

Di Giacinto, Valter

Conference Paper

Differential regional effects of monetary policy: a geographical SVAR approach

42nd Congress of the European Regional Science Association: "From Industry to Advanced Services - Perspectives of European Metropolitan Regions", August 27th - 31st, 2002, Dortmund, Germany

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Di Giacinto, Valter (2002) : Differential regional effects of monetary policy: a geographical SVAR approach, 42nd Congress of the European Regional Science Association: "From Industry to Advanced Services - Perspectives of European Metropolitan Regions", August 27th - 31st, 2002, Dortmund, Germany, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/115709>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Differential Regional Effects of Monetary Policy: A Geographical SVAR Approach

V. Di Giacinto

Banca d'Italia, L'Aquila Regional Economic Research Unit

E-mail: digiacinto.valter@insedia.interbusiness.it

May 21, 2002

Abstract

The study of possible asymmetric effects of monetary policy at a spatially disaggregated scale has recently received increasing attention in the literature due to the introduction of EMU. Different econometric approaches have been proposed to quantify the differences in monetary policy transmission, ranging from large-scale simultaneous equations models to structural vector autoregressions (SVAR). The current paper mainly builds on the SVAR approach and extends it by incorporating geographical information in model's specification, making use of techniques commonly employed in spatial econometrics. While, to capture spatial interactions it would be necessary to adopt a VAR specification modelling jointly the given set of regions, this is generally not feasible using standard VAR models due to the shortage of degrees of freedom. In the proposed specification, information on spatial proximity is used to derive parameter constraints that make the joint estimation feasible for panels of moderate or large dimension. Having introduced the model, with specific reference to the issue of parameter identification, the paper deals with parameter estimation, that is complicated by the simultaneous spatial dependence structure and by the need to impose a complex set of parameter constraints. Finally, to test the model's empirical performance, the paper presents an application to the analysis of the differential monetary policy effects on the US states. Based on the estimation results, geographical heterogeneity in the impulse response function found out in previous studies appears to be confirmed.

1 Introduction¹

The adoption of a single currency by twelve European countries has fostered the attention on the possibility that, given the diversity in the industrial, legal and financial structures across the EMU member countries, the common monetary policy shocks can have asymmetric effects. The measurement of output response to monetary policy decisions has been recently addressed in a number of cross country studies (see, among others, Dornbush, Favero and Giavazzi (1998) Carlino and DeFina (2000), Cecchetti (1999)) by means of a variety of

¹The views expressed are the author's and do not necessarily reflect those of the Banca d'Italia.

econometric techniques, ranging from large-scale econometric models to structural vector autoregressive (SVAR) models (see Guiso et al. (1999) for a review). At a sub-national scale, Carlino and DeFina (1998, 1999) and Horvath (2000) made use of structural VARs to measure regional response to monetary policy in U.S. regions and states, while de Lucio and Izquierdo (1999) conduct a similar exercise on Spanish regions. As noted by Carlino and DeFina, when modeling regional responses to aggregate shocks, it is important to account for feedback effects among regions. The simple estimation of a standard macro-type SVAR model to each region, as done in many cross-country studies, would hence result in serious misspecification, since indirect effects of policy actions, operating, e.g., through trade and financial linkages among regions, would be neglected. In Carlino and DeFina's approach, interdependence across states is dealt with by allowing the lagged output in remaining regions to enter the VAR equation relating to each region/state. However no contemporaneous feedback is allowed, i.e., simultaneous propagation of economic disturbances among regions is excluded. On the contrary, de Lucio and Izquierdo contribution, while ruling out lagged feedback effects among regions, do allow for contemporaneous correlation among the VAR model residuals in their preferred specification, consisting of a set of regional macro-type SVARs jointly estimated using Seemingly Unrelated Regression (SUR) techniques.

The econometric methodology proposed in the paper mainly builds on Carlino and DeFina contributions, trying to overcome some difficulties implicit in their approach. In particular, when analyzing the case of the 48 contiguous U.S. states, given the fact that fitting a single VAR model to large cross-sections data is not feasible due to the rapid shortage of degrees of freedom, the authors are forced to estimate a separate VAR model for each state. As a consequence, a different central bank reaction function is identified and different monetary policy shocks are estimated for each state, while, on theoretical grounds, monetary policy shocks should be the same for all states.

Another possible shortcoming of the Carlino-DeFina methodology relates to the identification scheme adopted, that rules out any contemporaneous interdependence among states by means of a set of overidentifying restrictions imposed on the contemporaneous VAR coefficients matrix. As a consequence, state specific shocks are assumed to affect only the state of origin contemporaneously, although they are allowed to spill over into other regions with a one-quarter lag. This assumption, however, is likely to be too restrictive in practice, especially when no monthly time series are available and the model is fitted using quarterly data.

In an attempt to overcome these difficulties, the SVAR approach of Carlino and DeFina (1999) is extended by incorporating information on geographical proximity in model specification. In particular, based on Tobler's first law of geography (Tobler, 1979), it is assumed that information coming from the nearest neighbouring units is the most relevant in predicting the process at a given location, in the sense that, once it is acquired, the addition of information from remaining locations does not improve the predictive performance of the model. Following the standard approach in spatial econometrics (see Anselin, 1988, chapter 3) and on the grounds of earlier contributions in spatial time series analysis (Martin Oeppen (1975), Pfeifer and Deutsch (1980), Pfeifer and Bodily (1990)), a priori information on the spatial connectivity structure underlying the observed data is made operational within the VAR model by means of a sequence spatial weights matrices, defined according to a proper spatial weighting scheme. A related approach is proposed by Lesage and Pan (1995), that use spatial contiguity information to define Bayesian priors for the parameters of a BVAR

model in a regional forecasting exercise.

Through the sequence of spatial weights matrices, a set of parameter restrictions are imposed on the VAR coefficients matrices that, eliminating the degrees of freedom constraint incurred by VAR models as the cross-sectional dimension of the model increases, allow, on the one hand, the identification and estimation of a single monetary policy shocks series for all regions. On the other hand, spatial constraints are used to model contemporaneous interdependence among regions, while preserving a number of restrictions sufficient for structural parameters identification.

The remaining part of the paper is structured as follows. Section 2 introduces the model, with specific reference to the issue of structural parameter identification. Section 3 then deals with parameter estimation. The simultaneous spatial dependence structure and the set of overidentifying restrictions on model parameters makes the usual OLS estimators inappropriate in this case and, following Amisano and Giannini (1997), FIML (Full Information Maximum Likelihood) estimators are derived assuming a gaussian distribution. In section 4 model's empirical performance is assessed through an empirical application to the measurement of monetary policy transmission in the U.S. states. Section 5 concludes with a brief summary.

2 The model

In this section the proposed structural VAR model is introduced on the basis of the Carlino and DeFina specification. As already noted, a single VAR model is specified including all the given geographical units (also referred to as regions in the following). Three sets of variables enter the VAR system. The first set, denoted as $\mathbf{x}_t = [x_{1t}, x_{2t}, \dots, x_{Kt}]'$, represent K macroeconomic control variables usually entering monetary VAR models, i.e. measures of inflation, commodity prices and other variables capturing supply or demand shocks that are relevant for monetary policy analysis.

The second set of variables, denoted as $\mathbf{y}_t = [y_{1t}, y_{2t}, \dots, y_{Nt}]'$, includes the stacked values of the output variable measured on the N regions. The third set is comprised of a single variable, the monetary policy instrument, denoted as r_t . To simplify notation, without loss of generality, a zero mean data generating process will be assumed throughout.

Setting $\mathbf{z}_t = [\mathbf{x}'_t, \mathbf{y}'_t, r_t]'$, the SVAR model has the following expression

$$\mathbf{C}_0 \mathbf{z}_t = \mathbf{C}_1 \mathbf{z}_{t-1} + \dots + \mathbf{C}_p \mathbf{z}_{t-p} + \mathbf{u}_t \quad (1)$$

where $\mathbf{u}_t = [u_{1t}^x, \dots, u_{Kt}^x, u_{1t}^y, \dots, u_{Nt}^y, u_t^r]$ is an orthogonal multivariate white-noise series, i.e. $E(\mathbf{u}_t) = 0$, $E(\mathbf{u}_t \mathbf{u}'_{t-h}) = \Omega = \text{diag}([\sigma_{x1}^2, \dots, \sigma_{xK}^2, \sigma_{y1}^2, \dots, \sigma_{yN}^2, \sigma_r^2])'$ if $h = 0$ and $E(\mathbf{u}_t \mathbf{u}'_{t-h}) = 0$ elsewhere.

It is well known (see., e.g., Hamilton (1994), chapter 11) that, in addition to the orthogonality of the structural disturbances, structural parameter identification requires proper constraints be imposed on the $(R \times R)$ matrix \mathbf{C}_0 , ($R = K + N + 1$). In Carlino and DeFina's approach, identification is achieved through the following set of assumptions. (1) Region-specific shocks affect only the region of origin contemporaneously, although they are allowed to spill over into other regions in future periods. (2) Monetary policy actions and shocks to macro variables are assumed to affect regional income no sooner that with

a period lag. (3) Macro control variables are not affected contemporaneously by shocks in the remaining variables in the model neither affect each other.

With these assumptions the \mathbf{C}_0 matrix takes the following form

$$\mathbf{C}_0 = \begin{bmatrix} \mathbf{I}_{K+N} & \mathbf{0} \\ \mathbf{a} & 1 \end{bmatrix} \quad (2)$$

where \mathbf{I}_{K+N} is the identity matrix of dimension $K + N$ and \mathbf{a} is a $(1 \times K + N)$ vector of unrestricted coefficients relating the current value of the monetary instrument to the contemporaneous values of remaining variables in the model.

Under these assumptions, the number of constraints imposed on matrix \mathbf{C}_0 exceeds the level strictly necessary for structural parameter identification, actually imposing a number of overidentifying restrictions in model specification. More specifically, while the order condition requires $R(R + 1)/2$ restrictions be imposed on coefficients in \mathbf{C}_0 , with $R = K + N + 1$, $R^2 - (R - 1)$ constraints are imposed in expression (2), introducing $(R^2 - 3R + 2)/2$ overidentifying restrictions.

In the present paper, a similar methodological position is taken, by assuming that the \mathbf{C}_0 matrix possess the following block triangular structure

$$\mathbf{C}_0 = \begin{bmatrix} \mathbf{I}_K & & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_0^{yy} & \\ -\mathbf{C}_0^{rx} & -\mathbf{C}_0^{ry} & 1 \end{bmatrix} \quad (3)$$

where \mathbf{C}_0^{rx} is a $(1 \times K)$ vector of unrestricted coefficients relating the policy instrument to the contemporaneous values of the macro variables x , and where

$$\mathbf{C}_0^{ry} = a_0^{ry} \boldsymbol{\omega}' \quad (4)$$

In expression (4) a_0^{ry} is a scalar parameter to be estimated and $\boldsymbol{\omega}$ is a vector of N fixed coefficients representing the average weight, over the time period analyzed, associated to the each single regional output series in the national aggregate. The restrictions on \mathbf{C}_0^{ry} coefficients, relating the monetary instrument to current values of the regional output series, although not strictly required for identification, are dictated by the need to keep the number of free coefficients in the policy instrument equation fixed as the number of regions increases, thus making it estimable also for large spatial cross-sections. On theoretical grounds, the restrictions are motivated by the assumption that only aggregate output enters the central bank information set and, hence, the monetary instrument response function.

Differently from the reference model, that assumes $\mathbf{C}_0^{yy} = \mathbf{I}_N$, in the proposed specification the \mathbf{C}_0^{yy} matrix, modeling simultaneous regional interdependence, has the following structure

$$\mathbf{C}_0^{yy} = \mathbf{I}_N - \sum_{k=1}^{\lambda_0} \Phi_{0k} \mathbf{W}_k \quad (5)$$

where $\Phi_{0k} = \text{diag} \left([\phi_{0k}^1, \phi_{0k}^2, \dots, \phi_{0k}^N]' \right)$ and \mathbf{W}_k is the $N \times N$ spatial weights matrix of lag order k (see Anselin and Smirnov (1996) for a proper definition) with typical element $w_k(i, j) > 0$ if locations i and j are contiguous at spatial lag k , and $w_k(i, j) = 0$ elsewhere. The definition of spatial lags assumes a preliminary hierarchical ordering, based on decreasing geographical proximity, has been attributed to the given set of spatial units. Such ordering is usually performed according to one of the two following procedures (see, e.g., Arbia, 1989):

1. in a first step an undirected N -node graph is defined whose nodes represent geographical units and arcs connect units that are nearest neighbours; subsequently two spatial units are defined to be k -th order neighbours if the minimal path connecting them on the graph is composed of k edges;
2. letting $d(i, j)$ denote the distance in space between units i and j , a sequence of non overlapping distance intervals $(0, d_1], (d_1, d_2], \dots, (d_{l-1}, d_l]$ is introduced and two units are defined to be spatial neighbours of order k if $d_{k-1} < d(i, j) \leq d_k$.

A common specification in spatial time series analysis (see, e.g., Pfeifer and Deutsch, 1980), assumes equal scaled weights, with non zero spatial weights taking the value $w_k(i, j) = N_i^{-1}(k)$, and where $N_i(k)$ denotes the number of k -th order spatial neighbours of unit i . As in Cliff and Ord (1981), more complex specifications for the weights matrices can be defined by weighting spatial connections according to the length of the common border or the geographical distance between locations.

Identification of the simultaneous coefficients matrix will be discussed with reference to the standard triangular identification scheme, that is well known to provide an exactly identified structure (see, e.g., Hamilton (1994), chapter 11), by imposing the order condition that the number of free structural coefficients in \mathbf{C}_0^{yy} does not exceed the number of free coefficients assuming a recursive structure, that is equal to $(N^2 - N)/2$. Given the fact that, in the proposed specification, the only unrestricted coefficients are the diagonal elements of the Φ_{0k} matrices, $k = 1, \dots, \lambda_0$, the number of free parameters in (5) is equal to $N\lambda_0$. Consequently, identification is achieved when the following condition holds

$$\lambda_0 \leq (N^2 - N)/2N. \quad (6)$$

imposing a bound on the upper summation limit λ_0 in (5), i.e. on the maximum spatial order entering the simultaneous spatial component of the model. Given that the numerator dominates the denominator in the RHS of (6), the constraint becomes progressively less binding as the number of locations in the sample increases, thus being actually relevant only for spatial samples of very small size (say $N < 10$). As a consequence, in most applications the model will still incorporate a set of overidentifying restrictions, although in smaller number with respect to the reference model, that allows no free parameter in \mathbf{C}_0^{yy} , ruling contemporaneous spatial transmission of region specific disturbances.

At the root of the proposed spatial identification scheme there is the assumption that the impact of location specific shocks is highest on nearest neighbouring regions and decays progressively as geographical distance increases, eventually vanishing as the distance separating locations diverges. However, it must be stressed that, although the spatial spill over effects are assumed to decrease with distance in space, the rate of decay varies according to the spatial order of the model and the values of autoregressive coefficients, thus accommodating a large number of empirical situations.

Two types of restrictions are imposed on \mathbf{C}_h matrices ($h = 1, \dots, p$) in expression (1), motivated by the necessity of reducing the number of free coefficients in the model to make it estimable as the cross-sectional dimension of the sample increases. More specifically, the following partition of the \mathbf{C}_h matrix is introduced, where blocks are singled out with

reference to the three sets of variable entering the model

$$\mathbf{C}_h = \begin{bmatrix} \mathbf{C}_h^{xx} & \mathbf{C}_h^{xy} & \mathbf{C}_h^{xr} \\ \mathbf{C}_h^{yx} & \mathbf{C}_h^{yy} & \mathbf{C}_h^{yr} \\ \mathbf{C}_h^{rx} & \mathbf{C}_h^{ry} & \mathbf{C}_h^{rr} \end{bmatrix} \quad (7)$$

and, on the one hand, spatial restrictions are imposed on blocks \mathbf{C}_h^{yy} , that have structure

$$\mathbf{C}_h^{yy} = \sum_{k=0}^{\lambda_h} \Phi_{hk} \mathbf{W}_k \quad (8)$$

where $\Phi_{hk} = \text{diag}([\phi_{hk}^1, \phi_{hk}^2, \dots, \phi_{hk}^N]')$.

On the other hand, coefficients \mathbf{C}_h^{yx} and \mathbf{C}_h^{yr} , relating the control variables and the monetary instrument to past values of the regional output series, are constrained, as in (4), to have values

$$\mathbf{C}_h^{xy} = \mathbf{a}_h^{xy} \boldsymbol{\omega}' \quad (9)$$

$$\mathbf{C}_h^{ry} = a_h^{ry} \boldsymbol{\omega}' \quad (10)$$

where \mathbf{a}_h^{xy} and a_h^{ry} are, respectively, a k -dimensional vector and a scalar to be estimated. All remaining blocks are left unrestricted, as in the standard VAR specification.

Model's reduced form, obtained by left multiplying both sides of (1) by \mathbf{C}_0^{-1} , has expression

$$\mathbf{z}_t = \tilde{\mathbf{C}}_1 \mathbf{z}_{t-1} - \dots - \tilde{\mathbf{C}}_p \mathbf{z}_{t-p} + \tilde{\mathbf{u}}_t \quad (11)$$

with $\tilde{\mathbf{C}}_h = \mathbf{C}_0^{-1} \mathbf{C}_h$, ($h = 1, \dots, p$), and $\tilde{\mathbf{u}}_t = \mathbf{C}_0^{-1} \mathbf{u}_t$.

Denoting as $\tilde{\mathbf{C}}(B)$ the matrix polynomial $(\mathbf{I}_N - \tilde{\mathbf{C}}_1 B^1 - \dots - \tilde{\mathbf{C}}_p B^p)$, where B is the usual Box-Jenkins backward shift operator, stationarity of the process defined by expression (11) requires that the following condition be satisfied

$$\left| \mathbf{I}_N - \sum_{h=1}^p \tilde{\mathbf{C}}_h z^h \right| \neq 0 \text{ for } |z| \leq 1. \quad (12)$$

When expression (12) holds and $E(\tilde{\mathbf{u}}_t' \tilde{\mathbf{u}}_t) \leq c$, ($t = 0, \pm 1, \pm 2, \dots$), for some finite constant c , the process

$$\mathbf{y}_t = \tilde{\mathbf{C}}(B)^{-1} \tilde{\mathbf{u}}_{t-h} = \tilde{\mathbf{C}}(B)^{-1} \mathbf{C}_0^{-1} \mathbf{u}_{t-h} = \sum_{h=1}^{\infty} \Psi_h \mathbf{u}_{t-h} \quad (13)$$

is well defined, in the sense that the infinite sum exists in mean square (Lutkepohl, 1991, p. 490) and expression (13) gives the Wold's representation of the process in terms of the structural disturbance vector \mathbf{u}_t .

The $N \times N$ matrices Ψ_h give the coefficients of the impulse response function, i.e. the dynamic multipliers relating the observed variables to the unobservable random shocks driving the system.

3 Maximum likelihood estimation

Given the fact that the \mathbf{C}_0 matrix is not strictly triangular and \mathbf{C}_h matrices, ($h = 1, \dots, p$), must satisfy a set of parameter constraints, the multivariate LS estimator (Lutkepohl, 1991, pp. 62-65) is not applicable to the proposed specification. Following Amisano and Giannini (1997), consistent estimators of model parameters will be derived by applying the FIML method. Under the assumption that \mathbf{z}_t is jointly normal, the distribution of \mathbf{z}_t , conditional on past observations $\mathbf{z}_{t-1}, \mathbf{z}_{t-2}, \dots$, will be gaussian with mean $\hat{\mathbf{z}}_t = \tilde{\mathbf{C}}_1 \mathbf{z}_{t-1} - \dots - \tilde{\mathbf{C}}_p \mathbf{z}_{t-p}$ and covariance matrix $\tilde{\Omega} = E(\tilde{\mathbf{u}}_t \tilde{\mathbf{u}}_t') = \mathbf{C}_0^{-1} \Sigma_u \mathbf{C}_0^{-1'}$ and has expression

$$\begin{aligned} f(\mathbf{z}_t | \mathbf{z}_{t-1}, \mathbf{z}_{t-2}, \dots) &= \text{const.} - \frac{1}{2} \log(|\Sigma_{\tilde{u}}|) + \\ &- \frac{1}{2} \left(\mathbf{z}_t - \tilde{\mathbf{C}}_1 \mathbf{z}_{t-1} - \dots - \tilde{\mathbf{C}}_p \mathbf{z}_{t-p} \right)' \tilde{\Omega}^{-1} \left(\mathbf{z}_t - \tilde{\mathbf{C}}_1 \mathbf{z}_{t-1} - \dots - \tilde{\mathbf{C}}_p \mathbf{z}_{t-p} \right) = \\ &= \text{const.} + \log(|\mathbf{C}_0|) - \frac{1}{2} \log(|\Omega|) - \frac{1}{2} \mathbf{u}_t' \Omega^{-1} \mathbf{u}_t \end{aligned}$$

where $\mathbf{u}_t = (\mathbf{C}_0 \mathbf{z}_t - \mathbf{C}_1 \mathbf{z}_{t-1} - \dots - \mathbf{C}_p \mathbf{z}_{t-p})$.

Given the block triangular structure of \mathbf{C}_0 , it follows that $|\mathbf{C}_0| = |\mathbf{C}_0^{xx}| \cdot |\mathbf{C}_0^{yy}| \cdot c_0^{rr} = |\mathbf{C}_0^{yy}|$ and the sample distribution, conditional on p pre-sample values of \mathbf{z}_t and assuming T consecutive observations are collected over time, will have expression

$$\begin{aligned} f(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_T | \mathbf{z}_0, \dots, \mathbf{z}_{-p+1}) &= \\ &= \text{const.} + T \log(|\mathbf{C}_0^{yy}|) - \frac{T}{2} \log(|\Omega|) - \frac{1}{2} \sum_{t=1}^T \mathbf{u}_t' \Omega^{-1} \mathbf{u}_t = \\ &= \text{const.} + T \log(|\mathbf{C}_0^{yy}|) - \frac{T}{2} \log(|\Omega|) - \frac{1}{2} \text{tr}(\mathbf{U}' \Omega^{-1} \mathbf{U}). \end{aligned} \quad (14)$$

In the following it will be shown how the sum of squares component in expression (14) can be computed from the residuals of a properly defined linear regression model. To this purpose the following set of positions is introduced first

$$\begin{aligned} \mathbf{Z} &= [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_T]; \quad \underline{\mathbf{z}} = \text{vec}(\mathbf{Z}'); \\ \mathbf{X} &= [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_T]; \quad \underline{\mathbf{x}} = \text{vec}(\mathbf{X}'); \\ \mathbf{Y} &= [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T]; \quad \underline{\mathbf{y}} = \text{vec}(\mathbf{Y}'); \\ \mathbf{r} &= [r_1, r_2, \dots, r_T]'; \\ \mathbf{U} &= [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_T]; \quad \underline{\mathbf{u}} = \text{vec}(\mathbf{U}'); \\ \mathbf{U}_x &= [\mathbf{u}_{x1}, \mathbf{u}_{x2}, \dots, \mathbf{u}_{xT}]; \quad \underline{\mathbf{u}}_x = \text{vec}(\mathbf{U}'_x); \quad \mathbf{u}_{xt} = [u_{1t}^x, \dots, u_{Kt}^x]'; \\ \mathbf{U}_y &= [\mathbf{u}_{y1}, \mathbf{u}_{y2}, \dots, \mathbf{u}_{yT}]; \quad \underline{\mathbf{u}}_y = \text{vec}(\mathbf{U}'_y); \quad \mathbf{u}_{yt} = [u_{1t}^y, \dots, u_{Nt}^y]'; \\ \mathbf{u}_r &= [u_1^r, \dots, u_t^r]'; \\ \mathbf{Z}_{-h} &= [\mathbf{z}_{-h+1}, \dots, \mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_{T-h}] \\ \mathbf{X}_{-h} &= [\mathbf{x}_{-h+1}, \dots, \mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{T-h}] \\ \mathbf{Y}_{-h} &= [\mathbf{y}_{-h+1}, \dots, \mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{T-h}] \end{aligned}$$

$$\begin{aligned}\mathbf{r}_{-h} &= [r_{-h+1}, \dots, r_0, r_1, \dots, r_{T-h}]' \\ \overline{\mathbf{C}}_0 &= -(\mathbf{C}_0 - \mathbf{I}_K)\end{aligned}$$

Based on the above given positions and given the fact that $vec(\mathbf{Z}') = vec([\mathbf{X}', \mathbf{Y}', \mathbf{r}])$, in the following a linear regression form will be derived separately for the three sets of variables, and will subsequently be reassembled to obtain the regression system for the whole model.

Letting \mathbf{C}_h^x denote the $K \times R$ matrix $[\mathbf{C}_h^{xx}, \mathbf{C}_h^{xy}, \mathbf{C}_h^{xr}]$, $h = 1, \dots, p$, and considering that $\overline{\mathbf{C}}_0^x = \mathbf{0}$ equations referring to x variables, considering the whole sample period, become

$$\begin{aligned}\mathbf{X} &= \overline{\mathbf{C}}_0^x \mathbf{Z} + \mathbf{C}_1^x \mathbf{Z}_{-1} + \dots + \mathbf{C}_p^x \mathbf{Z}_{-p} + \mathbf{U}_x \\ \mathbf{X}' &= \mathbf{Z}'_{-1} \mathbf{C}_1^{x'} + \dots + \mathbf{Z}'_{-p} \mathbf{C}_p^{x'} + \mathbf{U}'_x\end{aligned}$$

and can be written in linear regression form as

$$\begin{aligned}\underline{\mathbf{x}} &= (\mathbf{I}_K \otimes \mathbf{Z}'_{-1}) vec(\mathbf{C}_1^{x'}) + \dots + (\mathbf{I}_K \otimes \mathbf{Z}'_{-p}) vec(\mathbf{C}_p^{x'}) + \underline{\mathbf{u}}_x = \\ &= (\mathbf{I}_K \otimes [\mathbf{Z}'_{-1}, \dots, \mathbf{Z}'_{-p}]) vec([\mathbf{C}_1^x, \dots, \mathbf{C}_p^x]') + \underline{\mathbf{u}}_x.\end{aligned}\quad (15)$$

Considering that coefficients $\mathbf{C}_h^{xy'}$ are restricted by the assumption that macro variables x respond only to aggregate shocks, the system can be conveniently simplified by separating explanatory variables in the three groups of variables x , y and r , writing

$$\begin{aligned}\underline{\mathbf{x}} &= (\mathbf{I}_K \otimes [\mathbf{X}'_{-1}, \dots, \mathbf{X}'_{-p}]) vec([\mathbf{C}_1^{xx}, \dots, \mathbf{C}_p^{xx}]') + \\ &+ (\mathbf{I}_K \otimes [\mathbf{Y}'_{-1}, \dots, \mathbf{Y}'_{-p}]) vec([\mathbf{C}_1^{xy}, \dots, \mathbf{C}_p^{xy}]') + \\ &+ (\mathbf{I}_K \otimes [\mathbf{r}_{-1}, \dots, \mathbf{r}_{-p}]) vec([\mathbf{C}_1^{xr}, \dots, \mathbf{C}_p^{xr}]') + \underline{\mathbf{u}}_x.\end{aligned}$$

Given the fact that $\mathbf{C}_h^{xy'} = \boldsymbol{\omega} \mathbf{a}_h^{xy'}$ one obtains

$$(\mathbf{I}_K \otimes \mathbf{Y}'_{-h}) vec(\mathbf{C}_h^{xy'}) = vec(\mathbf{Y}'_{-h} \mathbf{C}_h^{xy'}) = vec(\mathbf{Y}'_{-h} \boldsymbol{\omega} \mathbf{a}_h^{xy'}) = vec(\overline{\mathbf{y}}_{-h} \mathbf{a}_h^{xy'})$$

where

$$\overline{\mathbf{y}}_{-h} = [\overline{y}_{-h+1}, \dots, \overline{y}_0, \overline{y}_1, \dots, \overline{y}_{T-h}]'$$

and \overline{y} denotes the aggregate output series. As a consequence

$$(\mathbf{I}_K \otimes [\mathbf{Y}'_{-1}, \dots, \mathbf{Y}'_{-p}]) vec([\mathbf{C}_1^{xy}, \dots, \mathbf{C}_p^{xy}]') = (\mathbf{I}_K \otimes [\overline{\mathbf{y}}_{-1}, \dots, \overline{\mathbf{y}}_{-p}]) vec([\mathbf{a}_1^{xy}, \dots, \mathbf{a}_p^{xy}]')$$

and, finally, the regression system for x variables can be written as

$$\begin{aligned}\underline{\mathbf{x}} &= (\mathbf{I}_K \otimes [\overline{\mathbf{Z}}'_{-1}, \dots, \overline{\mathbf{Z}}'_{-p}]) vec([\overline{\mathbf{C}}_1^x, \dots, \overline{\mathbf{C}}_p^x]') + \underline{\mathbf{u}}_x = \\ &= \overline{\Xi}_x \boldsymbol{\beta}_x + \underline{\mathbf{u}}_x\end{aligned}\quad (16)$$

where

$$\begin{aligned}\bar{\mathbf{Z}}_{-h} &= [\bar{\mathbf{z}}_{-h+1}, \dots, \bar{\mathbf{z}}_0, \bar{\mathbf{z}}_1, \dots, \bar{\mathbf{z}}_{T-h}], \quad \bar{\mathbf{z}}_t = [\mathbf{x}'_t, \bar{y}_t, r_t]' \\ \bar{\mathbf{C}}_h^x &= [\mathbf{C}_h^{xx}, \mathbf{a}_h^{xy}, \mathbf{C}_h^{xr}], \quad h = 1, \dots, p.\end{aligned}$$

The sub-system relating to the regional output series can be given the following expression

$$\mathbf{Y}' = \mathbf{Z}'\bar{\mathbf{C}}_0^{yy'} + \mathbf{Z}'_{-1}\mathbf{C}_1^{yy'} + \dots + \mathbf{Z}'_{-p}\mathbf{C}_p^{yy'} + \mathbf{U}'_y$$

or, in linear regression form,

$$\begin{aligned}\underline{\mathbf{y}} &= (\mathbf{I}_N \otimes \mathbf{Z}') \text{vec}(\bar{\mathbf{C}}_0^{yy'}) + \dots + (\mathbf{I}_N \otimes \mathbf{Z}'_{-p}) \text{vec}(\mathbf{C}_p^{yy'}) + \underline{\mathbf{u}}_y = \\ &= (\mathbf{I}_N \otimes [\mathbf{Z}', \mathbf{Z}'_{-1}, \dots, \mathbf{Z}'_{-p}]) \text{vec}([\bar{\mathbf{C}}_0^{yy'}, \dots, \mathbf{C}_p^{yy'}]') + \underline{\mathbf{u}}_y\end{aligned}\quad (17)$$

where \mathbf{C}_h^{yy} is the $N \times R$ matrix $[\mathbf{C}_h^{yx}, \mathbf{C}_h^{yy}, \mathbf{C}_h^{yr}]$, $i = 1, \dots, p$, $\bar{\mathbf{C}}_0^{yy} = [\mathbf{0}, \bar{\mathbf{C}}_0^{yy}, \mathbf{0}]$ and $\bar{\mathbf{C}}_0^{yy} = \sum_{k=1}^{\lambda_0} \Phi_{0k} \mathbf{W}_k$.

Separating on the RHS the components relative to the three sets of explanatory variables, expression (17) becomes

$$\begin{aligned}\underline{\mathbf{y}} &= (\mathbf{I}_N \otimes [\mathbf{X}'_{-1}, \dots, \mathbf{X}'_{-p}]) \text{vec}([\mathbf{C}_1^{yx}, \dots, \mathbf{C}_p^{yx}]') + \\ &\quad + (\mathbf{I}_N \otimes [\mathbf{Y}'_{-1}, \mathbf{Y}'_{-2}, \dots, \mathbf{Y}'_{-p}]) \text{vec}([\bar{\mathbf{C}}_0^{yy}, \mathbf{C}_1^{yy}, \dots, \mathbf{C}_p^{yy}]') + \\ &\quad + (\mathbf{I}_N \otimes [\mathbf{r}'_{-1}, \dots, \mathbf{r}'_{-p}]) \text{vec}([\mathbf{C}_1^{yr}, \dots, \mathbf{C}_p^{yr}]') + \underline{\mathbf{u}}_y \\ &= \Xi_y^x \beta_y^x + \Xi_y^y \beta_y^y + \Xi_y^r \beta_y^r + \underline{\mathbf{u}}_y \\ &= [\Xi_y^x, \Xi_y^y, \Xi_y^r] [\beta_y^{x'}, \beta_y^{y'}, \beta_y^{r'}]' + \underline{\mathbf{u}}_y\end{aligned}$$

Considering the expression for \mathbf{C}_h^{yy} , $h = 1, \dots, p$, given by (8) one obtains

$$\text{vec}(\mathbf{C}_h^{yy'}) = \text{vec}\left(\sum_{k=0}^{\lambda_h} \mathbf{W}'_k \Phi_{hk}\right) = \sum_{k=0}^{\lambda_h} (\mathbf{I}_N \otimes \mathbf{W}'_k) \text{vec}(\Phi_{hk})$$

and consequently

$$(\mathbf{I}_N \otimes \mathbf{Y}'_{-h}) \text{vec}(\mathbf{C}_h^{yy'}) = \sum_{k=0}^{\lambda_h} (\mathbf{I}_N \otimes \mathbf{Y}'_{-h} \mathbf{W}'_k) \text{vec}(\Phi_{hk}) \quad (18)$$

The RHS of (18), involving a matrix of dimensions $NT \times N^2$, can be conveniently reduced to an $NT \times N$ dimension considering the diagonal structure of matrix Φ_{hk} . The product $(\mathbf{I}_N \otimes \mathbf{Y}'_{-h} \mathbf{W}'_k) \text{vec}(\Phi_{hk})$ performs, in fact, a linear combination of the columns of matrix $(\mathbf{I}_N \otimes \mathbf{Y}'_{-h} \mathbf{W}'_k)$ with weights given by the vector $\text{vec}(\Phi_{hk})$. Since only to N of the N^2 columns is associated non zero weight, the linear combination can be computed only with

reference to these columns. Since Φ_{hk} is an N -dimensional diagonal matrix, the only non zero elements in $vec(\Phi_{hk})$ will occupy positions $(i-1)*N+i$, $i=1, \dots, N$. Consequently expression (18) can be written as

$$\begin{aligned} & \sum_{k=0}^{\lambda_h} (\mathbf{I}_N \otimes \mathbf{Y}'_{-h} \mathbf{W}'_k) vec(\Phi_{hk}) = \\ & = \sum_{k=0}^{\lambda_h} [(\mathbf{S}_1 \otimes \mathbf{Y}'_{-h} \mathbf{W}'_k \mathbf{s}_1), \dots, (\mathbf{S}_N \otimes \mathbf{Y}'_{-h} \mathbf{W}'_k \mathbf{s}_N)] \phi_{hk} = \sum_{k=0}^{\lambda_h} \mathbb{Y}_{hk} \phi_{hk} \end{aligned}$$

where $\phi_{hk} = [\phi_{1k}^1, \phi_{1k}^2, \dots, \phi_{2k}^N]'$ and, to select the relevant columns in matrix $(\mathbf{I}_N \otimes \mathbf{Y}'_{-h} \mathbf{W}'_k)$ the selection operators \mathbf{S}_i and \mathbf{s}_i have been employed, with \mathbf{S}_i an $N \times N$ matrix with all zero elements except the i -th element on the main diagonal that is equal to 1 and \mathbf{s}_i an N -dimensional vector with zero entries except the i -th element that is set equal to 1.

At this stage, setting $\mathbb{Y}_h = [\mathbb{Y}_{h0}, \dots, \mathbb{Y}_{h\lambda_h}]$ and $\phi_h = [\phi_{h0}, \dots, \phi_{h\lambda_h}]$ we have

$$(\mathbf{I}_N \otimes \mathbf{Y}'_{-h}) vec(\mathbf{C}_h^{yy'}) = \mathbb{Y}_h \phi_h, \quad h = 1, \dots, p.$$

Following the same argument it can be shown that

$$(\mathbf{I}_N \otimes \mathbf{Y}') vec(\overline{\mathbf{C}}_0^{yy'}) = \mathbb{Y}_0 \phi_0$$

with $\mathbb{Y}_0 = [\mathbb{Y}_{01}, \dots, \mathbb{Y}_{0\lambda_0}]$ and $\phi_0 = [\phi_{01}, \dots, \phi_{0\lambda_0}]$ and finally, setting $\mathbb{Y} = [\mathbb{Y}_0, \dots, \mathbb{Y}_p]$ and $\phi = [\phi'_0, \dots, \phi'_p]'$, and rearranging the order of the variables, the regression system for y variables becomes

$$\underline{\mathbf{y}} = [\mathbb{Y}, \Xi_y^x, \Xi_y^r] [\phi', \beta_y^{x'}, \beta_y^{r'}]' + \underline{\mathbf{u}}_y \quad (19)$$

To derive the ML estimators it convenient to separate the simultaneous and predetermined variables. To this purpose we set

$$\begin{aligned} \tilde{\Xi}_y &= [\Xi_{y0}, \tilde{\Xi}_{y1}] \\ \tilde{\beta}_y &= [\beta_y^{0'}, \tilde{\beta}_y^{1'}]' \end{aligned}$$

where $\Xi_{y0} = \mathbb{Y}_0$, $\tilde{\Xi}_{y1} = [\mathbb{Y}_1, \dots, \mathbb{Y}_p, \Xi_y^x, \Xi_y^r]$, $\beta_y^0 = \phi_0$, $\tilde{\beta}_y^1 = [\phi'_1, \dots, \phi'_p, \beta_y^{x'}, \beta_y^{r'}]'$. With these positions expression (19) becomes

$$\underline{\mathbf{y}} = \tilde{\Xi}_y \tilde{\beta}_y + \underline{\mathbf{u}}_y = \Xi_{y0} \beta_y^0 + \tilde{\Xi}_{y1} \tilde{\beta}_y^1 + \underline{\mathbf{u}}_y \quad (20)$$

Since matrix dimensions in (20) can become large when N increases, to facilitate the computation of ML estimates it is convenient to put matrix $\tilde{\Xi}_y^1$ in block diagonal form by reshuffling the columns grouping variables by equation; this can be done by introducing the regressors matrix for generic equation i ,

$$\xi_y^i = [\underline{y}_{10}^i, \dots, \underline{y}_{1\lambda_1}^i, \dots, \underline{y}_{p0}^i, \dots, \underline{y}_{p\lambda_p}^i, \mathbf{X}'_{-1}, \dots, \mathbf{X}'_{-p}, \mathbf{r}_{-1}, \dots, \mathbf{r}_{-p}]$$

where $\underline{y}_{hk}^i = (\mathbf{S}_i \otimes \mathbf{Y}'_{-h} \mathbf{W}'_k \mathbf{s}_i)$, and corresponding coefficient vector

$$\boldsymbol{\beta}_y^1 = \left[\phi_{10}^i, \dots, \phi_{1\lambda_1}^i, \dots, \phi_{p0}^i, \dots, \phi_{p\lambda_p}^i, \mathbf{C}_1^{yx}(i, \cdot), \dots, \mathbf{C}_p^{yx}(i, \cdot), \mathbf{C}_1^{yr}(i), \dots, \mathbf{C}_p^{yr}(i) \right]'$$

where $\mathbf{C}_h^{yx}(i, \cdot)$ is the i -th row of matrix \mathbf{C}_h^{yx} and $\mathbf{C}_h^{yr}(i)$ is the i -th element of \mathbf{C}_1^{yr} . With this positions it is immediate to see that

$$\tilde{\Xi}_{y1} \tilde{\boldsymbol{\beta}}_y^1 = \Xi_{y1} \boldsymbol{\beta}_y^1$$

where Ξ_{y1} is the block diagonal matrix

$$\Xi_y^1 = \begin{bmatrix} \boldsymbol{\xi}_y^1 & & & \\ & \boldsymbol{\xi}_y^2 & & \\ & & \ddots & \\ & & & \boldsymbol{\xi}_y^N \end{bmatrix}.$$

Finally, writing $\Xi_y = [\Xi_{y0}, \Xi_{y1}]$ and $\boldsymbol{\beta}_y = [\boldsymbol{\beta}_y^{0'}, \boldsymbol{\beta}_y^{1'}]'$, expression (20) becomes

$$\underline{\mathbf{y}} = \Xi_y \boldsymbol{\beta}_y + \underline{\mathbf{u}}_y = \Xi_y^0 \boldsymbol{\beta}_y^0 + \Xi_y^1 \boldsymbol{\beta}_y^1 + \underline{\mathbf{u}}_y \quad (21)$$

that in the following will be used to derive the ML estimators.

The regression system for the monetary policy instrument series is straightforward to derive and has expression

$$\mathbf{r}' = [\mathbf{Z}', \mathbf{Z}'_{-1}, \dots, \mathbf{Z}'_{-p}] [\overline{\mathbf{C}}_0^r, \mathbf{C}_1^r, \dots, \mathbf{C}_p^r]' + \mathbf{u}_r'$$

where $\overline{\mathbf{C}}_0^r = [\mathbf{C}_0^{rx}, \mathbf{C}_0^{ry}, 0]$ and $\mathbf{C}_h^r = [\mathbf{C}_h^{rx}, \mathbf{C}_h^{ry}, \mathbf{C}_h^{rr}]$, $h = 1, \dots, p$.

To incorporate constraints deriving from the assumption that the central bank sets the policy instrument on the basis of aggregate output the policy instrument equation can be written as

$$\mathbf{r} = \left[\mathbf{Z}^{*'}, \mathbf{Z}'_{-1}, \dots, \mathbf{Z}'_{-p} \right] \boldsymbol{\beta}_r + \mathbf{u}_r = \Xi_r \boldsymbol{\beta}_r + \mathbf{u}_r \quad (22)$$

where $\mathbf{Z}^* = [\mathbf{z}_1^*, \mathbf{z}_2^*, \dots, \mathbf{z}_T^*]$, $\mathbf{z}_t^* = [\mathbf{x}'_t, \bar{y}_t]'$ and $\boldsymbol{\beta}_r = [\mathbf{C}_0^{rx}, a_0^{ry}, \mathbf{C}_1^{rx}, a_1^{ry}, \mathbf{C}_1^{rr}, \dots, \mathbf{C}_p^{rx}, a_p^{ry}, \mathbf{C}_p^{rr}]'$.

Reassembling the three sub-systems, regression equations for the complete model can be written compactly as

$$\underline{\mathbf{z}} = \Xi \boldsymbol{\beta} + \underline{\mathbf{u}} \quad (23)$$

where $\boldsymbol{\beta} = [\boldsymbol{\beta}'_x, \boldsymbol{\beta}'_y, \boldsymbol{\beta}'_r]'$ and

$$\Xi = \begin{bmatrix} \Xi_x & 0 & 0 \\ 0 & \Xi_y & 0 \\ 0 & 0 & \Xi_r \end{bmatrix}. \quad (24)$$

Using (23) the conditional likelihood function of model parameters β and Σ_u expression becomes

$$\begin{aligned}
l(\beta, \Sigma_u | \mathbf{z}_0, \dots, \mathbf{z}_{-p+1}) &= \text{const.} + T \log(|\mathbf{C}_0^{yy}|) - \frac{T}{2} \log(|\Omega|) + \\
&\quad - \frac{1}{2} \text{vec}(\mathbf{U}')' (\Omega^{-1} \otimes \mathbf{I}_T) \text{vec}(\mathbf{U}) = \\
&= \text{const.} + T \log(|\mathbf{C}_0^{yy}|) - \frac{T}{2} \log(|\Omega|) + \\
&\quad - \frac{1}{2} (\underline{\mathbf{z}} - \Xi \beta)' (\Omega^{-1} \otimes \mathbf{I}_T) (\underline{\mathbf{z}} - \Xi \beta). \tag{25}
\end{aligned}$$

Considering the block diagonal structure of Ξ and given the fact that Ω is a diagonal matrix the log-likelihood function, apart from the constant term, can be expressed as the sum of three separate components, each function only of a subset of model parameters. More specifically we have

$$l(\beta, \Sigma_u | \mathbf{z}_0, \dots, \mathbf{z}_{-p+1}) = \text{const.} + l_x + l_y + l_r$$

where

$$l_x = l(\beta_x, \Omega_x | \mathbf{z}_0, \dots, \mathbf{z}_{-p+1}) = -\frac{T}{2} \log(|\Omega_x|) - \frac{1}{2} (\underline{\mathbf{x}} - \Xi_x \beta_x)' (\Omega_x^{-1} \otimes \mathbf{I}_T) (\underline{\mathbf{x}} - \Xi_x \beta_x)$$

$$\begin{aligned}
l_y &= l(\beta_y, \Omega_y | \mathbf{z}_0, \dots, \mathbf{z}_{-p+1}) = T \log(|\mathbf{C}_0^{yy}|) - \frac{T}{2} \log(|\Omega_y|) + \\
&\quad - \frac{1}{2} (\underline{\mathbf{y}} - \Xi_y \beta_y)' (\Omega_y^{-1} \otimes \mathbf{I}_T) (\underline{\mathbf{y}} - \Xi_y \beta_y)
\end{aligned}$$

$$l_r = l(\beta_r, \sigma_r^2 | \mathbf{z}_0, \dots, \mathbf{z}_{-p+1}) = -\frac{T}{2} \log(\sigma_r^2) - \frac{1}{2\sigma_r^2} (\mathbf{r} - \Xi_r \beta_r)' (\mathbf{r} - \Xi_r \beta_r)$$

$$\Omega_x = \text{diag}([\sigma_{x1}^2, \dots, \sigma_{xK}^2]') \text{ and } \Omega_y = \text{diag}([\sigma_{y1}^2, \dots, \sigma_{yN}^2]')$$

First order conditions for the maximum of the log-likelihood function have expression

$$\frac{\partial l}{\partial \beta_x} = \Xi_x' (\Omega_x^{-1} \otimes \mathbf{I}_T) \underline{\mathbf{u}}_x = 0 \tag{26}$$

$$\frac{\partial l}{\partial \beta_y^0} = g(\phi_0) + \Xi_y^{0'} (\Omega_y^{-1} \otimes \mathbf{I}_T) \underline{\mathbf{u}}_y = 0 \tag{27}$$

$$\frac{\partial l}{\partial \beta_y^1} = \Xi_y^{1'} (\Omega_y^{-1} \otimes \mathbf{I}_T) \underline{\mathbf{u}}_y = 0 \tag{28}$$

$$\frac{\partial l}{\partial \beta_r} = \frac{1}{\sigma_r^2} \Xi_r' \mathbf{u}_r = 0 \tag{29}$$

where $g(\phi_0) = -[tr(\mathbf{C}_0^{yy} \mathbf{W}_1), \dots, tr(\mathbf{C}_0^{yy} \mathbf{W}_{\lambda_0})]$.

In the first set of normal equations

$$\Xi_x' (\Omega_x^{-1} \otimes \mathbf{I}_T) = \Xi_x' (\Omega_x^{-1} \otimes \mathbf{I}_T) \Xi_x \beta_x$$

setting $\mathbf{M}_x = [\bar{\mathbf{Z}}'_{-1}, \dots, \bar{\mathbf{Z}}'_{-p}]$ it follows that

$$\begin{aligned}
(\mathbf{I}_K \otimes \mathbf{M}_x)' (\Omega_x^{-1} \otimes \mathbf{I}_T) (\mathbf{I}_K \otimes \mathbf{M}) &= (\mathbf{I}_K \otimes \mathbf{M})' (\Omega_x^{-1} \otimes \mathbf{I}_T)'_x (\mathbf{I}_K \otimes \mathbf{M}) \boldsymbol{\beta}_x \\
(\Omega_x^{-1} \otimes \mathbf{M}'_x) \underline{\mathbf{x}} &= (\Omega_x^{-1} \otimes \mathbf{M}'_x \mathbf{M}_x) \boldsymbol{\beta}_x \\
\boldsymbol{\beta}_x &= (\Omega_x^{-1} \otimes \mathbf{M}'_x \mathbf{M}_x)^{-1} (\Omega_x^{-1} \otimes \mathbf{M}'_x) \underline{\mathbf{x}} = \\
&= [\mathbf{I}_K \otimes (\mathbf{M}'_x \mathbf{M}_x)^{-1} \mathbf{M}_x] \underline{\mathbf{x}}
\end{aligned} \tag{30}$$

the last expression stating the well known result from SUR model literature (see Greene (2000), chapter 15) that ML estimators coincide with OLS estimators applied to each single equation when the matrix of regressors is the same for all equations.

From the corresponding normal equations, the ML estimator of the $\boldsymbol{\beta}_r$ vector is also seen to coincide with the simple OLS estimator

$$\hat{\boldsymbol{\beta}}_r = (\mathbf{M}'_r \mathbf{M}_r)^{-1} \mathbf{M}_r \mathbf{r} \tag{31}$$

where $\mathbf{M}_r = [\mathbf{Z}^{*'}, \bar{\mathbf{Z}}'_{-1}, \dots, \bar{\mathbf{Z}}'_{-p}]$.

Normal equations relating to the $\boldsymbol{\beta}_y^1$ parameters have the form

$$\begin{aligned}
\Xi_y^{1'} (\Omega_y^{-1} \otimes \mathbf{I}_T) \underline{\mathbf{y}} &= \Xi_x^{1'} (\Omega_y^{-1} \otimes \mathbf{I}_T) \Xi_y \boldsymbol{\beta}_y = \\
&= \Xi_x^{1'} (\Omega_y^{-1} \otimes \mathbf{I}_T) (\Xi_y^0 \boldsymbol{\beta}_y^0 + \Xi_y^1 \boldsymbol{\beta}_y^1)
\end{aligned}$$

from which the following expression for the ML estimator ensues

$$\hat{\boldsymbol{\beta}}_y^1 = [\Xi_y^{1'} (\Omega_y^{-1} \otimes \mathbf{I}_T) \Xi_y^1]^{-1} \Xi_y^{1'} (\Omega_y^{-1} \otimes \mathbf{I}_T) (\underline{\mathbf{y}} - \Xi_y^0 \boldsymbol{\beta}_y^0). \tag{32}$$

Given the block diagonal structure of Ξ_y^1 and the diagonal structure of Ω_y , the $\hat{\boldsymbol{\beta}}_y^1$ estimates can be computed equations by equation (i.e. separately for each region) by means of the much simpler expression

$$\hat{\boldsymbol{\beta}}_{yi}^1 = (\boldsymbol{\xi}_y^{i'} \boldsymbol{\xi}_y^i)^{-1} \boldsymbol{\xi}_y^{i'} (\underline{\mathbf{y}}_i - \underline{\mathbf{y}}_i^0) \tag{33}$$

where $\underline{\mathbf{y}}_i = (\mathbf{s}'_i \otimes \mathbf{I}_T) \underline{\mathbf{y}}$ and $\underline{\mathbf{y}}_i^0 = (\mathbf{s}'_i \otimes \mathbf{I}_T) \Xi_y^0 \boldsymbol{\beta}_y^0$, ($i = 1, \dots, N$), i.e. by the OLS regression of $(\underline{\mathbf{y}}_i - \underline{\mathbf{y}}_i^0)$ on $\boldsymbol{\xi}_y^{i'}$.

Analogously to the case of the mixed regressive-spatial autoregressive model (Ord (1975); Anselin (1988)), while the normal equations yield a closed form solution for the ML estimator of $\boldsymbol{\beta}_y^1$, this is not directly usable since it involves the unknown parameters $\boldsymbol{\beta}_y^0$.

Due to the inclusion of the term $g(\boldsymbol{\phi}_0)$ no closed form solutions for $\boldsymbol{\beta}_y^0$ obtains from first order conditions and iterative search methods will be required to maximize the log-likelihood function.

As a first step in the maximization of the l_y function with respect to $\boldsymbol{\beta}_y$ one can concentrate Ω_y out of the likelihood by substituting the corresponding ML estimators for $\sigma_{y1}^2, \dots, \sigma_{yN}^2$. From first order conditions

$$\begin{aligned}
\frac{\partial l}{\partial \sigma_{yi}^2} &= -\frac{T}{2} (\sigma_{yi}^2)^{-1} + \frac{1}{2} \mathbf{u}'_y (\Omega_y^{-1} \mathbf{S}_i \Omega_y^{-1} \otimes \mathbf{I}_T) \underline{\mathbf{u}}_y = \\
&= -\frac{T}{2} (\sigma_{yi}^2)^{-1} + \frac{1}{2} (\sigma_{yi}^2)^{-2} \mathbf{u}'_{yi} \underline{\mathbf{u}}_{yi} = 0
\end{aligned}$$

where $\mathbf{u}_{yi} = (\mathbf{S}_i \otimes \mathbf{I}_T) \underline{\mathbf{u}}_y$, $i = 1, \dots, N$, the usual ML variance estimator obtains

$$\hat{\sigma}_{yi}^2 = \frac{1}{T} \mathbf{u}'_{yi} \mathbf{u}_{yi} \quad (34)$$

and the concentrated likelihood, apart from a constant, will thus have expression

$$l_y^c = T \log (|\mathbf{C}_0^{yy}|) - \frac{T}{2} \sum_i^N \log (\hat{\sigma}_{yi}^2). \quad (35)$$

To compute the ML estimates of coefficients β_y the following step-wise optimization routine can be employed:

- given a starting value for β_y^0 compute β_y^1 using (32);
- with the value obtained for β_y^1 maximize (35) with respect to β_y^0 ;
- iterate until convergence.

4 An empirical illustration

To assess model performance in empirical applications, an analysis of the regional effects of monetary policy in the U.S. was conducted, the choice being motivated both by the availability of spatially disaggregated quarterly output series covering a long time span and by the willingness to compare model's predictions with the findings of other authors that have previously thoroughly analyzed the topic. Particular emphasis will be given to comparisons with the Carlino and the DeFina (1999) analysis of differential monetary policy effects in the U.S. states, also referred to as the reference model in the following.

Analogously to the reference specification, the quarterly personal income series by the Bureau of Economic Analysis for the 48 conterminous U.S. states plus the District of Columbia was used to represent state output while the Federal Funds rate was assumed as the monetary policy instrument. State income series were deflated using the aggregate Consumer Price Index (CPI) series.

The period analyzed ranges from the second quarter of 1958 to the fourth quarter of 2000, and, as in the reference model, first differenced data including lags up to order 4 were used for estimation. In a first specification the set of macro control variables x included only an energy price variable, defined as the ratio of the Producer Price Index for fuels and related products to the total Producer Price Index.

Regarding the spatial structure of the model, following a standard approach (see, e.g., Rey and Montoury, 1999) the first order spatial weights matrix was defined based on the existence of a common border. Higher order matrices were then obtained applying the graph based approach reviewed in section 2. Binary weights matrices were subsequently row-normalized. As a preliminary choice, subject to possible revision, the spatial lag order was set to 1 for each temporal lag, ($\lambda_0 = \lambda_1 = \dots = \lambda_4 = 1$).

Based on ML estimates significant contemporaneous feedback effects among states were detected, with all ϕ_{01}^i coefficients significant except in 3 cases and equal on average to 0.57, with an average standard error of 0.11.

To evaluate model's ability in tracking the correlation features of the data, table 1 reports the space-time autocorrelation function (see Pfeifer and Deutsch (1980) for a definition) for the growth rate of the state income series and for the corresponding model residuals. While the original series displays substantial correlation, especially high between spatially lagged contemporaneous series and decaying progressively as the spatial lag increases, the residual series appears essentially uncorrelated. Given the absence of residual correlation the preliminary lag order specification was accepted and the impulse response function was computed based on the estimated autoregressive coefficient matrices.

TABLE 1: Space-time autocorrelation function for the state income series and model residuals.

Growth rate of quarterly state personal income series

Temporal lag	Spatial lag									
	0	1		2		3		4		
0	1.000	0.759	***	0.692	***	0.635	***	0.613	***	
1	0.119	***	0.151	***	0.151	***	0.140	***	0.137	***
2	0.130	***	0.115	***	0.105	***	0.095	***	0.097	***
3	0.190	***	0.209	***	0.206	***	0.197	***	0.200	***
4	0.080	***	0.084	***	0.096	***	0.088	***	0.079	***
5	-0.004		-0.014	*	-0.024	***	-0.029	***	-0.034	***
6	0.058	***	0.046	***	0.037	***	0.027	***	0.023	***
7	0.011		-0.014	*	-0.022	***	-0.030	***	-0.035	***
8	-0.049	***	-0.070	***	-0.078	***	-0.090	***	-0.101	***

Model 1 residual series

Temporal lag	Spatial lag								
	0	1	2	3	4				
0	1.000	0.048	***	0.007	0.030	***	0.008		
1	-0.018	-0.013	*	0.015	**	-0.010	*	0.015	***
2	0.019	0.015	**	-0.027	***	-0.005		0.016	***
3	-0.011	0.003		-0.005		-0.016	**	0.000	
4	-0.004	-0.004		-0.014	**	0.004		0.007	
5	-0.005	0.006		-0.009		-0.006		-0.007	
6	-0.001	-0.006		0.021	***	0.003		0.010	*
7	0.004	-0.007		-0.005		0.005		-0.001	
8	-0.014	-0.001		-0.013	**	-0.013	**	-0.014	**

- * Significant at the 0.05 level.
- ** Significant at the 0.01 level.
- *** Significant at the 0.001 level.

TABLE 1. Continued

Model 2 residual series

Temporal lag	Spatial lag				
	0	1	2	3	4
0	1.000	0.021 ***	0.016 ***	0.011 *	-0.003
1	-0.024 *	0.015 **	-0.032 ***	-0.051 ***	0.024 ***
2	-0.017	-0.006	-0.018 ***	-0.012 **	-0.007
3	0.002	0.006	0.014 **	-0.004	0.007
4	-0.016	0.007	0.000	-0.012 **	0.010 *
5	-0.001	0.003	0.003	-0.003	0.001
6	-0.001	0.010	0.011 *	0.001	0.002
7	0.002	-0.010	-0.001	-0.001	0.005
8	0.004	0.011 *	0.008	0.000	-0.013 **

* Significant at the 0.05 level.

** Significant at the 0.01 level.

*** Significant at the 0.001 level.

The cumulative response of real income to a 1 percentage point shock to the Federal Funds rate is plotted in Figure 1 separately for the different BEA regions. In each figure the aggregate response, computed as an income weighted average of the state level responses, is plotted as a reference value. Average response appears to be well behaved, declining progressively after the policy shock and reaching a long run level of -0.95 percentage points after about 12 quarters, a value comparable with the -1.16 found with the reference model. The dispersion of long run (12 quarters cumulated) state responses, as measured by the coefficient of variation (computed as the ratio of the standard deviation to the weighted average), is equal to 0.50, a value larger than the one obtained based on long run responses reported by Carlino and Defina (0.40) and confirming the evidence of considerable differences in regional responses to monetary policy. Moreover, long run responses appear to be in overall agreement with the corresponding estimates under the reference model, the correlation coefficient between the two being equal to 0.48.

Among the states Arizona, Florida, Georgia, Oregon and Ohio display the largest responses while Montana, New Mexico, Texas and Colorado appear the least sensitive. At the regional level, Southeast and Great Lakes present the majority of states responding more than the average, while in the Southwest, Rocky Mountain and Plains regions income response to monetary policy shocks appears generally smaller than average. The Mideast displays the least variation, with all states responding close to the national aggregate. All these features appear overall consistent with the findings of Carlino and the DeFina, and, in particular, the different behaviour between core and non core regions appears to be confirmed.

FIGURE 1: Cumulative Impulse Response of the Growth Rate of State Real Personal Income to Federal Funds Rate shock, grouped by BEA Region (Model 1).

Fig. 1a - New England

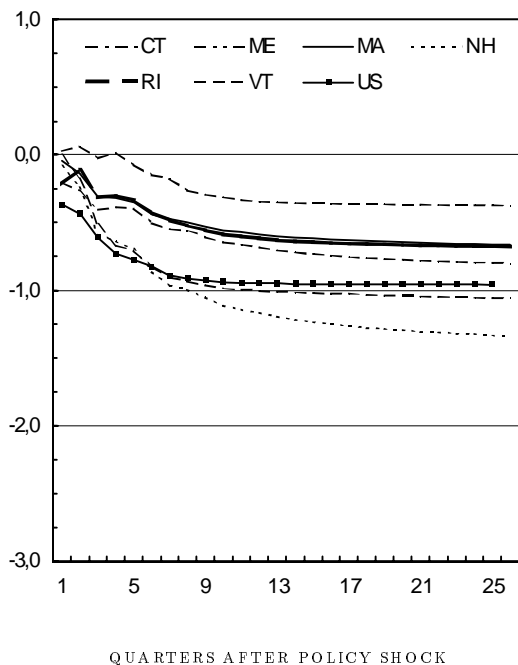


Fig. 1b - Mid East

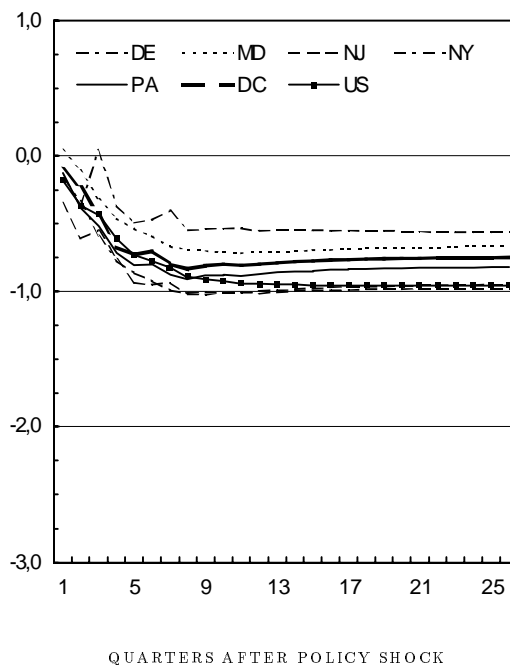


Fig. 1c - Great Lakes

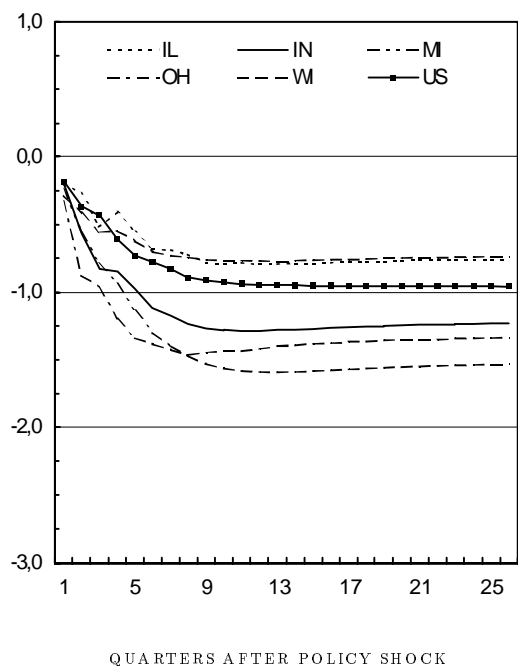


Fig. 1d - Plains

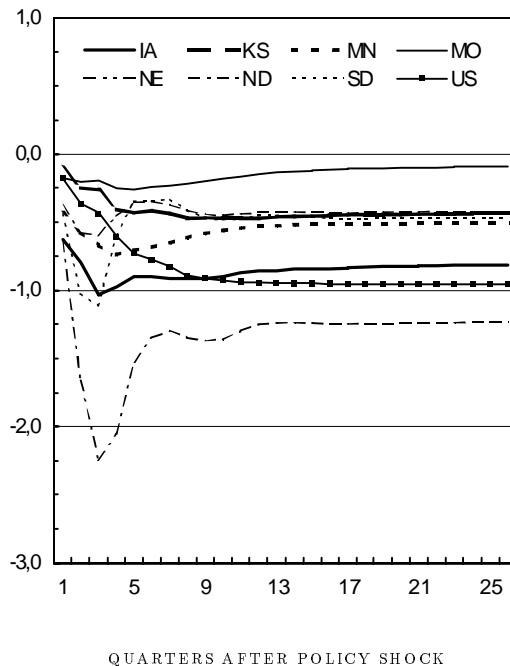


FIGURE 1: Continued.

Fig. 1e - Southeast I

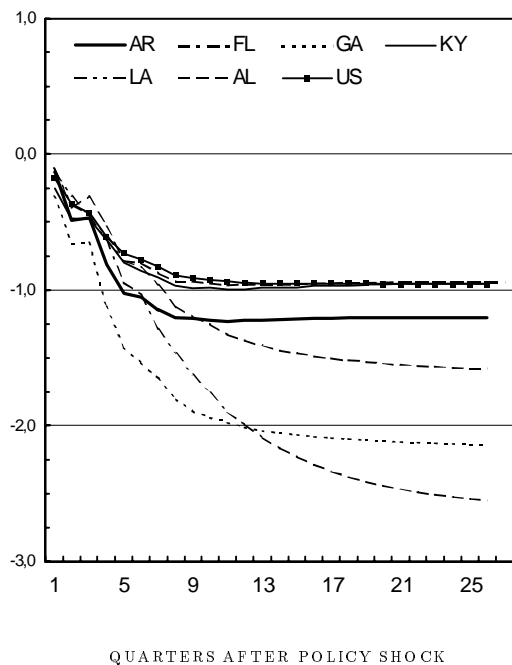


Fig. 1f - Southeast II

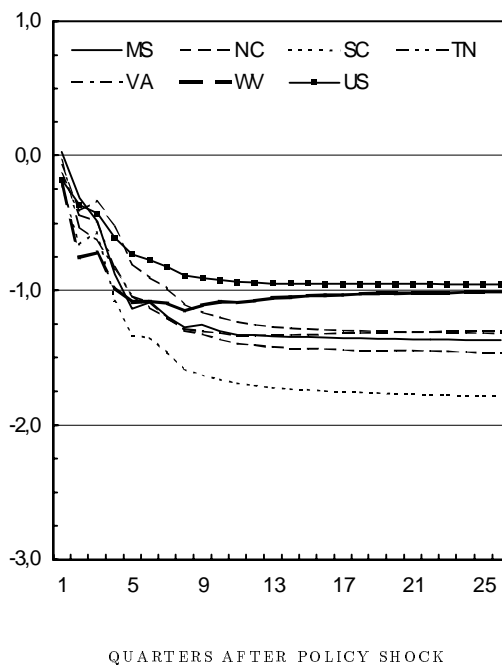


Fig. 1g - Southwest

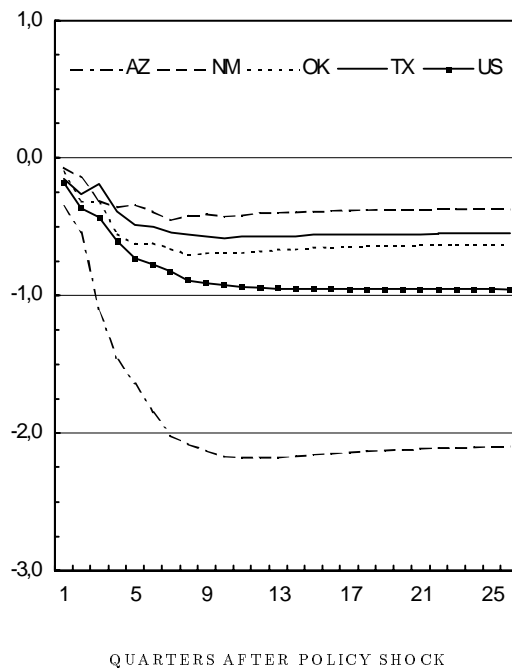


Fig. 1h - Rocky Mountain

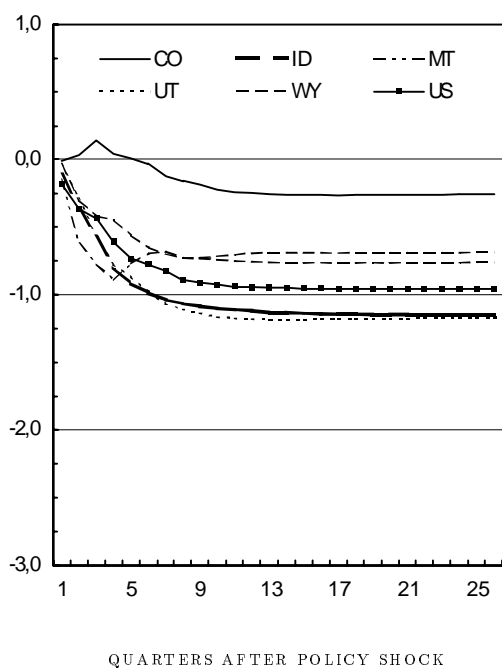
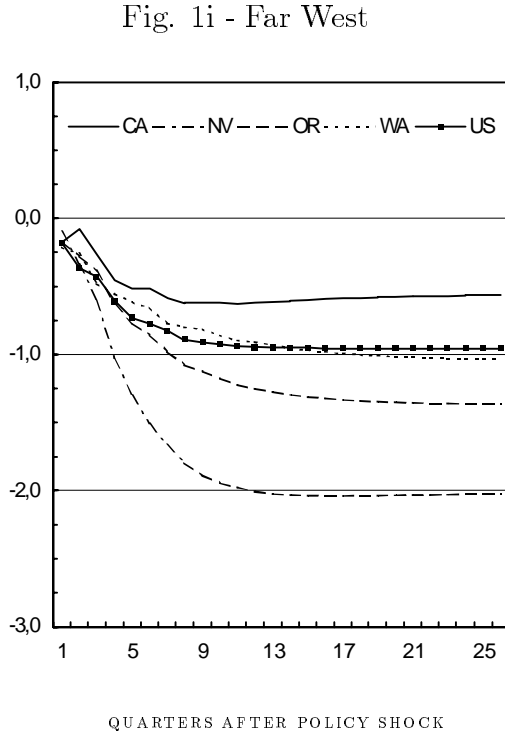


FIGURE 1: Continued.



To check the robustness of the empirical findings, a second specification was estimated, by augmenting the set of variables employed to control for macroeconomic disturbances affecting state economies and monetary policy decisions. More specifically, following the reference model, the CPI and a leading indicator (the OECD composite leading indicator series) were added to the energy price variable. The lag structure and remaining features of the model were left unchanged.

Residual correlation, displayed in table 1, is still very small and estimated contemporaneous feedback effects among states appear to be overall unaffected, ϕ_{01}^i coefficients being significant except in 2 cases, with an average value of 0.56 and an average standard error of 0.11. Cumulative output responses to a shock to the Federal Funds Rate are plotted in Figure 2. The average response is still well behaved and reaches the long run level of -0.74 after about 10 quarters, still a reasonable value, although evidencing a milder monetary policy effect with respect both to the reference model and the baseline specification. The dispersion of state-level responses appears to be intermediate between Model 1 and the reference model, the coefficient of variation being equal to 0.45.

Differently from Model 1 individual states response functions display a damped pseudo-oscillatory pattern, making the graphic evaluation of the responses less immediate. However the results appear to be in overall agreement with the ones obtained for Model 1. The correlation of long run responses is, in fact, equal to 0.47 and most qualitative findings are essentially unchanged, the highest sensitivity still being recorded in states belonging to the Southeast and Great Lakes regions, while the Mideast and Great Lakes regions still display responses generally close to the national average. A different result is observed for the Plains region, where the majority of the states now respond more than the average. Southwest and Plains continue to display the largest dispersion of state responses around the aggregate average.

FIGURE 2: Cumulative Impulse Response of the Growth Rate of State Real Personal Income to Federal Funds Rate shock, grouped by BEA Region (Model 2).

Fig. 2a - New England

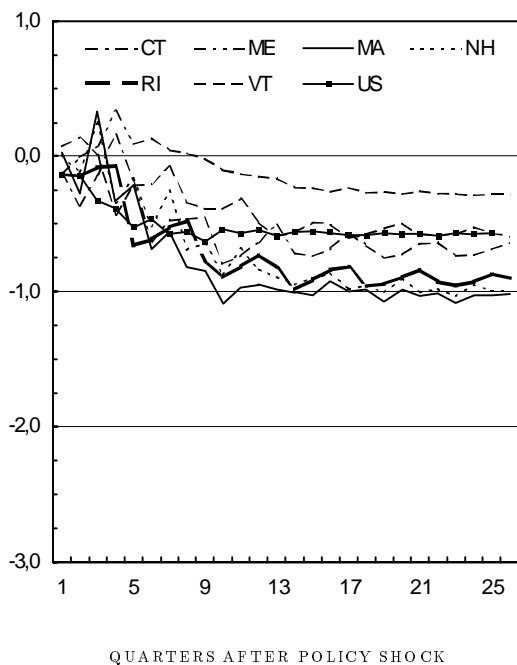


Fig. 2b - Mid East

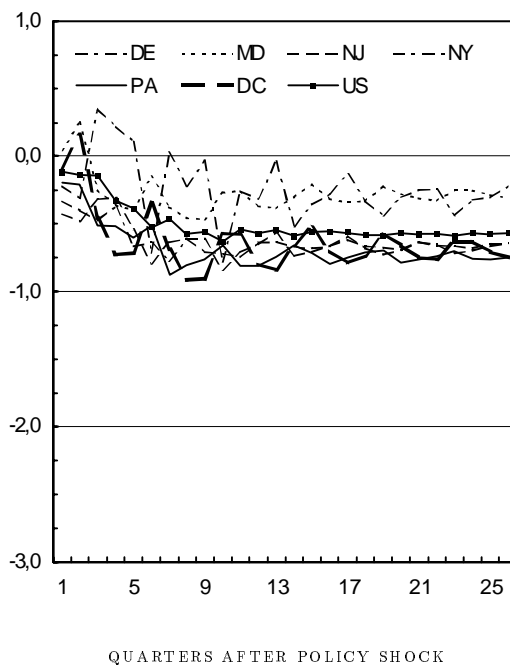


Fig. 2c - Great Lakes

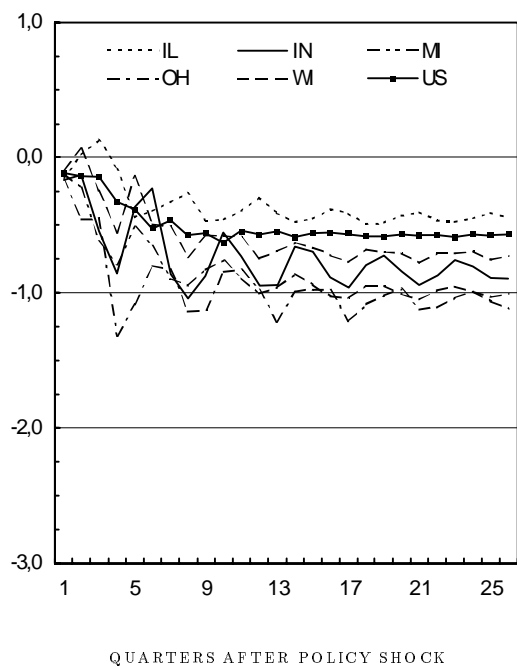


Fig. 2d - Plains

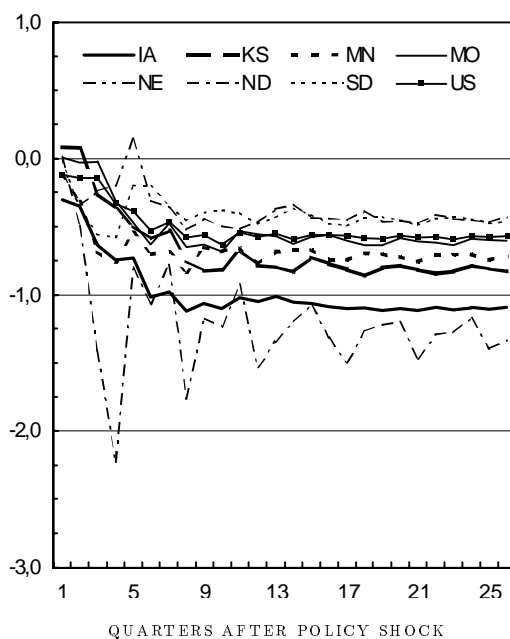


FIGURE 2: Continued.

Fig. 2e - Southeast I

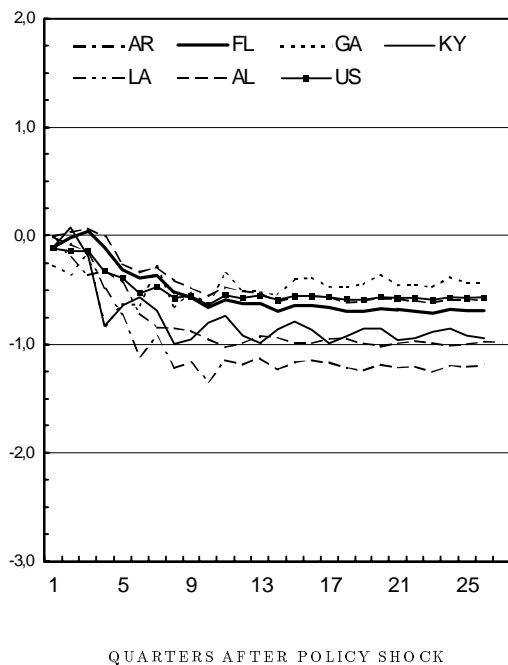


Fig. 2f - Southeast II

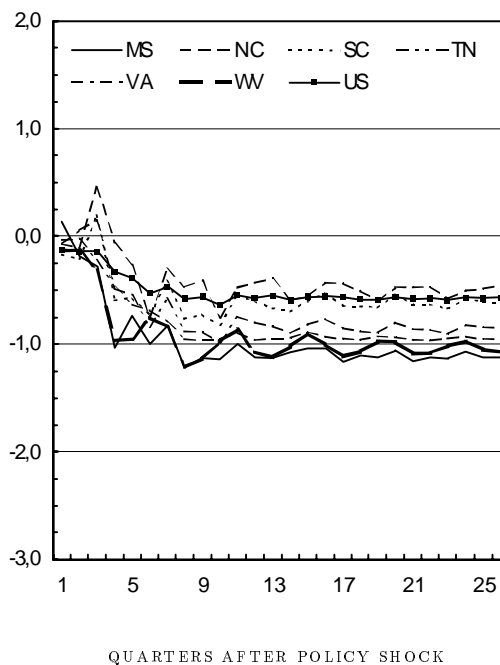


Fig. 2g - Southwest

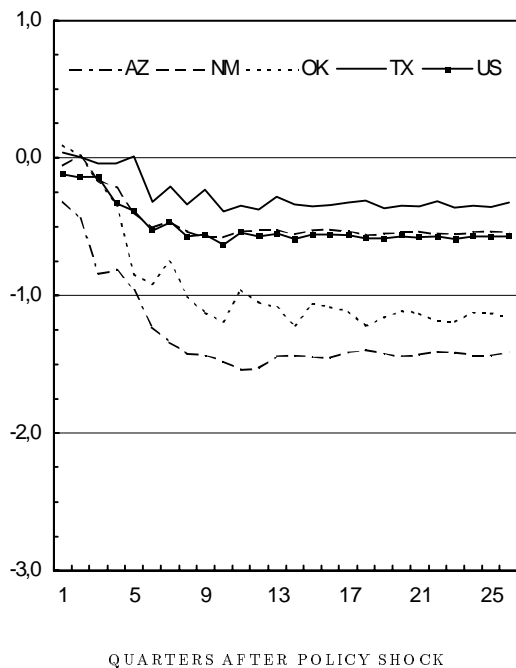


Fig. 2h - Rocky Mountain

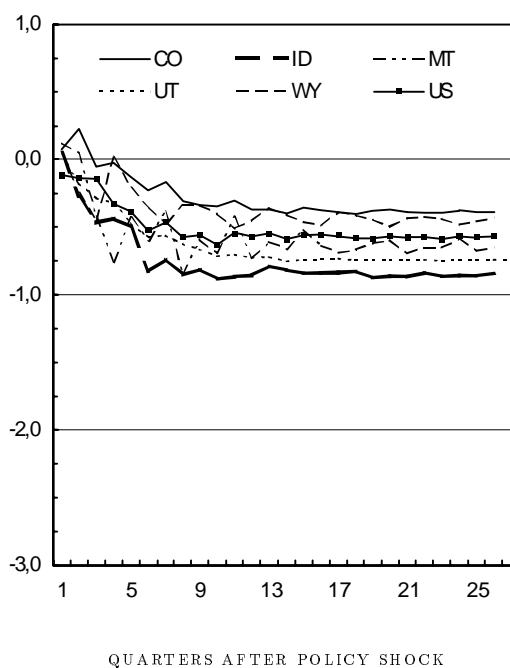
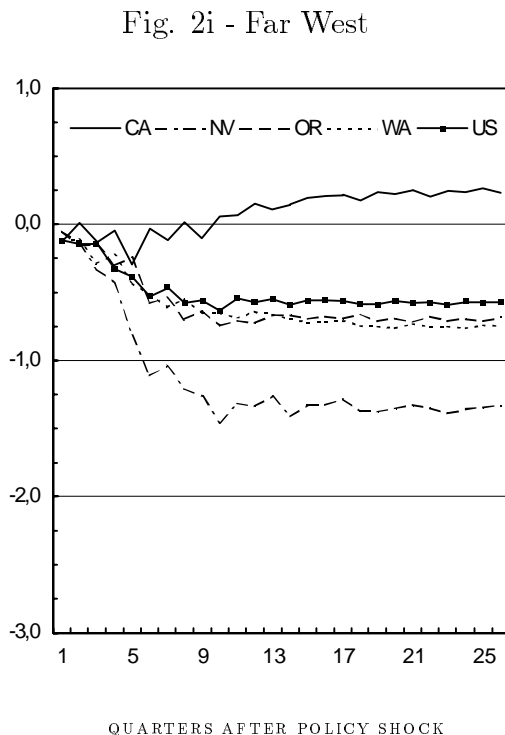


FIGURE 2: Continued.



5 Conclusions

In the paper the structural VAR model specification proposed in the literature for the analysis of regional effects of monetary policy shocks has been extended incorporating information on geographical proximity by means of techniques developed in fields of spatial time series analysis and spatial econometrics. A priori constraints deriving from geographical information, apart from making possible the estimation of a single VAR model for moderate or large spatial cross-sections (the size of the sample being limited only by computational feasibility), allow for the identification and estimation of simultaneous spatial feedback effects, a feature generally neglected in regional applications of monetary SVARs.

As an empirical verification of model's performance an analysis of the differential response of real income to monetary policy was conducted using data for the U.S. states. A baseline specification, controlling for the aggregate supply side shocks stemming from energy price fluctuations, displayed substantial contemporaneous spatial feedback effects, thus suggesting the effective need to model this channel of transmission of economic disturbances. The estimated income responses to a 1 percentage point shock to the monetary policy rate appeared to be generally well behaved and in overall agreement with the findings of previous studies, with a slightly larger spatial dispersion of long run dynamic multipliers.

Following the formulation proposed by Carlino and DeFina (1999), a second specification was estimated by augmenting the set of macroeconomic variables with the Consumer Price Index and a leading indicator series. While quantitative results present differences, qualitative findings appear overall confirmed. In particular, the enhanced control for common macroeconomic shocks does not alter the magnitude of the contemporaneous spatial interaction coefficients, confirming the importance of simultaneous feedback effects when

estimation is not carried out on very high frequency data.

Overall, while a larger set of empirical applications is required to fully evaluate the utility of the proposed methodology, first empirical findings appear to validate the use of geographical constraints in regional monetary VAR modeling, with respect to both the plausibility of results and general agreement with the findings of previous studies.

References

- [1] Amisano, G. and C. Giannini (1997), *Topics in Structural VAR Econometrics*, Berlin, Springer.
- [2] Anselin, L. (1988), *Spatial econometrics: models and applications*, Dordrecht, Kluwer Academic Publishers.
- [3] Anselin L., and O. Smirnov (1996), "Efficient Algorithms for Constructing Proper Higher Order Spatial Lag Operators", *Journal of Regional Science*, Vol. 36, No. 1, pp. 67-89.
- [4] Arbia, G. (1989) *Spatial data configuration in statistical analysis of regional economics and related problems*, Dordrecht, Kluwer Academic Publishers.
- [5] Carlino G. A. and R. DeFina (1998), "The Differential Regional Effects of Monetary Policy", *Review of Economics and Statistics*, Vol. 80, No. 4, pp. 572-587.
- [6] Carlino G. A. and R. DeFina (1999), "The Differential Regional Effects of Monetary Policy: Evidence from the U.S. States", *Journal of Regional Science*, Vol. 80, No. 4, pp. 572-587.
- [7] Carlino G. A. and R. DeFina (2000), "Monetary Policy and the U.S. States and Regions: Some Implications for European Monetary Union", in von Hagen J. and Waller C. J. (eds.), *Regional Aspects of Monetary Policy in Europe*, Kluwer Academic Publisher, Dordrecht, pp. 45-67.
- [8] Cecchetti, S. G. (1999), "Legal Structure, Financial Structure and the Monetary Policy Transmission Mechanism", National Bureau of Economic Research, working paper, No. 7151.
- [9] Cliff, A. D. and J. K. Ord (1981), *Spatial Processes: Models and Applications*, London, Pion.
- [10] de Lucio J. J. and M. Izquierdo (1999), "Local Responses to a Global Monetary Policy: The Regional Structure of Financial Systems", FEDEA, Documento de trabajo, No. 99-14.
- [11] Greene, W.H. (2000) *Econometric analysis-fourth edition*, Prentice Hall Int.
- [12] Guiso L., A. K. Kashyap, F. Panetta and D. Terlizzese (1999), "Will a Common European Monetary Policy Have Asymmetric Effects?", *Economic Perspectives*, Vol. 23, No. 4, pp. 56-75.

- [13] Hamilton, J.D. (1994), *Time series analysis*, Princeton University Press.
- [14] Horvath, M. (2000). "Empirical Evidence on Common Money and Uncommon Regions in the United States", in von Hagen J. and Waller C. J. (eds.), *Regional Aspects of Monetary Policy in Europe*, Kluwer Academic Publisher, Dordrecht, pp. 3-43.
- [15] Lutkepohl, H. (1991), *Introduction to Multiple Time Series Analysis*, Berlin, Springer-Verlag.
- [16] Lesage J. P. and Zheng Pan (1995), "Using Spatial Contiguity as Bayesian Prior Information in Regional Forecasting Models", *International Regional science Review*, Vol. 18, No. 1, pp. 33-53.
- [17] Martin, R. L. and J. E. Oeppen (1975), "The identification of regional forecasting models using space-time correlation functions", *Trans. Inst. British Geographers*, 66, 95-118.
- [18] Ord, J. K. (1975), "Estimation Methods for Models of Spatial Interaction" *Journal of the American Statistical Association*, Vol. 70, No. 1, pp. 120-126.
- [19] Pfeifer, P. E. and S. E. Bodily (1990), "A Test of Space-Time ARMA modelling and Forecasting of Hotel Data", *Journal of Forecasting*, Vol. 9, pp. 255-272.
- [20] Pfeifer, P. E. and S. J. Deutsch (1980), "A Three-Stage Iterative Procedure for Space-Time Modeling", *Technometrics*, 22, pp. 35-47.
- [21] Rey, S. J. and B. D. Montouri (1999), "US Regional Income Convergence: A Spatial Econometric Perspective", *Regional Studies*, Vol. 3, No. 2, pp. 143-156.
- [22] Tobler W. (1979), "Cellular Geography", in Gale S. and Olsson G. (eds.) *Phylosophy in Geography*, Reidel, Dordrecht, pp. 379-386.