Cross Equation Effects of Misspecification: A partial estimation approach for DSGE models *

Kai Christoffel, DG Research, European Central Bank

January 31, 2005

Abstract

Fully specified DSGE models are increasingly successful in explaining observed macroeconomic data. Thinking about the specification of a certain equation in a DSGE approach has the drawback of imposing many implicit priors on the specification of the remaining equations. Mis-specifications in one block can have effects on the structural parameter estimates of the remaining equations. One resort from this problem is to use a VAR as an auxiliary model and to impose the structural equations stepwise on the unrestricted VAR. In a linear framework, we can interpret the unrestricted equations as an approximation of the solution process of the structural model. Once the model contains unobservable variables the solution process does not have a finite VAR representation anymore and the VAR approximation to the solution process is misspecified. The method of indirect inference allows to correct for mis-specification in the auxiliary model. The approach is illustrated with the example of the basic New Keynesian Phillips Curve and an extended version containing unobservable variables. In a Monte Carlo exercise the estimation properties of Kalman filter based maximum likelihood and indirect inference are evaluated for both models.

Keywords: Simulation Based Estimation, Simultaneous Equation Systems, DSGE Modelling
JEL classification: C13, C15, C51

---

*The opinions expressed are those of the author and do not necessarily reflect views of the European Central Bank.

1 Introduction

The methods and the modelling approach from the Real Business Cycle literature in combination with various forms of rigidities in the New Keynesian tradition provide a successful strategy for macroeconomic modelling. Dynamic Stochastic General Equilibrium (DSGE) models, with rich specifications as in Christiano, Eichenbaum, and Evans (2001) and Smets and Wouters (2003) are able to provide a good fit to macroeconomic data. However, there is much less consensus on the preferred method to bring the models to the data. Most of the available estimation procedures remain problematic. Due to various identification problems and the focus on selected moments only limited information methods seem less suitable for the estimation of DSGE models. With increasing dimension of the parameter space direct full information methods suffer from multiple local optima in the likelihood and small sample problems in typical macroeconomic applications. The actual implementation of the estimation procedure usually includes various priors on the size of the structural parameters. These priors occur either in the implicit form of starting values in the classical estimation procedure or as modelled priors in a Bayesian approach (Schorfheide (2000)). The simultaneous interdependence between the estimation results on different equations is intrinsic to the full information method. A misguided prior on one parameter can have an influence on the estimation result for another parameter. Furthermore model misspecification in one block of the model can have an effect on the estimation results for other blocks. Especially in situations where the researcher is interested in a certain specification embedded into a larger model, the model and parameter priors affect the estimation results for the specification of interest.

This paper works with an approach aiming to estimate the specification of interest in an otherwise unrestricted system. We use the vector autoregressive representation of the solution process of (a class of) rational expectation models to build a hybrid model. This hybrid model consists of some structural and some reduced form equations. Using the corresponding equations of an estimated VAR as an approximation of the solution process the method allows to estimate a subset of structural parameters in a full information setting without having to estimate all structural parameters jointly.

From a statistical point of view it is fair to say that the fully structural model as a description of the underlying data generating process is false, at least in some respects. Nevertheless the full model is needed to estimate some parameters of interest. In order to avoid misspecification from the part of the model that is not in the central interest of the research these equations are replaced by an instrumental or auxiliary model. The problematic structural model is (partially) replaced by an auxiliary model which is easier to estimate. Gourieroux, Monfort, and Renault (1993) analyze indirect inference on a structural model via an auxiliary model. Complementary to their approach in this paper the method of indirect inference is used to estimate the equations of the model in a stepwise procedure. While the partial approach has the merit of reducing the cross-equation effect of misspecification the method of indirect inference allows to correct for misspecification in the auxiliary model. This approach is related to the work by Dridi, Guay, and Renault (2003) and Dridi and Renault (2000) formalizing the calibration method of Cechetti, Lam, and Mark (1993). The basic idea of their approach is to divide the set of parameters into deep parameters that are of interest to the researcher and nuisance parameter needed only to be able to estimate the structural parameters. They start that from the assumption that "the model is false" and try to extract some elements of truth (deep parameters) from the false model. The auxiliary model is then used to examine the nuisance parameters, before the estimation of the structural parameters.

A comparison of the small sample properties in a partial approach of indirect inference and maximum likelihood sheds some light on the relative advantages of the indirect inference in this setting. There is a branch of literature on the estimation properties of indirect inference in comparison with other estimation techniques. Michaelides and Ng (2000) use a rational expectation model with speculative storage as a benchmark to compare the simulated method of moments (SMM by Duffie and Singleton (1993)), with Efficient Method of Moments (EMM by Gallant and Tauchen (1996)) and Indirect Inference by Gourieroux, Monfort, and Renault (1993)). Ruge-Murcia (2003) compares the estimation properties of different techniques using a one sector Real Business cycle model with indivisible labor as in Hansen (1985). He includes maximum likelihood in a classical (Ireland (2003b)) and in a Bayesian interpretation (Chang, Gomes, and Schorfheide (2002)) as well as simulated methods of moments, generalized method of moments (GMM) and Indirect Inference into the comparison. Both comparisons are explicitly focused on the estimation properties in a small system.

The proposed method of partial estimation is applicable to larger models. The estimation of larger systems with a higher number of structural parameters is less troublesome if the estimation can be performed equation by equation without leaving the full information setting.

Section 2 gives an overview on the chosen partial estimation approach, followed by a short excursion on a statistical example. In section 4 the framework of the partial estimation strategy for structural models is introduced, describing the relation between structural model, solution process and the hybrid combining structural and solution process equations. Section 5 states the used estimators. The economic model used is described in section 6, section 7 reports the results from a Monte Carlo exercise on the properties of the two estimators applied to the economic model. Finally section 8 concludes.

2 The Estimation Strategy

Various methods and approaches are used to evaluate the performance of macroeconomic models with rational expectations.
A first class of evaluation strategies works with a well defined set of moments a successful model has to match. Prominent examples of this approach can be found in real business cycle literature such as King, Plosser, and Rebelo (1988), where basic correlations and cross-correlations are chosen as the moments to match. In another branch of applications theoretical moments are generated to estimate the model by GMM (Clarida, Gali, and Gertler (1998)). Focusing on the dynamic responses of the model to different shocks Christiano, Eichenbaum, and Evans (2001) estimate the model parameters using the corresponding impulse response functions as the moments to match. The definition of key moments to match is an indispensable tool to construct empirically successful models. To test and evaluate the stochastic process implied by the structural model however it is necessary to switch to full information methods.

Maximum likelihood based full information methods have recently been used to estimate rational expectation models in a classical (Ireland (2003a) and Kim (2000)) or in Bayesian interpretation (Schorfheide (2000) or Smets and Wouters (2002)). Using a Kalman filter based calculation of the likelihood functions these approaches are able to estimate models containing unobservable variables. The main merit as well as mayor difficulty in the full information estimation of structural model lies in the formulation of a full stochastic process for all involved variables. Small changes in the assumed nature of the underlying shock processes can have a considerable impact on the structural estimation results. In addition to the uncertainty attached to the model specification the choice of starting values in the numerical estimation procedure or the choice of priors in a bayesian approach have an impact on the estimation results. The high dimension of parameters to estimate makes it very hard to distinguish local from global optima.

Another branch in macroeconometrics uses an auxiliary model to evaluate theoretical models. The predominance of vector autoregressions (VAR) in the

representation of the unrestricted version of the data generating process is mirrored in its manyfold applications [1]. One way to combine the full information approach with the VAR approach is offered by Fuhrer, Moore, and Schuh (1995) and Fuhrer (2000). This approach is related to the earlier rational expectation literature in macroeconomics where it is common practice to test the implication of economic models by a likelihood ratio test of restricted to unrestricted model. If restrictions are imposed on one or more equations of the otherwise unrestricted model the common interpretation for the unrestricted equation is to represent the information set on which expectations are formed. Fuhrer (2000) starts with estimating a VAR as an approximation to the data generating process. In a second step he imposes the structural representation of a habit formation consumption equation onto the VAR and estimates the corresponding structural parameters by maximum-likelihood.
In a linear framework however, we can interpret the unrestricted equations as an approximation to the solution process of the structural model. This interpretation allows to estimate a structural model by imposing the model restrictions block- or equationwise on an unrestricted VAR.
More complex models involving unobservable variables or autocorrelated structural shocks do not have a finite VAR representation in the observable variables in general. Using a finite VAR as an approximation to the non modelled equation results in a misspecification possibly affecting the estimation results for the structural parameters. The approach of indirect inference Gourieroux, Monfort, and Renault (1993) provides a method to base estimation on a (misspecified) auxiliary model.

To analyze the estimation properties of Indirect Inference in a partial estimation approach to structural models the following steps have to be taken.

First it must be shown that the principle of indirect inference can be extended to a situation where the auxiliary model is also used to approximate a part of the true data generating process directly. In Section 3 the properties of indirect inference in a partial approach are investigated in a Monte Carlo Exercise using a purely statistical example. Second it must be shown that this approach is applicable to the estimation of structural economic models. Section 4 relates the partial approach to the estimation of structural economic models by using the straightforward relation between the structural form and its solution process. With this relation it is possible to build hybrid models consisting of some structural equations and some solution process equation with the property that the hybrid models

[1] Christiano, Eichenbaum, and Evans (2001); Rotemberg and Woodford (1998)

have the same solution process as the fully structural model.
This equivalence of the solution process of structural model and hybrid model allows for a partial estimation strategy, where the solution process equations are approximated by an unrestricted VAR.

3 Partial Indirect Inference on a Statistical Model

Before we turn to the estimation of a structural economic model, we analyze the properties of the partial approach in a statistical experiment which is related to the estimation of DSGE models. Once the economic model contains unobservable variables and or autocorrelated shocks the reduced form of the structural model does not have a finite VAR representation in observable variables anymore. In fact the reduced form of the model has a VARMA representation where the moving average terms capture the dynamics of the unobservable variables.

We assume that the data generating process is given by a VARMA model and try to estimate this model using a VAR as auxiliary model in the indirect inference approach. This situation is comparable to the situation of estimating a model containing unobservable variables using a VAR as auxiliary model.

For the statistical experiment we assume that the data generating process is given by the following VARMA representation:

$$\boldsymbol{y}_t = \boldsymbol{A}_1\boldsymbol{y}_{t-1} + \cdots + \boldsymbol{A}_p\boldsymbol{y}_{t-p} + \boldsymbol{\varepsilon}_t \tag{1}$$

where the error terms are autocorrelated according to

$$\boldsymbol{\varepsilon}_t = \boldsymbol{u}_t + \boldsymbol{M}_1\boldsymbol{u}_{t-1} + \cdots + \boldsymbol{M}_q\boldsymbol{u}_{t-q} \tag{2}$$

Instead of a direct estimation of the VARMA model we can use the method of Indirect Inference. In a first step an auxiliary model with parameters $\hat{\boldsymbol{\beta}}_T$ is estimated. The auxiliary model is required to give a good description of the statistic properties of the underlying DGP but need not be well specified. A VAR is misspecified for a VARMA DGP but the inclusion of a sufficient number of lags will provide a good fit to the data. Formally this procedure defines a *binding function* (Gourieroux and Monfort (1992)) from the structural onto the auxiliary parameters.

$$b(DGP) : \boldsymbol{\theta} \rightarrow b(DGP, \boldsymbol{\theta}) = \beta_0 \tag{3}$$

If the binding function was non and one-to-one we could retrieve the structural parameters from the inversion of the binding function. Since this is generally not the case we have to rely on simulations. Therefore in a second step the VARMA model is simulated and the auxiliary parameters $\hat{\boldsymbol{\beta}}_{ST}$ are estimated on the simulated data.

The vector of structural parameters ($\boldsymbol{\theta}$) is determined by minimizing the distance between both auxiliary parameters.

$$\boldsymbol{\theta} \;=\; arg \min_{\boldsymbol{\theta}} [\hat{\boldsymbol{\beta}}_T - \hat{\boldsymbol{\beta}}_{ST}(\boldsymbol{\theta})]' \boldsymbol{\Omega} [\hat{\boldsymbol{\beta}}_T - \hat{\boldsymbol{\beta}}_{ST}(\boldsymbol{\theta})] \tag{4}$$

where $\boldsymbol{\Omega}$ is the optimal weighting matrix. [2]

With the same logic we can estimate a subset of structural parameters in the partial approach. Assume that the researcher is interested in the VARMA specification of the second equation only. The estimated auxiliary model can then be used to approximate the first equation. With this approximation some structural parameters are replaced by a combination of auxiliary parameters. This implies that the respective parameters in the binding function are mapped onto themselves.

Table 1 gives the results for different specifications of a VAR as auxiliary model from the following Monte Carlo exercise. The data generating process is given by a two variable VARMA(1,1) model. We estimate a VAR with different lag lengths and replace the first equation of the VARMA by the corresponding equation from the auxiliary VAR. Using the method of indirect inference we retrieve the VARMA parameters of the second equation. We repeat this exercise using maximum likelihood directly on the hybrid model.

The most important result can be seen in the property that the mean of the simulation exercise converges to the true value for all parameters in the indirect inference approach. The moving average term is slightly underestimated for a medium sample size using a VAR(3) as auxiliary model. This bias is due to the fact that the VAR(3) does not capture the full degree of autocorrelation present in the data. The partial approach in a maximum likelihood setting produces unbiased

[2] Sections **??** and **??** give a more detailed description of the method of indirect inference. Compare also Gourieroux, Monfort, and Renault (1993) for the estimation properties of indirect inference in a related univariate exercise.

Table 1: Misspecification in large samples: VARMA with VAR as auxiliary model, samplesize=950, 400 path simulations

		$\boldsymbol{A}(2,1) = 0.6$		$\boldsymbol{A}(2,2) = 0.3$		$\boldsymbol{M}(2,2) = 0.9$	
		M.L.	Ind.Inf.	M.L.	Ind.Inf.	M.L.	Ind.Inf.
VAR(2)*	mean	0.601	0.600	0.294	0.300	0.924	0.886
	std. dev.	0.025	0.034	0.026	0.032	0.064	0.061
	RMSE	0.025	0.038	0.026	0.032	0.068	0.062
VAR(3)	mean	0.598	0.599	0.299	0.304	0.923	0.890
	std. dev.	0.024	0.071	0.026	0.053	0.061	0.075
	RMSE	0.024	0.071	0.026	0.053	0.065	0.075
VAR(4)	mean	0.600	0.604	0.293	0.314	0.925	0.904
	std. dev.	0.024	0.045	0.027	0.042	0.064	0.095
	RMSE	0.023	0.045	0.028	0.044	0.069	0.095

Notes: The bivariate DGP is given by $\boldsymbol{y}_t = \boldsymbol{A}\boldsymbol{y}_{t-1} + \boldsymbol{u}_t + \boldsymbol{M}\boldsymbol{u}_{t-1}$ where a VAR with various lag lengths is chosen as the auxiliary model. The first equation is approximated by the corresponding auxiliary model equation. Note that $\boldsymbol{M}(2,1) = \boldsymbol{M}(1,2) = 0$.

estimates for all autoregressive parameters. The moving average term seems to be bias upwards. Taken a closer look at the density function in figure 1 we can see that there is a high density on $\widehat{\boldsymbol{M}}(2,2)$ around 1.1.[3]

Looking at the standard deviations and the root mean squared error (RMSE) maximum likelihood is outperforming the indirect inference approach. The differences between maximum likelihood and indirect inference are smallest for the moving average term in the large sample using a VAR(3) as auxiliary model. The increased RMSE of the point estimates from indirect inference comes from the additional variance in the weighting matrix.

4 Structural form and solution process

The intuition for the approach is straightforward: the indirect estimation procedures is based on a possibly misspecified model which gives a good description of

[3] Note that this problem is also present in direct maximum likelihood estimation of the VARMA model. The partial approach reduces this undesired effect as can be seen from table 3.

Table 2: Misspecification in medium samples: VARMA with VAR as auxiliary model, samplesize=150, 400 path simulations

		$\boldsymbol{A}(2,1) = 0.6$		$\boldsymbol{A}(2,2) = 0.3$		$\boldsymbol{M}(2,2) = 0.9$	
		M.L.	Ind.Inf.	M.L.	Ind.Inf.	M.L.	Ind.Inf.
	mean	0.592	0.593	0.290	0.300	0.940	0.854
VAR(2)*	std. dev.	0.064	0.084	0.065	0.078	0.085	0.158
	RMSE	0.065	0.085	0.065	0.078	0.094	0.164
	mean	0.597	0.595	0.288	0.301	0.947	0.872
VAR(3)	std. dev.	0.065	0.125	0.067	0.108	0.088	0.143
	RMSE	0.065	0.124	0.067	0.108	0.098	0.145
	mean	0.590	0.595	0.289	0.306	0.947	0.901
VAR(4)	std. dev.	0.067	0.104	0.064	0.094	0.090	0.148
	RMSE	0.067	0.104	0.063	0.094	0.100	0.148

Notes: The bivariate DGP is given by $\boldsymbol{y}_t = \boldsymbol{A}\boldsymbol{y}_{t-1} + \boldsymbol{u}_t + \boldsymbol{M}\boldsymbol{u}_{t-1}$ where a VAR with various lag lengths is chosen as the auxiliary model. The first equation is approximated by the corresponding auxiliary model equation. Note that $\boldsymbol{M}(2,1) = \boldsymbol{M}(1,2) = 0$.

the underlying data generating process. In a multivariate framework the auxiliary model can also be used to approximate the statistical properties of some equations. The partial approach might be advantageous in situation where the researcher is interested in specific equation that is embedded into a system of equations. One example for such a situation is the estimation of dynamic stochastic general equilibrium models.

DSGE models are characterized by an explicit micro-foundation and the important role of the rational expectation hypothesis. These characteristics define several challenges for the empirical implementation. First, the micro-founded model dynamics are completely defining the stochastic process. Second, the high degree of abstraction in the micro-foundation of the model is typically producing rather stylized models. Third, the rational expectations hypothesis implies a predominance of forward looking elements. The strong interrelation among the equations within the system implies that misspecification in one equation can have a substantial effect on the estimation results for the remaining equations.

The application of the partial approach in a DSGE setting differs from the statistical example given in section 3. The existence of forward looking variables makes it necessary to perform the estimation on the reduced form of the structural model. The partial approach implies that some reduced form equations are replaced by the approximation given by the auxiliary model. The relation between structural model, reduced form and auxiliary model and its implications for structural parameter estimation is given in section 4.1. Once the structural model contains unobservable variables the argumentation is changed as described in section 4.2.

For both approaches it can be shown that the solution process of the structural model has a vector autoregressive representation.[4] Furthermore unconstrained VARs are generally perceived to provide an adequate description of the part of aggregate data that is relevant for empirical monetary policy research. We will therefore use a VAR as auxiliary model in the following sections.

4.1 Structural model and auxiliary model: standard case

If the auxiliary model nests the solution process of the structural model we have a natural relation between structural form, reduced form and auxiliary model.
The structural model can be written in the following form

$$\sum_{i=1}^{\theta} \boldsymbol{B}_i \,\mathrm{E}_t\, \boldsymbol{x}_{t+i} \;=\; \sum_{i=-\tau}^{0} \boldsymbol{C}_i \boldsymbol{x}_{t+i} + \boldsymbol{\zeta}_t \tag{5}$$

where $\boldsymbol{x}_t$ denotes the vector of m (endogenous) model variables and $\boldsymbol{\zeta}_t$ stands for a vector of uncorrelated structural shocks[5]. The entries of the structural coefficient matrices $\boldsymbol{B}_i$ and $\boldsymbol{C}_i$ as well as the number of forward and backward looking terms are derived from the specification of the model equations.
Employing the usual solution schemes as in Blanchard and Kahn (1980) we can solve the model for its solution path.

$$\mathrm{E}_t\, \boldsymbol{x}_{t+k} \;=\; \sum_{i=-\tau}^{-1} \boldsymbol{D}_i \,\mathrm{E}_t\, \boldsymbol{x}_{t+k+i} \tag{6}$$

From the form of the solution path it is natural to chose an unconstrained VAR as an auxiliary model, nesting the reduced form of the structural model.[6]

[4]For the case of unobservable variables this relation is only true if the set of state variables is extended.

[5]Obviously, not every structural relation must be subject to a structural shock, implying that some entries of $\boldsymbol{\zeta}_t$ can be zero

[6]Equation 6 is an expectational identity. Note however that the errors of the reduced form contain a structure implied by the structural model.

Furthermore this observation allows us to estimate subsystems of the full structural model. Assume that $\boldsymbol{x}_t$ is partitioned into two subgroups $\boldsymbol{x}_t^1$ of dimension $(m_1 \times 1)$ and $\boldsymbol{x}_t^2$ of dimension $(m_2 \times 1)$, where $m_1 + m_2 = m$. Let's assume that the analysis focuses on the structural parameters of the subgroup $\boldsymbol{x}_t^1$ while the dynamics are determined by the full system $\boldsymbol{x}_t$. Assume for example that the demand side is the focus of the analysis. Accordingly the equations are ordered such that all demand side equations belong to subgroup $\boldsymbol{x}_t^1$ while the supply side variables are captured by $\boldsymbol{x}_t^2$.
We can construct a hybrid consisting of reduced form and structural form equations as follows.

$$\begin{aligned}
&\sum_{i=2}^{\theta} \begin{pmatrix} \boldsymbol{B}_i^{11} & \boldsymbol{B}_i^{12} \\ \mathrm{O} & \mathrm{O} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_{t+i}^1 \\ \boldsymbol{x}_{t+i}^2 \end{pmatrix} + \begin{pmatrix} \boldsymbol{B}_1^{11} & \boldsymbol{B}_1^{12} \\ \mathrm{O} & \boldsymbol{I}_{(m_2)} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_{t+1}^1 \\ \boldsymbol{x}_{t+1}^2 \end{pmatrix} \\
&= \sum_{i=-\tau}^{0} \begin{pmatrix} \boldsymbol{C}_i^{11} & \boldsymbol{C}_i^{12} \\ \boldsymbol{D}_i^{21} & \boldsymbol{D}_i^{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_{t+i}^1 \\ \boldsymbol{x}_{t+i}^2 \end{pmatrix} + \begin{pmatrix} \mathrm{O} & \mathrm{O} \\ \boldsymbol{S}^{21} & \boldsymbol{S}^{22} \end{pmatrix} \begin{pmatrix} \zeta_t^1 \\ \zeta_t^2 \end{pmatrix}
\end{aligned} \tag{7}$$

where the $\boldsymbol{S}$ matrices follow from the solution of the system of differential equations (5).

The solution process of this hybrid is again given by (6). Using the specification in (7) it is possible to estimate the subset of structural parameters contained in $\boldsymbol{x}^1$ without having to estimate the structural parameters contained in $\boldsymbol{x}^2$. By approximating the solution process for $\boldsymbol{x}^2$ by the corresponding equations of the auxiliary model we are able to abstain from the joint estimation of all structural parameters while still keeping track of the joint dynamics.
In section 7.1 this approach is demonstrated using the model of section 6 for the case of serially uncorrelated structural shocks.

4.2 Structural model and auxiliary model: unobservable variables

With the introduction of latent or unobservable variables the auxiliary model no longer nests the structural model. In this context it is useful to switch to a state space representation.
If we want to explain the joint distributions of the observable variables the structural relation explaining this distribution might be dependent on unobservable variables like the habit formation level in consumption models or measures for

installed capital in investment models. We start with a decomposition of the complete set of variables

$$\boldsymbol{x}_t = \begin{pmatrix} \boldsymbol{\xi}_t \\ \boldsymbol{y}_t \end{pmatrix} \tag{8}$$

where $\boldsymbol{\xi}_t$ denotes a $(r \times 1)$ vector of possibly unobserved state variables and $\boldsymbol{y}_t$ denotes a $(n \times 1)$ vector of observed variables. We assume that the variables evolve according to the following law of motion[7]:

$$\begin{aligned} \boldsymbol{\xi}_{t+1} &= \boldsymbol{F}\boldsymbol{\xi}_t + \boldsymbol{v}_{t+1} && (9) \\ \boldsymbol{y}_t &= \boldsymbol{A}'\boldsymbol{x}_t + \boldsymbol{H}'\boldsymbol{\xi}_t + \boldsymbol{w}_t && (10) \end{aligned}$$

It can be shown that $\boldsymbol{y}_t$ is distributed normal $\mathrm{N}(\boldsymbol{\mu}_{\theta_0}, \boldsymbol{\Sigma}_{\theta_0})$ with $\boldsymbol{\mu} = \boldsymbol{A}'_{\theta_0}\boldsymbol{x}_t + \boldsymbol{H}'_{\theta_0}\boldsymbol{\xi}_{\theta_0;t|t-1}$ and $\boldsymbol{\Sigma}_{\theta_0} = \boldsymbol{H}'_{\theta_0}\boldsymbol{P}_{\theta_0;t|t-1}\boldsymbol{H}_{\theta_0} + \boldsymbol{R}_{\theta_0}$

The argumentation given in section 4.1 carries over to the more general case, though we have to make some slight modifications. In the following we assume that the statistical auxiliary model for the observable variables can be represented as a vector autoregression:

$$\boldsymbol{y}_t = \sum_{i=-\tau}^{-1} \boldsymbol{G}_i \boldsymbol{y}_{t+i} + \boldsymbol{\varepsilon}_t \tag{11}$$

Where $\boldsymbol{y}_t$ is a n-dimensional vector.

There are two ways to represent the observable variables: the state space formulation based on the structural model and the statistical auxiliary model. In contrary to the approach in section 4.1 we have to distinguish between the set of variables of the structural equations and the set of variables that are actually observed. In the following we assume that the observable variables are chosen in a way such that there is a one-to-one relation between structural variables and observable variables. Each observable variable is uniquely attributed to one structural variable. All structural variables not assigned to an observable variable are modelled as state equations. With these assumptions it is possible to combine the statistical auxiliary model with the structural model to allow for the estimation of subsystems.

[7] For higher order autoregressive terms it is necessary to augment the state vector containing the lagged terms also.

In a first step the variables are partitioned into observable variables and state variables as in (14) [8]. As in section 4.1 we can partition the observable variables into two different type of structural variables. Starting with the representation of the structural model given in equation (5), let us assume that $\boldsymbol{y}_t$ is partitioned into two subgroups $\boldsymbol{y}_t^1$ of dimension $(n_1 \times 1)$ and $\boldsymbol{y}_t^2$ of dimension $(n_2 \times 1)$, where $n_1 + n_2 = n$. The analysis focuses on the structural parameters of the subgroup $\boldsymbol{y}_t^2$ while the dynamics are determined by the full system $\boldsymbol{x}_t$.

$$\boldsymbol{x}_t = \begin{pmatrix} \boldsymbol{\xi}_t \\ \boldsymbol{y}_t^1 \\ \boldsymbol{y}_t^2 \end{pmatrix} \tag{12}$$

As in equation (7) we can formulate a hybrid model consisting of equations from the statistical model and structural equations, where the latter are allowed to feature unobservable variables.

$$\sum_{i=2}^{\theta} \begin{pmatrix} \boldsymbol{B}_i^{\xi,\xi} & \boldsymbol{B}_i^{\xi,y_1} & \boldsymbol{B}_i^{\xi,y_2} \\ \boldsymbol{B}_i^{y_1,\xi} & \boldsymbol{B}_i^{y_1,y_1} & \boldsymbol{B}_i^{y_1,y_2} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} \end{pmatrix} \begin{pmatrix} \boldsymbol{\xi}_{t+i} \\ \boldsymbol{y}_{t+i}^1 \\ \boldsymbol{y}_{t+i}^2 \end{pmatrix} + \begin{pmatrix} \boldsymbol{B}_1^{\xi,\xi} & \boldsymbol{B}_1^{\xi,y_1} & \boldsymbol{B}_1^{\xi,y_2} \\ \boldsymbol{B}_1^{y_1,\xi} & \boldsymbol{B}_1^{y_1,y_1} & \boldsymbol{B}_1^{y_1,y_2} \\ \mathbf{O} & \mathbf{O} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{\xi}_{t+1} \\ \boldsymbol{y}_{t+1}^1 \\ \boldsymbol{y}_{t+1}^2 \end{pmatrix}$$
$$= \sum_{i=-\tau}^{0} \begin{pmatrix} \boldsymbol{C}_i^{\xi,\xi} & \boldsymbol{C}_i^{\xi,y_1} & \boldsymbol{C}_i^{\xi,y_2} \\ \boldsymbol{C}_i^{y_1,\xi} & \boldsymbol{C}_i^{y_1,y_1} & \boldsymbol{C}_i^{y_1,y_2} \\ \mathbf{O} & \boldsymbol{G}_i^{y2} & \boldsymbol{G}_i^{y2} \end{pmatrix} \begin{pmatrix} \boldsymbol{\xi}_{t+i} \\ \boldsymbol{y}_{t+i}^1 \\ \boldsymbol{y}_{t+i}^2 \end{pmatrix} + \begin{pmatrix} \mathbf{O} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} \\ \boldsymbol{S}^{31} & \boldsymbol{S}^{32} & \boldsymbol{S}^{33} \end{pmatrix} \begin{pmatrix} \zeta_t^{\xi} \\ \zeta_t^{y1} \\ \zeta_t^{y_2} \end{pmatrix} \tag{13}$$

We can solve this formulation and transfer it into state space form.
There is a conceptual difference between the interpretation of (7) and (13). In the case without latent variables we can interpret the auxiliary model as a an approximation to the solution process. If the sample size grows we will be able to extract the true solution process from the auxiliary model. In contrary to this we know that the solution process of (**??**) will depend on latent variables. The statistical auxiliary model (11) also provides an approximation to the data generating process but will not be able to produce the parameters of the solution process.

Instead of using this Kalman generated likelihood function we can make use of an indirect approach. From a statistical point of view we could try to evaluate the distribution of the observable variables without relying on a structural model or unobservable variables. We introduce a statistical model chosen according to its ability to give an appropriate description of the data and to give rise to a tractable

[8] Note at this point that the dynamics of the system cannot be captured by (16) and (15) because the model is still in its structural form.

likelihood function[9]. Assume that the statistical auxiliary model can be parameterized with the parameter $\boldsymbol{\beta}$ and results in a normal distribution $\mathrm{N}(\boldsymbol{\mu}_{\beta}, \boldsymbol{\Sigma}_{\beta})$.
If we assume that the data generating process is given by the structural model (5) then for an infinite sample size the distribution measured by the statistical model coincides with the true distribution generated by the structural model. This implies that the parameter of the auxiliary model estimated on the observed data and the parameter of the auxiliary model estimated on data simulated from the structural model under the true (structural) parameter $\boldsymbol{\theta}_0$ comply with each other. Since the value of the parameter of the auxiliary model estimated on the simulated data depends on the vector of structural parameters we can define a binding function (Gourieroux, Monfort, and Renault (1993)) mapping from the structural parameter into the auxiliary parameter. If this binding function was known and bijective we could retrieve the structural parameter estimated from this function and the auxiliary parameter. Since the binding function fulfills these properties only in special cases we can implicitly define the inverse of the binding function by choosing those structural parameters minimizing the distance between the auxiliary parameters estimated on observed and simulated data [10].
The estimation procedure makes use of the statistical auxiliary model capturing the dynamics of the structural model. As in the case of the model without unobservable variables we can therefore combine structural and statistical equations to focus on a subset of structural parameters to estimate.

5 Estimation Technique

5.1 Maximum Likelihood Estimation

The estimation results from the partial indirect inference approach are compared to the results from standard maximum likelihood estimation. In section 7.1 we are working with a model having a solution with an autoregressive representation in observed variables, implying a standard maximum likelihood estimation procedure. All other estimation procedures involve unobservable variables. For these

[9]Tractability refers to the desirable characteristic to compute the likelihood function analytically or with fast reliable numerical methods. The appropriateness of the statistical model should be evaluated along statistical specification tests.

[10]Since the binding function makes use of an asymptotical argument a large number of simulated values must be drawn from the structural model to generate the true distribution of the data generating process measured by the auxiliary model.

models the following Kalman filter based maximum likelihood procedure is employed. In a first step the model is written in state space form where the complete set of model variables $\boldsymbol{x}_t$ is divided into state variables and measure variables.

$$\boldsymbol{x}_t = \begin{pmatrix} \boldsymbol{\xi}_t \\ \boldsymbol{y}_t \end{pmatrix} \tag{14}$$

where $\boldsymbol{\xi}_t$ denotes a $(r \times 1)$ vector of possibly unobserved state variables and $\boldsymbol{y}_t$ denotes a $(n \times 1)$ vector of observed variables. We assume that the variables evolve according to the following law of motion[11]:

$$\begin{aligned} \boldsymbol{\xi}_{t+1} &= \boldsymbol{F}\boldsymbol{\xi}_t + \boldsymbol{v}_{t+1} && (15) \\ \boldsymbol{y}_t &= \boldsymbol{A}'\boldsymbol{x}_t + \boldsymbol{H}'\boldsymbol{\xi}_t + \boldsymbol{w}_t && (16) \end{aligned}$$

It can be shown that $\boldsymbol{y}_t$ is distributed normal $\mathrm{N}(\boldsymbol{\mu}_{\theta_0}, \boldsymbol{\Sigma}_{\theta_0})$ with $\boldsymbol{\mu} = \boldsymbol{A}'_{\theta_0}\boldsymbol{x}_t + \boldsymbol{H}'_{\theta_0}\boldsymbol{\xi}_{\theta_0;t|t-1}$ and $\boldsymbol{\Sigma}_{\theta_0} = \boldsymbol{H}'_{\theta_0}\boldsymbol{P}_{\theta_0;t|t-1}\boldsymbol{H}_{\theta_0} + \boldsymbol{R}_{\theta_0}$

Compare Hamilton (1994). Where $\boldsymbol{\xi}_{t|t-1}$ denotes the conditional expectation based on $t-1$ information and $\boldsymbol{P}_{t|t-1}$ denotes the variance of the conditional distribution of the state vector. Furthermore $\boldsymbol{Q}$ and $\boldsymbol{R}$ denote the covariance matrices of the state equation and the observation equation respectively, $\boldsymbol{Q} = \mathrm{E}(\boldsymbol{v}_{t+1}\boldsymbol{v}'_{t+1})$ and $\boldsymbol{R} = \mathrm{E}(\boldsymbol{w}_t\boldsymbol{w}'_t)$. The Kalman filter makes use of the following updating equations:

$$\begin{aligned} \widehat{\boldsymbol{\xi}}_{t+1|t} &= \boldsymbol{F}\widehat{\boldsymbol{\xi}}_{t|t-1} \\ &\quad + \boldsymbol{F}\boldsymbol{P}_{t|t-1}\boldsymbol{H}\left(\boldsymbol{H}'\boldsymbol{P}_{t|t-1}\boldsymbol{H} + \boldsymbol{R}\right)^{-1}\left(\boldsymbol{y}_t - \boldsymbol{A}'\boldsymbol{x}_t - \boldsymbol{H}'\widehat{\boldsymbol{\xi}}_{t|t-1}\right) \\ \boldsymbol{P}_{t+1|t} &= \boldsymbol{F}\boldsymbol{P}_{t|t-1}\boldsymbol{F}' \\ &\quad - \boldsymbol{F}\boldsymbol{P}_{t|t-1}\boldsymbol{H}\left(\boldsymbol{H}'\boldsymbol{P}_{t|t-1}\boldsymbol{H} + \boldsymbol{R}\right)^{-1}\boldsymbol{H}'\boldsymbol{P}_{t|t-1}\boldsymbol{F}' + \boldsymbol{Q} \end{aligned}$$

5.2 Indirect Inference estimator

- we define the estimation of the auxiliary model as the criterion

$$\max_{\boldsymbol{\beta} \in \boldsymbol{B}} \boldsymbol{Q}_T(\boldsymbol{y}_T, \boldsymbol{x}_T, \boldsymbol{\beta}) \tag{17}$$

[11] For higher order autoregressive terms it is necessary to augment the state vector containing the lagged terms also.

with

$$\widehat{\boldsymbol{\beta}}_T = arg \max_{\boldsymbol{\beta} \in \boldsymbol{B}} \boldsymbol{Q}_T(\boldsymbol{y}_T, \boldsymbol{x}_T, \boldsymbol{\beta}) \tag{18}$$

- we assume that the limit converges

$$\begin{aligned}
\lim_{T\to\infty} \boldsymbol{Q}_T(\boldsymbol{y}_T, \boldsymbol{x}_T, \boldsymbol{\beta}) &= \boldsymbol{Q}_\infty(\boldsymbol{F}_0, \boldsymbol{A}_0, \boldsymbol{H}_0, \boldsymbol{V}_0, \boldsymbol{W}_0, \beta) && (19)\\
&= \boldsymbol{Q}_\infty(\boldsymbol{V}_0, \boldsymbol{W}_0, \boldsymbol{\theta}_0, \boldsymbol{\beta}) && (20)
\end{aligned}$$

  where $(\boldsymbol{F}_0, \boldsymbol{A}_0, \boldsymbol{H}_0) = \boldsymbol{\theta}_0$

- we assume that $\boldsymbol{\beta}_0$ is the unique maximum of the limit of the criterium

$$\boldsymbol{\beta}_0 = arg \max_{\boldsymbol{\beta} \in \boldsymbol{B}} \boldsymbol{Q}_\infty(\boldsymbol{V}_0, \boldsymbol{W}_0, \boldsymbol{\theta}_0, \boldsymbol{\beta}) \tag{21}$$

- furthermore we introduce the binding function mapping from the structural model into the criterion

$$\boldsymbol{b}(\boldsymbol{V}, \boldsymbol{W}, \boldsymbol{\theta}) = arg \max_{\boldsymbol{\beta} \in \boldsymbol{B}} \boldsymbol{Q}_\infty(\boldsymbol{V}, \boldsymbol{W}, \boldsymbol{\theta}, \boldsymbol{\beta}) \tag{22}$$

  and

$$\boldsymbol{\beta}_0 = \boldsymbol{b}(\boldsymbol{V}_0, \boldsymbol{W}_0, \boldsymbol{\theta}_0) \tag{23}$$

  If the binding function is known and one to one we could deduce the structural parameters from this relation. In more general terms the binding function is difficult to compute. In these cases it is useful to replace the analytical binding function by a functional estimator based on simulations of the structural model. [12]

  $(\boldsymbol{y}^H)_T^1$ denotes H path simulations of length T taken from (16) under θ and a joint random seed.[13] We can define

$$\widetilde{\boldsymbol{\beta}}_T(\boldsymbol{\theta}) = arg \max_{\boldsymbol{\beta} \in \boldsymbol{B}} \boldsymbol{Q}_T((\boldsymbol{y}^H)_T^1, \boldsymbol{x}_T^1, \boldsymbol{\beta}) \tag{24}$$

[12] For the further reasoning we have to make the following two assumptions: $b(\boldsymbol{V}_0, \boldsymbol{W}_0, .)$ is one to one and $\frac{\partial b}{\partial \boldsymbol{\theta}'}(\boldsymbol{V}_0, \boldsymbol{W}_0, .)$ is of full column rank

[13] At this point we abstract from initial value problems.

as the solution of the maximization problem of the criterion on simulated data. For T tending to infinity the solution ($\widetilde{\boldsymbol{\beta}}$) tends to the solution of the limit problem:

$$\max_{\boldsymbol{\beta}\in\boldsymbol{B}} \boldsymbol{Q}_\infty(\boldsymbol{V}_0, \boldsymbol{W}_0, \boldsymbol{\theta}, \boldsymbol{\beta}) \tag{25}$$

$$\lim_{T\to\infty} \widetilde{\boldsymbol{\beta}}_T(\boldsymbol{\theta}) = \boldsymbol{b}(\boldsymbol{V}_0, \boldsymbol{W}_0, \boldsymbol{\theta}) \tag{26}$$

the binding function defines a mapping from the auxiliary parameter (solution of the simulated criterion function) into the structural parameters.

$$\boldsymbol{b} : \boldsymbol{\beta} \to \boldsymbol{\theta} \tag{27}$$

- the auxiliary criterion might be misspecified. The only requirement on the specification issue is that
  - $\boldsymbol{\beta}_0$ is the unique maximum of the limit problem under the true vector of structural parameters
  - the solution of the limit on the simulated series tends to the solution of the theoretical limit problem
- this implies that the auxiliary model can be misspecified as long as the $\boldsymbol{\beta}_0$ is the unique maximum
- if we combine this approach with the partial idea we have a partition of the set of structural parameters into structural parameters of interest and auxiliary parameters
- the approximation to the solution process can be misspecified, because the corresponding auxiliary parameters are mapped onto itself. The solution process of the non-modelled part are approximated and measured with respect to the same auxiliary model.
- the explanatory contribution of the approximated variables to the dynamics of the structural variables is also measured with respect to the auxiliary model such that possible misspecifications do not affect structural estimation results

6 A workhorse of macro-economics: the New Keynesian Model

The New Keynesian model is often represented in a system of three equations capturing the intertemporal optimality conditions of the households (”New IS-Curve”), the pricing behaviour of firms behaving as monopolisitc competitors faced with price adjustment costs (”New Keynesian Phillips Curve”) and monetary policy rule of the central bank (”Taylor-Rule”).

”New IS Curve” in terms of output gap (y_t)

$$y_t = E_t y_{t+1} - \frac{1}{\sigma}(r_t - E_t \pi_{t+1}) + \epsilon_{yt} \tag{28}$$

”New Keynesian Phillips Curve”

$$\pi_t = \lambda y_t + \delta E_t \pi_{t+1} + \epsilon_{\pi t} \tag{29}$$

”Taylor-Rule”

$$r_t = \rho r_{t-1} + (1-\rho)\beta E_t \pi_{t+t} + (1-\rho)\gamma y_t + \epsilon_{rt} \tag{30}$$

The empirical performance of models like (28)-(30) suffers from the lack of persistence generated by the model. Instead of or in addition to internal propagation mechanism some authors propose to introduce external propagation mechanisms in the form of autocorrelated structural shocks.

The budget constraint of the representative consumer can be written as

$$M_{t-1} + B_{t-1} + T_t + W_t h_t + D_t \geq P_t C_t + B_t / r_t + M_t \tag{31}$$

where M_{t-1} denotes the money holdings stemming from the last period and B_{t-1} stand for the bonds household hold at the beginning of period t. Furthermore the household receives a lump-sum monetary transfer from the central bank T_t, as well as labor income $W_t h_t$ and profits D_t. Household expenditure is given by consumption expenditure $P_t C_t$ and acquired money and bond holdings.

The utility function of the representative consumer can be written as follows:

$$\mathrm{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t u \left(c_t, \frac{M_t}{P_t}, h_t; \xi_t \right) \right\} \tag{32}$$

In this specification ξ_t represents a vector of structural disturbances to the utility function. Furthermore we assume utility to be additive separable in consumption and liquidity and the following specification for the utility function.

$$u\left(c_t, \frac{M_t}{P_t}; \xi_t\right) = a_t \frac{c_t^{1-\sigma}}{1-\sigma} + \ln(\frac{M_t}{P_t} - (1/\eta) h_t^{\eta}) \tag{33}$$

In this setting a_t is the only disturbance to the utility function following the dynamics specified by an AR:

$$a_t = \rho_a a_{t-1} + \epsilon_{a,t} \tag{34}$$

Furthermore a cost-push shock is introduced by allowing for a autocorrelated shock to the demand elasticity for each intermediate good.

$$\left[\int_0^1 Y_t(i)^{\theta_t - 1/\theta_t} di\right]^{\theta_t/(\theta_t - 1)} \tag{35}$$

with

$$\theta_t = (1-\rho_\theta)\ln(\theta) + \rho_\theta \theta_{t-1} + \epsilon_{\theta,t} \tag{36}$$

.

Since the intermediate goods are imperfect substitutes in the production of the final good, intermediate producer face some monopolistic pricing power. In the price setting process firms face a cost of nominal price adjustment of following form following Rotemberg (1982).

$$\frac{\phi}{2}\left[\frac{P_t(i)}{\pi P_{t-1}(i)} - 1\right]^2 Y_t \tag{37}$$

Labor h_t is the only factor for the production of the intermediate good of firm i $Y_t(i)$.

$$Z_t h_t(i) \geq Y_t(i) \tag{38}$$

where Z_t follows the process:

$$\ln(Z_t) = (1-\rho_z)\ln(z) + \rho_z \ln(Z_{t-1}) + \epsilon_{z,t} \tag{39}$$

Following Ireland (2003b) this specification results in a modified IS Curve and Phillips-curve relation in terms of the output gap (y_t)

$$y_t \;=\; E_t y_{t+1} - \frac{1}{\sigma}(r_t - E_t\pi_{t+1}) + \frac{1-\rho_a}{\sigma}a_t \tag{40}$$

$$\pi_t \;=\; \beta E_t\pi_{t+1} + \frac{(\theta-1)\sigma}{\phi}y_t - \frac{(\theta-1)}{\phi}z_t - \frac{1}{\phi}\theta_t \tag{41}$$

The monetary policy rule remains as specified in (30)

$$r_t \;=\; \rho_r r_{t-1} + (1-\rho_r)\beta E_t\pi_{t+t} + (1-\rho_r)\gamma y_t \tag{42}$$

It is straightforward to see that this formulation nests a model that is much easier to deal with. If we assume that the persistence parameters of the shock processes (ρ_θ, ρ_z and ρ_a) are equal to zero we have a model containing observable variables only as in equations (28), (29) and (30). The error terms can be interpreted as measurement errors and are scaled versions of $\epsilon_{a,t}, \epsilon_{\theta,t}$ and $\epsilon_{z,t}$. This model is used as an example in section 4.1 while the more general version is discussed in section 4.2.

7 Comparison of estimation properties

7.1 Well specified model (to be completed)

The following setting for the estimation experiment was chosen. The researcher knows that the observed data is generated by the model specified in (28), (29) and (30). The values of the structural parameters however are unobserved. The researcher chooses to estimate the value of the degree of risk-aversion (σ) without having to make an assessment on the value of the remaining parameters. His estimation strategy is to start with estimating an

unrestricted, non-structural system on the observed data. In a second step he imposes the restrictions from the economic model on the demand side by estimating equation (28) in the otherwise unrestricted system. From the reduced form (48) it is straightforward to chose a VAR with one lag as the representation for the unconstrained model.
In an initial step the structural system is solved and simulated under the true vector of structural parameters for a drawn vector of random disturbances. In a second step a VAR is estimated on the realization of the data generating process.

$$\begin{pmatrix} \pi_t \\ y_t \\ r_t \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} * \begin{pmatrix} \pi_{t-1} \\ y_{t-1} \\ r_{t-1} \end{pmatrix} + \boldsymbol{\varepsilon}_t \tag{43}$$

Taking the VAR(1) representation of equation (29) and (30) as an approximation to the solution process of the model the following hybrid model is build.

$$\begin{aligned} \pi_t &= a_{11} * \pi_{t-1} + a_{12} * y_{t-1} + a_{12} * r_{t-1} + \epsilon_{\pi,t} \\ y_t &= E_t y_{t+1} - \frac{1}{\sigma}(r_t - E_t \pi_{t+1}) \\ r_t &= a_{31} * \pi_{t-1} + a_{32} * y_{t-1} + a_{33} * r_{t-1} + \epsilon_{r,t} \end{aligned} \tag{44}$$

The structural parameters of the hybrid model are estimated using maximum likelihood and indirect inference. The results are reported in table xxx. The standard deviation of the point estimates is rather large even though we have a big sample. The high standard deviation of the estimators comes from the fact that the structural model used as the DGP is singular in the non-stochastic version. In the stochastic version of the model a large proportion of fluctuations is actually driven by the error terms and not by the structural model parameters. From a statistical point of view

7.2 Unobservable Variables and mis-specification (to be completed)

8 Conclusion

The combination of a rich economic DSGE specifications together with a set of autocorrelated error processes allows to explain a large part the observed data. Parameter estimates and specification evaluations in this approach are affected by the strong simultaneous effects between the equations. Mis-specifications in one equation or a misguided prior on some parameter estimated can have an effect on the estimation results for other equations. Imposing the restrictions from a DSGE specification on an otherwise unrestricted system allows to circumvent the cross-equation effect of mis-specifications. This paper states the explicit relation between the structural model, solution process and hybrid model consisting of some structural and some solution process equations. If all variables are observable the VAR provides a good approximation to the solution process leading to an estimation approach based on the hybrid model.

If the system contains unobservable variables the finite VAR approximation is no longer appropriate. The estimation problem can be compared to the estimation of a single equation in a VARMA model using a VAR as auxiliary model. The method of indirect inference allows to correct for this bias. Using a statistical example in a Monte Carlo experiment the estimation results of the indirect inference approach are compared to maximum likelihood results.

In a further step the Monte Carlo experiment is extended to the partial estimation of two DSGE models. The first DSGE model does not contain unobservable and the partial version can be estimated by maximum likelihood. The second DSGE model incorporates unobservable variables in the form of autocorrelated structural error processes. For both model the estimation results of a maximum likelihood approach are compared to an indirect inference approach. First results suggest that the increased robustness of the indirect inference estimator in a partial setting is counterweighted by the inferior small sample properties when compared to a (partial) maximum likelihood approach.

References

BLANCHARD, O. J., AND C. M. KAHN (1980): "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, 48.5, 1305–1311.

CECHETTI, S. G., P.-S. LAM, AND N. C. MARK (1993): "The Equity Premium and the Risk Free Rate: Matching the Moments," *Journal of Monetary Economics*, 31, 21–45.

CHANG, Y., J. GOMES, AND F. SCHORFHEIDE (2002): "Learning-by-Doing as a propagation mechanism," *American Economic Review*, 92, 1498–1520.

CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (2001): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *NBER-Working Paper*, w8403.

CLARIDA, R., J. GALI, AND M. GERTLER (1998): "Monetary Policy Rules and Macroeconomic Stability: Evidence and some Theory," *NBER Working Paper*, 6442.

DRIDI, R., A. GUAY, AND E. RENAULT (2003): "Indirect Inference and Calibration of Dynamic Stochastic General Equilibrium Models," *mimeo*.

DRIDI, R., AND E. RENAULT (2000): "Semiparametric Indirect Inference," *LSE-Workingpaper*, EM/00/392.

DUFFIE, D., AND K. SINGLETON (1993): "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica*, 61(4), 929–952.

FUHRER, J., G. R. MOORE, AND S. D. SCHUH (1995): "Estimating the linear-quadratic inventory model: Maximum likelihood versus generalized method of moments," *Journal of Monetary Economics*, 35, 115–157.

FUHRER, J. C. (2000): "Habit Formation in Consumption and Its Implications for Monetary-Policy Models," *American Economic Review*, 90(3), 367–390.

GALLANT, A. R., AND G. TAUCHEN (1996): "Which Moments to Match?," *Econometric Theory*, 12, 657–681.

GOURIEROUX, C., AND A. MONFORT (1992): “Testing, encompassing and simulating dynamic econometric models,” *CREST-Department de la Recherche INSEE*, DP 9124.

GOURIEROUX, C., A. MONFORT, AND E. RENAULT (1993): “Indirect Inference,” *Journal of Applied Econometrics*, V.8, S85–S118.

HAMILTON, J. D. (1994): *Time Series Analysis*. Princeton Press, 3rd edn.

HANSEN, G. D. (1985): “Indivisible Labor and the Business cycle,” *Journal of Monetary Economics*, 16, 309–327.

IRELAND, P. (2003a): “A Method for Taking Models to the Data,” *Journal of Economic Dynamics and Control*, 28, 1205–1226.

——— (2003b): “Technology Shocks in the New Keynesian Model,” *Working Paper*.

KIM, J. (2000): “Constructing and Estimating a realistic optimizing model of monetary policy,” *Journal of Monetary Economics*, 45, 329–359.

KING, R. G., C. I. PLOSSER, AND S. T. REBELO (1988): “Production Growth and Business Cycles: 1. The Basic Neoclassical Model,” *Journal of Monetary Economics*, 21, 195–232.

MICHAELIDES, A., AND S. NG (2000): “Estimating the Rational Expectations Model of Speculative Storage: A Monte Carlo Comparison of Three Simulation Estimators,” *Journal of Econometrics*, 96(2), 231–266.

ROTEMBERG, J. J. (1982): “Sticky Prices in the United States,” *Journal of Political Economics*, 90, 1187–1211.

ROTEMBERG, J. J., AND M. WOODFORD (1998): “ An Optimization - Based Econometric framework for the Evaluation of Monetary Policy: Expanded Version,” *NBER: Technical Working Paper*, 223.

RUGE-MURCIA, F. J. (2003): “Methods to Estimate Dynamic Stochastic General Equilibrium Models,” *Discussion Paper*, June 2003.

SCHORFHEIDE, F. (2000): “Loss-Function based evaluation of DSGE models,” *Journal of Applied Econometrics*, 15(6), 645–670.

SMETS, F., AND R. WOUTERS (2002): "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area," *NBER-ISOM*, June 14-15.

——— (2003): "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1(5), 1123–1175.

9 Appendix

9.1 Solving the model: a numerical example

For sake of disposition of the approach the model given in equations (28) to (30) is solved for a numerical example.
The interest rate is the only backward looking variable. A first guess for the expectation formation is therefore:

$$\begin{aligned} E_t\pi_{t+1} &= \phi r_t \\ E_t y_{t+1} &= \theta r_t \end{aligned} \tag{45}$$

Using this guess we end up with the following relations:

$$\phi = \frac{\rho\left[\phi\delta + \lambda\theta - \frac{\lambda}{\sigma} + \frac{\lambda\phi}{\sigma}\right]}{1-(1-\rho)\left[\beta\phi + \gamma\theta - \frac{\gamma}{\sigma} + \frac{\gamma\theta}{\sigma}\right]} \tag{46}$$

$$\theta = \frac{\rho\left[\theta - \frac{1}{\sigma} + \frac{\phi}{\sigma}\right]}{1-(1-\rho)\left[\beta\phi + \gamma\theta - \frac{\gamma}{\sigma} + \frac{\gamma\theta}{\sigma}\right]} \tag{47}$$

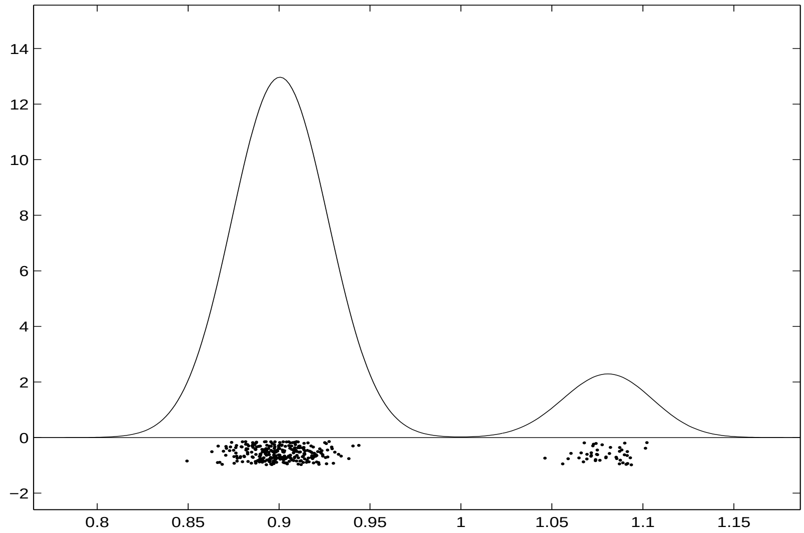
Since there is no closed form solution to these equations we proceed with an numerical example. Following Clarida, Gali, and Gertler (1998) we assume the following values for the structural parameters:

$$\begin{aligned} \beta &= 0.8 \\ \sigma &= 1 \\ \delta &= 0.99 \\ \lambda &= 0.3 \\ \rho &= 1.75 \\ \gamma &= 0.4 \end{aligned}$$

With these parameter values and the above relations (46 and 47) we can write the system in backward looking form:

$$E_t \begin{pmatrix} \pi_{t+1} \\ y_{t+1} \\ r_t \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1.4152 \\ 0 & 0 & 1.2275 \\ 0 & 0 & 0.7473 \end{pmatrix} * \begin{pmatrix} \pi_t \\ y_t \\ r_{t-1} \end{pmatrix} \tag{48}$$

Figure 1: **Density Function of partial Maximum Likelihood**



Notes: The nonparametric density function was estimated using a gaussian kernel

If we add shocks to the model we can simulate the observable representation (48). The shock in the ”New IS Curve” can be interpreted as a preference shock, while the shock to the NKPC is a cost push shock. The shock in the central bank reaction function can be interpreted as all those interest rate decision that deviate from the estimated rule.

Table 3: Partial versus full estimation of VARMA model, samplesize=950, 400 path simulations

	Direct Maximum Likelihood			*Partial Maximum Likelihood*		
parameter	mean	std. dev	RMSE	mean	std. dev	RMSE
$\boldsymbol{A}(1,1) = 0.8$	0.797	0.027	0.029	$n.e.$	$n.e.$	$n.e.$
$\boldsymbol{A}(1,2) = -0.1$	-0.102	0.019	0.019	$n.e.$	$n.e.$	$n.e.$
$\boldsymbol{M}(1,1) = 0.2$	0.198	0.039	0.041	$n.e.$	$n.e.$	$n.e.$
$\boldsymbol{A}(2,1) = 0.6$	0.599	0.025	0.023	0.600	0.024	0.023
$\boldsymbol{A}(2,2) = 0.3$	0.294	0.028	0.028	0.293	0.027	0.028
$\boldsymbol{M}(2,2) = 0.9$	0.967	0.088	0.119	0.925	0.064	0.069

Notes: A VAR(4) was chosen as the auxiliary model in the partial approach. As described in section 3 the parameters of the first equation were not estimated (n.e.).