Econometrica* **44**(1): 43–52.

- Newey, W. K. and West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* **55**(3): 703–708.
- Rabanal, P. and Rubio-Ramirez, J. F. (2005). Comparing New Keynesian Models of the Business Cycle: A Bayesian Approach, *Journal of Monetary Economics* **52**(6): 1151–1166.
- Ruge-Murcia, F. J. (2007). Methods to Estimate Dynamic Stochastic General Equilibrium Models, *Journal of Economic Dynamics and Control* **31**(8): 2599–2636.
- Schorfheide, F. (2011). Estimation and Evaluation of DSGE Models: Progress and Challenges, *NBER Working Papers 16781*, National Bureau of Economic Research.
- Smets, F. and Wouters, R. (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach, *American Economic Review* **97**(3): 586–606.
- Tauchen, G. (1986). Statistical Properties of Generalized Method-of-Moments Estimators of Structural Parameters Obtained from Financial Market Data, *Journal of Business and Economic Statistics* **4**(4): 397–416.
- Theil, H. (1971). *Principles of Econometrics*, John Wiley and Sons, inc.