Identification of Monetary Policy Shocks: a Graphical Causal Approach

Alessio Moneta*

February 28, 2005

Abstract

This paper develops a structural VAR methodology based on graphical models to identify the monetary policy shocks and to measure their macroeconomic effects. The advantage of this procedure is to work with testable overidentifying models, whose restrictions are derived by the partial correlations among residuals plus some institutional knowledge. This permits to test some restrictions on the reserve market used in several approaches existing in the literature. The main findings are that neither VAR innovations to federal funds rate nor innovations to nonborrowed reserves are good indicators of monetary policy shocks.

JEL classification: C32, C49, E52.
Keywords: Monetary Policy, Structural VARs, Graphical Models, Identification, Causation.

Palavras-chave: política monetária, VAR estrutural, modelos gráficos, identificação, causalidade.

*Correspondence address: Laboratory of Economics and Management, Sant’Anna School of Advanced Studies, Piazza Martiri della Libertà 33, 56127 Pisa, Italy; Tel.: +39-050-883591; fax: +39-050-883344; email: amoneta@sssup.it. I am grateful to Marco Lippi and Peter Spirtes for helpful advice, with the usual disclaimers.
Alessio Moneta

Scuola Superiore Sant’Anna
Piazza Martiri della Libertà 33
56127 Pisa, Italy.

Tel.: +39-050-883591;
fax: +39-050-883344;
email: amoneta@sssup.it.
1 Introduction

A monetary policy shock is the portion of variation in central bank policy, that is not caused by the systematic responses to variations in the state of the economy. It is an exogenous shock, which may reflect innovations to the preferences of the members of the monetary authority (e.g. Federal Open Market Committee), measurement errors of the same members, and any other conceivable variation orthogonal to macroeconomic innovations. Vector Autoregressive (VAR) models have been extensively used to isolate and study the effects of a monetary policy shock. A VAR is given by

\[ Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \ldots + A_p Y_{t-p} + u_t, \quad t = 1, \ldots, T \]  

where \( Y_t \) is a \( k \times 1 \) vector of data at date \( t = 1, \ldots, T \); \( A_i \) are coefficients matrices of size \( k \times k \); and \( u_t \) is the one-step ahead prediction error with variance-covariance matrix \( \Sigma_u \). Equation (1) is a “reduced form” model: it merely summarizes the statistical properties of the data. To study the dynamic effects of a monetary policy innovation, one needs an “identified” model, namely a model that has an economic interpretation. The problem of identifying a VAR consists in decomposing the prediction error \( u_t \) into economically meaningful or “fundamental” innovations. Suppose that there is a vector \( \nu_t \) of fundamental innovations of size \( k \times 1 \), which are mutually independent. Therefore \( E[\nu_t \nu'_t] = D \) is a diagonal matrix. What is needed is to find a matrix \( \Gamma \) such that \( u_t = \Gamma \nu_t \). It turns out that:

\[ \Sigma_u = E[u_t u'_t] = \Gamma E[\nu_t \nu'_t] \Gamma' = \Gamma D \Gamma'. \]  

The problem is that the \( k(k + 1)/2 \) non-zero elements, which can be obtained from the estimate of \( \Sigma_u \), are not sufficient to specify \( \Gamma \) and the diagonal of \( D \). Therefore, one needs further restrictions to achieve identification. In the literature, there exist three methods to impose the necessary restrictions. The first one consists in decomposing \( \Sigma_u \) according to the Choleski factorization, so that \( \Sigma_u = PP' \), where \( P \) is lower-triangular, defining a diagonal matrix \( V \) with the same diagonal as \( P \) and choosing \( \Gamma = PV^{-1} \). This is equivalent to impose a recursive ordering of the variables, called a “Wold causal chain”, as in Sims (1980). The second method consists in deriving from theoretical and institutional knowledge some “structural” relationships between the
fundamental innovations $\nu_{t,i}$, $i = 1, \ldots, k$ and the one-step ahead prediction errors $u_{t,i}$, $i = 1, \ldots, k$, as in Bernanke (1986), Blanchard and Watson (1986) and Sims (1986). The third method consists in separating transitory from permanent components of the innovations, as in Blanchard and Quah (1989) and King et al. (1991).

Any of these methods deals with a high degree of arbitrariness (for a criticism see Faust and Leeper (1997)). Indeed, imposing a set of restrictions corresponds to ascribing a particular causal relation, which is often difficult to be justified, to the variables. To address this problem some authors try several identification schemes and derive stylized facts from them. Thus, Christiano et al. (1999) state that “there is considerable agreement about the qualitative effects of a monetary policy shock in the sense that inference is robust across a large subset of the identification schemes that have been considered in the literature. The nature of this agreement is as follows: after a contractionary monetary policy shock, short term interest rate rise, aggregate output, employment, profits and various monetary aggregate fall, the aggregate price responds very slowly, and various measures of wages fall, albeit by very modest amounts”.

These conclusions are often considered as “facts” and if a particular identification scheme does not accomplish them, it is sometimes seen as rejectable. Uhlig (1999) has persuasively argued that the way these restrictions are used may render the inference procedure circular (“we just get out what we have stuck in”) and proposes to identify the effects of a monetary policy shock on output by directly imposing sign restrictions on the dynamic responses of prices, nonborrowed reserves and interest rate to the same shock. Gordon and Leeper (1994), Bernanke and Mihov (1998) and Bagliano and Favero (1998) emphasize the importance of taking account of different monetary policy regimes.

Another point of controversy is the choice of the indicator of the stance of policy. Bernanke and Blinder (1992) propose VAR innovation to the federal funds rate as measure of the monetary policy shock, basing their argument on prior information about the Fed’s operating procedures. However, Christiano and Eichenbaum (1992) have made the case for using the quantity of nonborrowed reserves as indicator of monetary policy. On the other hand,
Strongin (1995) argues that the central bank has to accommodate in the short run total reserves demand, therefore monetary policy shocks are the shocks to nonborrowed reserves orthogonal to shocks to total reserves.

Since there is no consensus on which of the various measures is more appropriate to capture the stance of policy, many authors check the robustness of their results using a variety of indicators (see e.g. Christiano et al. (1999)). In this paper, in the spirit of Bernanke and Mihov (1998), the indicator of monetary policy stance is not assumed but rather is derived from an estimated model of the Fed’s operating procedure. We employ a structural VAR model, but before imposing the restrictions derived from institutional knowledge, we narrow the number of acceptable causal structures using graphical models. The idea is that causal relationships can be inferred from the set of vanishing partial correlations among the variables that constitute such (unobserved) relationships. Graphs form a rigorous language for the “calculus” and representation of causation. This method, which is an extension of the method used in a previous paper (Moneta (2003)) associates a graph to the (unobserved) causal structure of the model and addresses the problem of identification as a problem of causal discovery from vanishing partial correlations. In particular, we infer the class of acceptable causal structures among contemporaneous variables from all the correlations and partial correlations among the residuals.

The goal of this paper is twofold. On the one hand, we want to analyze what are the effects of a monetary policy shock, when in the structural VAR the identification assumptions are derived by means of graphical models, using US macroeconomic and policy variables. The results are consistent with the stylized facts of Christiano et al. (1997) and Christiano et al. (1999), at least for the entire sample 1965-1996. However, the subsample 1979-1996 yields dynamic responses to monetary policy shocks which are qualitatively different from those stylized facts. On the other hand, we want to investigate which shock embeds better the exogenous monetary shock, in the spirit of Bernanke and Mihov (1998). The results cast some doubts on the practice of using the shock to the federal funds or the shock to nonborrowed reserves as a measure of an exogenous monetary policy shock, while they bring some support on the conjecture of Strongin (1995), that a good measure of mon-
etary policy innovation is the shock to nonborrowed reserves orthogonal to the shock to total reserves.

The rest of the paper is structured as follows. The second section describes briefly the identification procedure of the structural VAR. The third section presents a standard model of the market for commercial bank reserves and central bank behaviors, which is a slight extension of the model used by Bernanke and Mihov (1998). The fourth section describes the data and the estimation procedure. The fifth section shows the application of the identification procedure. The sixth section summarizes the main empirical results. Conclusions are drawn in the seventh section.

2 Structural VAR and Graphical Models

Suppose that a structural model of the monetary transmission mechanism can be represented as:

$$ A \left( \begin{array}{c} X_t \\ M_t \end{array} \right) = \sum_{i=1}^{p} \Gamma_i \left( \begin{array}{c} X_{t-i} \\ M_{t-i} \end{array} \right) + B \left( \begin{array}{c} \nu^X_t \\ \nu^M_t \end{array} \right), $$(3)

where $X_t$ is a vector of macroeconomic (non-policy) variables (e.g., output and prices) and $M_t$ is a vector of variables (partially) controlled by the monetary policy-maker (e.g., interest rates and monetary aggregates). $Y_t \equiv (X'_t, M'_t)'$ is a vector of length $k$, whose components are indicated as $(y_{1t}, \ldots, y_{kt})'$, and $\nu_t \equiv (\nu^X_t, \nu^M_t)'$.

The Structural VAR methodology suggests first to estimate the reduced form:

$$ \left( \begin{array}{c} X_t \\ M_t \end{array} \right) = \sum_{i=1}^{p} C_i \left( \begin{array}{c} X_{t-i} \\ M_{t-i} \end{array} \right) + \left( \begin{array}{c} u^X_t \\ u^M_t \end{array} \right), $$(4)

where $C_i = A^{-1} \Gamma_i$ and

$$ u_t \equiv \left( \begin{array}{c} u^X_t \\ u^M_t \end{array} \right) = A^{-1} B \nu_t $$

Then, one has to face the problem of recovering the structural shocks from the estimated residuals. If we call $\Sigma_u$ the covariance matrix of the estimated
residuals, the identification problem in this context consists in inferring $A$ and $B$ from $\Sigma_u$. The model is overidentified if more than $k + k(k + 1)/2$ restrictions are imposed. In this case the validity of the restrictions can be tested via a likelihood ratio test statistic asymptotically distributed as a $\chi^2$ with a number of degrees of freedom equal to the number of overidentifying restrictions (see Sims (1980), p. 17 and Doan (2000), p. 287).

The idea here is to use graphical models to strongly narrow the number of acceptable contemporaneous causal structures. Then, one can further discriminate using institutional knowledge, jointly with $\chi^2$ tests on overidentifying restrictions. The advantage of this method with respect of the standard structural VAR approach is that eliminating the implausible causal structures significantly lowers the degree of arbitrariness. The method applied here is an extension of a search procedure that was developed in a previous paper (Moneta (2003)) and that we are going to describe briefly here, referring to the appendix for more detailed terminology.

Statistical models represented by graphs, in particular directed acyclic graphs (DAGs), have been proved to be useful to represent causal hypotheses and to encode independence and conditional independence constraints implied by those hypotheses (Pearl (2000), Spirtes et al. (2000), Lauritzen (2001), Lauritzen and Richardson (2002), see appendix for a definition of graph and directed acyclic graph). In this framework, algorithms have been developed to recover some features of the causal graph from (conditional) independence relations among the variables which constitute the unobserved causal structure. A set of algorithms starts from the assumption of direct causation, ruling out the possibility of feedbacks, loops and confounder (e.g. PC algorithm in Spirtes et al. (2000)). A more sophisticated version of it allows for latent variables (e.g. FCI algorithm in Spirtes et al. (2000)). An algorithm developed by Richardson and Spirtes (1999) deals with the problem of inferring features of the causal graph under the assumption that it may be cyclic (feedbacks and loops are allowed), but there are no latent common causes.

Here we apply a general algorithm, which is basically the first common part of the algorithms mentioned above. The algorithm has, as input, the covariance matrix among the VAR residuals and produces, as output, an
undirected graph, which represents a class of possible causal relationships among the contemporaneous variables of the VAR model. The algorithm starts connecting all the contemporaneous variables \((y_{1t}, \ldots, y_{kt})^t\) in a complete graph and progressively eliminates most of the edges among variables which are not associated neither by a causal link, nor by a feedback link, nor by a latent variable, which we interpret as a common shock (for a definition of edges, undirected and complete graph, see appendix)\(^1\).

The procedure is based, first, on the fact that in a VAR model partial correlations among residuals are equivalent to partial correlations among contemporaneous variables, conditioned on all the past variables (see Proposition 1 in appendix).

Second, there are two fundamental assumptions relating causes and probability. Given a graph \(G\) and any three vertices \(A, B, C\) belonging to \(G\): (i) **Causal Independence Assumption**: if \(A\) does not cause \(B\), and \(B\) does not cause \(A\), and there is no third variable that causes both \(A\) and \(B\), then \(A\) and \(B\) are independent; (ii) **Causal Faithfulness Assumption**: if \(\text{corr}(A, B|C)\) is zero then \(A\) and \(B\) are \(d\)-separated by \(C\) on the graph \(G\). For the definition of \(d\)-separation see appendix.

Assuming normality of the error terms, the search algorithm described in appendix permits to infer an undirected graph, which represents a pattern of directed graphs (feedbacks and loops are allowed), from Wald tests on vanishing partial correlations among the residuals.

### 3 A Model of the Reserve Market

The undirected graph resulting from the search algorithm permits to narrow considerably the class of causal structure, but seldom this is enough for a

---

\(^1\) There are some (recognizable) cases in which an edge in the output of the algorithm does not correspond to the presence of an edge in the causal graph representing the data generating process (for details see Moneta (2004), p. 43). So, the output graph may contain more edges than the unobserved “true” graph. In general, the algorithm gives a graph that represents a class of causal structures, not a unique causal structure. It is possible to show that if there is a “true” causal structure which has generated the data, such causal structure is included in the class of causal structures represented by the output graph of the algorithm (see Moneta (2004) for details).
reliable identification. Background knowledge about the way the central bank sets the monetary policy is very useful at this stage. More detailed institutional assumptions, associated with causal hypotheses, can then be tested.

Following Christiano et al. (1999), a general model of the monetary authority behavior can be written as:

$$S_t = f(\Omega_t) + \nu_t^s,$$

where $S_t$ is the instrument of the monetary authority, say the federal funds rate or some monetary aggregate, $f$ is a linear function that represents the central bank’s feedback rule, $\Omega_t$ is the monetary authority’s information set and $\nu_t^s$ is a monetary policy shock.

Bernanke and Mihov (1998) model the banks’ total demand for reserves as:

$$TR_t = f_{TR}(\Omega_t^X) - \alpha FFR_t + \nu_t^d,$$

where $\Omega_t^X$ is the information set that comprehends only current and past macroeconomic variables. According to (7), the demand for total reserves $TR_t$ depends on $\Omega_t^X$ and is affected negatively by the federal funds rate ($FFR_t$). The demand for borrowed reserves is:

$$BR_t = f_{BR}(\Omega_t^X) + \beta(FFR_t - DISC_t) - \gamma NBR_t + \nu_t^b,$$

where $BR_t$ is the portion of reserves that banks choose to borrow at the discount window. According to (8), $BR_t$ is affected positively by the federal funds rate - discount rate differential and negatively by the nonborrowed reserves ($NBR_t$). Bernanke and Mihov (1998) assume that innovation to the discount rate is zero, which means that fluctuations in the discount rate are completely anticipated, so that $DISC_t$ does not enter in (8).

As far as the parameter $\gamma$ is concerned, Christiano et al. (1999) give two reasons for including $NBR_t$ in equation (8). The first one is that, if we would be willing to include expected valued in the equation describing demand for $BR_t$, we should include expected values of $FFR_t$ among the variables affecting $BR_t$, because of the existence of nonprice costs of borrowing at the Federal Reserve discount window. (These costs rise for banks that use too
much or too frequently the discount window). Indeed, for example, when \( E_t(FFR)_{t+1} \) is high, banks want \( BR_t \) to be low so that they can take full advantage of the high expected federal funds rate in next period without having to suffer large nonprice penalties at the discount window. Since \( NBR_t \) helps forecast future values of \( FFR_t \), it should enter on equation (8).

The second reason is that a bank that possesses a large amount of \( NBR_t \) and is using the discount window is simply trying to profit from the spread between the federal funds rate and discount rate. In that case, the bank would suffer a higher nonprice marginal cost of borrowing. So, \( NBR_t \) should enter the equation describing demand for \( BR_t \).

However, Bernanke and Mihov (1998) assume \( \gamma = 0 \), presumably in order to achieve overidentified and testable identification scheme.\(^2\)

The following equation describes the behavior of the Federal Reserve:

\[
NBR_t = f_{NBR}(\Omega^X_t) + \phi^d \nu^d_t + \phi^b \nu^b_t + \nu^*_t. \quad (9)
\]

According to (9), the Fed, by means of open-market operations, can change the amount of nonborrowed reserves supplied to banks in response to shocks to the total demand for reserves and to the demand for borrowed reserves. The coefficients \( \phi^d \) and \( \phi^b \) denote the strength of the responses and \( \nu^*_t \) represents the monetary policy shock.

Since \( TR_t = NBR_t + BR_t \), we can derive from (8) (omitting the discount rate) the following equation:

\[
FFR_t = -\frac{1}{\beta} f_{BR}(\Omega^X_t) + \frac{1}{\beta} TR_t + \gamma \frac{1}{\beta} NBR_t - \frac{1}{\beta} \nu^b_t. \quad (10)
\]

From (7), (9) and (10), we can derive restrictions on the contemporaneous variables, which correspond to zero coefficients on the matrix \( A \) of equations (3) and (5), as we will see in section 5.

\(^2\)This is an important limitation, as underlined by Christiano et al. (1999). Indeed, in the Bernanke and Mihov (1998) approach, one can always interpret an alleged rejection of an identification scheme as evidence against the maintained hypothesis \( \gamma = 0 \) and save the identification scheme. An advantage of our approach is that, thanks to the pre-selection of graphical models, we do not need to assume \( \gamma = 0 \) to reach overidentification and we can assess whether \( \gamma = 0 \) is in fact rejected or not by the data.
4 Data and Estimation

The data set used is the same as that of Bernanke and Mihov (1998) and consists of 6 series of monthly US data (1965:1-1996:12)\textsuperscript{3}. We refer to the non-policy macroeconomic variables as $GDP_t$: real GDP (log); $PGDP_t$: GDP deflator (log); $PSCCOM_t$: Dow-Jones index of spot commodity prices (log). The policy variables are: $TR_t$: total bank reserves (normalized by 36-month moving average of total reserve); $NBR_t$: nonborrowed reserves plus extended credit (idem); $FFR_t$: federal funds rate.

We estimate the model both in the vector error correction model parameterization and in levels (equation-by-equation OLS).\textsuperscript{4} Since the results of the two estimations are very close, we report the results of the level estimation, in order to have a clear comparison with the results of Christiano et al. (1999) and Bernanke and Mihov (1998), who estimate the model in level. The number of lags used to estimate the VAR is 13 in the full sample (the same used by Bernanke and Mihov (1998)). The covariance matrix among the residuals obtained by OLS estimation is the following:

$$\Sigma_u = \begin{bmatrix}
322 & 4 & -31 & -4 & 3473 & 42 \\
4 & 26 & 16 & 15 & 650 & 26 \\
-31 & 16 & 1682 & 660 & -19341 & -388 \\
-4 & 15 & 660 & 802 & 7652 & -46 \\
3473 & 650 & -19341 & 7652 & 2236763 & 10530 \\
42 & 26 & -388 & -46 & 10530 & 2670 \\
\end{bmatrix} \times 10^{-7}.$$

5 Identification of the Structural Shocks

The search algorithm mentioned in section 2 and reported in appendix is employed to derive the class of admissible causal structures among the contemporaneous variables of the structural model (represented in equation (3)). The input of the search algorithm is the covariance matrix among residuals

\textsuperscript{3}The data set was downloaded from Ben Bernanke’s home page.

\textsuperscript{4}The estimation via VECM parameterization does not imply any difference in the way the identification problem is faced, since, once the covariance matrix among the residuals is estimated, the model is reconverted in levels.
and the output is a pattern of directed graphs, which is represented by an undirected graph. Figure 1 displays the output of the algorithm for the full sample.

[ Include Figure 1 here ]

The graph in Figure 1 has to be read according to the following criterion. An undirected edge between any two vertices $A$ and $B$ of the graph corresponds to one or more of the following alternatives: (i) there is a direct causal relationship from $A$ to $B$; (ii) there is a direct causal relationship from $B$ to $A$; (iii) there is a common shock affecting both $A$ and $B$. It should be emphasized that a lack of edges between any two variables does not mean that there is no correlation at all (in fact there is usually correlation through lagged variables affecting both), but just that all of the three options listed above are excluded. Thus, since in the graph of Figure 1 there is no edge starting from PGDP, prices (measured by the GDP deflator) do not affect instantaneously any other variable, prices are not affected instantaneously by any other variable, and that there is no common shock affecting contemporaneously PGDP and any other variable.

From the mere analysis of correlations and partial correlations it is very difficult to infer which structure, among (i), (ii) and (iii), holds. One needs to incorporate background knowledge, which, if it implies overidentifying constraints, can be tested.

One set of assumptions can be derived by prior knowledge about the nature of interaction between policy variables and macroeconomic non-policy variables. A common interpretation of a contemporaneous association between macroeconomic variables and policy variables is that the monetary authority monitors continuously prices and output, so that there are causal effects from non-policy macroeconomic variables to policy variables within

---

5 The significance level employed in the vanishing partial correlation tests is 0.05. We remove time subscripts. A variable without time subscripts should be interpreted as a variable at time $t$.

6 Actually, there is also a fourth possibility. If there are feedbacks and common shocks in the data generating graph, the output of the algorithm may contain edges that do not correspond to any of the associations mentioned. Such configurations are however recognizable, as mentioned in footnote 1 (see Moneta (2004), p. 43 for more details).
the period (one month). A further possible identifying assumption (used e.g. in Bernanke and Mihov (1998)) is that there is no feedback from policy variables to the economy within the period. This corresponds to ruling out causal relationships from any of the policy variables $NBR$, $BR$ and $FFR$ to any of the macroeconomic variables $GDP$, $PGDP$ and $PSCCOM$. In the following we will consider and assess the hypothesis of orthogonality of the policy shock to the macroeconomic variables, which we call recursiveness assumption, against the alternative hypothesis of correlation between policy shock and these variables (non-recursiveness assumption). Under the recursiveness assumption, the undirected edges of Figure 1 between $GDP$ and $FFR$ and between $NBR$ and $PSCCOM$ are interpreted as directed edges from $GDP$ and $PSCCOM$ to $FFR$ and $NBR$ respectively. Under the non-recursiveness assumption the same edges are interpreted as bi-directed.

The length of “the period” is crucial here. For example, the assumption of no feedback from policy variables to the economy is more difficult to defend at the quarterly frequency and easier to defend at the weekly frequency. The opposite occurs with the assumption of causal effects from the economy to policy variables, which is more reliable at low than at high frequencies. Notice also that we do not use, consistently with the studies quoted, real-time data, and that measurement errors, which are common in the first data releases, are embedded in the exogenous monetary shock. How the identification results would change with the use of real-time data is an interesting open research question.

The scheme of identification associated with the recursiveness assumption should in general be distinguished from the recursive VAR identification scheme, which is derived by the Choleski factorization of the residuals covariance matrix and is associated with a Wold casual chain.

Indeed, it is possible to show that under the Faithfulness condition, the recursiveness assumption implies an absence of contemporaneous direct causes from non-policy to policy variables and an absence of a direct cause from the policy shock to non-policy variables (and an absence of any latent variable caused by the policy shock and causing non-policy variables). Moreover, under the Causal Independence condition, the non-recursiveness assumption implies that either policy variables cause non-policy variables within the period, or that the policy shock is a common shock affecting both non-policy and policy variables (or that there is a latent variable affected by policy variables and affecting non-policy variables). To put it in another way, the economic content of the recursiveness assumption is that non-policy variables do not respond within the period to realization of the policy shock, while the economic content of the non-recursiveness assumption is just the opposite.
Institutional knowledge can be used to impose identifying restrictions concerning the interactions among the policy variables (NBR, TR and FFR). From the considerations of section 3 about the model of the reserve market, it results that the relationships between VAR residuals and structural disturbances can be represented as follows, as far as the monetary policy block is concerned:

\begin{align*}
    u_{TR} &= -\alpha u_{FFR} + \nu^d, \\ 
    u_{FFR} &= \frac{1}{\beta} u_{TR} + \frac{\gamma - 1}{\beta} u_{NBR} - \frac{1}{\beta} \nu^b, \\ 
    u_{NBR} &= \phi^d \nu^d + \phi^b \nu^b + \nu^s.
\end{align*}

The system of equations (11)-(12)-(13) corresponds to a set of causal restrictions, as illustrated below. The restrictions on the relationships among macroeconomic variables (GDP, PGDP and PSCCOM) and on the relationships between macroeconomic variables and policy variables, that are derived by the graph output of the search algorithm, are numerous enough, so that the system can be identified. We also consider further restrictions on the policy block, which correspond to five alternative models, the same considered by Bernanke and Mihov (1998), with the difference that we allow $\gamma$ to be different from zero.

The first case we consider is the model of equations (11), (12), (13), without further restrictions, which we call Model 0. The graph connected to this model is represented in Figure 2. In this case the monetary policy shock $\nu^s$ is related with the VAR residuals in the following way:

\begin{equation}
    \nu^s = -(\phi^d + \phi^b) u_{TR} + (1 + \phi^b(1 - \gamma)) u_{NBR} - (\alpha \phi^d - \beta \phi^b) u_{FFR},
\end{equation}

which is obtained solving (11)-(12)-(13) for $\nu^s$.

The second case is Model 0 plus the restrictions $\alpha = 0$. The graph for this model, which we call model $\alpha = 0$, is represented in Figure 3. It corresponds to assuming that the demand for total reserves is inelastic in the short run. Strongin (1995) presents institutional arguments to support this assumption.

In the third case we impose to Model 0 the restrictions $\phi^d = \frac{1}{1 - \gamma}$ and $\phi^b = -\frac{1}{1 - \gamma}$. This corresponds to the assumption that the central bank uses
NBR to neutralize borrowing and demand shocks and targets the federal funds rate. Indeed the monetary policy shock turns out to be proportional to the innovation to the federal funds rate:

\[ \nu^s = \frac{1}{1 - \gamma} (\alpha + \beta) u_{FFR}. \]  

(15)

The graph related with this model, which we call model \textit{FFR}, is represented in Figure 4. Bernanke and Blinder (1992) presented arguments in support of the federal funds rate as a measure of policy instrument.

In the next case the following restrictions are imposed: \( \phi^d = 0 \) and \( \phi^b = 0 \). In this case the monetary policy shock coincides with the VAR innovation to the nonborrowed reserves: \( \nu^s = u_{NBR} \). The graph related with this model, which we call model \textit{NBR}, is represented in Figure 5. The argument that innovations to nonborrowed reserves primarily reflect shocks to monetary policy was defended by Christiano and Eichenbaum (1995) and Christiano et al. (1996).

The fifth case we consider is the model which imposes the restrictions \( \alpha = 0 \) and \( \phi^b = 0 \) on Model 0. The implied monetary policy shock is \( \nu^s = -\phi^d u_{TR} + u_{NBR} \). This corresponds to assuming that shocks to total reserves are purely demand shocks (\( \nu^d \)), which the central bank has to accommodate immediately (either through open-market operations or the discount window). Therefore monetary policy shocks (\( \nu^s \)) are the shocks to NBR orthogonal to \( \nu^d \). Moreover, this specification, defended by Strongin (1995), does not consider the possibility that the central bank reacts to borrowing shocks. Figure 6 represents the graph associated with this model, which is called Model \textit{NBR/TR}.

The last case we consider is the model which corresponds to assuming that the central bank targets borrowed reserves. The restrictions imposed are \( \phi^d = \frac{\beta}{\beta + \alpha} \) and \( \phi^b = \frac{\alpha}{\beta + \alpha} \). This implies that

\[ \nu^s = \frac{\alpha + \beta}{\beta + \alpha \gamma} (u_{TR} - u_{NBR}) = -\frac{\alpha + \beta}{\beta + \alpha \gamma} u_{BR}. \]  

(16)

Figure 7 represents the graph associated with this model, which is called Model \textit{BR}.

[ Include Figures 3-7 ]
The set of restrictions implied by each model corresponds to a set of restrictions on the matrices $A$ and $B$ of equation (3). Equation (5) can be written as:

$$ Au_t = Bv_t. $$

(17)

Imposing the restrictions of Model 0 (without recursiveness), we can write (17) as:

$$ \begin{bmatrix} 1 & 0 & 0 & 0 & a_{45} & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & a_{36} \\ 0 & 0 & 0 & 1 & a_{45} & 0 \\ a_{53} & 0 & a_{53} & a_{54} & 1 & 0 \\ 0 & 0 & a_{63} & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_{GDP} \\ u_{FGDP} \\ u_{NBR} \\ u_{TR} \\ u_{FFR} \\ u_{PSCCOM} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & b_{34} & b_{35} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{54} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \nu_{GDP} \\ \nu_{FGDP} \\ \nu^d \\ \nu^{d^2} \\ \nu^b \\ \nu_{PSCCOM} \end{bmatrix} $$

(18)

The restrictions on the elements of $A$ and $B$ for each model of the bank reserves are reported in Table 1. The relations among the parameters of equation (18) and $\alpha$, $\beta$, $\gamma$, $\phi^d$ and $\phi^b$ are the following: $\phi^d = b_{34}$, $\phi^b = b_{35}$, $\alpha = a_{45}$, $\beta = -\frac{1}{a_{54}}$, $\gamma = 1 + \frac{a_{53}}{a_{54}}$. The key to interpret equation (18) is looking at equation (3), where matrix $A$ represents the structural relations among contemporaneous variables.

Each model is estimated by maximum likelihood estimation, using a RATS procedure based on the hill-climbing BFGS method (see Doan (2000) for details).\(^{10}\) Estimates of the parameters of model 0 are displayed in Table 1.\(^{11}\) Notice that the theoretical restrictions that we impose do not include interactions with non-policy variables. The only restrictions about such interactions are derived from the graph-search procedure (with the exception of the general restriction embedded in the recursiveness assumption). This is because, first, we have more reliable and precise background knowledge about policy variables, than about the relations between policy and non-policy variables. Indeed, we may call the theoretical knowledge about policy variables “institutional knowledge”, because is more based on assumptions about the procedures followed by the banking system, than on economic theory. Second, each set of restrictions on the policy variable comprises a precise interpretation of the exogenous monetary policy shock. Since we can easily test each set of restrictions, we can get information as to which measure represents better the exogenous monetary shock.

The results of the restrictions of Model $FFR$ and Model $BR$ with the non-recursiveness assumption should be taken with caution, because they do not take into account policy parameters that enter in the equation of the monetary policy shock via non-policy variables. This does not change, however, the substance of the results (see next section). I
2 and 3. Estimates of the parameters $\alpha$, $\beta$, $\gamma$, $\phi^d$ and $\phi^b$ for each model are reported in Table 4. Each model is overidentified and produces a likelihood ratio test for the restrictions. In the same table $p$ values for such tests are also reported, that indicate whether a model has been rejected or not.

We do not have space here to give specific comments on the estimates of $\gamma$, $\alpha$, $\beta$, $\phi^d$ and $\phi^b$ (for a detailed analysis see Moneta (2004)). The substance of these results is reported in the next section.

We have also investigated the robustness of the results across subsamples. We do not have space here to report the results (Moneta (2004) contains a wider illustration of the results and the method to deal with small samples), whose substance is reported in the next section.

[ Include Tables 1-4 ]

6 Main Results

Our analysis permits to give some answers to the following questions.

What happens after a monetary policy shock?

If we consider the full sample 1965-1996, the qualitative responses of output, prices and interest rate are consistent with the stylized facts presented by Christiano et al. (1999) and with the results of Bernanke and Mihov (1998). After an expansionary monetary policy shock, output has an uncertain behavior in the first 2-3 months, then it increases rapidly, reaching its peak around the 15th month. The response of output in the long run is almost null, that means that money is close to being neutral in the long run.\footnote{This does not mean that the equation, in which GDP is dependent variable, is stationary (indeed it contains a unit root, according to the standard tests).} Price level responds slowly in the first year, after that increases. Short term interest rate falls immediately (showing the so called “liquidity effect”), but after 10-12 months returns to its previous value. Impulse response functions of GDP, GDP deflator and federal funds rate to a monetary policy shock are displayed in Figures 8-9 for the full sample. The impulse response functions are calculated for those models which have passed the likelihood ratio test.
(whose results are displayed in Table 4). The results about the effects of a monetary policy shock are quite robust across different assumptions about the central banks operating procedures and are approximately repeated in the sub-sample 1965:1-1979:9. However, in the sub-sample 1979:10-1996:12 we obtain slightly different results. Output still rises after a monetary policy shock, but much more moderately. Price levels responds positively, especially at the beginning, but very slowly. The impulse response functions for the two sub-samples are displayed in the Figures 10-11.

Which indicator most accurately captures the monetary policy shock?

Generally speaking, neither VAR innovation to the federal funds rate, nor VAR innovation to nonborrowed reserves turns out to be good indicator of monetary policy shocks. This is in the spirit of some results of Strongin (1995), Thornton (2001) and Sarno et al. (2002). Bernanke and Blinder (1992) employ innovations to interest rate as indicators of monetary policy shocks and obtain results consistent with the stylized facts (output and money rise in response to a positive monetary policy shock). Table 4, however, shows that the different specifications of \( FFR \) model always fail the likelihood ratio test (with one exception). Indeed there are some problems in measuring monetary policy with federal funds rate. First, as argued by Strongin (1995), “without any demonstrated empirical linkage between Federal Reserve actions and interest rate movements, it is unclear how innovations in interest rates can be reasonably be attributed to monetary policy.” Second, there could be non-policy omitted variables which explain movements in interest rates. Third, Sarno et al. (2002) argue that the practice of identifying monetary policy shocks as shocks to federal funds rate should be taken with caution, because of the “information equivalence hypothesis” (all interest rates contain the same information about monetary policy and the other aggregate shocks that determine the state of the economy).

Christiano and Eichenbaum (1992) suggest that innovation in nonborrowed reserves is the correct measure of monetary policy. Analogously to what happens with the federal funds rate, Table 4 shows that the different specifications of \( NBR \) model are always rejected by the data. This result corroborates the argument of Strongin (1995) that a significant proportion of the movements in nonborrowed reserves is due to the Fed’s accommodation
of innovations in the demand for reserves, rather than policy-induced supply innovation. Indeed a good indicator of the monetary policy shock seems to be the measure suggested by Strongin (1995), which is the part of innovation to nonborrowed reserves orthogonal to innovation to total reserves. Table 4 shows that the $NBR/TR$ model is rejected by the data only in the case of $\gamma = 0$.\footnote{Furthermore, the model $NBR/TR$ is the only model which is never rejected in the sub-samples. Model $BR$ is also not rejected in the full sample, but the estimated of $\gamma$ obtained are significantly negative (also in the sub-samples). This fact casts doubt on the reliability of this model. These results can be seen in detail in Moneta (2004).}

Does the recursiveness assumption hold?

The recursiveness assumption is about the orthogonality of the policy shock to the macroeconomic variables. It implies that policy variables do not influence macroeconomic non-policy variables within the period and that the monetary policy shock is not affecting simultaneously the two sets of variables (ruling out latent variables affected by the policy shock and affecting non-policy variables). We do not obtain strong results about this issue, even though the empirical evidence does not reject the hypothesis of non-recursiveness (see Table 4). The assumption of recursiveness, however, does not bring much difference for the only scope of measuring the effects of monetary policy shocks.

[ Include figures 8-11. ]

7 Conclusions

This paper proposed a method to identify the exogenous monetary policy disturbances in a VAR model. Since the crucial issue to identify a VAR is to differentiate between correlation and causation, graphical models permitted to impose overidentifying restrictions on the contemporaneous causal structure, in particular on the relationships among macroeconomic variables and between macroeconomic variables and policy variables. These restrictions have the advantage of being derived from the statistical properties of the data, without using auxiliary assumptions. Once we have narrowed the set of possible contemporaneous causal relationships among the variables which
constitute the VAR, we have imposed restrictions derived from institutional and theoretical knowledge. Since the number of possible contemporaneous causal relationships is a finite (and relatively narrow) number, it was possible to check the robustness of our results under alternative schemes of the Fed’s operating procedure and to appraise the alternative measures of monetary policy shocks used in the literature. The empirical results cast doubt on the soundness of those researches which assume that VAR innovations to federal reserve rate or nonborrowed reserves are good indicators of exogenous monetary policy shocks.

Appendix: Graphical Models Terminology

**Graphs.** A graph is an ordered pair \( G = (V, E) \), where \( V \) is a nonempty set of vertices, and \( E \) is a subset of the set \( V \times V \) of ordered pair of vertices, called the edges of \( G \). If one or both of the ordered pairs \( (V_1, V_2) \), \( (V_2, V_1) \) belong to \( E \), \( V_1 \) and \( V_2 \) are said to be adjacent. If both ordered pairs \( (V_1, V_2) \) and \( (V_2, V_1) \) belong to \( E \), we say that we have an undirected edges between \( V_1 \) and \( V_2 \), and write \( V_1 \rightarrow V_2 \). We also say that \( V_1 \) and \( V_2 \) are neighbors. If all the edges of a graph are undirected, we say that it is an undirected graph. If \( (V_1, V_2) \) belongs to \( E \), but \( (V_2, V_1) \) does not belong to \( E \), we call the edge directed, and write \( V_1 \rightarrow V_2 \). We also say that \( V_1 \) is a parent of \( V_2 \) and that \( V_2 \) is a child of \( V_1 \). If all the edges of a graph are directed, we say that it is a directed graph. A path of length \( n \) from \( V_0 \) to \( V_n \) is a sequence \( \{V_0, \ldots, V_n\} \) of distinct vertices such that \( (V_{i-1}, V_i) \in E \) for all \( i = 1, \ldots, n \). A directed path is a path such that \( (V_{i-1}, V_i) \in E \), but \( (V_i, V_{i-1}) \notin E \) for all \( i = 1, \ldots, n \). A cycle is a directed path with the modification that the first and the last vertex are identical, so that \( V_0 = V_n \). A graph is complete if every pair of its vertices are adjacent. A directed acyclic graph (DAG) is a directed graph which contains no cycles. Given a directed graph, the set of the vertices \( V_i \) such that there is a directed path from \( V_i \) to \( V_j \) are the ancestors of \( V_j \) and the set of vertices \( V_i \) such that there is a directed path from \( V_j \) to \( V_i \) are the descendants of \( V_j \). The graph \( G_A = (A, E_A) \) is called a subgraph of \( G = (V, E) \) if \( A \subseteq V \) and \( E_A \subseteq E \cap (A \times A) \). Besides, if \( E_A = E \cap (A \times A) \), \( G_A \) is called the subgraph of \( G \) induced by the vertex set \( A \).
**D-separation.** In a directed graph $G$ a vertex $X$ is a *collider* on a path $\alpha$ if and only if there are two distinct edges on $\alpha$ both containing $X$ and both directed on $X$. In a directed graph $G$ a vertex $X$ is *active* on a path $\beta$ relative to a set of vertices $Z$ of $G$ if and only if: (i) $X$ is not a collider on $\beta$ and $X \notin Z$; or (ii) $X$ is a collider on $\beta$, and $X$ or a descendant of $X$ belongs to $Z$. A path $\beta$ is active relative to $Z$ if and only if every vertex on $\beta$ is active relative to $Z$. In a directed graph $G$ two vertices $X$ and $Y$ are *d-separated* by $Z$ if and only if there is no active path between $X$ and $Y$ relative to $Z$. $X$ and $Y$ are *d-connected* by $Z$ if and only if $X$ and $Y$ are not d-separated by $Z$.

**Proposition 1:** Let $u_{1t}, \ldots, u_{kt}$ be the residuals of $k$ OLS regressions of $y_{1t}, \ldots, y_{kt}$ on the same vector $J_{t-1} = (y_{1(t-1)}, \ldots, y_{k(t-1)}, \ldots, y_{1(t-p)}, \ldots, y_{k(t-p)})$. Let $u_{it}$ and $u_{jt}$ ($i \neq j$) be any two distinct elements of $\{u_{1t}, \ldots, u_{kt}\}$, $U_t$ any subset of $\{u_{1t}, \ldots, u_{kt}\} \setminus \{u_{it}, u_{jt}\}$ and $Y_t$ the corresponding subset of $\{y_{1t}, \ldots, y_{kt}\} \setminus \{y_{it}, y_{jt}\}$, so that $u_{gt}$ is in $U_t$ iff $y_{gt}$ is in $Y_t$, for $g = 1, \ldots, k$. Then it holds that:

$$\text{corr}(u_{it}, u_{jt}|U_t) = \text{corr}(y_{it}, y_{jt}|Y_t, J_{t-1}).$$

Proof in Moneta (2003).

From Proposition 1 and Faithfulness Condition it follows that if $\text{corr}(u_{ht}, u_{it}|u_{jt}, \ldots, u_{it}) = 0$, then $\text{corr}(y_{ht}, y_{it}|y_{jt}, \ldots, y_{it}, J_{t-1}) = 0$ (where $J_{t-1}$ is defined as in Proposition 1) and $y_{ht}$ and $y_{it}$ are d-separated by $y_{jt}, \ldots, y_{it}, J_{t-1}$ in the graph induced by $y_{1t}, \ldots, y_{kt}, J_{t-1}$. Then it follows quite intuitively (see for a rigorous proof Proposition 2 in Moneta (2003)) that $y_{ht}$ and $y_{it}$ are d-separated by $y_{jt}, \ldots, y_{it}$ in the sub-graph induced by $y_{1t}, \ldots, y_{kt}$. The search algorithm is displayed in Table 5.

[Include Table 5 here]

**References**


Table 1
Restrictions on A and B. Each of the six models considered implies zero restrictions on some elements of the matrices A and B of equation (17). Each model has four versions, depending on the recursiveness assumption and the assumption on \( \gamma \) (in one case \( \gamma \) is free, in another is zero).

<table>
<thead>
<tr>
<th></th>
<th>recursiveness</th>
<th>rec. + ( \gamma = 0 )</th>
<th>non-recursiveness</th>
<th>non-rec. + ( \gamma = 0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 0</td>
<td>( a_{15} = 0 )</td>
<td>( a_{15} = 0 )</td>
<td>( a_{15} = 0 )</td>
<td>( a_{53} = -a_{54} )</td>
</tr>
<tr>
<td></td>
<td>( a_{63} = 0 )</td>
<td>( a_{63} = 0 )</td>
<td>( a_{63} = 0 )</td>
<td>( a_{53} = -a_{54} )</td>
</tr>
<tr>
<td>( \alpha = 0 )</td>
<td>( a_{15} = 0 )</td>
<td>( a_{15} = 0 )</td>
<td>( a_{45} = 0 )</td>
<td>( a_{53} = -a_{54} )</td>
</tr>
<tr>
<td></td>
<td>( a_{63} = 0 )</td>
<td>( a_{63} = 0 )</td>
<td>( a_{63} = 0 )</td>
<td>( a_{45} = 0 )</td>
</tr>
<tr>
<td></td>
<td>( a_{45} = 0 )</td>
<td>( a_{45} = 0 )</td>
<td>( a_{45} = 0 )</td>
<td>( a_{45} = 0 )</td>
</tr>
<tr>
<td>FRR</td>
<td>( a_{15} = 0 )</td>
<td>( a_{15} = 0 )</td>
<td>( b_{34} = -a_{54}/a_{53} )</td>
<td>( b_{34} = -a_{54}/a_{53} )</td>
</tr>
<tr>
<td></td>
<td>( a_{63} = 0 )</td>
<td>( a_{63} = 0 )</td>
<td>( a_{63} = 0 )</td>
<td>( a_{63} = 0 )</td>
</tr>
<tr>
<td></td>
<td>( b_{34} = -\frac{a_{54}}{a_{53}} )</td>
<td>( b_{34} = 1 )</td>
<td>( b_{35} = a_{54}/a_{53} )</td>
<td>( b_{35} = \frac{a_{54}}{a_{53}} )</td>
</tr>
<tr>
<td></td>
<td>( b_{35} = \frac{a_{54}}{a_{53}} )</td>
<td>( b_{35} = -1 )</td>
<td>( b_{35} = \frac{a_{54}}{a_{53}} )</td>
<td>( b_{35} = \frac{a_{54}}{a_{53}} )</td>
</tr>
<tr>
<td>NBR</td>
<td>( a_{15} = 0; \ a_{63} = 0 )</td>
<td>( a_{15} = 0 )</td>
<td>( b_{34} = 0 )</td>
<td>( b_{34} = 0 )</td>
</tr>
<tr>
<td></td>
<td>( b_{34} = 0 )</td>
<td>( a_{63} = 0 )</td>
<td>( b_{35} = 0 )</td>
<td>( a_{35} = 0 )</td>
</tr>
<tr>
<td></td>
<td>( b_{35} = 0 )</td>
<td>( a_{63} = 0 )</td>
<td>( b_{35} = 0 )</td>
<td>( a_{35} = 0 )</td>
</tr>
<tr>
<td>NBR/TR</td>
<td>( a_{15} = 0; \ a_{63} = 0 )</td>
<td>( a_{15} = 0 )</td>
<td>( a_{45} = 0 )</td>
<td>( a_{53} = -a_{54} )</td>
</tr>
<tr>
<td></td>
<td>( a_{45} = 0 )</td>
<td>( a_{63} = 0 )</td>
<td>( a_{45} = 0 )</td>
<td>( a_{45} = 0 )</td>
</tr>
<tr>
<td></td>
<td>( b_{35} = 0 )</td>
<td>( a_{63} = 0 )</td>
<td>( b_{35} = 0 )</td>
<td>( a_{35} = 0 )</td>
</tr>
<tr>
<td>BR</td>
<td>( a_{15} = 0; \ a_{63} = 0 )</td>
<td>( a_{15} = 0 )</td>
<td>( b_{34} = \frac{1}{1-a_{45}(a_{53}+a_{54})} )</td>
<td>( b_{34} = \frac{1}{a_{45}(a_{53}+a_{54})} )</td>
</tr>
<tr>
<td></td>
<td>( b_{34} = \frac{1}{1-a_{45}(a_{53}+a_{54})} )</td>
<td>( a_{35} = -a_{54} )</td>
<td>( b_{35} = \frac{1}{a_{45}(a_{53}+a_{54})} )</td>
<td>( b_{35} = 1 )</td>
</tr>
<tr>
<td></td>
<td>( b_{35} = \frac{1}{a_{45}(a_{53}+a_{54})-1} )</td>
<td>( b_{35} = 1 )</td>
<td>( b_{35} = a_{45}a_{54} )</td>
<td>( b_{35} = -a_{45}a_{54} )</td>
</tr>
</tbody>
</table>

25
Table 2. Estimation Model 0 recursive (full sample):  
ML estimation by BFGS. Convergence in 84 Iterations.  
Log Likelihood Unrestricted 9620.7327.  
Chi-Squared(8) 13.3605. Significance Level 0.1000.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Coeff</th>
<th>Std Error</th>
<th>T-Stat</th>
<th>Signif</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_{36}$</td>
<td>0.1360</td>
<td>0.0386</td>
<td>3.5241</td>
<td>0.0004</td>
</tr>
<tr>
<td>$a_{45}$</td>
<td>0.0013</td>
<td>0.0056</td>
<td>0.2411</td>
<td>0.8094</td>
</tr>
<tr>
<td>$a_{51}$</td>
<td>-8.7989</td>
<td>3.6417</td>
<td>-2.4161</td>
<td>0.0156</td>
</tr>
<tr>
<td>$a_{53}$</td>
<td>30.6270</td>
<td>10.0566</td>
<td>3.0454</td>
<td>0.0023</td>
</tr>
<tr>
<td>$a_{54}$</td>
<td>-37.0140</td>
<td>13.2387</td>
<td>-2.7959</td>
<td>0.0051</td>
</tr>
<tr>
<td>$b_{34}$</td>
<td>0.7608</td>
<td>0.2393</td>
<td>3.1791</td>
<td>0.0014</td>
</tr>
<tr>
<td>$b_{35}$</td>
<td>-0.1810</td>
<td>0.2790</td>
<td>-0.6490</td>
<td>0.5163</td>
</tr>
</tbody>
</table>

Table 3. Estimation Model 0 non-recursive (full sample):  
ML estimation by BFGS. Convergence in 233 Iterations.  
Observations 371. Log Likelihood 9614.7285.  
Log Likelihood Unrestricted 9620.7327.  
Chi-Squared(6) 12.0084.  
Significance Level 0.0617.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Coeff</th>
<th>Std Error</th>
<th>T-Stat</th>
<th>Signif</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_{15}$</td>
<td>-0.0021</td>
<td>0.0007</td>
<td>-2.9976</td>
<td>0.0027</td>
</tr>
<tr>
<td>$a_{36}$</td>
<td>0.0000</td>
<td>0.0035</td>
<td>0.0017</td>
<td>0.9986</td>
</tr>
<tr>
<td>$a_{45}$</td>
<td>-0.0180</td>
<td>0.0019</td>
<td>-9.0292</td>
<td>0.0000</td>
</tr>
<tr>
<td>$a_{51}$</td>
<td>0.0007</td>
<td>0.4462</td>
<td>0.0017</td>
<td>0.9985</td>
</tr>
<tr>
<td>$a_{53}$</td>
<td>68.9252</td>
<td>2.1444</td>
<td>32.1408</td>
<td>0.0000</td>
</tr>
<tr>
<td>$a_{54}$</td>
<td>-55.4529</td>
<td>2.6953</td>
<td>-20.5736</td>
<td>0.0000</td>
</tr>
<tr>
<td>$a_{63}$</td>
<td>0.2305</td>
<td>0.0658</td>
<td>3.5036</td>
<td>0.0004</td>
</tr>
<tr>
<td>$b_{34}$</td>
<td>0.8046</td>
<td>0.0697</td>
<td>11.5274</td>
<td>0.0000</td>
</tr>
<tr>
<td>$b_{35}$</td>
<td>-0.8045</td>
<td>0.0504</td>
<td>-15.9625</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
Table 4
Parameter Estimates (full sample):

<table>
<thead>
<tr>
<th>Model</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \gamma )</th>
<th>( \phi^d )</th>
<th>( \phi^b )</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recursive assumption:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 0</td>
<td>0.0014</td>
<td>0.0270</td>
<td>0.1726</td>
<td>0.7609</td>
<td>-0.1811</td>
<td>0.1000</td>
</tr>
<tr>
<td>( \alpha = 0 )</td>
<td>0</td>
<td>0.0290</td>
<td>0.1216</td>
<td>0.8148</td>
<td>-0.1939</td>
<td>0.1444</td>
</tr>
<tr>
<td>( FFR )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0118</td>
</tr>
<tr>
<td>( NBR )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0253</td>
</tr>
<tr>
<td>( NBR/TR )</td>
<td>0</td>
<td>0.0358</td>
<td>0.2019</td>
<td>0.8148</td>
<td>0</td>
<td>0.1659</td>
</tr>
<tr>
<td>( BR )</td>
<td>-0.0051</td>
<td>0.0532</td>
<td>-0.2510</td>
<td>0.9767</td>
<td>-0.0929</td>
<td>0.1189</td>
</tr>
<tr>
<td>Model 0, ( \gamma = 0 )</td>
<td>-0.0022</td>
<td>0.0326</td>
<td>0</td>
<td>0.8861</td>
<td>-0.2220</td>
<td>0.1282</td>
</tr>
<tr>
<td>( \alpha = 0, \gamma = 0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0002</td>
</tr>
<tr>
<td>( FFR, \gamma = 0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0000</td>
</tr>
<tr>
<td>( NBR, \gamma = 0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0000</td>
</tr>
<tr>
<td>( NBR/TR, \gamma = 0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0450</td>
</tr>
<tr>
<td>( BR, \gamma = 0 )</td>
<td>-0.0035</td>
<td>0.0436</td>
<td>0</td>
<td>1.0000</td>
<td>-0.0794</td>
<td>0.1223</td>
</tr>
<tr>
<td><strong>Non-recursiveness:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 0</td>
<td>-0.0180</td>
<td>0.0180</td>
<td>-0.2429</td>
<td>0.8046</td>
<td>-0.8045</td>
<td>0.0617</td>
</tr>
<tr>
<td>( \alpha = 0 )</td>
<td>0</td>
<td>0.0270</td>
<td>0.0977</td>
<td>0.8165</td>
<td>-0.2716</td>
<td>0.0903</td>
</tr>
<tr>
<td>( FFR )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0428</td>
</tr>
<tr>
<td>( NBR )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0435</td>
</tr>
<tr>
<td>( NBR/TR )</td>
<td>0</td>
<td>0.0358</td>
<td>0.1999</td>
<td>0.8165</td>
<td>0</td>
<td>0.1070</td>
</tr>
<tr>
<td>( BR )</td>
<td>-0.0055</td>
<td>0.0552</td>
<td>-0.3175</td>
<td>0.9696</td>
<td>-0.0957</td>
<td>0.0952</td>
</tr>
<tr>
<td>Model 0, ( \gamma = 0 )</td>
<td>-0.0269</td>
<td>0.0112</td>
<td>0</td>
<td>0.5927</td>
<td>-0.9691</td>
<td>0.1013</td>
</tr>
<tr>
<td>( \alpha = 0, \gamma = 0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1109</td>
</tr>
<tr>
<td>( FFR, \gamma = 0 )</td>
<td>-0.0036</td>
<td>0.0119</td>
<td>0</td>
<td>1.0000</td>
<td>-1.0000</td>
<td>0.0868</td>
</tr>
<tr>
<td>( NBR, \gamma = 0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0000</td>
</tr>
<tr>
<td>( NBR/TR, \gamma = 0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0261</td>
</tr>
<tr>
<td>( BR, \gamma = 0 )</td>
<td>-0.0035</td>
<td>0.0433</td>
<td>0</td>
<td>1.0000</td>
<td>-0.0805</td>
<td>0.0890</td>
</tr>
</tbody>
</table>

Notes: The estimates are functions of the ML estimates of the coefficients of the matrices \( A \) and \( B \) of equation (17), obtained by the RATS procedure based on the BFGS method (see Doan (2000) for details). The last column gives p-values from likelihood ratio tests of overidentifying restrictions. If p-value > 0.05, the restrictions implied by the particular model cannot be rejected at the 5 percent level of significance. We do not report estimates of the models which have been rejected.
Table 5: Search algorithm.

A.)
From the estimated covariance matrix among the VAR residuals test all the possible partial correlations among the residuals (using the Wald test procedure described in Moneta (2003)).

B.)
Form the complete undirected graph \( C \) on the vertex set \( y_{1t}, \ldots, y_{kt} \). Let \( \text{Adjacencies}(C, y_{lt}) \) be the set of vertices adjacent to \( y_{lt} \) in \( C \) and let \( \text{Sepset} \ (y_{ht}, y_{lt}) \) be any set of vertices \( S \) so that \( y_{ht} \) and \( y_{lt} \) are d-separated by \( S \);

C.)
\[ n = 0 \]
repeat :
repeat :
select an ordered pair of variables \( y_{ht} \) and \( y_{lt} \) that are adjacent in \( C \) such that \( \text{Adjacencies}(C, y_{ht}) \setminus \{y_{lt}\} \) has cardinality greater than or equal to \( n \), and a subset \( S \) of \( \text{Adjacencies}(C, y_{ht}) \setminus \{y_{lt}\} \) of cardinality \( n \), and if \( y_{ht} \) and \( y_{lt} \) are d-separated by \( S \) in \( G_{Y_{t}} \), delete edge \( y_{ht} \rightarrow y_{lt} \) from \( C \);
until all ordered pairs of adjacent variables \( y_{ht} \) and \( y_{lt} \) such that \( \text{Adjacencies}(C, y_{ht}) \setminus \{y_{lt}\} \) has cardinality greater than or equal to \( n \) and all subsets \( S \) of \( \text{Adjacencies}(C, y_{ht}) \setminus \{y_{lt}\} \) of cardinality \( n \) have been tested for d-separation;
\[ n = n + 1; \]
until for each ordered pair of adjacent variables \( y_{ht}, y_{lt}, \text{Adjacencies}(C, y_{ht}) \setminus \{y_{lt}\} \) is of cardinality less than \( n \);

Note:
Adapted from common steps of PC-FCI-CCD algorithms of Spirtes et al. (2000) and Richardson and Spirtes (1999).
Figure 1
Contemporaneous structure for the full sample (time subscripts are removed for convenience).

Figure 2
Model 0. It is the model of equations (11), (12) and (13), without further restrictions.

Figure 3
Model \( \alpha = 0 \). The demand for total reserves is inelastic in the short run, so that there is no causal effect from \( FFR \) to \( TR \).

Figure 4
Model \( FFR \). The weights refer to the case \( \gamma = 0 \). The Fed fully offsets shocks to total reserves demand and borrowing demand and targets the federal fund rate.
Figure 5
Model NBR. Nonborrowed reserves respond only to policy shocks, so that borrowing and demand shocks do not affect NBR.

Figure 6
Model NBR/TR. Monetary policy shocks are shocks to NBR orthogonal to demand shocks.

Figure 7
Model BR. The weights refer to the case $\gamma = 0$. The Fed targets borrowed reserves $(TR - NBR)$. 
Figure 8: Responses of GDP, PGDP, and FFR to one-standard-deviation monetary shock for the sample 1965:1-1996:12 and with the recursiveness assumption.
Figure 9: Responses of GDP, PGDP, and FFR to one-standard-deviation monetary shock for the sample 1965:1-1996:12 and without the recursiveness assumption
Figure 10: Sample 1965:1-1979:9. The graphics on the first line refer to models identified under the recursiveness assumptions, the graphics on the second line under the non-recursiveness assumption.
Figure 11: Sample 1979:10-1996:12. The graphics on the first line refer to models identified under the recursiveness assumptions, the graphics on the second line under the non-recursiveness assumption.
Resumo

Esse artigo desenvolve uma metodologia VAR estrutural baseada em modelos gráficos para identificar os choques de política monetária e medir os seus efeitos macroeconómicos. A vantagem desse procedimento é trabalhar com modelos sobre-identificados testáveis, cujas restrições são derivadas das correlações parciais entre os resíduos adicionando-se algum conhecimento institucional. Isso permite testar algumas restrições sobre o mercado de reserva usadas em várias abordagens existentes na literatura. Os principais resultados são que nem as inovações VAR ligadas a federal funds rate nem as ligadas às reservas não-empréstáveis (nonborrowed reserves) são bons indicadores de choques de política monetária.

Résumé

Cet article élabore une méthodologie des VAR structurels basée sur les modèles de graphes pour identifier les chocs de la politique monétaire et mesurer leurs effets macroéconomiques. L’avantage de cette procédure est la possibilité d’utiliser des modèles suridentifiés dont les restrictions sont dérivées par des corrélations partielles des résidus, en plus des connaissances institutionnelles. Ceci permet de tester certaines restrictions relatives au marché des réserves qui ont été utilisées par de nombreuses approches dans la littérature. Les principaux résultats indiquent que ni les innovations VAR introduites sur le taux des fonds fédéraux ni celles introduites sur les réserves non empruntées (non-borrowed reserves) sont de bons indicateurs des chocs de la politique monétaire.