SFB 649 Discussion Paper 2008-060

Matching Theory and Data: Bayesian Vector Autoregression and Dynamic Stochastic General Equilibrium Models

Alexander Kriwoluzky*

* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

Matching Theory and Data: Bayesian Vector Autoregression and Dynamic Stochastic General Equilibrium Models*

Alexander Kriwoluzky
Humboldt-Universität zu Berlin

September 22, 2008

Abstract

This paper shows how to identify the structural shocks of a Vector Autoregression (VAR) while at the same time estimating a dynamic stochastic general equilibrium (DSGE) model that is not assumed to replicate the data generating process. It proposes a framework to estimate the parameters of the VAR model and the DSGE model jointly: the VAR model is identified by sign restrictions derived from the DSGE model; the DSGE model is estimated by matching the corresponding impulse response functions.

JEL classification: C51.

Keywords: Bayesian Model Estimation, Vector Autoregression, Identification.

*I am indebted to Harald Uhlig, Chris Sims and Bartosz Mackowiak for their comments and suggestions. Furthermore, I thank Martin Kliem for helpful discussions. Part of the research was conducted while the author visited Princeton University funded by the German Academic Exchange Service. Further grants from the DEKA Bank and the SFB 649 are gratefully acknowledged.

1 Introduction

How can the dynamic effects of sudden changes in one economic variable on other variables be estimated? In modern macroeconomics, these effects are estimated via a Vector Autoregression (VAR) model. However, in order to apply a VAR model, additional assumptions have to be made. This chapter suggests to derive these additional assumptions from a dynamic stochastic general equilibrium (DSGE) model. To take the dependence of the identifying restrictions on the structural parameters of the DSGE model into account, it is estimated by matching the impulse response functions of the DSGE model and the VAR model. Since the VAR model is identified by restrictions derived from the DSGE model, whose parametrization in turn is estimated using the identified VAR model, it is necessary to describe the joint distribution of both. This paper develops a methodology to do so.

The methodology is in the spirit of Altig, Christiano, Eichenbaum, and Linde (2002), DelNegro and Schorfheide (2004) and Sims (2006b) who also propose to derive the identifying restrictions directly from a DSGE model. It differs from them in the following aspects: First of all the VAR model is identified by sign restrictions derived from the DSGE model. Furthermore, similar to DelNegro and Schorfheide (2004) and Sims (2006b), I take a Bayesian perspective and estimate the joint posterior distribution, where, in contrast to those authors and comparable to Altig, Christiano, Eichenbaum, and Linde (2002), the structural parameters of the DSGE model are estimated by matching the corresponding impulse response functions.

The suggested methodology has the considerable advantage to the existing literature that the structural parameters of the DSGE model can be estimated while it is not assumed to be a proper representation of the data generating process. One challenge in DSGE model estimation is that the DSGE model has to be fully stochastically specified, i.e. the number of shocks has to correspond to the number of observable variables. Deriving sign restrictions from the impulse response functions of the DSGE model to identify the shocks of interest in the VAR has the advantage, as mentioned in Uhlig (2005a), that it is not necessary for the complete number of structural shocks of the VAR model to be identified. Therefore, the number of structural shocks of the DSGE model need not correspond to the number of observable variables (variables in the VAR model) either, as it is required for Bayesian model estimation or the DSGE-VAR methodology by DelNegro and Schorfheide (2004). The researcher can solely concentrate on the shock of interest. The DSGE model serving as an identifying prior for the VAR model has the advantage that different assumptions the investigator wants to discriminate between can be build into the DSGE model. A further advantage of this approach is that it allows for nonlinear restrictions, for example combined zero restrictions in the initial periods and restrictions on the sign of the response afterwards.

Since the identification of the VAR model relies on the solution of the DSGE model, which in turn depends on its structural parameters, it becomes crucial to take the uncertainty about the structural parameters of the DSGE model into account. This is done by estimating the parameters using an approach similar to Christiano, Eichenbaum, and Evans (2005) and to the simulation based model estimation in Mertens and Ravn (2008). The posterior distribution of the parameters of the DSGE model is derived by matching the impulse responses of the DSGE model with those of the VAR model. This procedure allows to abandon the assumption of the DSGE model to be a proper representation of the true data generating process. Furthermore, features and lags only included to fit the data can be dropped.

This paper is organized in the following way: The next section shortly reviews the corresponding literature. The third section sets up the framework in general and the fourth section describes the probability distributions and the suggested algorithm in detail. The methodology is illustrated by a Monte Carlo Experiment in the fifth section. I simulate data from a fiscal theory of the price level (FTPL) model and reestimate the parameters of the FTPL model and the impulse response functions of the VAR model. The experiment shows that the true impulse response is indeed found. The last section concludes.

2 Related Literature

Solutions and suggestions to resolve the identification problem in a VAR model are manifold. Excellent surveys were written by Christiano, Eichenbaum, and Evans (1999) and Rubio-Ramirez, Waggoner, and Zha (2005). The most closely related approaches to the methodology presented here are to identify the VAR model by sign restrictions (Uhlig, 2005a; Faust, 1998) or by probabilistic restrictions (Kociecki, 2005). Identification employing sign restrictions attempts to restrict the sign of the impulse response functions of some variables, while the variable of interest is unrestricted. In Kociecki (2005), a prior distribution for the impulse response functions is formulated and transformed into a prior distribution for the coefficients of the structural VAR model. Both approaches depend on the availability of a priori knowledge on the behavior of some impulse response functions.

With regard to explicitly basing the identifying assumptions on DSGE models, two strands of literature have emerged recently. One derives the identifying assumptions from a DSGE model (Altig, Christiano, Eichenbaum, and Linde (2002), DelNegro and Schorfheide (2004) and Sims (2006b)) the other suggests, once the DSGE model is large enough, to estimate the DSGE model and to thereby directly infer on the impulse responses (as in Smets and Wouters (2003)).

The estimation of DSGE models has lately become very popular despite its problematic issues: First, not all parameters of the DSGE model can be identified (see Canova and Sala (undated) and Beyer and Farmer (2006)). Second, the DSGE model already puts a lot of structure on the impulse responses a priori, i.e. it often does not allow to investigate the sign of a response and might therefore not be appropriate as a research tool. Finally, not all economists might feel comfortable with the assumption that the DSGE model is a proper representation of the data generating process. Instead, as mentioned in Christiano, Eichenbaum, and Evans (2005), the DSGE model is suited best to replicate the implied dynamics in the data, i.e. the impulse response functions.

The methodology presented in this paper is in the spirit of the former strand of the literature, i.e. it bases the identification of the VAR model on restrictions derived from the DSGE model. It differs from the existing literature in the following aspects. Altig, Christiano, Eichenbaum, and Linde (2002) and DelNegro and Schorfheide (2004) employ the rotation matrix of the DSGE model to identify the VAR model. To do so, the DSGE model has to be fully stochastically specified. In the case of DelNegro and Schorfheide (2004), additional dummy observations derived from the model are used to augment the VAR model as suggested originally by Ingram and Whiteman (1994). While one can control for the prior weight of the dummy observations, one cannot, in either approach, control for the prior weight of the implied dynamics of the DSGE. The methodology proposed here differs from them by not employing the implied rotation matrix of the DSGE model to identify the VAR model and therefore not requiring the DSGE model to be fully stochastically specified.

Sims (2006b) extends the idea to augment the VAR model with dummy observations in a more general framework. In his framework, the tightness of the prior can be varied across frequencies and the number of structural shocks does not need to equal the number of observations. The main difference to Sims (2006b) is that I suggest to employ the implied sign and shape restrictions (as described in Uhlig (2005a)).

As such the methodology laid out in this paper complements the existing literature: The pure sign restriction approach as in Uhlig (2005a), the probabilistic restrictions as in Kociecki (2005) and the DSGE-VAR of DelNegro and Schorfheide (2004) arise as special cases of the presented framework.

3 Framework

In this section I first set up the VAR model and the DSGE model. I continue by presenting the central idea of the methodology and relate it to existing and nested approaches.

3.1 The VAR model

The structural VAR model containing m variables is given by:

$$A^{-1}Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \dots A_l Y_{t-l} + \epsilon_t, t = 1, \dots, T \tag{1}$$

Y_t is a $m \times 1$ vector at date $t = 1 + l, \dots, T$, A and A_i are coefficient matrices of size $m \times m$ and ϵ an *i.i.d.* one-step ahead forecast error, distributed: $\epsilon \sim \mathcal{N}(0, I_{m \times m})$. The reduced form of the VAR is then defined as:

$$Y_t = B_1 Y_{t-1} + B_2 Y_{t-2} + \dots B_l Y_{t-l} + u_t, t + 1, \dots, T \tag{2}$$

with $B_i = AA_i$, $u_t = A\epsilon_t$ and $u \sim \mathcal{N}(0, \Sigma)$.

The factorization $\Sigma = A'A$ does not have a unique solution, which leads to an identification problem of A, since only the reduced form can be estimated. Impulse response functions of the VAR model are computed from the companion form of the reduced form:

$$Y = XB + U \tag{3}$$

where

$$\underset{T \times m}{Y} = \begin{bmatrix} Y_1 & \dots & Y_T \end{bmatrix}'$$

$$\underset{(m*l) \times 1}{X_t} = \begin{bmatrix} Y'_{t-1} & Y'_{t-2} & \dots & Y'_{t-l} \end{bmatrix}'$$

$$\underset{T \times (m*l)}{X} = \begin{bmatrix} X_1 & X_2 & \dots & X_T \end{bmatrix}'$$

$$\underset{(m*l) \times m}{B} = \begin{bmatrix} B_1 & B_2 & \dots & B_l \end{bmatrix}'$$

$$\underset{T \times m}{U} = \begin{bmatrix} u_1 & \dots & u_T \end{bmatrix}'$$

Given additional restrictions to yield a unique factorization of Σ, an impulse vector a is a vector contained in the impulse matrix A. The impulse response function of a VAR model to an impulse vector a_i at horizon k φ^V_{ik} is defined as:

$$\varphi^V_{ik} = \Gamma^k \breve{a}_i, k = 0, 1, ...K \tag{4}$$

with

$$\Gamma = \begin{bmatrix} \multicolumn{2}{c}{B'} \\ I_{m(l-1)} & 0_{m(l-1),m} \end{bmatrix}$$

and

$$\breve{a}_i = \begin{bmatrix} \breve{a}'_i & 0_{1,m(l-1)} \end{bmatrix}'.$$

3.2 The DSGE model

The fundamental solution of the DSGE model is given by[1]:

$$\hat{x}_t = T(\tilde{\theta})\hat{x}_{t-1} + R(\tilde{\theta})z_t \tag{5}$$

where z is a vector collecting the structural shocks of the DSGE model, while $T(\tilde{\theta})$ and $R(\tilde{\theta})$ are matrices one obtains after solving a DSGE model with standard solution techniques.

The impulse response functions of the variables in x to a structural shock i at horizon k φ^D_{ik} are given by:

$$\varphi^D_{i,0} = R(\tilde{\theta})z_i, k = 0 \tag{6}$$

$$\varphi^D_{i,k} = T(\tilde{\theta})\varphi^D_{k-1,i}, k = 1, 2, ...K \tag{7}$$

The vector of structural parameters of the DSGE model defined in the way above does not contain any variances or covariances of a measurement error or any error term emerging from confronting the DSGE model with the data, only the variances of the structural shocks. When the DSGE model is estimated by matching the corresponding impulse response functions, there is an additional error term. Its variance covariance matrix is denoted by Ω and is also be estimated. The vector comprising the vector of deep parameters $\tilde{\theta}$ and the vectorized Ω is defined as $\theta = [\ \tilde{\theta} \quad vec(\Omega)\ ]'$.

3.3 The idea in a nutshell

In order to identify the VAR model impulse response functions of the DSGE model are employed. Impulse response functions of the DSGE model φ^D can be restrict the distribution of VAR model parameters either as probabilistic restrictions or as sign and shape restrictions[2]. The parameter distribution of the VAR model is therefore conditional on the impulse response function of the DSGE model, i.e. its parameter vector ($p(A, B|\theta)$).

On the other hand the DSGE model is estimated by matching the impulse response functions of the VAR model φ^V and of the DSGE model φ^D, This is a conditional distribution of the structural parameters of the DSGE θ model given the impulse response function of the VAR model ($p(\theta|\varphi^V)$).

---

[1] $\widehat{x}_t$ denotes the percentage deviation of the generic variable x_t from a deterministic steady state x chosen as approximation point.

[2] Sign and shape restrictions put zero probability weight on the parameter space of the VAR model for which the restrictions are not satisfied.

Since the aim of the exercise is to evaluate the joint posterior distribution of the parameters of the VAR model and the DSGE model, given a matrix with time series observations Y, it is necessary to connect both conditional distributions above. The joint distribution $p(\theta, \varphi^V|Y)$ can be decomposed in different ways, depending on whether the DSGE model is employed to identify the VAR model or whether it is not. In the latter case the joint posterior is given by:

$$p(\varphi^V, \theta|Y) \propto p(\varphi^V|Y)p(\theta|\varphi^V) \tag{8}$$

This formula can be justified twofold: In case the DSGE model is estimated by matching the corresponding impulse response functions and not time series observations, the distribution of θ conditional on φ^V and Y is equal to the distribution of θ conditional on φ^V only[3]. The second justification is shown by Smith (1993) and DelNegro and Schorfheide (2004) and discussed below, when setting the framework in a broader context.

In case the likelihood of the VAR model impulse response functions depends on restrictions from the DSGE model, $p(\theta, \varphi^V|Y)$ is given as:

$$p(\varphi^V, \theta|Y) \propto p(\varphi^V|\theta, Y)p(\theta|Y) \tag{9}$$

The framework presented in this paper is based on the argument that both distributions are at least proportionally equal:

$$p(\varphi^V|Y)p(\theta|\varphi^V) \propto p(\varphi^V|\theta, Y)p(\theta|Y) \tag{10}$$

and can be approximated sufficiently well by Monte Carlo Markov Chain Methods. Since the impulse response function of the VAR model is a function of A and B, the middle term has to be replaced by:

$$p(\varphi^V|\theta, Y) = p(A, B|\theta, Y)J(\varphi^V \to A, B) \tag{11}$$

where J denotes the Jacobian.

Note that the conditional distributions of interest are on different sides of the proportionally sign in (10). It is therefore possible to employ a Gibbs sampling algorithm, i.e. to draw from two conditional distributions in order to evaluate the joint distribution. But it is not possible to draw from $p(\theta|\varphi^V)$ directly. Instead, I suggest a Metropolis step between, i.e. drawing θ from a proposal density, in combination with the Gibbs sampling algorithm to approximate $p(\theta|\varphi^V)$.

[3]It then holds:

$$p(\theta|\varphi^V, Y)p(\varphi^V|Y) = p(\varphi^V|Y)p(\theta|\varphi^V)$$

3.4 Nested approaches

Taking a broader perspective, several closely related methodologies evolve as special cases of this approach: The pure sign restriction approach of Uhlig (2005a), the DSGE-VAR methodology of DelNegro and Schorfheide (2004) and the case of probabilistic restrictions of Kociecki (2005).

The latter arises in case the restrictions derived from the DSGE model are constant across the parameter space. Then, it is possible to generate a prior distribution for the impulse response functions of the VAR model from the DSGE model and use it as a prior for the parameters of the VAR model. Since, as pointed out by Kociecki (2005), the sign restriction approach is a special case of the probabilistic approach, this methodology is also nested. The sign restriction approach arises if the prior distribution for some impulse response function exhibits a very small variance, i.e. determines the sign of this impulse. It is equivalent to using an indicator function placing zero probability weight on VAR model parameter regions whenever the a priori sign restrictions are not satisfied. Therefore, in case the DSGE model determines constant sign restrictions across the parameter space it is not necessary to draw from the conditional distribution of θ one only needs to draw from $p(A, B|\theta, Y)$.

The DSGE-VAR methodology arises once the framework is written completely in terms of the parameters instead of the impulse response functions of the VAR model and in case the DSGE model is fully stochastically specified.

$$p(A, B|Y)p(\theta|A, B) \propto p(A, B|\theta, Y)p(\theta|Y) \tag{12}$$

The right hand side is the formula used to evaluate the joint posterior distribution of $p(A, B, \theta|Y)$: Since the DSGE model is fully stochastically specified it is possible to derive an analytical solution of the marginal posterior of θ. The decomposition on the left hand side legitimates again the decomposition used in (8): The posterior distribution of the parameters of the VAR does not depend on the vector of structural parameters of the DSGE model. As argued in DelNegro and Schorfheide (2004) and Smith (1993), A and B can then be used to learn about the parameters θ.

4 The joint posterior distribution

The joint posterior distribution is evaluated by drawing from two conditional distributions the distribution of the VAR model parameters conditional on restrictions derived from the DSGE model and the distribution of the DSGE model parameters conditional on the impulse response functions of the VAR model. This section describes these distributions in detail. Afterwards the algorithm to approximate the joint posterior distribution is set up.

4.1 The conditional distribution of the VAR model parameters

The methodology laid out in section 3.3 was written in terms of impulse response functions (φ). The transformation in equation 11 is not straightforward. It involves a highly nonlinear mapping between the impulse response functions and the structural coefficients of the VAR model. Kociecki (2005) shows how a prior distribution of impulse response functions can be used as a prior for the parameters of the VAR model. Those restrictions are called probabilistic restrictions. A summary of his results is given in appendix A.3. One advantage of probabilistic restrictions is that it is not necessary to specify prior distributions for all shocks. This is achieved by choosing very large prior variances for the impulse response functions. On the other side, by choosing very small variances, the probabilistic restrictions are equivalent to the sign restriction approach. Since, to the best of my knowledge, (it is not possible to draw the impulse matrix A of the VAR model for a reasonable large set of variables), I use Kociecki's framework, but in the sense of the sign restriction approach. For each realization of the impulse response function of the DSGE model the corresponding sign restrictions are put on the impulse response function of the VAR model. The distribution of the VAR model parameters is then conditional on the impulse response functions of the DSGE model, similar to Uhlig (2005a), where the posterior distribution of the VAR parameters is multiplied with an indicator function that puts zero probability in parameter regions whenever the restrictions derived from the DSGE model are not satisfied.

The sign restriction approach is applied in the following way: The impulse matrix $\breve{A}$ is defined as a sub matrix of A of size $m \times n$ where n is the number of structural shocks under consideration, i.e. the structural shock of interest as well as other shocks necessary to distinguish this shock. These shocks have to be included into the DSGE model as well. In order to indicate that the restrictions put on A rely on the model and therefore its parameter vector θ, I write $\breve{A}(\theta)$. Given a number of rowvectors q_j forming an orthonormal matrix Q and the lower Cholesky decomposition of Σ, $\tilde{A}$, $\breve{A}(\theta)$ is defined as: $\breve{A}(\theta) = \tilde{A}Q(\theta)$.

As shown by Uhlig (1997), the prior distribution for B and Σ can be specified choosing appropriate B_0, N_0, S_0, v_0 as:

$$
\begin{aligned}
vec(B)|\Sigma &\sim \mathcal{N}(vec(B_0), \Sigma \otimes N_0^{-1}) && (13)\\
\Sigma &\sim \mathcal{IW}(v_0 S_0, v_0) && (14)
\end{aligned}
$$

Denote the maximum likelihood estimates of Σ and B as $\tilde{\Sigma} = \frac{1}{T}(Y - X\hat{B})'(Y - X\hat{B})$

and $\hat{B} = (X'X)^{-1}X'Y$. The posterior is then given as[4]:

$$
\begin{aligned}
vec(B)|\Sigma &\sim \mathcal{N}(vec(B_T), \Sigma \otimes N_T^{-1}) && (15)\\
\Sigma &\sim \mathcal{IW}(v_T S_T, v_T) && (16)
\end{aligned}
$$

where

$$
\begin{aligned}
N_T &= N_0 + X'X && (17)\\
B_T &= N_T^{-1}(N_0 B_0 + X'X\hat{B}) && (18)\\
S_T &= \frac{v_0}{v_T} S_0 + \frac{T}{v_T}\tilde{\Sigma} - \frac{1}{v_T}(B_0 - \hat{B})' N_0 N_T^{-1} X'X(B_0 - \hat{B}) && (19)\\
v_T &= v_0 + T && (20)
\end{aligned}
$$

Drawing from a joint posterior of B, Σ and $\breve{A}(\theta)$ is conducted in the following steps:

1. The impulse responses of the DSGE determine the restrictions put on $\breve{A}(\theta)$.
2. Draw B and Σ from the posterior (15) and (16).
3. Calculate $\tilde{A}$ and draw $Q(\theta)$ from a uniform distribution such that $\breve{A}(\theta) = \tilde{A}Q(\theta)$ fulfills the sign restriction.

4.2 The conditional distribution of the DSGE model parameters

Since the DSGE model is not assumed to be a proper representation of the data generating process, the structural parameters are not estimated by matching the data Y. Instead, the DSGE model is assumed to replicate the implied dynamics of the data, i.e. the impulse response functions of the VAR model. This induces to match a given realization of the impulse response function of the VAR model to the i-th shock at horizon k, $\varphi_{i,k}^V$:

$$\varphi_{i,k}^V = \varphi_{i,k}^D(\tilde{\theta}) + \omega_{i,k}. \tag{21}$$

Stacking the impulse response functions over $1, .., K$ periods together yields:

$$\varphi_i^V = \varphi_i^D(\tilde{\theta}) + \omega_i \tag{22}$$

with all vectors of dimension $m * k \times 1$. The error term ω_i has the property $E[\omega_i \omega_i'] = \Omega_{\omega_i}$, which was part of the vector θ.

Since the structural shocks are assumed to be independent, the probability of $p(\theta|\varphi^V)$ can be written as:

$$p(\theta|\varphi^V) = p(\theta|\varphi_1^V, \varphi_2^V, \cdots \varphi_i^V) = p(\theta|\varphi_1^V)p(\theta|\varphi_2^V)\cdots p(\theta|\varphi_i^V) \tag{23}$$

[4] A formal derivation is given in appendix A.2

For each shock i the likelihood $l_i\left(\tilde{\theta},\Omega_{\omega_i}|\varphi_i^V\right)$is given by:

$$l_i\left(\tilde{\theta},\Omega_{\omega_i}|\varphi_i^V,\right) = -\frac{Km}{2}ln(2\pi)-\frac{1}{2}ln(|\Omega_{\omega_i}|)-\frac{1}{2}(\omega_i)'\left(\Omega_{\omega_i}\right)^{-1}(\omega_i) \tag{24}$$

Combining this likelihood with a prior distribution for θ yields a posterior distribution. For the vector of deep model parameters $\tilde{\theta}$ the prior distribution picked include normal-, beta- or gamma distribution. A candidate prior distribution for Ω_{ω_i} is the Inverted-Wishart distribution.

Besides the choice of the prior distribution there is an additional issue to be taken into account. Estimation of the whole matrix easily leads to a large amount of parameters. I therefore adopt an auxiliary prior assumption concerning ω:

$$E[\omega_i\omega_i'] = \tilde{\Omega}_{\omega_i}\otimes I_K \tag{25}$$

This necessitates that variances and covariances are constant over the impulse response horizon and that error terms of different impulse response horizons are uncorrelated. Since the impulse response functions of the VAR model are likely to be correlated, that assumption is a problematic one. But there are also good reasons for it: First, it reduces the parameter space of the variance covariance matrix substantially and circumvents the likely case that the number of parameters to be estimated is larger than the parameters of the VAR which parameterize the impulse response functions. Second, the formulation of the kernel of the likelihood in 24 is similar to indirect inference criterions used in the literature (Christiano, Eichenbaum, and Evans (2005)), where equivalent assumptions concerning the error terms are made. Third, those assumptions are analogue to interpreting the estimation as a seemingly unrelated equation estimator, which is an efficient estimator for a multivariate equation system. In case the errors are not normal, the estimator is referred to as Quasi-Maximum-Likelihood estimator(see White (1982) and Gourieroux, Monfort, and Trognon (1984)), which is not an efficient, but consistent estimator. Equation 24 can then be rewritten as:

$$l_i\left(\tilde{\theta},\tilde{\Omega}_{\omega_i}|\varphi_i^V,\right) = -\frac{Km}{2}ln(2\pi)-\frac{1}{2}ln(|\tilde{\Omega}_{\omega_i}\otimes I_K|)-\frac{1}{2}(\omega_i)'\left(\tilde{\Omega}_{\omega_i}\otimes I_K\right)^{-1}(\omega_i) \tag{26}$$

and the corresponding estimator for $\tilde{\Omega}_{\omega_i}$ is given by:

$$\tilde{\Omega}_{\omega_i} = \frac{S_i(\theta)}{T}$$

with

$$S_i(\theta) = \begin{bmatrix} \omega_{1,i}'\omega_{1,i} & \dots & \omega_{1,i}'\omega_{m,i} \\ \vdots & \ddots & \vdots \\ \omega_{m,i}'\omega_{1,i} & \dots & \omega_{m,i}'\omega_{m,i} \end{bmatrix}.$$

Due to this additional auxiliary assumption I do not specify any further prior distribution for Ω.

4.3 Sampling algorithm

In order to evaluate the joint posterior distribution of the parameters of the DSGE model and the VAR model I propose a Gibbs sampling algorithm, i.e. drawing from from the conditional distributions laid out in detail in sections 4.1 and 4.2. Since it is not possible to draw directly from $p(\theta|\varphi^V)$, it is necessary to combine the Gibbs sampling algorithm with a Metropolis algorithm.

Since different θ imply different sign restrictions, the Gibbs sampler is build into the Metropolis algorithm, i.e. the probability of the parameters of the DSGE model is evaluated conditional on the impulse response function of the VAR model satisfying the restrictions derived from the DSGE model for this parameter vector. Otherwise there would always be a tendency in favor of the latter draw of the parameters of the DSGE model in the accept-reject-step, since it determined the restrictions of the VAR model before.

To initialize the algorithm I suggest to draw first from the prior distribution $p(\tilde{\theta})$, to derive for each draw the corresponding restrictions and to draw a φ^V satisfying the restrictions. This yields a wide range of possible restrictions from the DSGE model for the impulse response functions of the VAR model and therefore a wide range of possible impulse response functions of the VAR model. In order to compute the proposal density for the Markov chain I find the vector of deep parameters of the DSGE model that fits best the mean of the impulse response functions of the VAR model and compute the corresponding Hessian $\tilde{\Psi}^{-1}$ at this point. This initialization procedure accomplishes to get a complete picture of possible restrictions of the VAR model and its corresponding impulse response functions and to center the proposal density for the parameters of the DSGE model around the vector with the highest conditional probability given the mean of the impulse response functions of the VAR model.

Specifically, the sampling algorithm is implemented in the following way:

1. Draw $\tilde{\theta}_j$ from the proposal density $p(\tilde{\theta}_{j-1}, \tilde{\Psi})$
2. Given a realization of the vector of deep parameters of the DSGE model, compute $\varphi_j^D(\tilde{\theta}_j)$ - The signs of φ_j^D are used as sign restrictions on the VAR model.
3. Draw Σ_j from (16) and B_j from (15). Compute the lower Cholesky decomposition and find an $\check{A}_j = \tilde{A}_J Q_j$ fulfilling the sign restrictions from $\varphi_j^D(\tilde{\theta}_j)$. Compute φ_j^V.
4. Given φ_J^V compute $p(\theta_j|\varphi_j^V)$ by combining 24 with a prior distribution of θ.
5. Compute $D = \frac{p(\theta_j|\varphi^V(j))}{p(\theta_{j-1}|\varphi^V(j-1))}$. Accept θ_j and $\varphi^V(j)$ with probability $min[D, 1]$.

6. Start again at point 1.

To check whether the chain converged it is useful to employ a convergence check as described in Gelman, Carlin, Stern, and Rubin (2004). The idea is to derive a scale reduction measure, i.e. a factor expressing by which scale the precision of the estimate can be improved, if the number of iterations is increased.

5 Example

In order to illustrate the methodology suggested above I use a simple fiscal theory of the price level (FTPL) model as described in Leeper (1991) to identify the response of inflation to a monetary policy shock, i.e. an unexpected increase in the interest rate. The FTPL model is chosen since it can be reduced to two equations in real debt and inflation. It is the most simple DSGE model exhibiting different signs of the impulse response functions depending on two parameters only. Furthermore, the solution and properties of the FTPL model are well known across economists, which makes the example a very transparent.

I simulate data from the FTPL model and show using the methodology outlined above that the 'true' signs of the impulse response functions and the corresponding distribution of the parameters of the FTPL model are found, even if the chain is initialized with a wrong guess.

5.1 The FTPL model

The representative household maximizes its utility in consumption c and real money balances m:

$$U_t = log(c_t) + log(m_t) \tag{27}$$

subject to the budget constraint:

$$c_t + m_t + b_t + \tau_t = y + \frac{1}{\pi_t} m_{t-1} + \frac{R_{t-1}}{\pi_t} b_t \tag{28}$$

where b denotes bond holdings, τ lump sum taxes, y income, R nominal interest rates and π inflation. Small letters denote real variables, capital letters nominal variables.

The government has to finance government expenditures g by issuing bonds, collecting taxes and seignorage. The budget constraint is therefore given by:

$$b_t + m_t + \tau_t = g + \frac{M_{t-1}}{P_t} + R_{t-1} \frac{B_{t-1}}{P_t} \tag{29}$$

The monetary authority sets the nominal interste rate R following the interest rate rule:

$$R_t = \alpha_0 + \alpha\pi_t + z_t \tag{30}$$

where α_0 and α are policy coefficients. z denotes a monetary policy shock, specified as

$$\begin{aligned} z_t &= \rho_1 z_{t-1} + \epsilon_{1t} && (31) \\ \epsilon_{1t} &\sim N(0, \sigma_1) && (32) \end{aligned}$$

The fiscal authority sets taxes according to:

$$\tau_t = \gamma_0 + \gamma b_{t-1} + \psi_t \tag{33}$$

where again γ_0 and γ denote policy coefficients. The innovation in fiscal policy has the following characteristics:

$$\begin{aligned} \psi_t &= \rho_2 \psi_{t-1} + \epsilon_{2t} && (34) \\ \epsilon_{2t} &\sim N(0, \sigma_2) && (35) \end{aligned}$$

The model can be linearized and summarized by two equations[5]:

$$\begin{aligned} \tilde{\pi}_{t+1} &= \beta\alpha\tilde{\pi}_t + \beta z_t && (36) \\ \tilde{b}_t + \varphi_1\tilde{\pi}_t + \varphi_3 z_t + \psi_t &= (\beta^{-1} - \gamma)\tilde{b}_{t-1} - \varphi_4 z_{t-1} - \varphi_2\tilde{\pi}_{t-1} && (37) \end{aligned}$$

5.2 Dynamics of the FTPL model

The dynamics of the system depend on whether fiscal and monetary policy are active or passive, i.e. they depend on the policy parameters α and γ only. Different policy regimes emerge for:

- $|\alpha\beta| > 1$ and $|\beta^{-1} - \gamma| < 1$ for active monetary (AM) and passive fiscal policy (PF). This will be referred to as regime I.
- $|\alpha\beta| < 1$ and $|\beta^{-1} - \gamma| > 1$ for active fiscal (AF) and passive monetary policy (PM). This will be referred to as regime II.
- AM/AF and PF/PM. These cases are not considered here.

Both policy regimes imply different signs of the impulse response function for inflation and real debt. In regime I a monetary policy shock (an unanticipated increase in the nominal interest rate) will lead to a negative response of inflation and a positive response of real debt. A fiscal policy shock (an unanticipated increase in taxes)

[5] See appendix A.1 for a derivation.

will have no effect on inflation and decrease the real debt. In regime II, a monetary policy shock leads to an increase in inflation and an initial decrease in real debt. A fiscal policy shock has a negative effect on both variables. Impulse response functions for each shock, regime and variable are plotted in appendix A.4 together with the corresponding distributions of α and γ .

5.3 Specification and Identification of the VAR

I simulate data from the model over 200 periods with $\alpha = \gamma = 0$, i.e. the case of active fiscal and passive monetary policy. The VAR model consists of two variables, inflation π and real debt b, with no constant or time trend: $y_t = [\pi_t b_t]'$. The VAR model with one lag is given by:

$$\begin{aligned} y_t &= By_{t-1} + u_t \\ E[u_t u_t'] &= \Sigma \end{aligned}$$

Ordering the fiscal policy shock first and the monetary policy shock second, based on the model the following characteristics of the impulse matrix A have to hold:

- If regime I holds:
  - Fiscal policy shock: $A_{11} = 0$ $A_{21} < 0$.
  - Monetary policy shock: $A_{21} < 0$ and $A_{22} > 0$.
- If regime II holds:
  - Fiscal policy shock: $A_{11} < 0$ $A_{21} < 0$.
  - Monetary policy shock: $A_{21} > 0$ and $A_{22} < 0$.

Since the sign of the reaction of real debt to a monetary policy shock does not identify the shock in case of regime II, the monetary policy shock is ordered second, implying that both variables have to fulfill the sign restriction for a fiscal policy shock first. Then the sign of the response of real debt is restricted, while the response of inflation is left open.

5.4 A Monte Carlo Experiment

I choose the prior distribution of α and γ based on estimates of Davig and Leeper (2005):

The prior distribution is plotted in figure 1. The model fulfills the requirements to investigate the question how inflation reacts after a monetary policy shock: depending on the parameterization it allows for qualitatively different reactions of inflation,

Parameter	mean(I)	standard deviation(I)	mean(II)	standard deviation(II)
α	1.308	0.253	0.522	0.175
γ	0.0136	0.012	-0.0094	0.013

Table 1: Prior distribution for parameters of the model

and the DSGE model incorporates all other shocks necessary, here the fiscal policy shock, to distinguish the shock of interest. The corresponding impulse responses for each regime are plotted in the appendix A.4: Figures 2 and 3 provide Bayesian impulse response plots for draws from the prior distribution of regime I and figures 4 and 5 for draws from the prior distribution of regime II.

The plots indicate a break between regime I and II and not a smooth intersection. This means that determining a mode of the model parameters at the mean or mode of the impulse resonses of the VAR would be misleading - since those measures might not be appropriate. Therefore, I only take draws from the prior distribution of both regimes with equal probability of change. Since the data are simulated from regime II, the outcome to expect is the distribution of regime II, with the corresponding impulse responses of inflation and real debt for a fiscal policy shock and real debt for a monetary policy shock. Furthermore, inflation should rise in response to a monetary policy shock.

As figure 6 indicates this is indeed the case, even though I initialize the chain with a wrong guess. The posterior distribution of α and γ stems from regime II only. Figure 7 shows the response to a fiscal policy shock, figure 8 the response to a monetary policy shock. Inflation is indeed increasing.

6 Conclusion

This paper has shown how to identify the structural shocks of a Vector Autoregression (VAR) while at the same time estimating a dynamic stochastic general equilibrium (DSGE) model that is not assumed to replicate the data generating process. To this end it has presented a framework to jointly estimate the parameters of a VAR model and a DSGE model.

The VAR model is identified based on restrictions from the DSGE model, i.e identification relies on explicit restrictions derived from theory. Restrictions are formulated as sign restrictions. The DSGE model serves as a way to summarize the ideas economists have about the economy. Ideally it incorporates the assumptions the researcher wants to discriminate between. In any way it should be as agnostic as possible about the response of the variables of interest to the shock of interest.

The DSGE model is estimated by matching the impulse response functions of the

VAR and of the DSGE, i.e. their implied dynamics. Therefore, it need not be a representation of the data generating process. While the shock of interest has to be included, as well as other shocks necessary to distinguish it, the DSGE model need not be fully stochastically specified.

The methodology has been illustrated by a Monte Carlo experiment. Artificial data is simulated from a simple fiscal theory of the price level model in which fiscal policy is active and monetary policy passive. I use the methodology to investigate the sign of the response of inflation to a monetary policy shock. Depending on the policy regime, i.e. the reaction coefficients of the policy rules, the response can either be negative or positive. The prior distributions of the policy parameters are chosen such as to insure that both regimes and therefore both responses are equally likely. The estimation algorithm is initialized with a wrong guess. The estimated impulse response function of the VAR as well as the posterior distribution of the parameter of the DSGE model indicate that the methodology works correctly: The response of inflation shows the 'true' sign and the posterior distribution of the parameter of the DSGE model consists solely of policy coefficients from active fiscal and passive monetary policy.

References

ALTIG, D., L. J. CHRISTIANO, M. EICHENBAUM, AND J. LINDE (2002): "Technology Shocks and Aggregate Fluctuations," draft, Northwestern University.

BEYER, A., AND R. E. FARMER (2006): "Identification Problems in SDGE Models with an illustration to a small Macro model," Computing in Economics and Finance 2006 81, Society for Computational Economics.

CANOVA, F., AND L. SALA (undated): "Back to Square One: Identification Issues in DSGE Models," Working Papers 303, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University.

CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (1999): "Monetary policy shocks: What have we learned and to what end?," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1 of *Handbook of Macroeconomics*, chap. 2, pp. 65–148. Elsevier.

——— (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

DAVIG, T., AND E. M. LEEPER (2005): "Fluctuating Macro Policies and the Fiscal Theory," NBER Working Papers 11212, National Bureau of Economic Research, Inc.

DELNEGRO, M., AND F. SCHORFHEIDE (2004): "Priors from General Equilibrium Models for VARS," *International Economic Review*, 45(2), 643–673.

FAUST, J. (1998): "The robustness of identified VAR conclusions about money," International Finance Discussion Papers 610, Board of Governors of the Federal Reserve System (U.S.).

GELMAN, A., J. B. CARLIN, H. S. STERN, AND D. RUBIN (2004): *Bayesian Data Analysis*. Chapman & Hall.

GOURIEROUX, C., A. MONFORT, AND A. TROGNON (1984): "Pseudo Maximum Likelihood Methods: Theory," *Econometrica*, 52(3), 681–700.

INGRAM, B. F., AND C. H. WHITEMAN (1994): "Supplanting the 'Minnesota' prior: Forecasting macroeconomic time series using real business cycle model priors," *Journal of Monetary Economics*, 34(3), 497–510.

KOCIECKI, A. (2005): "Priors For Impulse Responses in Bayesian Structural VAR Models," Discussion paper, National Bank of Poland.

LEAMER, E. (1978): *Specification Searches*. John Wiley and Sons.

LEEPER, E. M. (1991): "Equilibria under 'active' and 'passive' monetary and fiscal policies," *Journal of Monetary Economics*, 27(1), 129–147.

MERTENS, K., AND M. O. RAVN (2008): "The Aggregate Effects of Anticipated and Unanticipated U.S. Tax Policy Shocks: Theory and Empirical Evidence," Economics Working Papers ECO2008/05, European University Institute.

RUBIO-RAMIREZ, J. F., D. WAGGONER, AND T. ZHA (2005): "Markov-switching structural vector autoregressions: theory and application," Working Paper 2005-27, Federal Reserve Bank of Atlanta.

SIMS, C. A. (2006b): "Making Macro Models Behave Reasonably," Working Papers 74, Princeton University, Department of Economics.

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

SMITH, A A, J. (1993): "Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions," *Journal of Applied Econometrics*, 8(S), S63–84.

UHLIG, H. (1997): "Bayesian Vector Autoregressions with Stochastic Volatility," *Econometrica*, 65(1), 59–74.

——— (2005a): "What are the effects of monetary policy on output? Results from an agnostic identification procedure," *Journal of Monetary Economics*, 52(2), 381–419.

WHITE, H. (1982): "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, 50(1), 1–25.

A Appendix

A.1 Description and solution of the DSGE model

A.1.1 DSGE Model Setup

$$\begin{aligned} U_t &= log(c_t) + log(m_t) && \text{(A-1)} \\ c_t + m_t + b_t + \tau_t &= y + \frac{1}{\pi_t} m_{t-1} + \frac{R_{t-1}}{\pi_t} b_t && \text{(A-2)} \end{aligned}$$

First-order conditions:

$$\begin{aligned} \frac{1}{R_t} &= \beta \frac{1}{\pi_{t+1}} && \text{(A-3)} \\ m_t &= c \frac{R_t}{R_t - 1} && \text{(A-4)} \end{aligned}$$

Government budget constraint:

$$b_t + m_t + \tau_t = g + \frac{M_{t-1}}{P_t} + R_{t-1} \frac{B_{t-1}}{P_t} \tag{A-5}$$

Monetary authority:

$$\begin{aligned} R_t &= \alpha_0 + \alpha \pi_t + \theta_t && \text{(A-6)} \\ \theta_t &= \rho_1 \theta_{t-1} + \epsilon_{1t} && \text{(A-7)} \\ \epsilon_{1t} &\sim N(0, \sigma_1) && \text{(A-8)} \end{aligned}$$

Fiscal authority:

$$\begin{aligned} \tau_t &= \gamma_0 + \gamma b_{t-1} + \psi_t && \text{(A-9)} \\ \psi_t &= \rho_2 \psi_{t-1} + \epsilon_{2t} && \text{(A-10)} \\ \epsilon_{2t} &\sim N(0, \sigma_2) && \text{(A-11)} \end{aligned}$$

A.1.2 Linearization

$$\bar{x}\hat{x}_t = \tilde{x}_t.$$

First equation:

$$\begin{aligned} R_t &= \alpha_0 + \alpha \pi_t + \theta_t \\ \pi_{t+1} &= \beta \alpha_0 + \beta \alpha \pi_t + \beta \theta_t \\ \tilde{\pi}_{t+1} &= \beta \alpha \tilde{\pi}_t + \beta \theta_t \end{aligned}$$

Second equation:

$$\begin{aligned}\bar{R}\hat{R}_t &= \frac{\bar{\pi}}{\beta}\hat{\pi}_{t+1}\\ \tilde{R}_t &= \frac{\tilde{\pi}_{t+1}}{\beta}\end{aligned}$$

$$\begin{aligned}m_t &= c\frac{R_t}{R_t-1}\\ \tilde{m}_t &= -\frac{c}{(\bar{R}-1)^2\beta}\tilde{\pi}_{t+1}\end{aligned}$$

$$\begin{aligned}\tilde{m}_t &= -\frac{c}{(\bar{R}-1)^2\beta}(\beta\alpha\tilde{\pi}_t+\beta\theta_t)\\ \tilde{m}_{t-1} &= -\frac{c\alpha}{(\bar{R}-1)^2}\tilde{\pi}_{t-1}-\frac{c}{(\bar{R}-1)^2}\theta_{t-1}\end{aligned}$$

$$\begin{aligned}b_t+m_t+\tau_t &= g+\frac{M_{t-1}}{P_t}+R_{t-1}\frac{B_{t-1}}{P_t}\\ \tilde{b}_t+\tilde{m}_t+\tilde{\tau}_t &= \frac{\tilde{m}_{t-1}}{\bar{\pi}}-\frac{\bar{m}}{\bar{\pi}^2}\tilde{\pi}_t+\frac{\bar{b}}{\bar{\pi}}\tilde{R}_{t-1}-\frac{\bar{R}\bar{b}}{\bar{\pi}^2}\tilde{\pi}_t+\frac{\bar{R}}{\bar{\pi}}\tilde{b}_{t-1}\end{aligned}$$

$$\tilde{b}_t-\frac{c\alpha}{(\bar{R}-1)^2}\tilde{\pi}_t+\frac{c\bar{R}}{(\bar{R}-1)\bar{\pi}^2}\tilde{\pi}_t+\frac{\bar{R}\bar{b}}{\bar{\pi}^2}\tilde{\pi}_t-\frac{c}{(\bar{R}-1)^2}\theta_t+\tilde{\tau}_t = \frac{\tilde{m}_{t-1}}{\bar{\pi}}+\frac{\bar{b}}{\bar{\pi}}\tilde{R}_{t-1}+\frac{\bar{R}}{\bar{\pi}}\tilde{b}_{t-1}$$

A.1.3 Simplifying the model

Define:

$$\begin{aligned}-\frac{c\alpha}{(\bar{R}-1)^2}+\frac{c\bar{R}}{(\bar{R}-1)\bar{\pi}^2}+\frac{\bar{R}\bar{b}}{\bar{\pi}^2} &= \frac{c}{(\bar{R}-1)}\left(-\frac{\alpha}{(\bar{R}-1)}+\frac{c}{\beta\bar{\pi}}\right)+\frac{\bar{b}}{\beta\bar{\pi}}=\varphi_1\\ -\frac{c}{(\bar{R}-1)^2} &= \varphi_3\end{aligned}$$

$$\begin{aligned}-\frac{1}{\bar{\pi}}\frac{c\alpha}{(\bar{R}-1)^2}+\frac{\bar{b}}{\bar{\pi}}\alpha &= -\frac{\alpha}{\bar{\pi}}\left[\frac{c}{(\bar{R}-1)^2}-\bar{b}\right]=-\varphi_2\\ -\frac{1}{\bar{\pi}}\frac{c}{(\bar{R}-1)^2}+\frac{\bar{b}}{\bar{\pi}} &= -\frac{1}{\bar{\pi}}\left[\frac{c}{(\bar{R}-1)^2}-\bar{b}\right]=\frac{-\varphi_2}{\alpha}=-\varphi_4\end{aligned}$$

This yields:

$$\tilde{b}_t + \varphi_1 \tilde{\pi}_t + \varphi_3 \theta_t - (\beta^{-1} - \gamma)\tilde{b}_{t-1} + \psi_t + \varphi_4 \theta_{t-1} + \varphi_2 \tilde{\pi}_{t-1} = 0 \tag{A-12}$$

$$\tilde{b}_t + \varphi_1 \tilde{\pi}_t + \varphi_3 \theta_t + \psi_t = (\beta^{-1} - \gamma)\tilde{b}_{t-1} - \varphi_4 \theta_{t-1} - \varphi_2 \tilde{\pi}_{t-1} \tag{A-13}$$

A.1.4 Calibration

Following Leeper (1991) the model is calibrated by setting:

$$\begin{aligned}
\beta &= 0.99 \\
\bar{c} &= 0.75 \\
\frac{\bar{b}}{\bar{y}} &= 0.4 \\
\bar{\pi} &= 3.43 \\
\rho_1 &= 0.8 \\
\rho_2 &= 0 \\
\sigma_1 &= 0.2 \\
\sigma_2 &= 0.2
\end{aligned}$$

A.2 Derivation of the posterior distribution of the BVAR

A.2.1 Prior distribution

$$\begin{aligned}
vec(B)|\Sigma &\sim \mathcal{N}(vec(B_0), \Sigma \otimes N_0^{-1}) && \text{(A-14)} \\
\Sigma &\sim \mathcal{IW}(v_0 S_0, v_0) && \text{(A-15)}
\end{aligned}$$

Σ is of size $m \times m$, N_0 of size $k \times k$, where $k = m * l$. The probability density function (p.d.f.) of $vec(B)$ is given by:

$$\begin{aligned}
p(B|B_0, \Sigma, N_0) &= (2\pi)^{-mk/2} |\Sigma \otimes N_0^{-1}|^{-1/2} \\
&\quad exp\left[-\frac{1}{2}\left(vec(B) - vec(B_0)\right)' \left(\Sigma^{-1} \otimes N_0\right)\left(vec(B) - vec(B_0)\right)\right] \\
&= (2\pi)^{-mk/2} |\Sigma|^{-k/2} |N_0|^{m/2} exp\{-\frac{1}{2} tr\left[\Sigma^{-1} (B - B_0)' N_0 (B - B_0)\right]\}
\end{aligned}$$

The p.d.f. of Σ is defined as:

$$p(\Sigma|v_0 S_0, v_0) = C_{IW}^{-1} |\Sigma|^{-\frac{1}{2}(v_0 + m + 1)} exp\left[-\frac{1}{2} tr\left(\Sigma^{-1} v_0 S_0\right)\right]$$

where:

$$C_{IW} = 2^{\frac{1}{2}v_0 m} \pi^{\frac{1}{4}m(m-1)} \prod_{i=0}^{m} \Gamma\left(\frac{v_0+1-i}{2}\right) |S_0|^{-\frac{1}{2}v_0}$$

A.2.2 Likelihood

For

$$vec(u) \quad \sim \quad \mathcal{N}(0, \Sigma \otimes I) \tag{A-16}$$

$$p(Y|B,\Sigma) \quad = \quad (2\pi)^{-Tm/2} |\Sigma|^{-T/2} exp\{-\frac{1}{2} tr\left[\Sigma^{-1}(Y-XB)'(Y-XB)\right]\} \tag{A-17}$$

The kernel can be rewritten as:

$$\begin{aligned} (Y-XB)'(Y-XB) \quad &= \quad (Y-XB-X\hat{B}+X\hat{B})'(Y-XB-X\hat{B}+X\hat{B}) \\ & \quad (Y-X\hat{B})'(Y-X\hat{B}) + (B-\hat{B})'X'X(B-\hat{B}) \end{aligned} \tag{A-18}$$

A.2.3 Posterior

$$\begin{aligned} p(\Sigma, B|Y) \quad = \quad & C_{IW}^{-1} |\Sigma|^{-\frac{1}{2}(v_0+m+1)} exp\left[-\frac{1}{2} tr\left(\Sigma^{-1} v_0 S_0\right)\right] \\ & \times (2\pi)^{-mk/2} |\Sigma|^{-k/2} |N_0|^{m/2} exp\{-\frac{1}{2} tr\left[\Sigma^{-1}(B-B_0)' N_0 (B-B_0)\right]\} \\ & \times (2\pi)^{-Tm/2} |\Sigma|^{-T/2} exp\{-\frac{1}{2} tr\left[\Sigma^{-1}(Y-X\hat{B})'(Y-X\hat{B})\right]\} \\ & \times exp\{-\frac{1}{2} tr\left[\Sigma^{-1}(B-\hat{B})'X'X(B-\hat{B})\right]\} \end{aligned} \tag{A-19}$$

Use the formula as stated in Leamer (1978)[6]:

$$\begin{aligned} (B-\hat{B})'X'X(B-\hat{B})\,(B-B_0)' N_0 (B-B_0) \quad = \quad & (B-B_T)'N_T(B-B_T) \\ & \times (B-B_0)'(X'X(N_T)^{-1}N_0)(B-B_0) \end{aligned} \tag{A-20}$$

where:

$$\begin{aligned} N_T \quad &= \quad N_0 + X'X \\ B_T \quad &= \quad N_T^{-1}(N_0 B_0 + X'X\hat{B}) \end{aligned}$$

[6] Appendix 1, T10

leads to:

$$
\begin{aligned}
p(\Sigma, B|Y) \quad = \quad & C_{IW}^{-1}|\Sigma|^{-\frac{1}{2}(v_0+m+1)}exp\left[-\frac{1}{2}tr\left(v_0\Sigma^{-1}S_0\right)\right] \\
& \times(2\pi)^{-mk/2}|\Sigma|^{-k/2}|N_0|^{m/2}exp\{-\frac{1}{2}tr\left[\Sigma^{-1}(B-B_0)'(X'X(N_T)^{-1}N_0)(B-B_0)\right]\} \\
& \times(2\pi)^{-Tm/2}|\Sigma|^{-T/2}exp\{-\frac{1}{2}tr\left[\Sigma^{-1}(Y-X\hat{B})'(Y-X\hat{B})\right]\} \\
& \times exp\{-\frac{1}{2}tr\left[\Sigma^{-1}(B-B_T)'N_T(B-B_T)\right]\}
\end{aligned}
\tag{A-21}
$$

$$
\begin{aligned}
p(\Sigma, B|Y) \quad = \quad & C_{IW}^{-1}|\Sigma|^{-\frac{1}{2}(T+v_0+m+1)} \\
& \times exp\left[-\frac{1}{2}tr\left(\Sigma^{-1}\left(\frac{v_0}{v_T}S_0+\frac{T}{v_T}\tilde{\Sigma}+\frac{1}{v_T}(B-B_0)'(X'X(N_T)^{-1}N_0)(B-B_0)\right)\right)\right] \\
& \times(2\pi)^{-m(T+k)/2}|\Sigma|^{-(T+k)/2}exp\{-\frac{1}{2}tr\left[\Sigma^{-1}(B-B_T)'N_T(B-B_T)\right]\}
\end{aligned}
\tag{A-22}
$$

A.3 Probabilistic restrictions

The algebra presented here summarizes the results of Kociecki (2005) and follows mostly his insights. The only innovation is to derive mean and standard deviations from a DSGE model.

A.3.1 Impulse response functions from the DSGE as a prior

Denote the impulse response functions in period k as in the paper φ_k^V. The matrix is, if all shocks are included, of size $m \times m$, where the the $i, j-th$ entry corresponds to the response of the $i-th$ variable to an innovation in the $j-th$ variable. The prior for the impulse responses has only to be specified for as many periods as lags are included into the VAR. The vectorized impulse responses are assumed tobe normally distributed:

$$
\begin{bmatrix} vec(\varphi_0) \\ vec(\varphi_1) \\ vec(\varphi_2) \\ \vdots \\ vec(\varphi_l) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} vec(\bar{\varphi}_0) \\ vec(\bar{\varphi}_1) \\ vec(\bar{\varphi}_2) \\ \vdots \\ vec(\bar{\varphi}_l) \end{bmatrix}, \begin{bmatrix} \bar{V}_{00} & \bar{V}_{01} & \bar{V}_{02} & \cdots & \bar{V}_{0l} \\ \bar{V}_{10} & \bar{V}_{11} & \bar{V}_{12} & \cdots & \bar{V}_{1l} \\ \bar{V}_{20} & \bar{V}_{21} & \bar{V}_{22} & \cdots & \bar{V}_{2l} \\ \vdots & & & & \\ \bar{V}_{l0} & \bar{V}_{l1} & \bar{V}_{l2} & \cdots & \bar{V}_{ll} \end{bmatrix}\right) \tag{A-23}
$$

The joint distribution can be decomposed into conditional distribution of φ_k given $\varphi_{k-1} \cdots \varphi_0$ and the marginal distribution of φ_0:

$$\begin{aligned} p(vec(\varphi_0)) &= \mathcal{N}(vec(\bar{\varphi}_0), \bar{\bar{V}}_{00}) && \text{(A-24)} \\ p(vec(\varphi_k)|vec(\varphi_{k-1}) \cdots vec(\varphi_0)) &= \mathcal{N}(\theta_k, \Delta_{kk}) && \text{(A-25)} \end{aligned}$$

where θ_k and Δ_{kk} abbreviate rather complex expressions:

$$\theta_k = vec(\bar{\varphi}) + \left[\begin{array}{c} \bar{V}_{k0} \cdots \bar{V}_{kk-1} \end{array} \right] \left[\begin{array}{ccc} \bar{V}_{00} & \cdots & \bar{V}_{0k-1} \\ \vdots & \ddots & \vdots \\ \bar{V}_{k-1,0} & \cdots & \bar{V}_{k-1,k-1} \end{array} \right]^{-1} \left[\begin{array}{c} vec(\varphi_0 - \bar{\varphi}_0) \\ \vdots \\ vec(\varphi_{k-1} - \bar{\varphi}_{k-1}) \end{array} \right] \tag{A-26}$$

$$\Delta_{kk} = \bar{V}_{kk} - \left[\begin{array}{ccc} \bar{V}_{k0} & \cdots & \bar{V}_{kk-1} \end{array} \right] \left[\begin{array}{ccc} \bar{V}_{00} & \cdots & \bar{V}_{0k-1} \\ \vdots & \ddots & \vdots \\ \bar{V}_{k-1,0} & \cdots & \bar{V}_{k-1,k-1} \end{array} \right]^{-1} \left[\begin{array}{c} \bar{V}_{0k} \\ \vdots \\ \bar{V}_{k-1,k} \end{array} \right] \tag{A-27}$$

A.3.2 Prior for structural coefficients

The prior distribution of the impulse response functions map into a prior distribution of the structural coefficients for the structural VAR defined in (1). The marginal prior distribution of A is given by:

$$p(A) \propto |det(A)|^{-2m} exp\{-0.5(vec(A^{-1}) - vec(\bar{\varphi}))' \bar{V}_{00}^{-1} (vec(A^{-1}) - vec(\bar{\varphi}))\} \tag{A-28}$$

The prior distibution of A_k conditional $A_{k-1} \cdots A$ is defined as:

$$p(A_k|A_{k-1} \cdots A) \propto |\bar{\bar{V}}_k| exp\{-0.5(vec(A_k) - vec(\bar{\bar{\varphi}}_k))' \bar{\bar{V}}_k (vec(A_k) - vec(\bar{\bar{\varphi}}_k))\} \tag{A-29}$$

with

$$\begin{aligned} vec(\bar{\bar{\varphi}}_k) &= (A' \otimes A)(\theta_k - vec(f(A \cdots A_{k-1}))) && \text{(A-30)} \\ vec(\bar{\bar{\varphi}}_1) &= (A \otimes A)\theta_1 && \text{(A-31)} \end{aligned}$$

and

$$vec(f(A \cdots A_{k-1})) = vec(B_1 \varphi_{k-1} + B_2 \varphi_{k-2} + \cdots B_{k-1} \varphi_k) \tag{A-32}$$

A.3.3 The Likelihood

The likelihood is again split into a marginal likelihood of A and conditional likelihoods of all other structural coefficient matrices.

$$vec(F) = \begin{bmatrix} vec(A_1) \\ vec(A_2) \\ \vdots \\ vec(A_l) \end{bmatrix} \tag{A-33}$$

with

$$vec(F) \sim N(vec(\hat{F}), (X'X \otimes I_m) \tag{A-34}$$

$$vec(\hat{F}) = (\mathbf{I}_l \otimes A_0)vec(\hat{\Pi}) \tag{A-35}$$

$$\hat{\Pi} = Y'X(X'X)^{-1} \tag{A-36}$$

A further definition:

$$(X'X)^{-1} \otimes I_m \equiv \begin{bmatrix} \Xi_{11} & \Xi_{12} & \cdots & \Xi_{1p} & \Xi_{1c} \\ \Xi_{21} & \Xi_{22} & \cdots & \Xi_{2p} & \Xi_{2c} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \Xi_{p1} & \Xi_{p2} & \cdots & \Xi_{pp} & \Xi_{pc} \\ \Xi_{c1} & \Xi_{c2} & \cdots & \Xi_{cp} & \Xi_{cc} \end{bmatrix} \tag{A-37}$$

The marginal distribution of A:

$$P(A \mid Y) \propto \mid det(A) \mid^T exp\left\{-\frac{1}{2}vec(A)'(Q \otimes I_m)vec(A)\right\} \tag{A-38}$$

The conditional distributions are then:

$$p(vec(A_k)|vec(A_{k-1} \cdots A, Y) = \mathcal{N}(\mu_k, \Sigma_k) \tag{A-39}$$

with:

$$\mu_k = vec(\hat{A}_k) + \begin{bmatrix} \Xi_{k1} & \cdots & \Xi_{k,k-1} \end{bmatrix} \begin{bmatrix} \Xi_{11} & \cdots & \Xi_{1,k-1} \\ \vdots & \ddots & \vdots \\ \Xi_{k-1,1} & \cdots & \Xi_{k-1,k-1} \end{bmatrix}^{-1} \begin{bmatrix} vec(A_1 - \hat{A}_1) \\ \vdots \\ vec(A_{k-1} - \hat{A}_{k-1}) \end{bmatrix} \tag{A-40}$$

$$\Sigma_k = \Xi_{kk} - \begin{bmatrix} \Xi_{k1} & \cdots & \Xi_{k,k-1} \end{bmatrix} \begin{bmatrix} \Xi_{11} & \cdots & \Xi_{1,k-1} \\ \vdots & \ddots & \vdots \\ \Xi_{k-1,1} & \cdots & \Xi_{k-1,k-1} \end{bmatrix}^{-1} \begin{bmatrix} \Xi_{1k} \\ \vdots \\ \Xi_{k-1,k} \end{bmatrix} \tag{A-41}$$

A.3.4 Posterior of structural coefficients

Decomposing the posterior in a similar way, the marginal posterior of A is given by:

$$\begin{aligned} p(A|Y) \propto |det(A)|^{T-2m(\rho+1)} \times exp\{-0.5vec(A)'(Q \otimes I_m)vec(A)\} \\ \times\{-0.5(vec(A^{-1}) - vec(\varphi_0))'\bar{V}_{00}^{-1}(vec(A^{-1}) - vec(\varphi_0))\} \\ \times \prod_{i=1}^{\rho} |\Sigma_k^{-1} + \bar{\bar{V}}_k^{-1}|^{-\frac{1}{2}} \end{aligned} \tag{A-42}$$

The conditional posterior distribution for A_k can be written as:

$$p(A_k|A_{k-1}\cdots A, Y) \propto exp\{-0.5(vec(A_k)-\tilde{\mu}_k)'\tilde{\Sigma}_k^{-1}(vec(A_k)-\tilde{\mu}_k)\} \qquad \text{(A-43)}$$

where:

$$\tilde{\Sigma}_k = (\Sigma_k^{-1} + \bar{\bar{V}}_k^{-1})^{-1} \qquad \text{(A-44)}$$

$$\tilde{\mu}_k = \tilde{\Sigma}_k(\Sigma_k^{-1}\mu_k + \bar{\bar{V}}_k^{-1} vec(\bar{\bar{\varphi}}_k)) \qquad \text{(A-45)}$$

μ_k and Σ_k represent the mean and variance of the likelihood estimates conditional on A defined similar to the prior distribution. The resulting impulse response functions of the VAR can then be employed to compute the conditional distribution $p(\theta|\varphi^V)$. As pointed out by Kociecki (2005), the probabilistic restrictions can be thought of as a generalization of the sign restriction approach. Both approaches presented are therefore not exclusive.

A.4 Figures

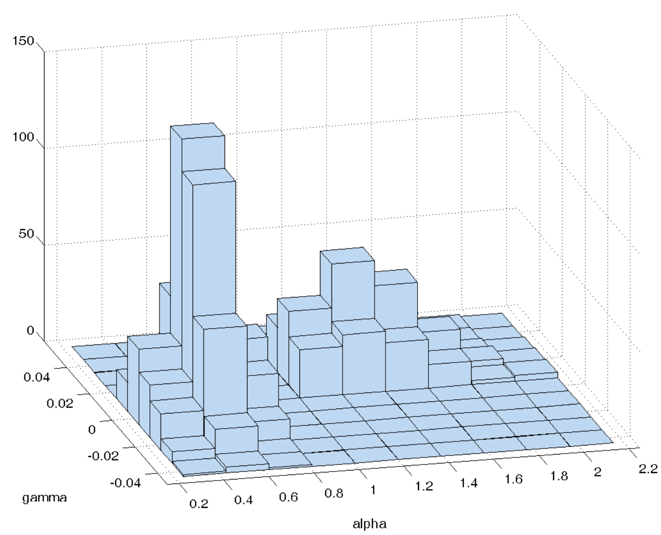

Figure 1: Prior distribution

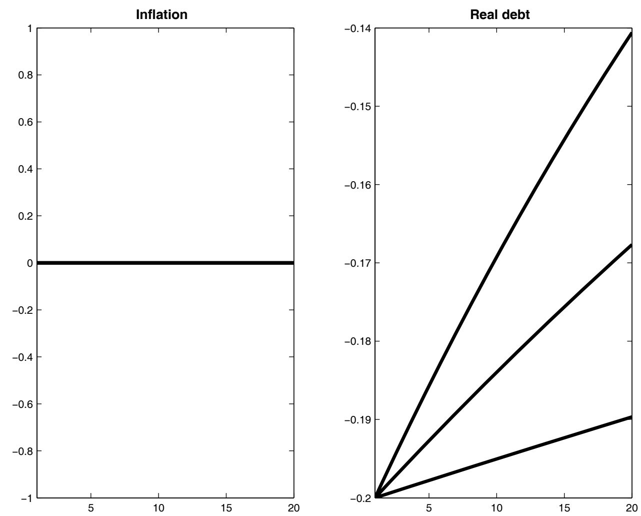

Figure 2: Bayesian IRF for a fiscal policy shock regimeI

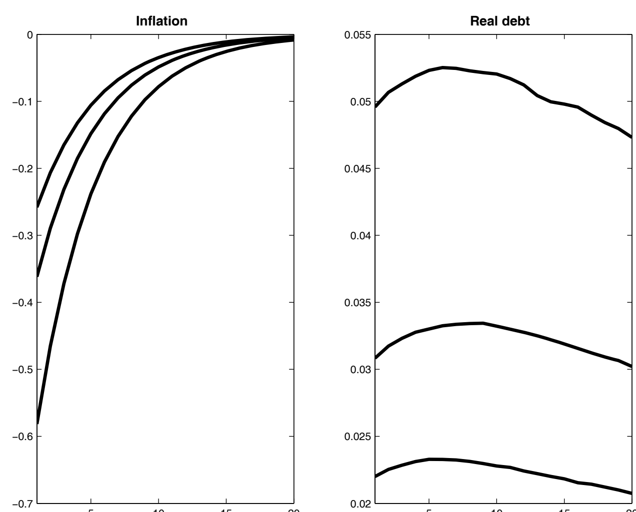

Figure 3: Bayesian IRF for a monetary policy shock regimeI

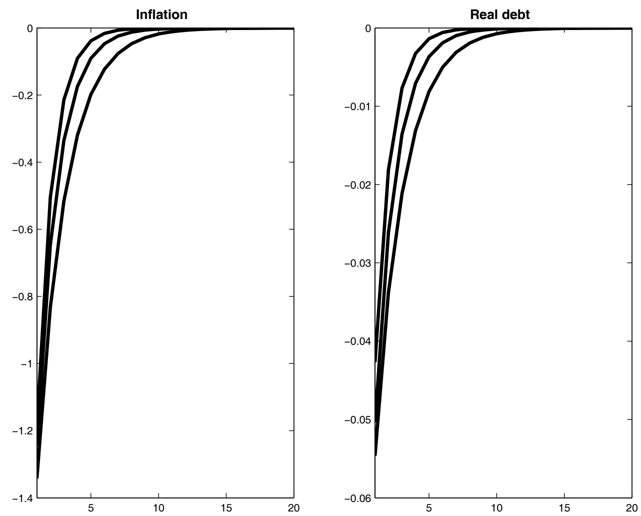

Figure 4: Bayesian IRF for a fiscal policy shock regimeII

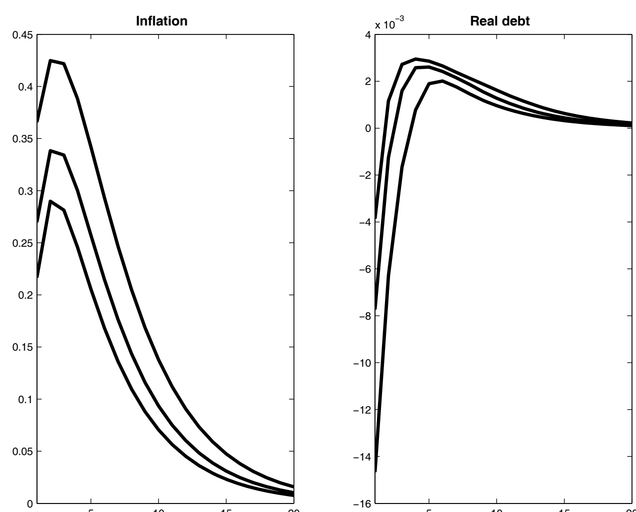

Figure 5: Bayesian IRF for a monetary policy shock regimeII

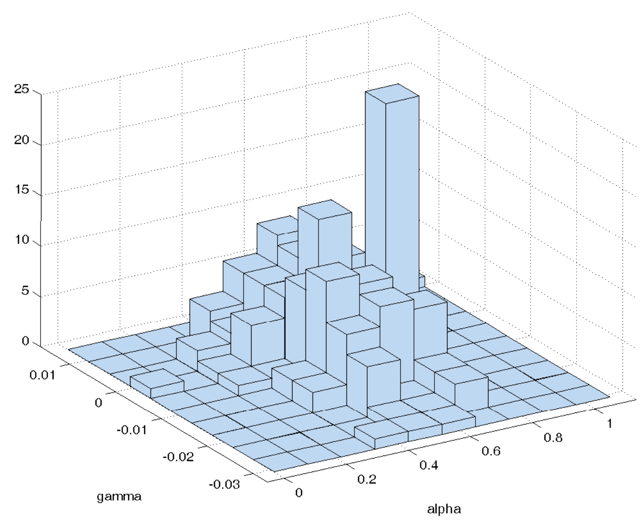

Figure 6: Posterior distribution of α and γ

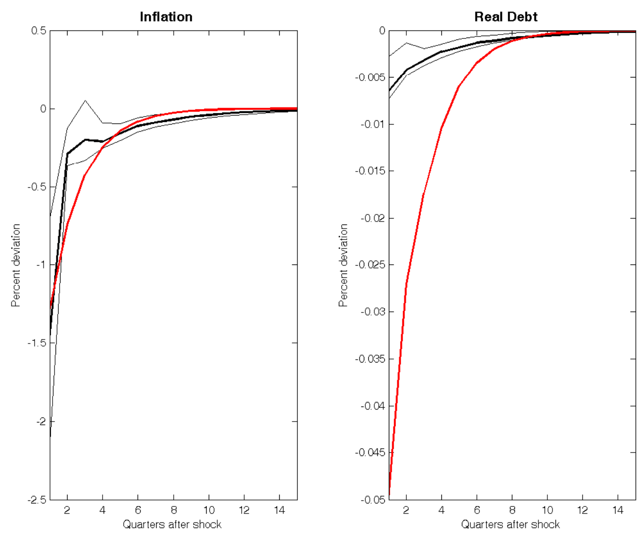

Figure 7: Posterior distribution (median and 16% and 84% probability bands) of impulse responses to a fiscal policy shock for the VAR model (black) and the median of the impulse responses of the DSGE model (red)

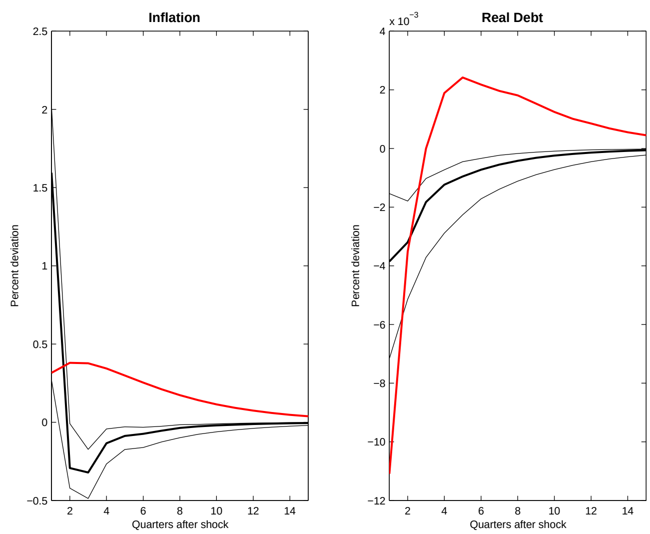

Figure 8: Posterior distribution (median and 16% and 84% probability bands) of impulse responses to a monetary policy shock for the VAR model (black) and the median of the impulse responses of the DSGE model (red)

SFB 649 Discussion Paper Series 2008

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Testing Monotonicity of Pricing Kernels" by Yuri Golubev, Wolfgang Härdle and Roman Timonfeev, January 2008.
002 "Adaptive pointwise estimation in time-inhomogeneous time-series models" by Pavel Cizek, Wolfgang Härdle and Vladimir Spokoiny, January 2008.
003 "The Bayesian Additive Classification Tree Applied to Credit Risk Modelling" by Junni L. Zhang and Wolfgang Härdle, January 2008.
004 "Independent Component Analysis Via Copula Techniques" by Ray-Bing Chen, Meihui Guo, Wolfgang Härdle and Shih-Feng Huang, January 2008.
005 "The Default Risk of Firms Examined with Smooth Support Vector Machines" by Wolfgang Härdle, Yuh-Jye Lee, Dorothea Schäfer and Yi-Ren Yeh, January 2008.
006 "Value-at-Risk and Expected Shortfall when there is long range dependence" by Wolfgang Härdle and Julius Mungo, Januray 2008.
007 "A Consistent Nonparametric Test for Causality in Quantile" by Kiho Jeong and Wolfgang Härdle, January 2008.
008 "Do Legal Standards Affect Ethical Concerns of Consumers?" by Dirk Engelmann and Dorothea Kübler, January 2008.
009 "Recursive Portfolio Selection with Decision Trees" by Anton Andriyashin, Wolfgang Härdle and Roman Timofeev, January 2008.
010 "Do Public Banks have a Competitive Advantage?" by Astrid Matthey, January 2008.
011 "Don't aim too high: the potential costs of high aspirations" by Astrid Matthey and Nadja Dwenger, January 2008.
012 "Visualizing exploratory factor analysis models" by Sigbert Klinke and Cornelia Wagner, January 2008.
013 "House Prices and Replacement Cost: A Micro-Level Analysis" by Rainer Schulz and Axel Werwatz, January 2008.
014 "Support Vector Regression Based GARCH Model with Application to Forecasting Volatility of Financial Returns" by Shiyi Chen, Kiho Jeong and Wolfgang Härdle, January 2008.
015 "Structural Constant Conditional Correlation" by Enzo Weber, January 2008.
016 "Estimating Investment Equations in Imperfect Capital Markets" by Silke Hüttel, Oliver Mußhoff, Martin Odening and Nataliya Zinych, January 2008.
017 "Adaptive Forecasting of the EURIBOR Swap Term Structure" by Oliver Blaskowitz and Helmut Herwatz, January 2008.
018 "Solving, Estimating and Selecting Nonlinear Dynamic Models without the Curse of Dimensionality" by Viktor Winschel and Markus Krätzig, February 2008.
019 "The Accuracy of Long-term Real Estate Valuations" by Rainer Schulz, Markus Staiber, Martin Wersing and Axel Werwatz, February 2008.
020 "The Impact of International Outsourcing on Labour Market Dynamics in Germany" by Ronald Bachmann and Sebastian Braun, February 2008.
021 "Preferences for Collective versus Individualised Wage Setting" by Tito Boeri and Michael C. Burda, February 2008.

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

022 "Lumpy Labor Adjustment as a Propagation Mechanism of Business Cycles" by Fang Yao, February 2008.
023 "Family Management, Family Ownership and Downsizing: Evidence from S&P 500 Firms" by Jörn Hendrich Block, February 2008.
024 "Skill Specific Unemployment with Imperfect Substitution of Skills" by Runli Xie, March 2008.
025 "Price Adjustment to News with Uncertain Precision" by Nikolaus Hautsch, Dieter Hess and Christoph Müller, March 2008.
026 "Information and Beliefs in a Repeated Normal-form Game" by Dietmar Fehr, Dorothea Kübler and David Danz, March 2008.
027 "The Stochastic Fluctuation of the Quantile Regression Curve" by Wolfgang Härdle and Song Song, March 2008.
028 "Are stewardship and valuation usefulness compatible or alternative objectives of financial accounting?" by Joachim Gassen, March 2008.
029 "Genetic Codes of Mergers, Post Merger Technology Evolution and Why Mergers Fail" by Alexander Cuntz, April 2008.
030 "Using R, LaTeX and Wiki for an Arabic e-learning platform" by Taleb Ahmad, Wolfgang Härdle, Sigbert Klinke and Shafeeqah Al Awadhi, April 2008.
031 "Beyond the business cycle – factors driving aggregate mortality rates" by Katja Hanewald, April 2008.
032 "Against All Odds? National Sentiment and Wagering on European Football" by Sebastian Braun and Michael Kvasnicka, April 2008.
033 "Are CEOs in Family Firms Paid Like Bureaucrats? Evidence from Bayesian and Frequentist Analyses" by Jörn Hendrich Block, April 2008.
034 "JBendge: An Object-Oriented System for Solving, Estimating and Selecting Nonlinear Dynamic Models" by Viktor Winschel and Markus Krätzig, April 2008.
035 "Stock Picking via Nonsymmetrically Pruned Binary Decision Trees" by Anton Andriyashin, May 2008.
036 "Expected Inflation, Expected Stock Returns, and Money Illusion: What can we learn from Survey Expectations?" by Maik Schmeling and Andreas Schrimpf, May 2008.
037 "The Impact of Individual Investment Behavior for Retirement Welfare: Evidence from the United States and Germany" by Thomas Post, Helmut Gründl, Joan T. Schmit and Anja Zimmer, May 2008.
038 "Dynamic Semiparametric Factor Models in Risk Neutral Density Estimation" by Enzo Giacomini, Wolfgang Härdle and Volker Krätschmer, May 2008.
039 "Can Education Save Europe From High Unemployment?" by Nicole Walter and Runli Xie, June 2008.
040 "Solow Residuals without Capital Stocks" by Michael C. Burda and Battista Severgnini, August 2008.
041 "Unionization, Stochastic Dominance, and Compression of the Wage Distribution: Evidence from Germany" by Michael C. Burda, Bernd Fitzenberger, Alexander Lembcke and Thorsten Vogel, March 2008
042 "Gruppenvergleiche bei hypothetischen Konstrukten – Die Prüfung der Übereinstimmung von Messmodellen mit der Strukturgleichungs-methodik" by Dirk Temme and Lutz Hildebrandt, June 2008.
043 "Modeling Dependencies in Finance using Copulae" by Wolfgang Härdle, Ostap Okhrin and Yarema Okhrin, June 2008.
044 "Numerics of Implied Binomial Trees" by Wolfgang Härdle and Alena Mysickova, June 2008.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

045 "Measuring and Modeling Risk Using High-Frequency Data" by Wolfgang Härdle, Nikolaus Hautsch and Uta Pigorsch, June 2008.
046 "Links between sustainability-related innovation and sustainability management" by Marcus Wagner, June 2008.
047 "Modelling High-Frequency Volatility and Liquidity Using Multiplicative Error Models" by Nikolaus Hautsch and Vahidin Jeleskovic, July 2008.
048 "Macro Wine in Financial Skins: The Oil-FX Interdependence" by Enzo Weber, July 2008.
049 "Simultaneous Stochastic Volatility Transmission Across American Equity Markets" by Enzo Weber, July 2008.
050 "A semiparametric factor model for electricity forward curve dynamics" by Szymon Borak and Rafał Weron, July 2008.
051 "Recurrent Support Vector Regreson for a Nonlinear ARMA Model with Applications to Forecasting Financial Returns" by Shiyi Chen, Kiho Jeong and Wolfgang K. Härdle, July 2008.
052 "Bayesian Demographic Modeling and Forecasting: An Application to U.S. Mortality" by Wolfgang Reichmuth and Samad Sarferaz, July 2008.
053 "Yield Curve Factors, Term Structure Volatility, and Bond Risk Premia" by Nikolaus Hautsch and Yangguoyi Ou, July 2008.
054 "The Natural Rate Hypothesis and Real Determinacy" by Alexander Meyer-Gohde, July 2008.
055 "Technology sourcing by large incumbents through acquisition of small firms" by Marcus Wagner, July 2008.
056 "Lumpy Labor Adjustment as a Propagation Mechanism of Business Cycle" by Fang Yao, August 2008.
057 "Measuring changes in preferences and perception due to the entry of a new brand with choice data" by Lutz Hildebrandt and Lea Kalweit, August 2008.
058 "Statistics E-learning Platforms: Evaluation Case Studies" by Taleb Ahmad and Wolfgang Härdle, August 2008.
059 "The Influence of the Business Cycle on Mortality" by Wolfgang H. Reichmuth and Samad Sarferaz, September 2008.
060 "Matching Theory and Data: Bayesian Vector Autoregression and Dynamic Stochastic General Equilibrium Models" by Alexander Kriwoluzky, September 2008.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

HUMBOLDT-UNIVERSITÄT · ZU BERLIN ·