UNIVERSITY OF SURREY

Discussion Papers in Economics

IDENTIFICATION OF MONETARY POLICY IN SVAR MODELS: A DATA-ORIENTED PERSPECTIVE

By

Matteo Fragetta
(University of Salerno, Italy)
&
Giovanni Melina
(University of London and University of Surrey)

DP 08/11

Department of Economics
University of Surrey
Guildford
Surrey GU2 7XH, UK
Telephone +44 (0)1483 689380
Facsimile +44 (0)1483 689548
Web www.econ.surrey.ac.uk
ISSN: 1749-5075

Identification of Monetary Policy in SVAR Models: A Data-Oriented Perspective*

Matteo Fragetta[a] and Giovanni Melina[†b,c]

[a] *University of Salerno, Italy*

[b] *Birkbeck, University of London, UK*

[c] *University of Surrey, UK*

27th May 2011

Abstract

This paper applies graphical modelling theory to recover identifying restrictions for the analysis of monetary policy shocks in a VAR of the US economy. Results are in line with the view that only high-frequency data should be assumed to be in the information set of the monetary authority when the interest rate decision is taken.

Keywords: Monetary policy; SVAR; Graphical modelling.

JEL Codes: E43; E52.

*We are grateful to Sergio Destefanis, Emanuel Gasteiger, Marco Reale, Lucio Sarno, Granville Wilson for helpful comments. Moreover, we thank Marco Reale and Granville Wilson for kindly sending some of their Matlab codes, and the developers of JMulTi for providing their software as an open source. Giovanni Melina acknowledges financial support from ESRC project RES-062-23-2451. The usual disclaimer applies.

†Corresponding author. Tel: +44 (0) 1483 689924; Fax: +44 (0) 1483 689 548.
E-mail addresses: mfragetta@unisa.it (Matteo Fragetta), g.melina@surrey.ac.uk (Giovanni Melina).

1 Introduction

The literature employing vector-autoregressions (VAR) to identify and estimate the effects of monetary policy shocks tipically distinguish among three sets of variables: (i) the information set, i.e. the set of variables known to the monetary authorities when the policy decision is taken; (ii) the policy instrument; (iii) the set of variables the value of which is known only after the policy is set. Such a distinction often suggests a block-recursive structure exploitable in identifying the VAR. Most of the existing empirical papers can be classified into two broad groups, which differ in the content of the information set of the monetary authority.

The first group of papers, that can be thought of following a "*workhorse*" approach, include, among many others, Christiano and Eichenbaum (1992), Christiano et al. (1996) as well as the influential paper by Christiano et al. (2005). These studies hold that the central bank has at its disposal sources of information about the economy well beyond the published data. In fact, policymakers have access to monthly or even daily estimates of a series of indicators on economic activity and prices sufficient to provide them with a clear and prompt indication of the state of the economy. Consistently with this argument, the assumption made is that, among other variables, the monetary authority is capable to observe the contemporaneous (within quarter) values of output and domestic prices (GDP deflator) at the time of the monetary policy decision.

The second group of papers can be thought of adopting an "*alternative*" approach. This approach is adopted for instance by Sims and Zha (1998), the extension proposed by Kim and Roubini (2000) with monthly data and international variables, and the macroecometric model of the UK proposed by Garratt et al. (2003). These papers argue that only high-frequency data should be assumed to be in the information set of the central bank. For example Sims and Zha (1998) use quarterly data and find it more reasonable to assume that only contemporaneous money supply and commodity prices are known to the central bank when the interest rate is set, since such indeces are released at monthly and daily frequencies, respectively. On the contrary, proper measures of variables such as the real GDP and the GDP deflator are assumed to be known to policymakers only with a lag.[1]

Both approaches make use of reasonable and convincing arguments, hence in principle there is no clear-cut reason why one should be preferred to the other. This makes the task of imposing *a-priori* short-term identifying restrictions contetious and complex. In fact, especially in small-scale VARs, conditional also on the degree of correlation betweeen reduced-form residuals, results depend (at least quantitatively) on the various possible timing restrictions imposed.

This paper applies Graphical Modelling (GM) theory to a small-scale VAR of the US economy to establish whether the data are informative on which of the two approaches is preferable. In fact, GM is a data-oriented tool as it allows one to obtain short-term identifying restrictions directly from statistical properties of the data. Reale and Wilson (2001) and Wilson and Reale (2008) show how the theory can be used in a VAR, while Oxley et al. (2009) and Fragetta and Melina (2011) are examples of how the method can be applied to macroeconomic analysis.

Results are in line with the "*alternative*" approach. In other words, GM suggests that only high-frequency data are in the information set of the central bank when it sets the interest rate. When it comes to impulse-response analysis, however, the two approaches generate similar responses to an interest rate shock, featuring only minor quantitative differences.

The remainder of the paper is structured as follows. Section 2 describes the econometric methodology. Section 3 presents the data. Section 4 illustrates the results. Finally, Section 5 concludes.

2 Econometric methodology

This section presents the econometric strategy adopted in the analysis. Subsection 2.1 illustrates the basic tools of graphical modelling theory, while Subsection 2.2 shows how these tools can be applied in the identification of a SVAR.

[1] For an extended survey of the literature see Christiano et al. (1999).

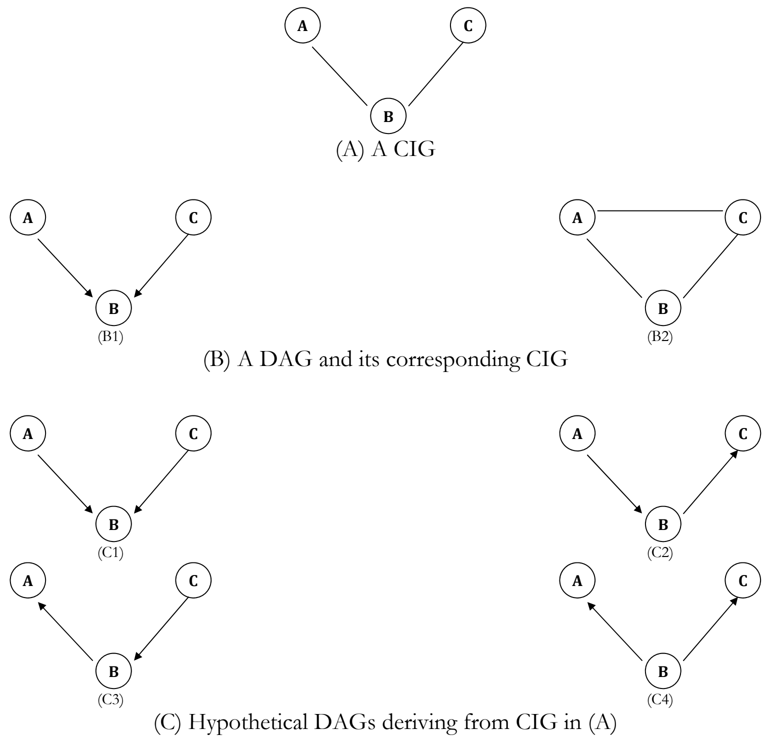


Figure 1: Conditional independence graphs and directed acyclic graphs

2.1 Graphical modelling

GM is a statistical approach aiming at uncovering statistical causality from partial correlations observed in the data, which can be interpreted as linear predictability in the context of least-square estimation. Primal contributions to the methodology are due to Dempster (1972) and Darroch et al. (1980).

The most informative object of the procedure is the *Direct Acyclic Graph* (DAG). Figure 1.C2 shows a typical and simple DAG, where nodes A, B and C represent random variables and the arrows connecting A and B, and B and C indicate the direction of a statistical causality. When *undirected edges* replace the arrows of a graph, a *Conditional Independence Graph* (CIG) is obtained. In a CIG, a link represents a significant partial correlation between any two random variables – conditioned on all the remaining variables of the model. Figure 1.A shows an example of a CIG. For instance, the edge connecting nodes A and B represents a significant partial correlation between A and B conditioned on C. A significant partial correlation implies conditional dependence if the variables are jointly distributed as a multivariate Gaussian distribution, hence the name CIG.

DAGs and CIGs imply a different definition of joint probability, however there is a correspondence between the two, represented by the so-called *moralization rule*, as firstly shown by Lauritzen and Spiegelhalter (1988). In fact, there is always a unique CIG deriving from a given DAG, obtained by transforming arrows into undirected edges and linking unlinked *parents* of a *common child* with a *moral edge*. In the DAG shown in Figure 1.B1, A and C are parents of B. In order to obtain the corresponding unique CIG arrows must be transformed into edges and a moral edge has to be added between parents A and C as in Figure 1.B2. Statistically, when both A and C determine B, a significant partial correlation due to moralization should be observed between A and C.

While there is a unique CIG deriving from a given DAG, the reverse is not true. What the econometrician can observe in the data is a CIG, where every edge can assume two possible directions. Therefore, for any given CIG, there are 2^n hypothetical DAGs, where n is the number of edges. Figure 1.C shows all the hypothetical DAGs corresponding to the CIG in Figure 1.A. The DAG in Figure 1.C1 is not compatible with the CIG, because the moralization rule requires a moral edge between A and C, which is not captured by the CIG.[2]

Any DAG, by definition, has to satisfy the principle of acyclicality. Therefore, the graph depicted in Figure 2 cannot be a DAG as it is clearly cyclic. The acyclicality in a DAG allows to completely determine the distribution of a set of variables and implies a recursive ordering of the variables themselves, where each element in turn depends on none, one or more elements. For example, in the DAG in Figure 1.C2, A depends on no other variables, B depends on A and C on B.

[2] In the process of obtaining plausible DAGs from an observed CIG, it may also be possible that some of the links captured by the CIG are due to moralization and hence must be eliminated in a corresponding DAG. Such *demoralization process*, in most cases, can be assessed by considering some quantitative rules. Let us suppose we observe a CIG such as the one in Figure 1.B2. If the true corresponding DAG were the one in Figure 1.B1, then the partial correlation between A and C, $\rho_{(A,C|B)}$, should be equal to $-\rho_{(A,B|C)} \times \rho_{(B,C|A)}$. In such a case, when tracing DAG 1.B1, the edge between A and C must be removed.

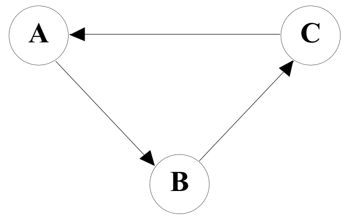


Figure 2: Directed cyclic graph

2.2 Identification of a SVAR with graphical modelling

GM theory can be applied to obtain identification of a structural VAR (SVAR), as shown by Reale and Wilson (2001) and Oxley et al. (2009) among others.

Any SVAR may be turned into a DAG where current and lagged variables are represented by nodes and causal dependence by arrows. After collecting the endogenous variables of interest in the k-dimensional vector X_t, the associated reduced-form, or canonical, VAR can be written as:

$$X_t = A(L)X_{t-1} + u_t, \tag{1}$$

where $A(L)$ is a polynomial in the lag operator L and u_t is a k-dimensional vector of reduced-form disturbances with $E[u_t] = 0$ and $E[u_t u_t'] = \Sigma_u$.

As reduced-form disturbances are correlated, in order to identify structural shocks, the reduced-form model has to be trasformed into a structural model. Pre-multiplying both sides of equation (1) by the $(k \times k)$ matrix A_0, yields the structural form:

$$A_0 X_t = A_0 A(L) X_{t-1} + Be_t. \tag{2}$$

The relationship between the structural disturbances e_t and the reduced-form disturbances u_t is described by the following:

$$A_0 u_t = Be_t, \tag{3}$$

where A_0 also describes the contemporaneous relations among the endogenous variables and B is a $(k \times k)$ matrix. In the structural model, disturbances are assumed to be uncorrelated with each other. In other words, the covariance matrix of the structural disturbances Σ_e is diagonal.

As it is, the model described by equation (2) is not identified because there may be possibly many matrices A and B that satisfy (2). Therefore, first matrix B can be restricted to be a $(k \times k)$ diagonal matrix. Second, in order to impose identifying restrictions on matrix A_0, graphical modeling theory can be applied to trace DAGs of the reduced-form residuals.

The acyclicality of DAGs implies a recursive ordering of the variables that makes A_0 a lower-triangular matrix. A_0 has generally zero elements also in its lower triangular part, hence, in general, the model is over-identified. The GM methodology has the distinctive feature that the variable ordering and any further restrictions come from statistical properties of the data.

First, in order to construct the CIG among contemporaneous variables conditioned on all the remaining contemporaneous and lagged variables, one can derive the sample partial correlation between the innovations, conditioned on the remaining innovations of the canonical VAR, calculated from the inverse $\hat{W}$ of the sample covariance matrix $\hat{V}$ of the whole set of innovations as in Greene (2003):

$$\hat{\rho}\left(u_{i,t}, u_{j,t} | \{u_{k,t}\}\right) = -\frac{\hat{W}_{ij}}{\sqrt{(\hat{W}_{ii}\hat{W}_{jj})}}, \tag{4}$$

where $\{u_{k,t}\}$ is the whole set of innovations excluding the two considered. Whenever a sample partial correlation is statistically significant a link is retained. Swanson and Granger (1997) have applied a similar strategy to sort out causal flows among contemporaneous variables, i.e. applying a residual orthogonalization of the innovations from a canonical VAR.

All possible DAGs (satisfying the moralization rule) which represent alternative competitive models are compared via likelihood based methods, such as the Akaike Information Criterion (AIC), the Hannan and Quinn Information Criterion (HIC) or the Schwarz Information Criterion (SIC), and choose the best-performing one.[3]

[3]In some cases, the distributional properties of the variables for different DAGs are likelihood equivalent, although the residual series are different. In such cases, it is possible to construct DAG models by considering only the lagged variables that play a significant role in explaining contemporaneous variables determined by the significant partial correlation. This can help, via comparison of information criteria, determine the best DAG for contemporaneous variables.

3 Data

The empirical analysis presented in the remainder of the paper employs quarterly US data over the period 1965:1-2007:4. The starting year coincides with that used by Christiano et al. (1999, 2005) while the end date falls in a pre-crisis quarter.

The model is a four-variable VAR including: (i) the log of real GDP, y_t; (ii) the effective federal funds rate (quaterly average), r_t; (iii) the log the GDP implicit price deflator, p_t, and (iv) the log of a commodity price index (producer price index), cp_t. The variables are representative of the real activity, monetary policy and price dynamics. Such a model specification represents a minimal setting similar to those adopted by Stock and Watson (2001) – for illustrative purposes – and by more recent contributions such as Primiceri (2005) and Koop et al. (2009). The addition of a commodity price proves helpful in ruling out the *price puzzle*.[4] The absence of monetary aggregates is due to a preference for parsimony coupled with the fading role of monetary aggregates in the conduct of monetary policy as empirically shown by Estrella and Mishkin (1997), among others, and theoretically explored by Woodford (2008).

A constant is included in the VAR and results are reported both for a VAR in levels, with and without a deterministic trend,[5] and for a VAR in which the logs of GDP, the GDP deflator and the commodity price index have been first differenced. The sampling properties of GM are valid regardless of the presence of unit roots in the data, as shown by Wilson and Reale (2008). In fact, we show below that the three model specifications give rise to the same CIGs and DAGs.

All series are extracted from the ALFRED database of the Federal Reserve Bank of St. Louis. The real GDP and the two price indices are seasonally adjusted by the source.

[4] The term *price puzzle* is due to Sims (1992). Christiano et al. (1999) show that omitting a commodity price index from the VAR specification delivers a rise in the price level that lasts several years after a contractionary monetary policy shock.

[5] We prefer to report results for both cases, as in the literature both options are explored. For instance, while Bernanke (1986) includes a deterministic trend in the level specification, Christiano et al. (2005) carry out the estimation including only the levels of the variables.

	u_t^r	u_t^y	u_t^p	u_t^{cp}
u_t^r	1.000			
u_t^y	0.141*	1.000		
u_t^p	0.113	-0.084	1.000	
u_t^{cp}	0.195**	-0.054	0.391***	1.000

(a) Model in first differences

	u_t^r	u_t^y	u_t^p	u_t^{cp}
u_t^r	1.000			
u_t^y	0.178**	1.000		
u_t^p	0.061	-0.120	1.000	
u_t^{cp}	0.214***	-0.026	0.435***	1.000

(b) Model in levels

	u_t^r	u_t^y	u_t^p	u_t^{cp}
u_t^r	1.000			
u_t^y	0.176**	1.000		
u_t^p	0.030	-0.126	1.000	
u_t^{cp}	0.214***	-0.016	0.437***	1.000

(c) Model in levels with deterministic trend

Note: *,** and *** denote significance at 0.10, 0.05 and 0.01 levels, respectively. The corresponding threshold values for the baseline model are 0.1270, 0.1504 and 0.1963, respectively.

Table 1: Estimated partial correlations of the series innovations

4 Results

DAGs of the VAR residuals are obtained by fitting the data to equation (1). The lag order is selected via the AIC.[6] Table 1 reports the estimated partial correlation matrices of the series innovations and their significance at 0.10, 0.05 and 0.01 levels. Both the matrix coming from the model in first differences and those coming from the model in levels (with and without trend) translate into the same CIG depicted in Figure 3. The three edges in the CIG cannot be moral, as moral edges link parents of a common child. The $2^3 = 8$ possible DAGs implied by the CIG are reported in Figure 4. The moralization rule implies that DAGs (A), (E), (G) and (H) can be discarded. The four remaining models are then compared via the likelihood-based information criteria mentioned in Section 2.

[6] The AIC typically selects a larger number of lags with respect to SIC and HIC, which we prefer based on the view that the consequences of overestimation of the order are less serious than underestimation (Kilian, 2001).

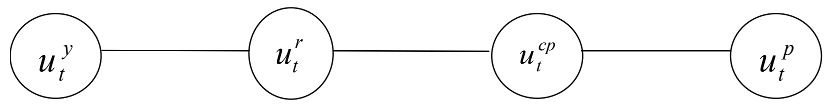


Figure 3: Sample CIG fitted to VAR residuals

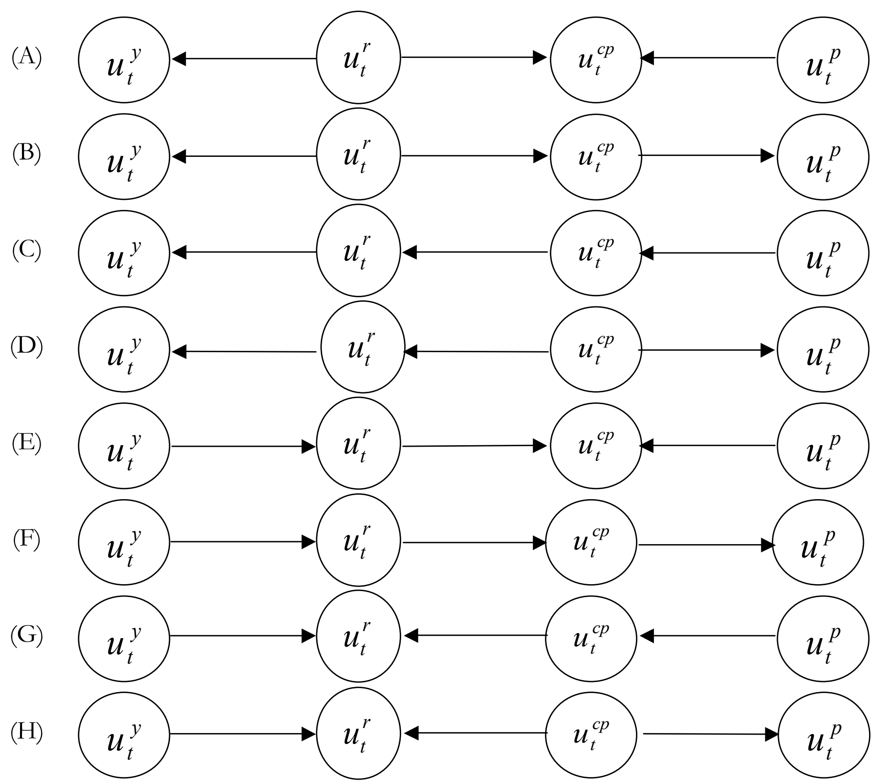


Figure 4: All possible DAGs deriving from the estimated CIG

Table 2 shows that the three information criteria for both model specifications are minimised by the model implied by DAG (C), which in turn implies that, within the same quarter, the Federal funds rate is not affected by shocks to the general price level and the real output,

Model	AIC	HIC	SIC
B	-418.56	-398.24	-368.48
C	-453.05	-433.17	-403.42
D	-358.26	-337.94	-308.19
F	-405.32	-385.00	-355.24

(a) Model in first differences

Model	AIC	HIC	SIC
B	-466.06	-445.74	-415.98
C	-521.79	-501.46	-471.71
D	-484.43	-464.11	-434.35
F	-471.50	-451.17	-421.42

(b) Model in levels

Model	AIC	HIC	SIC
B	-469.87	-444.47	-407.27
C	-525.34	-499.94	-462.74
D	-488.20	-462.79	-425.60
F	-463.66	-438.26	-401.06

(c) Model in levels with deterministic trend

Note: AIC = Akaike Information Criterion; HIC = Hannan-Quinn Information Criterion (HIC); SIC = Schwarz Information Criterion.

Table 2: Information criteria associated to feasible DAGs

	ϵ_t^r	ϵ_t^y	ϵ_t^p	ϵ_t^{cp}
ϵ_t^r	1.000			
ϵ_t^y	0.026	1.000		
ϵ_t^p	0.092	-0.112	1.000	
ϵ_t^{cp}	-0.043	-0.048	0.000	1.000

(a) Model in first differences

	ϵ_t^r	ϵ_t^y	ϵ_t^p	ϵ_t^{cp}
ϵ_t^r	1.000			
ϵ_t^y	0.022	1.000		
ϵ_t^p	0.036	-0.144	1.000	
ϵ_t^{cp}	-0.018	-0.020	0.000	1.000

(b) Model in levels

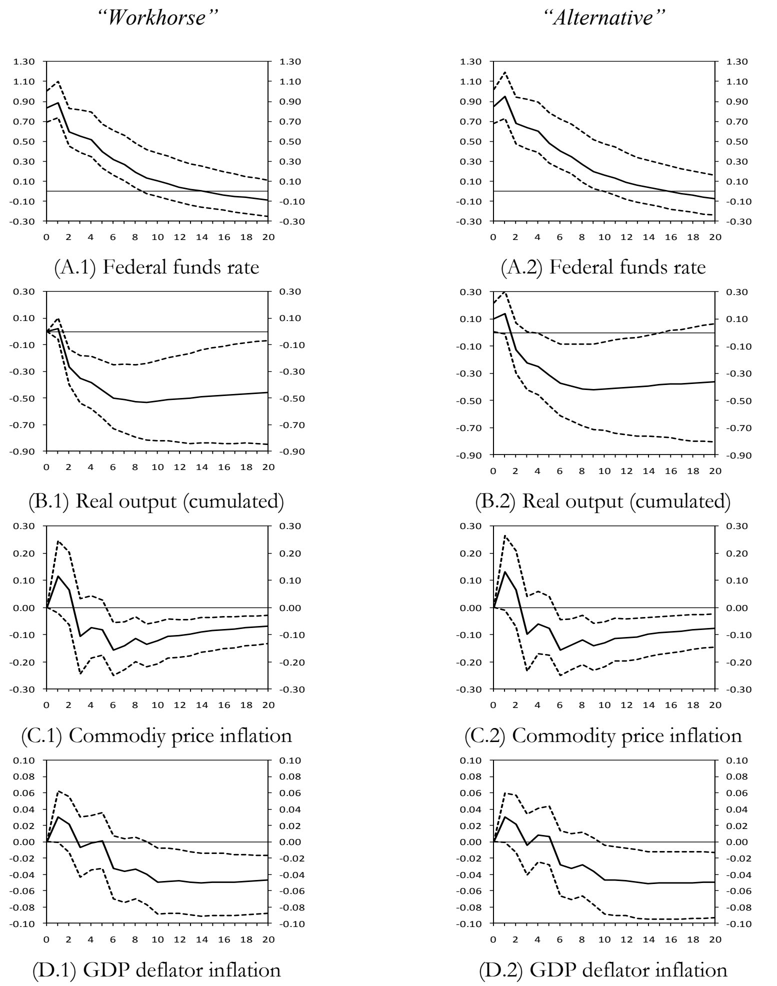
	ϵ_t^r	ϵ_t^y	ϵ_t^p	ϵ_t^{cp}
ϵ_t^r	1.000			
ϵ_t^y	0.020	1.000		
ϵ_t^p	0.035	-0.146	1.000	
ϵ_t^{cp}	-0.018	-0.010	0.000	1.000

Note: The two-standard-error band for a sample size of 204 is $\pm$ 0.1538

Table 3: Correlations between residuals of the DAGs fitted to the VAR estimated innovations

while it is affected by shocks to the commodity price.

In sum, GM selects only data available at high frequencies for the information set of the central bank, providing support for the "*alternative*" approach. A diagnostic check on the cross-correlations matrix of the resulting residuals reported in Table 3 unveils that all cross-



Note: Dashed lines represent 90% confidence intervals computed according to Hall (1992) algorithm with 2000 bootstrap replications. Responses are shown for a 20-quarter horizon.

Figure 5: Impulse responses to a Federal funds rate shock: "Workhorse" vs. "Alternative" (GM-consistent) identification

correlations lie within two standard errors from zero. In addition, DAG (C) implies three overidentying restrictions which are not rejected at any conventional significance level.

Figure 5 reports the impulse responses to a positive Federal funds rate shock obtained by adopting both the "*workhorse*" and the "*alternative*" identification approaches, the latter being consistent with GM. The two approaches generate impulse responses with small quantitative differences.

5 Conclusion

The empirical approaches aiming at identifying monetary policy shocks can be classified into two groups: the "*workhorse*" approach, which assumes that the central bank has sufficient information to accurately infer what contemporaneous real output and GDP deflators are when it takes the monetary policy decision; and the "*alternative*" approach, which assumes that only variables observed with high frequency, such as commodity prices, are in the information set of the central bank at the time of policy setting. This paper makes use of GM theory to identify a small-scale VAR of the US economy and finds that the application of such a data-based tool give rise to identifying restrictions consistent with the "*alternative*" approach. When impulse-response analysis is concerned, however, the "*workhorse*" approach and the model identified by imposing restrictions suggested by GM – coinciding with the "*alternative*" approach – generate responses to a Federal funds rate shock featuring small quantitative differences.

References

Bernanke, B. S. (1986). Alternative explanations of the money-income correlation. *Carnegie-Rochester Conference Series on Public Policy*, 25(1):49–99.

Christiano, L. J. and Eichenbaum, M. (1992). Liquidity effects, the monetary transmission mechanism, and monetary policy. *Economic Perspectives*, (Nov):2–14.

Christiano, L. J., Eichenbaum, M., and Evans, C. (1996). The effects of monetary policy shocks: Evidence from the flow of funds. *The Review of Economics and Statistics*, 78(1):16–34.

Christiano, L. J., Eichenbaum, M., and Evans, C. (1999). Monetary policy shocks: What have we learned and to what end? In Woodford, M. and Taylor, J. B., editors, *Handbook of Macroeconomics*, volume 1, chapter 2, pages 65–148. Elsevier, North Holland: Amsterdam.

Christiano, L. J., Eichenbaum, M., and Evans, C. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45.

Darroch, J. N., Lauritzen, S. L., and Speed, T. P. (1980). Markov fields and log-linear interaction models for contingency tables. *The Annals of Statistics*, 8(3):522–539.

Dempster, A. P. (1972). Covariance selection. *Biometrics*, 28(1):pp. 157–175.

Estrella, A. and Mishkin, F. S. (1997). Is there a role for monetary aggregates in the conduct of monetary policy? *Journal of Monetary Economics*, 40(2):279–304.

Fragetta, M. and Melina, G. (2011). The effects of fiscal policy shocks in SVAR models: A graphical modelling approach. *Scottish Journal of Political Economy*, 58(4):539–567.

Garratt, A., Lee, K., Pesaran, M. H., and Shin, Y. (2003). A long run structural macroeconometric model of the UK. *Economic Journal*, 113(487):412–455.

Greene, W. (2003). *Econometric analysis*. Prentice Hall, NJ: Upper Saddle River.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. Springer-Verlag, New York.

Kilian, L. (2001). Impulse response analysis in vector autoregressions with unknown lag order. *Journal of Forecasting*, 20(3):161–79.

Kim, S. and Roubini, N. (2000). Exchange rate anomalies in the industrial countries: A solution with a structural VAR approach. *Journal of Monetary Economics*, 45(3):561–586.

Koop, G., Leon-Gonzalez, R., and Strachan, R. W. (2009). On the evolution of the monetary policy transmission mechanism. *Journal of Economic Dynamics and Control*, 33(4):997 – 1017.

Lauritzen, S. L. and Spiegelhalter, D. J. (1988). Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 50(2).

Oxley, L., Reale, M., and Wilson, G. (2009). Constructing structural VAR models with conditional independence graphs. *Mathematics and Computers in Simulation*, 79(9):2910–2916.

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, 72(3):821–852.

Reale, M. and Wilson, G. (2001). Identification of vector AR models with recursive structural errors using conditional independence graphs. *Statistical Methods and Applications*, 10:49–65.

Sims, C. A. (1992). Interpreting the macroeconomic time series facts : The effects of monetary policy. *European Economic Review*, 36(5):975–1000.

Sims, C. A. and Zha, T. A. (1998). Does monetary policy generate recessions? *Federal Reserve Bank of Atlanta Working Paper*, 98-12.

Stock, J. H. and Watson, M. W. (2001). Vector autoregressions. *Journal of Economic Perspectives*, 15(4):101–115.

Swanson, N. R. and Granger, C. W. J. (1997). Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions. *Journal of the American Statistical Association*, 92(437):pp. 357–367.

Wilson, G. T. and Reale, M. (2008). The sampling properties of conditional independence graphs for I(1) structural VAR models. *Journal of Time Series Analysis*, 29(5):802–810.

Woodford, M. (2008). How important is money in the conduct of monetary policy? *Journal of Money, Credit and Banking*, 40(8):1561–1598.