European University Institute

Department of Economics

EUI Working Papers

ECO 2008/23

A Statistical Comparison of Alternative Identification Schemes for Monetary Policy Shocks

Markku Lanne and Helmut Lütkepohl

EUROPEAN UNIVERSITY INSTITUTE
DEPARTMENT OF ECONOMICS

A Statistical Comparison of Alternative Identification Schemes for Monetary Policy Shocks

MARKKU LANNE

and

HELMUT LÜTKEPOHL

EUI Working Paper **ECO** 2008/23

ISSN 1725-6704

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy

http://www.eui.eu/
http://cadmus.eui.eu/

May 20, 2008

A Statistical Comparison of Alternative Identification Schemes for Monetary Policy Shocks

Markku Lanne
Department of Economics,
P.O. Box 17 (Arkadiankatu 7), FI-00014 University of Helsinki, FINLAND
email: markku.lanne@helsinki.fi

and

Helmut Lütkepohl
Department of Economics, European University Institute
Via della Piazzola 43, I-50133 Firenze, ITALY
email: helmut.luetkepohl@eui.eu

Abstract. Different identification schemes for monetary policy shocks have been proposed in the literature. They typically specify just-identifying restrictions in a standard structural vector autoregressive (SVAR) framework. Thus, in this framework the different schemes cannot be checked against the data with statistical tests. We consider different approaches how to use the data properties to augment the standard SVAR setup for identifying the shocks. Thereby it becomes possible to test models which are just identified in a standard setting. For monthly US data it is found that a model where monetary shocks are induced via the federal funds rate is the only one which cannot be rejected when the data properties are used for identification.

Key Words: Mixed normal distribution, structural vector autoregressive model, vector autoregressive process

JEL classification: C32

1 Introduction

In a standard structural vector autoregressive (SVAR) approach, it is typical that just enough restrictions are imposed to just-identify the structural shocks. Clearly, any assumptions regarding certain restrictions imposed on a model may be incorrect and imposing false restrictions may lead to biased results and wrong conclusions. Therefore the desire to impose as few restrictions as possible is understandable. The drawback is, however, that different just-identified models cannot be compared by statistical tests. Comparisons therefore often rely on plausibility checks. For example, one may prefer one model to another one because the impulse responses of the former model have more plausible shapes.

Monetary policy is an active research area where different SVAR models coexist. For example, Christiano, Eichenbaum and Evans (1999) (henceforth CEE) review a number of identification schemes for SVAR models which have been used in the related literature to specify monetary policy shocks and analyze their impact on the economy. In some of these schemes, the structural restrictions just-identify the monetary policy shocks. Thus, in the standard setup, the different schemes cannot be checked against the data with statistical tests but a decision on which scheme to use is based on the subjective opinion of specific researchers. For example, CEE suppose that certain reactions to monetary policy shocks are widely accepted in the profession and therefore only identification schemes for monetary policy shocks should be considered which imply these widely accepted responses of the variables. Even if one accepts this strategy, it may not lead to a unique set of shocks.

Therefore, in this study we will use different approaches which use certain statistical properties of the data which can generate additional identifying information for the structural shocks. The first approach of this kind utilizes the change in the volatility of the shocks. It was proposed by Rigobon (2003), Rigobon and Sack (2003) and Lanne and Lütkepohl (2008a). It has been argued in the related literature that there has been a moderation of marcoeconomic fluctuations in the mid 1980s and our first approach uses this feature to identify monetary policy shocks. In this approach the change in the covariance structure of the model is assumed to have occurred at some prespecified point in time. An alternative approach was proposed by Lanne and Lütkepohl (2008b). They show how a nonnormal distribution of the residuals can be used as identifying information. In particular, they assume a mixed normal distribution of the residuals and show how this feature of the data can be utilized for identification purposes. Again statistical tests can be applied to check the normality of the data and, if rejected, allowing for a more general class of distributions is a natural next step. We will consider these two different approaches to compare a set of different identification schemes for US monetary policy.

More precisely, in our empirical analysis we use the same variables as CEE: log of real aggregate output (Y_t) and the log of its deflator (P_t), the smoothed change in an index of sensitive commodity prices ($PCOM_t$), the log of nonborrowed reserves plus extended credit (NBR_t), the log of total reserves (TR_t), the federal funds rate (FF_t) and the log of M1 (M_t). We use monthly US data only while CEE consider both monthly and quarterly data. As monthly data is available for most of the variables of interest here, it is worth utilizing the additional information in the more frequently observed series. Only output is not available in monthly form and therefore proxies as in CEE are used for this variable and its deflator (see Section 5). Our sampling period is 1965M7 - 1995M6 which is also the sampling period used by CEE. Although longer time series are now available, it may be worth considering exactly the same data as CEE to ensure that differences in the results are driven by the different methods used rather than different data.

Our study is structured as follows. In the next section the VAR model setup and the different sets of identifying restrictions considered by CEE will be presented. In Section 3 the 'statistical' identification strategies are presented. Estimation of the structural models is discussed in Section 4 and the results of the empirical analysis are presented in Section 5. Conclusions are drawn in Section 6.

2 The Model Setup

CEE consider a K-dimensional reduced form VAR(p) model of the type

$$Z_t = Dd_t + A_1 Z_{t-1} + \cdots + A_p Z_{t-p} + u_t, \tag{2.1}$$

where d_t is a deterministic term with coefficient matrix D, the A_j's ($j = 1, \ldots, p$) are $(K \times K)$ coefficient matrices and u_t is a white noise error term. They partition the K-dimensional vector of observable variables Z_t as

$$Z_t = \begin{bmatrix} X_{1t} \\ S_t \\ X_{2t} \end{bmatrix}, \tag{2.2}$$

where X_{1t} is $(k_1 \times 1)$, S_t is (1×1) and X_{2t} is $(k_2 \times 1)$. The vector X_{1t} contains variables whose contemporaneous values appear in the monetary authority's information set, i.e., variables orthogonal to the monetary policy shock. The variables in X_{2t} only appear with a lag in the information set and S_t is the monetary authority's policy instrument.

The structural shocks are usually obtained from the reduced form residuals by a linear transformation. If only specific structural shocks are of interest, it suffices to find a linear transformation of u_t which delivers these particular shocks of interest and possibly specifies the other shocks in an arbitrary way. In our case the monetary shocks are of primary interest. CEE show that a block triangular transformation $\varepsilon_t = A_0 u_t$, with

$$A_0 = \begin{bmatrix} \underset{(k_1\times k_1)}{a_{11}} & \underset{(k_1\times 1)}{0} & \underset{(k_1\times k_2)}{0} \\ \underset{(1\times k_1)}{a_{21}} & \underset{(1\times 1)}{a_{22}} & \underset{(1\times k_2)}{0} \\ \underset{(k_2\times k_1)}{a_{31}} & \underset{(k_2\times 1)}{a_{32}} & \underset{(k_2\times k_2)}{a_{33}} \end{bmatrix} \tag{2.3}$$

fully identifies the monetary policy shocks as the (k_1+1)th component of ε_t. Notice that the other elements of A_0 may be chosen such that $A_0^{-1}A_0^{-1\prime} = \Sigma_u$ and, hence, $\varepsilon_t \sim (0, I_K)$.

CEE consider three alternative schemes for just-identifying the monetary policy shocks:

1. *FF policy shock*: $X_{1t} = (Y_t, P_t, PCOM_t)'$, $X_{2t} = (NBR_t, TR_t, M_t)'$ and the federal funds rate is the policy instrument, that is, $S_t = FF_t$. This identification scheme is motivated by arguments presented in Bernanke and Blinder (1992), Sims (1986, 1992) and others.

2. *NBR policy shock*: $X_{1t} = (Y_t, P_t, PCOM_t)'$, $X_{2t} = (FF_t, TR_t, M_t)'$ and $S_t = NBR_t$. This scheme is based on work, for instance, of Christiano and Eichenbaum (1992).

3. *NBR/TR policy shock*: $X_{1t} = (Y_t, P_t, PCOM_t, TR_t)'$, $X_{2t} = (FF_t, M_t)'$ and $S_t = NBR_t$. CEE attribute this identification scheme to Strongin (1995).

As mentioned earlier, in a standard SVAR setting the implied zero restrictions on A_0 suffice to just-identify the monetary policy shocks. They do not provide over-identifying restrictions, however, which could be tested against the data. In the next section it will be explained how such over-identifying information can be obtained from the statistical properties of u_t which are usually not taken into account in a standard SVAR analysis.

3 Statistical Approaches for the Identification of Shocks

3.1 Identification via Heteroskedasticity

Our first approach to identify the structural shocks via specific statistical properties of the data assumes that there is at least one change in the volatility of the residuals and, hence, the residuals of the basic model (2.1) are heteroskedastic. As mentioned earlier, this approach has been used in SVAR analyses by Rigobon (2003), Rigobon and Sack (2003) and Lanne and Lütkepohl (2008a). In the first paper the relationship between the returns on different bonds is analyzed. Rigobon and Sack (2003) use the idea of identification via heteroskedasticity to investigate the relation between monetary policy and the stock market. Finally, Lanne and Lütkepohl (2008a) use this devise to compare different identification schemes for US monetary policy. The latter study is closely related to the one presented in the empirical section of the present paper. The model setup and sample period are slightly different, however.

To introduce the idea, let us assume that there is a single change in the volatility of the variables during the sample period. Hence, suppose we have a sample of size T, $Z_1, \ldots, Z_T$, and there is a change in the volatility of the shocks during the sample period, say in period T_B, so that

$$E(u_t u_t') = \begin{cases} \Sigma_1 & \text{for } t = 1, \ldots, T_B - 1, \\ \Sigma_2 & \text{for } t = T_B, \ldots, T. \end{cases} \tag{3.1}$$

From matrix theory it is well-known that the covariance matrices Σ_1 and Σ_2 can be diagonalized simultaneously, that is, there exists a $(K \times K)$ matrix W and a diagonal matrix $\Psi = \mathrm{diag}(\psi_1, \ldots, \psi_K)$ with positive diagonal elements ψ_i, $i = 1, \ldots, K$, such that $\Sigma_1 = WW'$ and $\Sigma_2 = W\Psi W'$ (e.g., Lütkepohl (1996, Section 6.1.2)). Here the diagonal elements of Ψ reflect the changes in the variances of the shocks after the possible change in volatility has occurred. In fact, a change in volatility has occurred if the ψ_i's are different from one. Lanne and Lütkepohl (2008b) show that W is unique except for sign changes if all ψ_i's are distinct and ordered in some way. Thus, if we choose $A_0^{-1} = W$, we get uniqueness of the shocks $\varepsilon_t = A_0 u_t$ (except for changes in sign) without the need for further identifying assumptions. Choosing $A_0 = W^{-1}$, the structural shocks have identity covariance matrix in the first regime and Ψ in the second regime, that is,

$$E(\varepsilon_t \varepsilon_t') = \begin{cases} I_K & \text{for } t = 1, \ldots, T_B - 1, \\ \Psi & \text{for } t = T_B, \ldots, T. \end{cases}$$

Hence, they are orthogonal in both regimes. In turn, requiring that they are orthogonal in both regimes suffices to identify our shocks uniquely if all ψ_i's are distinct. Notice that orthogonality of the structural shocks is a standard assumption in SVAR analysis.

The requirement that all ψ_i's are distinct is satisfied if the changes in volatility are not proportional in all variables. Even if the volatility in one of the shocks does not change at all, that is, one of the ψ_i's may be unity, the ψ_i's may of course be all distinct and this is the essential requirement for uniqueness of A_0. Any other restrictions for A_0, as for example formulated in Section 2, then become over-identifying. Since the change in variance is a testable assumption, we do not have to rely exclusively on information from economic theory or other sources to ensure identification of the structural shocks. Instead we can use the information in the data and apply statistical procedures to obtain identification.

To obtain uniqueness of W and, hence, of A_0 we may, for example, order the ψ_i's from smallest to largest. If restrictions are imposed on A_0, they may only be compatible with one specific ordering of the ψ_i's and not necessarily with an ordering according to size. Therefore, in estimations with restrictions on A_0 we do not impose any specific ordering on the ψ_i's but let the data decide which one is best. This is no problem in principle because local identification is ensured for any ordering of the ψ_i's, provided they are all distinct.

The fact that it is always possible to reverse the signs of all elements in a single column of $W = A_0^{-1}$ without affecting the likelihood is not a problem either in the present context because for asymptotic theory to work we only need local identification which is ensured despite the sign changes. For practical purposes changing all signs in a column of W just means to consider a negative shock if the shock is positive originally or vice versa.

It is also possible, of course, to accommodate more than one change in volatility. If there are $n + 1$ different regimes and the covariances in the different regimes are WW', $W\Psi_1W'$, $\dots$, $W\Psi_nW'$, where the Ψ_i's are all diagonal matrices, uniqueness of W (up to sign) is ensured, for example, if the diagonal elements in only one of the Ψ_i matrices are all distinct and we can again choose $A_0 = W^{-1}$.

So far we have assumed that only the residual covariance matrix changes and the other VAR parameters remain constant. This assumption was made for convenience because it ensures identification of the shocks. Obviously, statistical tests may be used to check the constancy of the other parameters as well. CEE and Bernanke and Mihov (1998a) argue that structural changes found by other authors in the data set underlying our empirical study may have been due to a change in the residual covariances only and not to a change in the whole dynamic structure. While identification of the shocks can also be achieved if other parameters vary as well, the impulse responses may be affected if there are changes in the other parameters. Such changes would therefore complicate our analysis.

If uniqueness of A_0 is ensured by residual heteroskedasticity, then all the restrictions from economic theories are over-identifying and, hence, can be tested. In particular, the just-identifying restrictions discussed in the previous section can be tested and we will do so in the empirical section. In other words, it can be checked whether they are compatible with orthogonality of the shocks across the different regimes. In the next subsection we will present another approach to serve the same purpose.

3.2 Mixed Normal Residuals

In our analysis of the reduced form VAR model for the time series described in Section 1, we will find strong evidence that the errors are not normally distributed. Hence, it makes sense to specify a more general distribution. A quite general class of distributions is given by a mixture of two normal distributions. Therefore we assume that u_t is a mixture of two serially independent normal random vectors such that

$$u_t = \begin{cases} e_{1t} \sim \mathcal{N}(0, \Sigma_1) \text{ with probability } \gamma, \\ e_{2t} \sim \mathcal{N}(0, \Sigma_2) \text{ with probability } 1-\gamma. \end{cases} \tag{3.2}$$

Here $\mathcal{N}(0, \Sigma)$ denotes a multivariate normal distribution with zero mean and covariance matrix Σ. The $(K \times K)$ covariance matrices Σ_1 and Σ_2 are assumed to be distinct and the mixture probability γ, $0 < \gamma < 1$, is a parameter of the model. If $\Sigma_1 = \Sigma_2$, u_t has a $\mathcal{N}(0, \Sigma_1)$ distribution and γ is not identified. Therefore we assume $\Sigma_1 \neq \Sigma_2$. We note that u_t has mean zero and covariance matrix $\Sigma_u = \gamma\Sigma_1 + (1-\gamma)\Sigma_2$, that is, $u_t \sim (0, \gamma\Sigma_1 + (1-\gamma)\Sigma_2)$.

This model was proposed by Lanne and Lütkepohl (2008b) who also show that a $(K \times K)$ matrix W and a diagonal matrix $\Psi = \text{diag}(\psi_1, \ldots, \psi_K)$, $\psi_i > 0$ $(i = 1, \ldots, K)$, exist such that $\Sigma_1 = WW'$ and $\Sigma_2 = W\Psi W'$. Thus, we may parameterize Σ_u as

$$\Sigma_u = W(\gamma I_K + (1-\gamma)\Psi)W'. \tag{3.3}$$

If all ψ_i's are distinct, then, for a given ordering of the ψ_i's, the matrix W in this decomposition is unique except that all signs of a column may be reversed. If we choose

$$A_0^{-1} = W(\gamma I_K + (1-\gamma)\Psi)^{1/2} \tag{3.4}$$

so that $\Sigma_u = A_0^{-1}A_0^{-1\prime}$, this decomposition of Σ_u is the unique one (apart from sign changes) which diagonalizes both Σ_1 and Σ_2 and also, of course, Σ_u. In other words, if we think of the two normal distributions which are mixed in (3.2) as representing two different regimes, the structural shocks are uncorrelated in both regimes. Hence, as in the heteroskedastic model, any additional restrictions on A_0 are over-identifying and, thus, testable. Note that this model differs from the previous one in that the allocation of time periods to the regimes is governed by a random mechanism whereas in Section 3.1 we have assumed that the regime changes occur at fixed, prespecified time points. Also, the residual distribution in (3.2) is not heteroskedastic.

4 Estimation

For our mixed normal model, maximum likelihood (ML) estimation is in principle the method of choice because we have made an assumption regarding the distribution of the residuals. Also for the heteroskedastic model using Gaussian ML estimation is useful because it results in estimators with desirable asymptotic properties even if the actual residual distribution is non-Gaussian. For both models the likelihood function and its normal equations are nonlinear in the parameters, however. Therefore maximizing the full likelihood function may be a formidable task if the dimension of the process, K, and/or the VAR order, p, are large, as in the case of our empirical example. Because the VAR coefficients can be estimated consistently by equation-wise OLS with standard asymptotic properties, it is in fact possible to estimate the structural parameters based on a "concentrated likelihood function" where the VAR coefficients are replaced by their OLS estimators. The resulting estimation methods for the two models of interest here will be discussed next. We denote by $\hat{u}_t = Z_t - \hat{D}d_t - \hat{A}_1 Z_{t-1} - \cdots - \hat{A}_p Z_{t-p}$ the residuals from estimating the reduced form VAR model (2.1) by equation-wise OLS.

4.1 Heteroskedastic Residuals

Let us focus on the case of two regimes with different residual covariance matrices as in (3.1) with $\Sigma_1 = WW'$ and $\Sigma_2 = W\Psi W'$ and define

$$\widetilde{\Sigma}_1 = \frac{1}{T_B - 1}\sum_{t=1}^{T_B-1} \hat{u}_t\hat{u}_t' \quad \text{and} \quad \widetilde{\Sigma}_2 = \frac{1}{T - T_B + 1}\sum_{t=T_B}^{T} \hat{u}_t\hat{u}_t'.$$

Replacing the VAR parameter estimators in the Gaussian log-likelihood function by their OLS estimators gives a "concentrated log likelihood function" of the form

$$\log L_H = -\frac{T_B - 1}{2}\left(\log\det(WW') + \operatorname{tr}\left\{\widetilde{\Sigma}_1(WW')^{-1}\right\}\right)$$
$$-\frac{T - T_B + 1}{2}\left(\log\det(W\Psi W') + \operatorname{tr}\left\{\widetilde{\Sigma}_2(W\Psi W')^{-1}\right\}\right). \tag{4.1}$$

Maximizing this function gives estimators $\widetilde{W}$ and $\widetilde{\Psi}$ of W and Ψ, respectively. Note, however, that these estimators are not full ML estimators even if the true residual distribution is Gaussian because the OLS estimators of the VAR coefficients from (2.1) are not ML estimators. They do not account for the heteroskedasticity in the residuals. We will use the estimators of the structural parameters obtained from maximizing (4.1) and refer the reader to Lanne and Lütkepohl (2008a) for further discussion of the estimation procedure and the properties of the estimators.

Any over-identifying restrictions imposed on the structural parameters can be tested by likelihood ratio (LR) type tests. Thus, if the identification schemes considered in Section 2 imply over-identifying restrictions we can check them against the data by LR type tests based on optimizing the objective function (4.1) with and without restrictions. Clearly, the resulting tests are not really LR tests because they are based on maximizing the pseudo concentrated likelihood in (4.1) rather than the fully maximized likelihood function. Still it can be seen from the discussion in Lanne and Lütkepohl (2008a) that they have the usual asymptotic properties of standard LR tests. Therefore we may use a χ^2 distribution with as many degrees of freedom (df) as there are zero restrictions imposed on A_0, provided all ψ_i's are distinct. If the latter condition is not satisfied, the asymptotic distribution of our pseudo LR test will still be χ^2 under general conditions. The number of df may be smaller than the number of zeros placed on A_0, however. In other words, our tests may be conservative when used with critical values from a χ^2 distribution with as many df as there are zero restrictions on A_0.

It is straightforward to extend the estimation procedure to the case of more than two regimes. We will not present the details to save space. In the empirical analysis models with up to three regimes will be used. It may also be worth noting that Rigobon (2003) has shown for his slightly less general setup that the time invariant parameters may be estimated consistently under usual assumptions even if the break times are fixed incorrectly.

4.2 Mixed Normal Residuals

For the mixed normal model (3.2) with $\Sigma_1 = WW'$ and $\Sigma_2 = W\Psi W'$ we estimate the parameters γ, Ψ and W by maximizing the pseudo concentrated likelihood function

$$L_{MN}(W,\Psi,\gamma) = \prod_{t=1}^{T} \hat{f}_{t-1}(Z_t), \tag{4.2}$$

where

$$\begin{aligned}\hat{f}_{t-1}(Z_t) &= \gamma \det(W)^{-1} \exp\left\{-\frac{1}{2}\hat{u}_t'(WW')^{-1}\hat{u}_t\right\} \\ &\quad + (1-\gamma)\det(\Psi)^{-1/2}\det(W)^{-1}\exp\left\{-\frac{1}{2}\hat{u}_t'(W\Psi W')^{-1}\hat{u}_t\right\}.\end{aligned}$$

Regarding tests of over-identifying restrictions for the structural parameters the same applies as for the heteroskedastic model. Thus, both model types allow us to test the restrictions for the A_0 parameters presented in Section 2 if at least some of the diagonal elements of Ψ are different and the model is a valid representation of the DGP.

5 Empirical Analysis

In the empirical analysis we use monthly US data for the period 1965M7 - 1995M6 which corresponds to the sample period used by CEE.[1] A similar sample period was also used in studies by Bernanke and Mihov (1998b) and Lanne and Lütkepohl (2008a). Our sample size is 360. We use nonfarm payroll employment as proxi for aggregate output and the implicit deflator of personal consumption expenditure as proxi for its deflator, as in CEE. The reduced form model is a 7-dimensional VAR(12) with an intercept.

The literature on the monetary transmission mechanism in the US presents evidence for a number of possible structural breaks during our sample period and in particular changes in the volatility of the shocks are diagnosed by different authors. Thus, our heteroskedastic model may be justified. Moreover, we have applied tests for nonnormality to the residuals of our model and have found clear evidence against Gaussian residuals. Therefore, considering a more general distribution class such as the mixed normal seems also reasonable. In the following we will consider both types of modelling assumptions. Thereby we will also be able to study the robustness of our main results with respect to the identifying assumptions.

[1]The data were obtained from L. Christiano's homepage http://www.faculty.econ.northwestern.edu/faculty/christiano/research.htm.

Table 1: Estimation Results for Parameters of VAR(12) Models with Heteroskedastic Errors for Sampling Period 1965M7 - 1995M6

	1 Regime		2 Regimes			
Parameter	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err
ψ_1	0.8000	0.1334	0.8735	0.2761	0.4623	0.0801
ψ_2	1.2588	0.2034	1.2161	0.4405	0.7458	0.1284
ψ_3	0.4927	0.0740	0.9791	0.2845	0.7992	0.1397
ψ_4	0.3063	0.0487	5.2158	1.1729	0.6916	0.1179
ψ_5	0.3866	0.0686	1.5875	0.4137	1.8507	0.3110
ψ_6	1.6341	0.2919	1.7791	0.5160	1.4448	0.2481
ψ_7	0.7189	0.1131	0.6639	0.3452	0.3619	0.0757

5.1 Results for Heteroskedastic Models

As mentioned earlier, different dates of possible volatility changes were found in the literature for similar models during our sample period. In the following we will only consider changes in 1979M10 and 1984M2. Bernanke and Mihov (1998b) and CEE agree on these dates. According to Bernanke and Mihov (1998b, p. 880) the choice of these two break dates is based on a "combination of historical and statistical evidence." Moreover, Lanne and Lütkepohl (2008a) present further statistical evidence for changes in the residual covariances of models similar to ours in these two months. Notice that the intermediate period 1979M10 - 1984M2 roughly corresponds to the Volcker era which is often regarded as special as far as monetary policy is concerned.

There is some disagreement in the literature regarding the type of structural break. Our assumption of a heteroskedastic model is supported by Bernanke and Mihov (1998b) and CEE. Assuming changes only in the disturbance covariance matrices and, hence, in the volatility of the structural shocks is not uncommon in the related literature (see, for example, Sims and Zha (2006)). In summary, our heteroskedastic model and our assumptions regarding the timing of changes in the volatility are not unconventional and have been confirmed by a variety of methods and authors.

For illustrative purposes and to check the robustness of our results we consider a model with just one change in the residual covariance in 1984M2 and one with two changes in 1979M10 and 1984M2. The evidence for a change in 1984M2 was somewhat stronger than for 1979M10 in the study by Lanne and Lütkepohl (2008a). It is therefore plausible to use 1984M2 as break date if only one break is considered.

The estimated ψ_i's for both models are presented in Table 1. These parameters are of particular interest because the shocks are fully identified

Table 2: LR Type Tests of Identification Schemes Based on Heteroskedastic Models

Regime Change in 1984M2

	H_0	df	mean loglik	LR	p-value
NBR/TR	$a_{12} = 0$	4	5.3441	3.4104	0.4916
	$a_{13} = 0$	8	5.3337	10.6070	0.2250
	$a_{23} = 0$	2	5.3485	0.3271	0.8491
	$a_{12} = a_{13} = a_{23} = 0$	14	5.2895	41.3633	0.0002
NBR	$a_{12} = 0$	3	5.3481	0.5916	0.8984
	$a_{13} = 0$	9	5.3388	7.0505	0.6319
	$a_{23} = 0$	3	5.3484	0.4176	0.9366
	$a_{12} = a_{13} = a_{23} = 0$	15	5.3164	22.6687	0.0914
FF	$a_{12} = 0$	3	5.3479	0.7726	0.8560
	$a_{13} = 0$	9	5.3362	8.8949	0.4470
	$a_{23} = 0$	3	5.3477	0.8700	0.8327
	$a_{12} = a_{13} = a_{23} = 0$	15	5.3320	11.7833	0.6954

Regime Change in 1979M10 and 1984M2

	H_0	df	mean loglik	LR	p-value
NBR/TR	$a_{12} = 0$	4	5.4155	26.3575	2.6802e–5
	$a_{13} = 0$	8	5.4233	20.9426	0.0073
	$a_{23} = 0$	2	5.4525	0.6334	0.7286
	$a_{12} = a_{13} = a_{23} = 0$	14	5.3660	60.8736	8.3617e–8
NBR	$a_{12} = 0$	3	5.4516	1.2528	0.7404
	$a_{13} = 0$	9	5.4119	28.8631	0.0007
	$a_{23} = 0$	3	5.4514	1.3433	0.7189
	$a_{12} = a_{13} = a_{23} = 0$	15	5.3991	37.7858	0.0010
FF	$a_{12} = 0$	3	5.4494	2.7770	0.4273
	$a_{13} = 0$	9	5.4402	9.1872	0.4202
	$a_{23} = 0$	3	5.4512	1.5451	0.6719
	$a_{12} = a_{13} = a_{23} = 0$	15	5.4284	17.4070	0.2951

by assuming orthogonality in both regimes if all the ψ_i's are distinct. Taking into account the standard errors in Table 1, it is not clear that all the ψ_i's are distinct in the two models. On the other hand, the estimates present strong evidence that at least some of the ψ_i's in each of the two models are different. That result is in fact sufficient to test the structural restrictions from Section 2. Notice that if some of the ψ_i's are distinct, at least some of the structural restrictions can be tested. Using the LR type tests mentioned in Section 4 for checking the restrictions, the number of degrees of freedom of the asymptotic χ^2 distributions may be lower than the number of zeros imposed by the different identification schemes. Hence, it would reduce the estimated p-values. Thus, the p-values of our tests based on the assumption that all ψ_i's are distinct would actually be conservative.

In Table 2 results for both models are presented with p-values based on the assumption of distinct ψ_i's. Given that these p-values are conservative, any model that can be rejected on the basis of the p-value in Table 2 can also be rejected if some of the ψ_i's are equal.

The restrictions $a_{12} = 0$ and $a_{13} = 0$ imply that the monetary policy shock is orthogonal to the elements of X_{1t}; a_{12} corresponds to the direct effect of S_t on X_{1t} and a_{13} to the indirect effect via the impact of the shock on X_{2t}. The restriction $a_{23} = 0$ is derived from the assumption that the monetary policy authority does not see X_{2t} when setting S_t. For a correct identification scheme none of the null hypotheses in Table 2 should be rejected. Clearly, this condition is not satisfied for the NBR/TR scheme. This scheme is rejected in both models, with one or two changes in covariance. At least the p-values for the tests of $a_{12} = a_{13} = a_{23} = 0$ are smaller than 1% and, hence, the NBR/TR scheme is clearly rejected even with our conservative tests. In fact, if two regime changes are allowed for, there are even more rejections and, hence, the evidence against the NBR/TR scheme is quite strong in our setup.

For the NBR identification scheme the situation is also quite clear in the model with two structural changes because both $a_{13} = 0$ and $a_{12} = a_{13} = a_{23} = 0$ produce very small p values below 1% and are, hence, rejected at common significance levels. The situation is different, however, if only one change in covariance is allowed for. In that case, $a_{12} = a_{13} = a_{23} = 0$ is the only restriction that can be rejected at the 10% level in the NBR scheme. Given that the model with two breaks is more credible and given that our tests are potentially asymptotically conservative, these results present considerable evidence also against the NBR identification scheme.

The situation is quite different for the FF identification scheme. Here none of the p-values is even close to a reasonable significance level for a usual test. In fact, all p-values are bigger than 20%. Thus, the FF scheme is the only one which can stand up against the data in our setup. Since our tests are potentially conservative if not all ψ_i's are distinct, it may well be that our tests do not have enough power to show that even this scheme is not compatible with the data. However, if anyone of the three identification schemes is consistent with the data, it is the FF scheme, at least in our testing framework.

One lesson to be learned from this exercise is that the model setup has a substantial impact on the results. Ignoring one of the changes in the residual covariance matrix can make a substantial difference. Nevertheless there is also a considerable robustness in our results. Even if the change in 1979M10 is ignored the results point in the same direction as in the model which allows for two changes. A sufficiently critical interpretation of the p-values would result in the same overall conclusions in both models.

Our general result is to some extent in line with findings by Lanne and Lütkepohl (2008a). Using a slightly different setup in which they also test restrictions on the deeper parameters of the monetary models, they do not find clear statistical evidence against either the NBR/TR or the FF model. However, based on other criteria they find the NBR/TR model to be the most plausible one. Clearly, the latter result is at variance with our tests presented in Table 2.

5.2 Results for Model with Mixed Normal Residuals

As mentioned earlier, there is substantial statistical evidence against Gaussian residuals. Therefore fitting a mixed normal distribution to the residuals using the method described in Section 4 becomes a plausible alternative to the approach used in the previous subsection. The estimated ψ_i's are given in Table 3. Clearly, taking into account the estimated standard errors, there is strong evidence that at least some ψ_i's are distinct. Therefore we proceed under this assumption in the following. Again, our tests of restrictions for the structural parameters may be conservative if some of the ψ_i's are in fact identical.

Interestingly, the results of the pseudo LR tests presented in Table 4 are fully in line with those from the heteroskedastic model. The NBR/TR and NBR models are strongly rejected because some of the p-values are smaller than 1%, whereas the FF model can not be rejected at common significance levels. Thus, if we let the data decide on the allocation of regimes rather than fixing the change dates as in the heteroskedastic model, produces basically the same conclusions regarding the different identification schemes for monetary policy shocks. Thus, our results are overall quite robust to variations in our identifying assumptions.

Table 3: Estimation Results for Parameters of VAR(12) Model with Mixed Normal Errors for Sampling Period 1965M7 - 1995M6

Parameter	Estimate	Std Err
ψ_1	0.3400	0.0887
ψ_2	0.6440	0.1581
ψ_3	1.6064	0.4045
ψ_4	2.0717	0.5018
ψ_5	3.1103	0.7102
ψ_6	5.1202	1.1495
ψ_7	7.0656	1.5924
γ	0.8014	0.0441

Table 4: LR Type Tests of Identification Schemes Based on Mixed Normal Model

	H_0	df	mean loglik	LR	p-value
NBR/TR	$a_{12} = 0$	4	5.3722	19.2931	0.0007
	$a_{13} = 0$	8	5.3953	3.2086	0.9206
	$a_{23} = 0$	2	5.3976	1.5938	0.4507
	$a_{12} = a_{13} = a_{23} = 0$	14	5.3707	20.3441	0.1197
NBR	$a_{12} = 0$	3	5.3847	10.6070	0.0141
	$a_{13} = 0$	9	5.3628	25.8216	0.0022
	$a_{23} = 0$	3	5.3974	1.7330	0.6296
	$a_{12} = a_{13} = a_{23} = 0$	15	5.3625	26.0513	0.0375
FF	$a_{12} = 0$	3	5.3979	1.4059	0.7042
	$a_{13} = 0$	9	5.3832	11.6023	0.2367
	$a_{23} = 0$	3	5.3987	0.8352	0.8410
	$a_{12} = a_{13} = a_{23} = 0$	15	5.3745	17.6575	0.2811

6 Conclusions

In this study we have compared three identification schemes for monetary policy shocks which cannot be tested in a standard SVAR framework because in that setting there are no over-identifying restrictions. We utilize the fact that the underlying reduced form VAR model has a potentially changing covariance structure and that the residuals are clearly nonnormal. These data features allow us to get additional identifying information and enables us to test the identification schemes for the monetary policy shocks against the data. Only one of the three identification schemes is not rejected in this framework. More precisely, a scheme where monetary shocks are induced via the federal funds rate is the only one which cannot be rejected in our framework. This result is robust with respect to the specific statistical setup used. It is obtained for both the heteroskedastic model and a model with mixed normal residuals.

References

Bernanke, B. S. and Blinder, A. (1992). The federal funds rate and the channels of monetary transmission, *American Economic Review* **82**: 901–921.

Bernanke, B. S. and Mihov, I. (1998a). The liquidity effect and long-run neutrality, *Carnegie-Rochester Conference Series on Public Policy* **49**: 149–194.

Bernanke, B. S. and Mihov, I. (1998b). Measuring monetary policy, *Quarterly Journal of Economics* **113**: 869–902.

Christiano, L. J. and Eichenbaum, M. (1992). Identification and the liquidity effect of a monetary policy shock, *in* A. Cukierman, Z. Hercowitz and L. Leiderman (eds), *Political Economy, Growth, and Business Cycles*, MIT Press, Cambridge, MA, pp. 335–370.

Christiano, L. J., Eichenbaum, M. and Evans, C. (1999). Monetary policy shocks: What have we learned and to what end?, *in* J. B. Taylor and M. Woodford (eds), *Handbook of Macroeconomics*, Vol. 1A, Elsevier, Amsterdam, pp. 65–148.

Lanne, M. and Lütkepohl, H. (2008a). Identifying monetary policy shocks via changes in volatility, *Journal of Money, Credit and Banking.*

Lanne, M. and Lütkepohl, H. (2008b). Structural vector autoregressions with nonnormal residuals, *Journal of Business & Economic Statistics.*

Lütkepohl, H. (1996). *Handbook of Matrices*, John Wiley & Sons, Chichester.

Rigobon, R. (2003). Identification through heteroskedasticity, *Review of Economics and Statistics* **85**: 777–792.

Rigobon, R. and Sack, B. (2003). Measuring the reaction of monetary policy to the stock market, *Quarterly Journal of Economics* **118**: 639–669.

Sims, C. A. (1986). Are forecasting models usable for policy analysis?, *Quarterly Review, Federal Reserve Bank of Minneapolis* **10**: 2–16.

Sims, C. A. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy, *European Economic Review* **36**: 975–1000.

Sims, C. A. and Zha, T. (2006). Were there regime switches in U.S. monetary policy?, *American Economic Review* **96**: 54–81.

Strongin, S. (1995). The identification of monetary policy disturbances: Explaining the liquidity puzzle, *Journal of Monetary Economics* **35**: 463–498.