Monetary Policy and Exchange Rate Shocks in Brazil: Sign Restrictions versus a New Hybrid Identification Approach^{*}

Elcyon Caiado Rocha Lima^{**} Alexis Maka^{***} Paloma Alves^{****}

Abstract

This paper analyzes the impacts of monetary policy, exchange rate, demand, and supply exogenous disturbances on the Brazilian economy using a structural vector autoregression model identified by two alternative methodologies. The first uses sign restrictions on impulse responses based on an open-economy macroeconomic model. The second (hybrid) is a new methodology that combines the first one with restrictions on the contemporaneous causal interrelationships among variables, derived by directed acyclic graphs.

A comparison of the results shows that while the effects of exchange rate shocks are nearly the same, the effects of monetary policy shocks depend on the methodology adopted. There is a strong response of the exchange rate to demand shocks and to shocks originating in the foreign exchange market. Exchange rate shocks have an important role in explaining short-run fluctuations of prices and output. We conclude that the exchange rate is an independent source of shocks and a shock absorber.

 $Keywords\colon$ Structural VAR, Hybrid Identification, Directed Acyclic Graphs, Sign Restrictions.

JEL Codes: C32, E31, E32, F31.

Brazilian Review of Econometrics v. 31, n° 1, pp. 97–136 May 2011

^{*}Submitted in April 2011. Revised in July 2011. Financial support from PNPD is gratefully acknowledged. The views expressed in this paper are those of the authors and do not necessarily represent those of the IPEA or of the Brazilian Ministry of Strategic Affairs.

^{**}Instituto de Pesquisa Econômica Aplicada (IPEA) e Universidade Estadual do Rio de Janeiro (UERJ), Rio de Janeiro, Brazil. E-mail: elcyon.rocha-lima@ipea.gov.br

^{****}Instituto de Pesquisa Econômica Aplicada (IPEA). E-mail: alexis.maka@ipea.gov.br

^{*****}Instituto de Pesquisa Econômica Aplicada (IPEA). E-mail: paloma.palmieri@ipea.gov.br

1. Introduction

Uncovering some stylized facts about short-run fluctuations of the Brazilian economy after the Real Plan, Céspedes et al. (2008) [CLM] find that not only do unanticipated movements of the exchange rate have large inflationary effect, but also large real effects. This suggests that the exchange rate is an important component in any model aimed at explaining the behavior of the Brazilian economy after the Real Plan. CLM's analysis is concerned mainly with the identification of the effects of monetary policy shocks; therefore, not much attention was given to the identification of exchange rate shocks or of any other shock.

In this article, we extend CLM's analysis by identifying several (monetary policy, exchange rate, demand, and supply) exogenous disturbances on the Brazilian economy using a structural vector autoregression (SVAR) model identified by two alternative methodologies. The first methodology uses sign restrictions on impulse responses of the shocks based on short-run dynamics of a stochastic open-economy macroeconomic model. The second methodology (hybrid) is a new methodology developed by us, which combines sign restrictions with restrictions on the contemporaneous causal interrelationships among variables, derived by directed acyclic graphs (DAGs). The hybrid identification strategy pursued in this article consists of two steps. In the first step, we use DAGs to select overidentifying restrictions on the contemporaneous coefficients based on the conditional independence relations between the variables. These overidentifying restrictions allow us to identify a monetary policy shock and to restrict the covariance matrix of the reduced-form residuals. In the second step, by maintaining the restriction on the covariance matrix of reduced-form residuals, we keep the identified monetary policy shock and impose sign restrictions on the impulse response functions of the other three shocks to identify the demand, supply, and exchange rate shocks.

The algorithms we employ to identify the exogenous shocks are based on the algorithm developed by Rubio-Ramírez et al. (2007), which is currently the most efficient algorithm when several independent exogenous shocks are identified by sign restrictions on impulse responses. We estimate our SVARs using the Bayesian method proposed by Sims and Zha (1998) and Waggoner and Zha (2003), which allows for consistent estimation of overidentified models, present in the hybrid methodology.

A comparison of the results of the two identification approaches shows that while the effects of exchange rate shocks are nearly the same, the effects of monetary policy shocks depend on the methodology adopted. We find a higher contribution of monetary policy shocks to output, prices, and exchange rate fluctuations when using sign restrictions only. There is a strong response of the exchange rate to demand shocks and to shocks originating in the foreign exchange market. Exchange rate shocks have an important role in explaining short-run fluctuations of prices and output. We conclude that the exchange rate is an independent source of shocks and a shock absorber.

The article is organized as follows. Section 2 presents a brief discussion of the identification methodologies that we combine in our hybrid approach. Section 3 explains the methodology used to identify and estimate the VARs. Section 4 describes the empirical model. Section 5 presents the hybrid identification procedure that combines short-run restrictions on the contemporaneous coefficients with sign restrictions on the impulse response functions. Section 6 shows an alternative identification procedure based on sign restrictions only. Finally, Section 7 offers some concluding remarks.

2. Identification of SVARs

Identifying restrictions are necessary in order to give a meaningful interpretation to residuals in vector autoregression (VARs) models. Without such restrictions, impulse response functions typically do not trace out the effects of exogenous structural disturbances such as monetary policy or exchange rate shocks. Instead, they typically pick up the effects of a linear combination of these structural shocks. Typical restrictions employed in the literature include constraints on the short-run or long-run impact of certain shocks on variables or informational delays (e.g., inflation is not contemporaneously observed by Central Banks when deciding interest rates).

The identification of structural shocks is, in general, a highly controversial enterprise because, by imposing different identifying assumptions, researchers may reach different conclusions about interesting economic questions (e.g., the sources of business cycle fluctuations). Criticisms about the nature of the identification process have repeatedly appeared in the literature. For example, Cooley and LeRoy (1985) criticize Cholesky decompositions because contemporaneous recursive structures are hard to obtain in general equilibrium models. Faust and Leeper (1997) argue that long-run restrictions are unsatisfactory as they may exclude structures that generate perfectly reasonable short-run dynamics but fail to satisfy long-run constraints by infinitesimal amounts. Cooley and Dwyer (1998) indicate that long-run restrictions may also incompletely disentangle permanent and transitory disturbances. Canova and Pina (2005) show that standard dynamic stochastic general equilibrium (DSGE) models almost never provide the zero restrictions employed to identify monetary disturbances in structural VAR systems and that misspecification of the features of the underlying economy can be substantial.

Following the growing interest in graphical models and, in particular, in those based on DAGs as a general framework to describe and infer causal relations, several authors have applied this methodology to identify SVARs. Swanson and Granger (1997) were the first to apply graphical models to identify contemporaneous causal order of a SVAR, followed by Bessler and Lee (2002), Demiralp and Hoover (2003) and Céspedes et al. (2008) [CLM]. The last three articles adopt a procedure that uses statistical properties of the sample – more specifically,

Brazilian Review of Econometrics 31(1) May 2011

conditional independence relations between the variables to select overidentifying restrictions to estimate structural VARs. These restrictions follow from DAGs estimated by the TETRAD software developed by Spirtes et al. (2000) using the covariance of reduced-form VAR disturbances as input. However, the use of DAGs for making causal inferences is subject to an important caveat: as Robins et al. (2003) have shown, causal procedures based on associations of non-experimental data under weak conditions are not uniformly consistent. That means that for any finite sample, there are no guarantees that the results of the causality tests will converge to the asymptotic (correct) results. Zhang (2002) and Zhang and Spirtes (2003) showed that, under the hypothesis that small partial correlations among variables indicate small direct causal effects, it is possible to guarantee convergence to the asymptotic correct results.¹

Recently, a new identification approach emerged based on the procedures of Faust (1998), Canova and De Nicoló (2002), and Uhlig (2005), where identification is achieved by restricting the sign (and/or shape) of structural responses. Such sign restrictions are attractive for several reasons. First, while (log)-linearized versions of DSGE models rarely deliver the m(m-1)/2 set of zero restrictions needed to recover m structural shocks, they contain a large number of sign restrictions usable for identification purposes. Second, sign restrictions make explicit restrictions that are often used implicitly by researchers when identifying VARs. Third, they can be robust in the sense that they hold across several structural models or parameterizations of the same model. However, these advantages come at a cost: identification of the structural shocks is not exact. There are multiple matrices defining the linear mapping from orthogonal structural shocks to VAR residuals. All of these matrices satisfy the sign restrictions and imply the same reducedform covariance matrix of VAR residuals. In other words, they are observationally equivalent and equally consistent with the economic theory imposed with sign restrictions. Paustian (2007) evaluates the sign restrictions method based on two DSGE models and concludes that sign restrictions can be a useful tool to recover structural shocks from VAR residuals. However, two conditions must be met for the method to unambiguously deliver the correct sign of unconstrained impulse responses. First, a sufficiently large number of restrictions must be imposed; more than what is typically employed in applied work. Second, the variance of the shock under study must be sufficiently large.

In an effort to overcome the limitations of the available methodologies, in this article we develop a new identification methodology that combines sign restrictions with restrictions on the contemporaneous causal interrelationships among variables, derived by DAGs.² The motivation for this hybrid strategy comes from

 $^{^1 \}rm Throughout$ this article, we assume that small partial correlations indicate small direct causal effects.

 $^{^2 \}rm Dungey$ and Fry (2009) propose a different hybrid identification that combines sign restrictions, cointegration and traditional exclusion restrictions within a system which explicitly models

the fact that the DAG and sign restriction approaches complement each other, so that their combination may be superior to each methodology taken isolatedly. While the DAG approach imposes restrictions that may identify exogenous shocks, the response of variables to these shocks may indicate that they are not the ones we are trying to identify. They may be linear combinations of the shocks we are interested in or parameter uncertainty may be responsible for the distortions in the responses. On the other hand, sign restrictions have economic justification, but may not impose enough restrictions to identify the shocks (as described in the previous paragraph). We believe that a combination of the available methodologies increases the chance that all shocks of interest are identified. The next section describes this new methodology.

3. Methodology³

Let y_t be the data vector – there are six variables in the model, therefore y_t has dimension $n \times 1(n = 6)$ for each period t:

$$y_t = [y_{1t} \ y_{2t} \ \cdots \ y_{nt}]'$$

where:

 $y_{1t} = \log(\text{Gross annualized Selic interest rate}),$

 $y_{2t} = \log(\text{Nominal exchange rate}(R\$/US\$)),$

 $y_{3t} = \log(\text{IPCA index}),$

 $y_{4t} = \log(180 \text{-day swap rate (PRE } \times \text{CDI} - \text{annualized considering 252 working days)}),$

 $y_{5t} = \log(\text{Industrial Production Index}), \text{ and}$

$$y_{6t} = \log(\mathrm{M1}).$$

The structural VAR model has the general form:

$$y'_t A' = \sum_{t=1}^p y'_{t-l} A'_l + z'_t D' + \varepsilon'_t, \quad \text{for} \quad t = 1, \cdots, T$$
 (1)

where:

 y_t is an $n \times 1$ column vector of endogenous variables at time t,

A and A_t are $n \times n$ parameter matrices;

Brazilian Review of Econometrics 31(1) May 2011

stationary and non-stationary variables and accounts for both permanent and temporary shocks. ³The methodology developed here builds on Rubio-Ramírez et al. (2007), Sims and Zha (1998), and Waggoner and Zha (2003).

D is an $n \times h$ parameter matrix,

 z_t is an $h \times 1$ column vector of seasonal dummies and constant term at time t,

 ε_t is an $n \times 1$ column vector of structural disturbances at time t;

p is the lag length, and

T is the sample size (p = 6 and T = 113).

The parameters of individual equations in (1) correspond to the columns of A', A'_l and D'.

The structural disturbances have a Gaussian distribution with $E(\varepsilon_t|y_1, \ldots, y_{t-1}, z_1, \ldots, z_T) = 0_{n \times 1}$ and the variance-covariance matrix of the structural disturbances is normalized to be an identity matrix, i.e. $E(\varepsilon_t \varepsilon'_t | y_1, \ldots, y_{t-1}, z_1, \ldots, z_T) = I_{n \times n}$.⁴ Right multiplying the structural form (1) by $(A')^{-1}$, we will obtain the usual representation of a reduced-form VAR with the reduced-form variance-covariance matrix being $\Omega = (A'A)^{-1}$.

Unlike typical unrestricted VAR models, Ω will be restricted when the contemporaneous parameter matrix A is overidentified.

The structural VAR models (1) can be rewritten in the compact form:

$$y_t'A' = x_t'F' + \varepsilon_t'$$

where

$$x'_{t_{1\times k}} = \begin{bmatrix} y'_{t-1}\cdots y'_{t-p}z'_t \end{bmatrix}, \qquad F_{n\times k} = \begin{bmatrix} A_1\cdots A_pD \end{bmatrix}$$

and k = np + h. We will refer to F' as lagged parameters even though F' may also contain exogenous parameters.

For $1 \leq i \leq n$ let a_i be the *i*th column of A', let f_i be the *i*th column of F' and let Q_i be an $n \times n$ matrix of rank q_i . The linear restrictions of interest can be summarized as follows:

$$Q_i a_i = 0, \qquad i = 1, \cdots, n \tag{2}$$

The restrictions given by (2) are said to be non-degenerate if there exists at least one non-singular matrix A' satisfying them. In this paper, all restrictions are assumed to be non-degenerate.

The VAR model is fitted with six lags, despite the small sample size and the fact that the Schwarz and Hannan-Quinn information criteria suggest a smaller

⁴There are two alternative normalizations of a structural VAR: assuming that matrix A has 1's in its diagonal and that the covariance matrix of structural residuals is diagonal or assuming that A has positive values in its diagonal and that the covariance matrix of structural residuals is equal to the identity matrix. The latter normalization is adopted here and is also employed, for example, by Blanchard and Quah (1989), and Faust (1998).

lag length (less than or equal to 3). When adopting the Litterman prior, it is recommended that the number of lags be at least superior to the one prescribed by classical procedures. When VAR models are large and degrees of freedom are low, the likelihood function itself can be ill-behaved and there is the wellknown tendency of estimates to become unreliable. To deal with these problems, Litterman (1984) introduces a widely used Bayesian prior distribution for reducedform models to down-weight models with large coefficients on distant lags and explosive dynamics. Sims and Zha (1998) incorporate Litterman's idea in the structural framework by specifying the prior distribution of a_i and f_i as

$$a_i \sim N\left(0, \bar{S}_i\right)$$
 and $f_i | a_i \sim N\left(\bar{P}_i a_i, \bar{H}_i\right)$ (3)

where \bar{H}_i is defined as a $k \times k$ diagonal, symmetric and positive definite (SPD) matrix:

$$\bar{H}_{i} = \begin{bmatrix} \frac{\lambda_{0}\lambda_{1}}{\sigma_{i}} & 0 & \cdots & 0 \\ 0 & \frac{\lambda_{0}\lambda_{1}}{\sigma_{i}} & 0 & \vdots & 0_{(36\times12)} \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \frac{\lambda_{0}\lambda_{1}}{\sigma_{i}6^{\lambda_{3}}} \\ & & & \frac{\lambda_{0}\lambda_{4}}{\sigma_{i}} & 0 & \cdots & 0 \\ 0_{(12\times36)} & 0 & \frac{\lambda_{0}\lambda_{4}}{\sigma_{i}} & 0 & \vdots \\ & & & \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \frac{\lambda_{0}\lambda_{4}}{\sigma_{i}} \\ & & & & (12\times12) \end{bmatrix}_{(48\times48)}$$

The standard deviation of the conditional prior of f_i (subset of parameters of equation i) for the coefficient on lag l of the variable j, is given by

$$\frac{\lambda_0\lambda_1}{\sigma_i l^{\lambda_3}}$$

where: the hyperparameter λ_0 controls the tightness of beliefs in A'; λ_1 controls what Litterman called overall tightness of beliefs around the random walk prior; λ_3 controls the rate at which prior variance shrinks for increasing lag length; λ_4 is the tightness for the constant term and seasonal dummies, i.e., for the last 12 rows of each column of F'. We give it a conditional prior mean of zero and a standard deviation controlled by $\lambda_0 \lambda_4$.

The parameters $\sigma_1, \ldots, \sigma_n$ (one for each equation) are scale factors, allowing for the fact that the units of measurement or scale of variation may not be uniform

Brazilian Review of Econometrics 31(1) May 2011

across variables. The scale factors are taken as the sample standard deviations of residuals from univariate autoregressive models, with lag length p, fit to the individual series in the sample.

The diagonal matrix S_i is an $n \times n$ SPD matrix, the individual elements in the *i*th column of A' are assumed to be independent, with prior standard deviations set to $\lambda_0/\hat{\sigma}_i$ (parameters defined above):

$$\bar{S}_{i_{(n\times n)}} = \begin{bmatrix} \frac{\lambda_0}{\hat{\sigma}_1} & 0 & 0 & \cdots & 0 \\ 0 & \frac{\lambda_0}{\hat{\sigma}_2} & 0 & \cdots & \vdots \\ \vdots & 0 & \frac{\lambda_0}{\hat{\sigma}_3} & & \\ & \vdots & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \frac{\lambda_0}{\hat{\sigma}_n} \end{bmatrix}$$

We use the following values for the hyperparameters:

| Hyperparameter | Value |
|----------------|-------|
| λ_0 | 0.5 |
| λ_1 | 0.25 |
| λ_3 | 1 |
| λ_4 | 0.5 |

 \bar{P}_i is a $k \times n$ matrix defined as:

$$\bar{P}_i = \left[\begin{array}{c} I_{(6\times 6)} \\ 0_{(42\times 6)} \end{array} \right]$$

The prior form summarized above represents a class of existing Bayesian priors that have been widely used for structural VAR models. Combining prior form (3) with restriction (2), we wish to obtain the functional form of the conditional prior distribution:

$$q\left(a_{i}, f_{i} | T_{i} a_{i} = 0\right) \tag{4}$$

In our case, the following matrices are the restricted A and A' matrices obtained by the application of the TETRAD software:

Brazilian Review of Econometrics 31(1) May 2011

$$A = \begin{bmatrix} A_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & A_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & A_{33} & 0 & A_{35} & 0 \\ A_{41} & A_{42} & 0 & A_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & A_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & A_{66} \end{bmatrix}$$
$$A' = \begin{bmatrix} A_{11} & 0 & 0 & A_{41} & 0 & 0 \\ 0 & A_{22} & 0 & A_{42} & 0 & 0 \\ 0 & 0 & A_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & A_{44} & 0 & 0 \\ 0 & 0 & A_{35} & 0 & A_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & A_{66} \end{bmatrix}$$

Then, we can obtain the T'_i 's matrices which satisfy the constraints for each column i of A':

$$T_{i_{(q_i \times n)}} a_{i_{(n \times 1)}} = 0_{(q_i \times 1)}$$

Each matrix T_i reproduces the restrictions present in column *i* of A', given by TETRAD. All element of T_i off the diagonal are zero. At the diagonal, there are zeros in the position of free parameters and ones in the position of parameters restricted to be equal to zero. Therefore, for example

$$T_{1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \text{ and } T_{6} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Let U_i be an $n \times q_i$ matrix whose columns form an orthonormal basis for the null space of T_i . Column a_i will satisfy restriction (2) if and only if there exists a $q_i \times 1$ vector b_i (q_i = number of free parameters in column *i* of matrix A') such that

$$a_i = U_i b_i \tag{5}$$

Column vector b_i contains the free parameters of column *i* of matrix A' given by TETRAD. For this matrix A' the U'_i s are given by,

Brazilian Review of Econometrics 31(1) May 2011

$$U' = \begin{bmatrix} U_1' \\ U_2' \\ U_3' \\ U_4' \\ U_5' \\ U_6' \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

For example,

$$a_{3} = \begin{bmatrix} 0\\0\\A_{33}\\0\\A_{35}\\0 \end{bmatrix} = U_{3}b_{3} = \begin{bmatrix} 0&0\\0&0\\1&0\\0&0\\0&1\\0&0 \end{bmatrix} \begin{bmatrix} A_{33}\\A_{35} \end{bmatrix}$$

The distributions of b_i and f_i are given by

$$b_i \sim N\left(0, \tilde{S}_i\right) \text{ and } f_i | b_i \sim N\left(\tilde{P}_i b_i, \tilde{H}_i\right)$$
 (6)

where

$$\tilde{H}_i = \bar{H}_i, \qquad \tilde{P}_i = \bar{P}_i U_i, \qquad \text{and} \qquad \tilde{S}_i = \left(U'_i \bar{S}_i^{-1} U_i\right)^{-1}$$

Note that \tilde{S}_i is a $q_i \times q_i$ SPD matrix, \tilde{H}_i is an $r_i \times r_i$ SPD matrix, and \tilde{P}_i is an $r_i \times q_i$ matrix. It can be verified that the prior distribution (6) for b_i is equivalent to prior distribution (4) for a_i . For the most part of this paper, we work directly with b_i with the understanding that the original parameters a_i can be easily recovered via linear transformations U_i .

Let $b = [b'_1 \dots b'_n]'$, $f = [f'_1 \dots f'_n]'$, $X = [x_1 \dots x_T]'$, and $Y = [y_1 \dots y_T]'$. The likelihood function for b and f(L((b, f) | X, Y)) is proportional to

Brazilian Review of Econometrics 31(1) May 2011

$$|\det [U_1b_1|\cdots|U_nb_n]|^T \exp\left(-\frac{1}{2}\sum_{i=1}^n b_i' U_i' Y' Y U_i b_i - 2f_i' X' Y U_i b_i + f_i' X' X f_i\right)$$
(7)

Combining the priors on b and f given by (6) with the likelihood function given by (7) leads to the following *joint posterior p.d.f.* function for b and f:

$$p(b_1,\cdots,b_n|X,Y)\prod_{i=1}^n p(f_i|b_i,X,Y)$$

where

$$p(b_1, \dots, b_n | X, Y) \propto |\det[U_1 b_1| \cdots | U_n b_n]|^T \exp\left(-\frac{T}{2} \sum_{i=1}^n b'_i S_i^{-1} b_i\right)$$
 (8)

$$p(f_i|b_i, X, Y) = \varphi(P_i b_i, H_i)$$
(9)

with

$$H_{i} = \left(X'X + \tilde{H}_{i}^{-1}\right)^{-1}$$
$$S_{i} = \left(\frac{1}{T}\left(U_{i}'Y'YU_{i} + \tilde{S}_{i}^{-1} + \tilde{P}_{i}'\tilde{H}_{i}^{-1}\tilde{P}_{i} - P_{i}'H_{i}^{-1}P_{i}\right)\right)^{-1}$$

Since (8) has an unknown distribution, we must take draws from the posterior distribution of b by Gibbs Sampling and, and given each draw of b, take draws of f from the Gaussian conditional distribution (9). The notation $\varphi(P_i b_i, H_i)$ in (9) denotes the Gaussian density with mean $P_i b_i$ and covariance matrix H_i .

In many works with VARs, only the likelihood function (i.e., proportional to the posterior density under a flat prior for b and f) is considered. Because (7) is the same as (8) and (9) when the prior variances (diagonal elements in \tilde{S}_i and \tilde{H}_i) approach infinity, the posterior density specified in (8) and (9) includes the likelihood as a special case.

To obtain small-sample inferences of b and f or for functions of them (e.g., impulse responses), it is necessary to simulate the joint posterior distribution of b and f. This simulation involves two consecutive steps. First, simulate draws of b from the marginal posterior distribution (8). Second, given each draw of b, simulate draws of f from the conditional posterior distribution (9). The second step is straightforward because it requires draws only from a multivariate normal distribution. The first step, as mentioned earlier, can be challenging when linear restrictions on A imply a restricted reduced-form covariance matrix.

Brazilian Review of Econometrics 31(1) May 2011

The following algorithm was designed to obtain a sample of the impulse response functions, which satisfy the sign restrictions.

Algorithm: The following steps compose the algorithm for simulating draws from the posterior distribution of b, f and, given these draws, draws of the impulse responses that satisfy the sign restrictions.

1. Get the values at the peak of the posterior density function.

i. Draw an independent standard normal $n \times n$ matrix \tilde{X} and let $\tilde{X} = \tilde{Q}\tilde{R}$ be the QR decomposition of \tilde{X} with the diagonal \tilde{R} normalized to be positive. j. Let $P = \tilde{Q}$ and generate the impulse responses $IRF^{(s)}$ from $A^{(s)}P$ and $B^{(s)}P = F^{(s)}A^{(s)-1}P$.

- k. If $IRF^{(s)}$ satisfies the sign restrictions keep it, otherwise discard it.
- l. If the number of accepted IRF is equal to 1000 stop.

5. Collect all the IRF that were not discarded in step 4.

In step 2 and 4 of the Algorithm, all simulations are carried out according Theorem 2 of Waggoner and Zha (2003). The central result of Theorem 2 states that drawing from the distribution of b_i conditional on $b_1, \ldots, b_{i-1}, b_n$ is equivalent to drawing from a multivariate Gaussian distribution and a special univariate distribution.

Brazilian Review of Econometrics 31(1) May 2011

For a fixed i_* , where $1 \leq i_* \leq n$. Let w be an non-zero $n \times 1$ vector perpendicular to each vector in $\{U_i b_i | i \neq i_*\}$. Since the restrictions are assumed to be non-degenerate, the n-1 vectors $U_i b_i$ for $i \neq i_*$ will almost surely be linearly independent and $U'_{i_*}w$ will be non-zero. Define $w_1 = T'_{i_*}U'_{i_*}w/||T'_{i_*}U'_{i_*}w||$, where T_{i_*} is a $q_{i_*} \times q_{i_*}$ matrix such that $T_{i_*}T'_{i_*} = S_{i_*}$, and choose w_2, \ldots, w_{q_*} so that $w_1, w_2, \ldots, w_{q_*}$ form an orthonormal basis for R^{q_*} . Then the random vector b_i conditional on $b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n$ can be represented as

$$b_i = \beta_1 U_i' T_i^{-1} w_1 + \sum_{j=2}^{q_i} \beta_j U_i' T_i^{-1} w_j$$

The random variable β_j , for $2 \leq j \leq q_i$, is normally distributed with mean zero and variance 1/T and is straightforward to simulate. The density function for β_1 , the special univariate distribution, is proportional to $|\beta_1|^T \exp(-T\beta_1^2/2)$. Waggoner and Zha (2003) show how to simulate from this latter distribution.

i. Hybrid Identification⁵

Suppose we want to keep the identification of the first shock obtained by TETRAD (the monetary policy shock). Then we have to modify matrix P employed in step 4-j of the previous algorithm. It will take the hybrid form:

$$P = \tilde{Q} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \tilde{Q}_{22} & \tilde{Q}_{32} & \cdots & \tilde{Q}_{62} \\ 0 & \tilde{Q}_{23} & \tilde{Q}_{33} & \cdots & \tilde{Q}_{63} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \tilde{Q}_{26} & \tilde{Q}_{36} & \cdots & \tilde{Q}_{66} \end{bmatrix}$$

where the submatrix,

$$Q_s = \begin{bmatrix} \tilde{Q}_{22} & \tilde{Q}_{32} & \cdots & \tilde{Q}_{62} \\ \tilde{Q}_{23} & \tilde{Q}_{33} & \cdots & \tilde{Q}_{63} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{Q}_{26} & \tilde{Q}_{36} & \cdots & \tilde{Q}_{66} \end{bmatrix}$$

is obtained by a draw of an independent standard normal $(n-1) \times (n-1)$ matrix \tilde{X} , and Q_s is obtained by the QR decomposition of $\tilde{X}(\tilde{X} = Q_s \tilde{R})$, with the diagonal \tilde{R} normalized to be positive).

Brazilian Review of Econometrics 31(1) May 2011

 $^{^{5}}$ The hybrid methodology adopted here shares some similarities with the one used by Dungey and Fry (2009). However, they use "Givens rotation" while we use the QR decomposition. Furthermore, unlike them, we use DAGs to impose restrictions on the contemporaneous causal interrelationships among variables.

4. Empirical Model

The model is estimated using monthly data and it is composed of the following variables: the short-term interest rate (SELIC), the nominal exchange rate (e), the price index (IPCA), the medium-term interest rate (SWAP), output (y), and a monetary aggregate (M1), a constant, and seasonal dummies.⁶ Following the results of Sims and Uhlig (1991) and Sims et al. (1990), we do not perform unit root tests or cointegration analysis.⁷ Based on the exchange rate regime and monetary policy operational procedures, we decided to start our sample period in 1999:03, going up to 2008:07. This is a period characterized by a free-floating exchange rate and explicit SELIC targeting. The lag length chosen is six months. The model identifies four independent sources of exogenous disturbances: monetary policy, demand, supply, and exchange rate shocks.

5. Model Identification: A Hybrid Approach

The hybrid identification strategy pursued in this article consisted of two steps. In the first step, we use DAGs to select overidentifying restrictions on the contemporaneous coefficients based on the conditional independence relations between the variables. These overidentifying restrictions allow us to identify a monetary policy shock and to restrict the covariance matrix of the reduced-form residuals. In the second step, by maintaining the restriction on the covariance matrix of reduced-form residuals, we keep the identified monetary policy shock and impose sign restrictions on the impulse response functions of other three shocks to identify the demand, supply, and exchange rate shocks.

Step 1: Selection of the Overidentifying Restrictions to Identify Monetary Policy Shocks 8

Spirtes et al. (2000) [SGS] developed algorithms for inferring causal relations from data that are embodied in a computer program used in this article, called TETRAD.⁹ The program assumes a multivariate normal distribution and takes the covariance matrix of the variables of the model as input,¹⁰ converting it into a correlation matrix and performing hypothesis tests on which the null hypothesis is a zero partial correlation.

 $^{^{6}\}mathrm{A}$ detailed description of the data and their sources can be found in Appendix I.

 $^{^{7}}$ They show that the classical unit root asymptotic is of little practical value and that the common practice of attempting to transform models into a stationary form by difference or cointegration operators, whenever it appears likely that the data are integrated, is, in many cases, unnecessary.

⁸For an introduction on how to use DAGs to identify VARs, see CLM.

 $^{^9{\}rm The}$ program is available for download at www.phil.cmu.edu/projects/tetrad/index.html. We used TETRAD III in this paper.

 $^{^{10}\}mathrm{In}$ our application, the input of TETRAD is the covariance matrix of the reduced-form VAR residuals.

Conditional independence is a key notion in multivariate analyses such as graphical modeling, where two vertices are connected if and only if the corresponding variables are not conditionally independent. To confirm the conditional independence, it is a common practice to check whether the partial correlation is close enough to zero. This is done because it is assumed that zero partial correlation suggests that the variables are conditionally independent, or nearly so. Under the assumption of multivariate normality, a test of zero correlation or zero partial correlation is also a test of independence or conditional independence. Moreover, if X, Y and Z are normally distributed, the partial correlation coefficient $\rho_{XY,Z}$ is zero if and only if X is independent of Y conditional on Z.

TETRAD begins with a 'saturated' causal graph, where any pair of nodes (variables) is joined by an undirected edge.¹¹ If the null hypothesis of zero partial correlation cannot be rejected – at, say, the 5% level, using Fisher's z test – the edge is deleted.¹² After examining all pairs of vertices, TETRAD moves on to triples, and so forth, orienting the edges left in the graph through the connection between probabilistic independence and graph theory. The final output of TETRAD is a set of observationally equivalent DAGs containing the proposed causal structure(s) of the model.

Robins et al. (2003) showed that the asymptotically consistent procedures of SGS are pointwise consistent, but not uniformly consistent.¹³ Furthermore, they also showed that there exists no causality test, based on associations of nonexperimental data under the conditions assumed by SGS, which is uniformly consistent. Therefore, for any finite sample, it is impossible to guarantee that the results of the SGS causality tests (or any other causality test) will converge to the asymptotic results.

Under the SGS model, it is sufficient to have a sample covariance between two variables, say, v_1 and v_2 , exactly equal to zero to deduce that v_1 is not a cause of v_2 . However, if the sample correlation between v_1 and v_2 is not exactly zero (as will almost always happen in finite samples) and the true model is unknown, as Robins et al. (2003) have shown, the acceptance or rejection of the null hypothesis of zero partial correlation is not unequivocally tied to the absence of causality. In other words, we do not know, in any finite sample, how close to zero a partial correlation has to be to indicate non-causality. When the sample correlation is

Brazilian Review of Econometrics 31(1) May 2011

¹¹An edge in a graph can be either directed (marked by a single arrowhead on the edge) or undirected (unmarked). Arrows represent causal relationships: if there is an arrow pointing from X_i to X_j it means that X_i has a direct causal effect on X_j .

 $^{^{12}}$ In the case of the normal distribution, the partial correlation coincides with the conditional correlation, which is another measure of conditional independence of two random variables. See Baba et al. (2004) for further details.

¹³A pointwise consistent test is guaranteed to avoid incorrect decision if the sample size can be increased indefinitely. However, pointwise consistency is only a guarantee for what happens in the limit, not in any finite sample size. A stronger form of consistency, uniform consistency, guarantees that it is possible to bound the decisions error rates to a finite number of observations.

not exactly zero, it is not possible to determine which significance level should be used to test for zero partial correlation when attempting to test for the presence of causality. The "significance level," used by TETRAD, cannot be interpreted as the probability of type I error for the pattern output, but merely as a search parameter. The higher this search parameter, the smaller the absolute value of the partial correlation that is taken as an indication of absence of causality. Intuitively, we are assuming that small partial correlations indicate small direct causal effect, but we do not know how small the absolute value of the correlation has to be for us to obtain the correct causal inferences for the sample data we are using. Nevertheless, we can test the sensitivity of the impulse response function of the model to different discrete values of this "parameter."¹⁴

Applying the TETRAD software at the 20% significance level, we obtain a graphical representation of the DAG containing the contemporaneous causal ordering of the variables, displayed in Figure 1.¹⁵ According to Figure 1, the SELIC and the exchange rate affect contemporaneously the SWAP rate, while the price index responds to output shocks within the current period. None of the variables affects contemporaneously the SELIC rate, even the price level and the output. The fact that output and prices have no contemporaneous effect on the SELIC rate may be associated with the difficulty in obtaining information on the current level of output and price level at the time policymakers have to make their decisions.¹⁶



It is interesting to compare Figure 1 with the patterns obtained by CLM in their alternative model that includes money (see Figures 20-22 of their article). The differences between our results and CLM's can be explained by three factors. First

112

 $^{^{14}}$ This is a bit more data-oriented than the usual procedure of changing the order of the Cholesky decomposition of reduced-form VAR residuals to identify the model.

¹⁵TETRAD suggests that the contemporaneous causality can go either from IPCA to output or from output to IPCA. In what follows, we chose the causality from output to IPCA, but the adoption of the alternative direction does not change the results.

¹⁶This is an identification assumption made, for example, by Sims and Zha (2006).

of all, our sample period is longer than that of CLM (CLM ends in 2004:12, while ours goes up to 2008:07). Second, the lag lengths are different (CLM analyze 1-3 lags, while we use six lags). Third, CLM employ a classical estimation procedure while ours is a Bayesian one.

The causal ordering between the variables of the VAR can be represented by matrix A, which establishes a relationship between reduced-form and structural form residuals. The DAG pictured in Figure 1 can be represented by the following matrix:

$$A = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & A_{35} & 0 \\ A_{41} & A_{42} & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

where A_{ij} are parameters to be estimated and the vector of endogenous variables that multiplies A is given by [SELIC, exchange rate, IPCA, SWAP, output, M1].

The contemporaneous causal ordering resulting from the application of DAGs implies restrictions on the covariance matrix of the reduced-form residuals, meaning that we now have an overidentified model. Structural VAR models that are overidentified can be consistently estimated only by Bayesian estimation methods that introduce these restrictions on the covariance matrix of reduced-form residuals. These restrictions are considered when Bayesian estimation methods are applied to the parameters of a structural VAR (and not to the parameters of a reduced-form VAR). The method developed by Sims and Zha (1998), and adopted in this article, is one of these methods.

Using the contemporaneous causal ordering of Figure 1 to identify the SVAR, we obtained the impulse response functions of economic variables to exogenous and independent shocks, displayed in Figure 2. We identify SELIC shocks as monetary policy shocks, leaving the exchange rate shocks to be identified by sign restrictions in the next step, when we also identify demand and supply shocks in order to better identify exchange rate disturbances.¹⁷

According to Figure 2, after a positive SELIC innovation the stock of M1 falls and output decreases temporarily, taking near 15 months to recover. The direction of the exchange rate response is not clear, but it is more likely that it will appreciate. The price level goes down, but it takes near six months until the price level starts to fall despite the contraction of economic activity. The price level temporarily increases in response to a positive SELIC shock, a result known in the literature as the "price puzzle," since it is at odds with the prediction of *most* theoretical models that restrictive monetary policy should reduce the price

 $^{^{17}}$ With contemporaneous restrictions only, it is not clear how one should identify demand and/or supply shocks, which may affect the identification of exchange rate shocks.

Brazilian Review of Econometrics 31(1) May 2011

level.¹⁸ Some authors [e.g., Sims (1998), Christiano et al. (1999)] have argued that the presence of the price puzzle should serve as an informal specification test of a VAR model: if such anomalous result is observed, then what one has labeled "monetary policy" probably has not been correctly identified. Sims (1992) suggested that the price puzzle might emerge in models that do not specify correctly the information available to the central bank. If policymakers observe variables that forecast inflation, but these variables are not included in the model, there will be apparently unpredictable changes in interest rates that are actually systematic responses to expected inflation. However, Barth and Ramey (2001) argue that a temporary price puzzle may not be a puzzle at all once one takes into account the possibility that the monetary transmission mechanism itself has cost effects. Prices should go up in the short run following an unanticipated monetary contraction if the cost effects of the monetary transmission mechanism dominate the demand effects.

Step 2: Imposing Sign Restrictions to Identify Demand, Supply and Exchange Rate Shocks

Having identified monetary policy shocks and restricted the covariance matrix of the reduced-form residuals using the contemporaneous causal order suggested by TETRAD, now we impose sign restrictions on the remaining impulse response functions in order to identify the demand, supply, and exchange rate shocks. We impose the sign restrictions for a four-month window. The sign restrictions used to identify the SVAR model are similar to those employed by Farrant and Peersman (2006) and can be justified by the short-run dynamics of a stochastic open-economy macroeconomic model, like the one presented in Appendix II.¹⁹ Table 1 summarizes the sign restrictions on the IRFs used to identify the demand, supply, and exchange rate shocks.²⁰

According to Table 1, positive demand shocks do not decrease the SELIC rate, the price level, output, and the stock of money, and do not imply a depreciation of the real exchange rate. A positive supply shock implies that prices do not increase and output does not go down. An unexpected depreciation of the nominal exchange rate is supposed to imply changes in the same direction of the real

¹⁸The temporary price puzzle appears as a common feature in several alternative identification schemes tested. With respect to the Cholesky decomposition of residuals, the following ordering eliminates the price puzzle: SWAP, SELIC, IPCA, output, M1, exchange rate. In Section 5, where we identify all shocks by imposing sign restrictions on the impulse response functions, we assume that prices do not increase in response to monetary policy shocks, meaning that the price puzzle is eliminated by construction.

¹⁹When imposing sign restrictions, we assume that $\frac{\partial q_t}{\partial \varepsilon_t^d} \leq 0$ and $\frac{\partial i}{\partial \varepsilon_t^d} \geq 0$, which is consistent with the conventional view about the effects of demand shocks.

²⁰The (log) real exchange rate is defined as $q_t = s_t + p_t^* - p_t$, where s_t is the (log of) nominal exchange rate, $p_t(p_t^*)$ is the (log of) domestic (foreign) price level. We assume that the foreign price level is constant, so that a restriction on the real exchange rate translates into a restriction on $s_t - p_t$.



exchange rate, and that the short-term interest rate, prices, output,²¹ and the stock of money do not go down after the exchange rate shock. The sign restrictions are supposed to hold for four months.

| | SELIC | IPCA | Output | M1 | Real Exchange |
|---------------|----------|----------|----------|----------|---------------|
| | | | | | Rate |
| Demand Shock | ≥ 0 | ≥ 0 | ≥ 0 | ≥ 0 | ≤ 0 |
| Supply Shock | | ≤ 0 | ≥ 0 | | |
| Exchange Rate | ≥ 0 |
| Shock | | | | | |

 Table 1

 Sign Restrictions Used to Identify the SVAR Model

The IRFs that result from the imposition of sign restrictions are presented in Figures 3-4, showing the median as well as the 68% probability bands for a horizon of 24 and 60 months following the shocks, respectively.

The impact of monetary policy shocks on the exchange rate is not significant; however, there is a strong response of the exchange rate to demand shocks and to shocks originating in the foreign exchange market, as can also be seen in the variance decomposition presented in Figure 5.²² The fact that the exchange rate moves in response to demand shocks suggests that the exchange rate acts as a shock absorber in order to restore the equilibrium of the economy. The fact that the exchange rate moves in response to shocks originating in the foreign exchange market, together with the finding that exchange rate shocks have an important role in explaining short-run fluctuations on prices and output, allows us to conclude that the exchange rate is an independent source of shocks, in addition to being a shock absorber.

With the hybrid identification, in response to an exchange rate shock, there is an initial expansionary effect on output during the first five months, followed by a

 22 While impulse response functions trace the effects of an exogenous shock to one endogenous variable on all variables of the VAR, variance decomposition separates the variation in an endogenous variable into the component shocks to the VAR. Thus, the variance decomposition provides information about the relative importance of each random innovation for the variables in the VAR.

²¹The nonnegative response of output to the exchange rate shock adopted is based on the stochastic Mundell-Fleming model presented in Appendix II. Alternative approaches have been followed in the literature to analyze whether, contrary to the conventional view, devaluations are contractionary. A well-known study by Edwards (1986) uses data on 12 developing countries for 1965-1980 and estimates a model of real output behavior. His results indicate that devaluations generate a small contractionary effect in the first year. In the second year, this effect is completely reversed. In the long run, devaluations are neutral. However, Edwards' approach is unable to distinguish between anticipated and unanticipated effects of exchange rate shocks. It also does not identify exogenous shocks. Therefore, it cannot be used to identify exchange rate unanticipated exogenous shocks.

contractionary period of near six months (Figure 3). However, with the identification by short-run restrictions coming out from the use of DAGs (Figure 2), in the first five months, the initial impact of exchange rate shocks on output is not clear, followed by a contraction of output that lasts for near 18 months. Additionally, the resulting currency depreciation is smaller under the hybrid identification.

The identification of structural shocks by sign restrictions is not exact, meaning that there are multiple models satisfying the sign restrictions. Therefore, the IRFs to demand, supply, and exchange rate shocks, exhibited in Figures 3-4, display not only sampling (parameter) uncertainty, but also model uncertainty. Figures 5-6 disentangle model uncertainty (the bold line) from parameter uncertainty (the distance between the dashed and the bold lines).²³ In doing so, when we simulate only models (using the Gibbs sampling procedure described in Section 3, but skipping items a-h), we consider the values of parameters at the peak of the posterior density function as "true." When simulating only parameters, we fix the identification given by TETRAD and simulate the parameters using Gibbs sampling procedure described in Section 3 (skipping item i).

6. An Alternative Identification Strategy Using only Sign Restrictions to Identify All Shocks

We now consider an alternative identification where we impose sign restrictions on the IRFs to all shocks, including the monetary policy shock. We maintain the previous restrictions summarized in Table 1 and, in addition, we assume that in response to a "contractionary" monetary policy shock, interest rates do not fall, and that output, prices, the stock of money, and the real exchange rate do not increase. These additional restrictions follow from the open-economy macroeconomic model presented in Appendix II.²⁴ Table 2 shows the sign restrictions on the IRFs used to identify the monetary policy, demand, supply, and exchange rate shocks. We impose the sign restrictions for a four-month window.

The IRFs based on the alternative identification that uses only sign restrictions to identify all shocks are presented in Figures 8-9, showing the median as well as the 68% probability bands for a horizon of 24 and 60 months following the shocks, respectively. The variance decomposition is presented in Figure 10. The main difference with respect to the IRFs and variance decomposition based on the hybrid identification relies on the effects of monetary policy shocks. Now, monetary policy disturbances have an important role as a source of short-run fluctuations of output, prices, and the exchange rate. Monetary policy shocks are still neutral in the long run, but now they have a long-run effect on the price level,

Brazilian Review of Econometrics 31(1) May 2011

 $^{^{23}{\}rm In}$ the case of monetary policy shocks, identified by short-run restrictions, there is no model uncertainty, only parameter uncertainty.

²⁴Here we consider that innovations to the SELIC rate are monetary policy shocks, whereas the model in the Appendix assumes that monetary policy shocks are represented by money supply innovations. Just bear in mind that money and interest rates are negatively related.



118





Figure 5 Variance Decomposition, with 68% probability bands



ŝ

ŝ 0

ŝ 0

ŝ

Ĺ

dAWS

ť

0

_-5 × 10⁻³

20

15

9

ŝ

x 10⁻³

20

15

9

ŝ

x 10⁻³

20

15

9

x 10⁻³

20

15

9

Ь

15

ŝ

t

Ŧ

A

Ħ

TU9TUO

ംറ്

ŝ

ĥ

ŝ

-5 -5 -5

15 5 15 15



Brazilian Review of Econometrics 31(1) May 2011

121

-0.02

20

15

9

ŝ

20

15

9

ŝ

20

15

9

20

15

9

ŝ

1

L 1

Simulating only parameters Simulating only models Simulating parameters and models

-0.02

-0.02

-0.02

0.02

20

15

9

0.02

20

15

9

ŝ

20

15

9

0.02

20

15

9 Ħ

ഹ

0

١W

ĥ

0 2 J 3

0.02

c

•



122

Monetary Policy and Exchange Rate Shocks in Brazil: Sign Restrictions

| | SELIC | IPCA | Output | M1 | Real Exchange |
|---------------|----------|----------|----------|----------|---------------|
| | | | | | Rate |
| Monetary | ≥ 0 | ≤ 0 | ≤ 0 | ≤ 0 | ≤ 0 |
| Police Shock | | | | | |
| Demand Shock | ≥ 0 | ≥ 0 | ≥ 0 | ≥ 0 | ≤ 0 |
| Supply Shock | | ≤ 0 | ≥ 0 | | |
| Exchange Rate | ≤ 0 | ≥ 0 | ≥ 0 | ≥ 0 | ≥ 0 |
| Shock | | | | | |

 Table 2

 Sign Restrictions Used to Identify the SVAR Model

and on the exchange rate. Figures 11-12 disentangle model uncertainty (the bold line) from parameter uncertainty (the distance between the dashed and the bold lines).

7. Concluding Remarks

In this article, we analyzed the impacts of several (monetary policy, exchange rate, demand, and supply) exogenous disturbances on the Brazilian economy using a structural vector autoregression (SVAR) model identified by two alternative methodologies. The first methodology used sign restrictions on impulse responses of the shocks based on short-run dynamics of a stochastic open-economy macroeconomic model. The second methodology (hybrid) is a new methodology developed by us, which combines sign restrictions with restrictions on the contemporaneous causal interrelationships among variables, derived by DAGs. The hybrid identification strategy pursued in this article consists of two steps. In the first step, we used DAGs to select overidentifying restrictions on the contemporaneous coefficients based on the conditional independence relations between the variables. These overidentifying restrictions allow us to identify a monetary policy shock and to restrict the covariance matrix of the reduced-form residuals. In the second step, by maintaining the restriction on the covariance matrix of reduced-form residuals, we kept the identified monetary policy shock and imposed sign restrictions on the impulse response functions of the other three shocks to identify the demand, supply, and exchange rate shocks.

The results of the hybrid identification show a delayed response of the price level to monetary policy shocks, consistent with the presence of some price rigidity. A comparison of the results of the two identification approaches shows that while the effects of exchange rate shocks are nearly the same, the effects of monetary policy shocks depend on the methodology adopted. We find a larger contribution of monetary policy shocks to output, prices, and exchange rate fluctuations when using sign restrictions only. There is a strong response of the exchange rate to

Brazilian Review of Econometrics 31(1) May 2011



Brazilian Review of Econometrics 31(1) May 2011









128

demand shocks and to shocks originating in the foreign exchange market. Exchange rate shocks have an important role in explaining short-run fluctuations of prices and output. We conclude that the exchange rate is an independent source of shocks and a shock absorber.

References

- Baba, K., Shibata, R., & Sibuya, M. (2004). Partial correlation and conditional correlation as measures of conditional independence. Australian & New Zealand Journal of Statistics, 46:657–664.
- Barth, M. & Ramey, V. (2001). The cost channel of monetary transmission. In *NBER Macroeconomics Annual*, pages 199–239. NBER.
- Bessler, D. & Lee, S. (2002). Money and prices: U.S. data 1869-1914 (A Study with Directed Graphs). *Empirical Economics*, 27:427–446.
- Blanchard, O. & Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review*, 79:655–673.
- Canova, F. & De Nicoló, G. (2002). Money matters for business cycle fluctuations in the G7. Journal of Monetary Economics, 49:1131–1159.
- Canova, F. & Pina, J. (2005). What VARs tell us about DSGE models? In Diebolt, C. & Krystou, C., editors, New Trends In Macroeconomics. Springer Verlag, New York, NY.
- Christiano, L., Eichenbaum, M., & Evans, C. (1999). Monetary policy shocks: What have we learned and to what end? In Taylor, J. & Woodford, M., editors, *Handbook of Macroconomics*, volume IA, pages 65–148. Elsevier.
- Cooley, T. & Dwyer, M. (1998). Business cycle analysis without much theory: A look at structural VARs. *Journal of Econometrics*, 83:57–88.
- Cooley, T. & LeRoy, S. (1985). Atheoretical macroeconometrics: A critique. Journal of Monetary Economics, 16:283–308.
- Céspedes, B., Lima, E., & Maka, A. (2008). Monetary policy, inflation, and the level of economic activity in Brazil after the Real Plan: Stylized facts from SVAR models. *Revista Brasileira de Economia*, 62:123–150.
- Demiralp, S. & Hoover, K. (2003). Searching for the causal structure of a vector autoregression. Oxford Bulletin of Economics and Statistics, 65 (supplement):745– 767.
- Dungey, M. & Fry, R. (2009). The identification of fiscal and monetary policy in a structural VAR. *Economic Modelling*. doi:10.1016/j.econmod.2009.05.001.

Brazilian Review of Econometrics 31(1) May 2011

- Edwards, S. (1986). Are devaluations contractionary? *Review of Economics and Statistics*, 68:501–508.
- Farrant, K. & Peersman, G. (2006). Is the exchange rate a shock absorber or a source of shocks? New empirical evidence. *Journal of Money, Credit, and Banking*, 38:939–961.
- Faust, J. (1998). On the robustness of identified VAR conclusions about money. Carnegie-Rochester Conference Series on Public Policy, 49:207–244.
- Faust, J. & Leeper, E. (1997). Do long run restrictions really identify anything? Journal of Business and Economic Statistics, 15:345–353.
- Litterman, R. (1984). The costs of intermediate targeting. Working Paper 254, Research Department, Federal Reserve Bank of Minneapolis.
- Paustian, M. (2007). Assessing sign restrictions. The B. E. Journal of Macroeconomics (Topics), 7.
- Robins, J., Sheines, R., Spirtes, P., & Wasserman, L. (2003). Uniform consistency in causal inference. *Biometrika*, 90:491–515.
- Rubio-Ramírez, J., Waggoner, D., & Zha, T. (2007). Structural vector autoregressions: Theory of identification and algorithms for inference. Mimeo.
- Sims, C. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review*, 36:975–1000.
- Sims, C. (1998). Comment on Glenn Rudebusch's "does measures of monetary policy in a VAR make sense?". *International Economic Review*, 39:933–941.
- Sims, C., Stock, J., & Watson, M. (1990). Inference in linear time series models with some unit roots. *Econometrica*, 58:113–144.
- Sims, C. & Uhlig, H. (1991). Understanding unit rooters: A helicopter tour. *Econometrica*, 59:1591–1599.
- Sims, C. & Zha, T. (2006). Does monetary policy generate recessions? Macroeconomic Dynamics, 10:231–272.
- Sims, C. A. & Zha, T. (1998). Bayesian methods for dynamic multivariate models. International Economic Review, 39:949–968.
- Spirtes, P., Glymour, C., & Scheines, R. (2000). Causation, Prediction and Search. MIT Press, 2nd. edition.

130

- Swanson, N. & Granger, C. (1997). Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions. *Journal of the American Statistical Association*, 92:357–367.
- Uhlig, H. (2005). What are the effects of monetary policy? Results from an agnostic identification procedure. *Journal of Monetary Economics*, 52:381–419.
- Waggoner, D. & Zha, T. (2003). A Gibbs simulator for restricted VAR models. Journal of Economic Dynamics and Control, 26:349–366.
- Zhang, J. (2002). Consistency in causal inference under a variety of assumptions. Master's thesis, Department of Philosophy, Carnegie Mellon University.
- Zhang, J. & Spirtes, P. (2003). Strong faithfulness and uniform consistency in causal inference. In Publishers, M. K., editor, *Proceedings of the 19th Annual Conference on Uncertainty in Artificial Intelligence*, pages 632–639.

A. Appendix I: Description of the Data

Short-term interest rate: SELIC interest rate – adjusted average rate of daily financing guaranteed by federal government securities, calculated in the Special Settlement and Custody System (SELIC) and published by the Central Bank of Brazil (BCB) – annualized rate.

Nominal exchange rate: R / US\$ – end of period buying rate – Source: BCB. Price index: IPCA price index – Source: IBGE.

Medium-term interest rate: 180-day swap rate (PRE \times CDI) – Source: Brazilian Mercantile & Futures Exchange – annualized rate.

Output: the industrial production index – three-month moving average – Source: IBGE.

Monetary Aggregate: M1 – working days average – Source: BCB.

B. Appendix II: A Stochastic Mundell-Fleming Model

In this Appendix, we present a stochastic rational expectations open-economy model with sticky prices, which can be used to justify the sign restrictions employed to identify the SVAR model. Let y_t^d be aggregate demand, s_t the nominal exchange rate (the domestic value of foreign currency), p_t the domestic price level, i_t the domestic nominal interest rate, mt the nominal money stock, and $E_t(X_t)$ the mathematical expectation of the random variable X_t conditioned on information at date t. Foreign variables are taken as given, so, without loss of generality, we set the foreign price level and foreign interest rate equal to zero ($p^* = 0$ and $i^* = 0$). The shadow values associated with the flexible-price equilibrium are denoted by a superscript 'flex.'

$$y_t^d = d_t + \eta q_t - \sigma \left[i_t - E_t \left(p_{t+1} - p_t \right) \right]$$
(A.1)

$$p_t = (1 - \theta) E_{t-1} p_t^{flex} + \theta p_t^{flex}$$
(A.2)

$$m_t^S - p_t = y_t - \lambda i_t \tag{A.3}$$

$$i_t = E_t \left(s_{t+1} - s_t \right) + c_t \tag{A.4}$$

Equation (A.1) is an open-economy IS equation where demand for output (y_t^d) depends on a demand shock (d_t) , is increasing in the real exchange rate $(q_t = s_t - p_t)$, and decreasing in the ex ante real interest rate. The sticky-price adjustment rule (A.2) states that the price level in period t is an average of the market-clearing price expected at t - 1 to prevail at $t(E_{t-1}p_t^{flex})$, and the price that would clear the market in period $t(p_t^{flex})$. Prices are instantaneously perfectly flexible if $\theta = 1$ and they are completely fixed one-period in advance if $\theta = 0$. Intermediate degrees of price flexibility are characterized by $0 < \theta < 1$. Equation (A.3) is a standard LM equation where the income elasticity of money demand is assumed to be 1. Capital market equilibrium is given by equation (A.4), where c_t is a term that reflects the exchange rate risk of the domestic currency.

The stochastic processes that drive the dynamics in this model – supply, demand, monetary policy, and exchange rate shocks – are given by: 25

$$y_t^s = y_{t-1}^s + \varepsilon_t^s \tag{A.5}$$

$$d_t = d_{t-1} + \varepsilon_t^d - \gamma \varepsilon_{t-1}^d \tag{A.6}$$

Brazilian Review of Econometrics 31(1) May 2011

²⁵Here we consider money supply innovations as monetary policy shocks, whereas in the text we assume that innovations to the SELIC rate are monetary policy shocks. Just bear in mind that money and interest rates are negatively related.

$$m_t = m_{t-1} + \varepsilon_t^m \tag{A.7}$$

$$c_t = c_{t-1} + \varepsilon_t^c \tag{A.8}$$

where $\varepsilon_t^s \sim N(0, \sigma_s^2), \varepsilon_t^d \sim N(0, \sigma_d^2), \varepsilon_t^m \sim N(0, \sigma_m^2), \varepsilon_t^c \sim N(0, \sigma_c^2), 0 < \gamma < 1.$ The long-run equilibrium or the steady state is not conveniently character-

ized in a stochastic environment because the economy is constantly being hit by shocks to the non-stationary exogenous state variables. Instead of a long-run equilibrium, we work with an equilibrium concept given by the solution formed under hypothetically fully flexible prices. Then, as long as there is some degree of price-level stickiness that prevents complete instantaneous adjustment, the disequilibrium can be characterized by the gap between sticky-price solution and the shadow flexible-price equilibrium. We apply a two-stage procedure for solving the equilibrium system (A.1) – (A.8). In the first stage, we solve for a flexible-price equilibrium that corresponds to this system. In the second stage, we use the flexprice equilibrium to arrive at a full-fledged solution for the mixed fix-flex-price system.

It is possible to show that the flexible-price equilibrium values of the model are given by:

$$y_t = y_{t-1} + \varepsilon_t^s \tag{A.9}$$

$$q_t^{flex} = \frac{y_t - d_t}{\eta} + \frac{\gamma \sigma}{\eta(\eta + \sigma)} \varepsilon_t^d + \frac{\sigma}{\eta} c_t \tag{A.10}$$

$$p_t^{flex} = m_t - y_t + \lambda c_t + \frac{\lambda \gamma}{(\eta + \sigma)(1 + \lambda)} \varepsilon_t^d$$
(A.11)

$$s_t^{flex} = m_t + \frac{(1-\eta)}{\eta} y_t - \frac{1}{\eta} d_t + \left(\frac{\sigma}{\eta} + \frac{\lambda\gamma}{(\eta+\sigma)(1+\lambda)}\right) \varepsilon_t^d \tag{A.12}$$

$$i_t^{flex} = \left(\frac{\gamma}{\eta + \sigma} - \frac{\lambda\gamma}{(\eta + \sigma)(1 + \lambda)}\right)\varepsilon_t^d + c_t \tag{A.13}$$

We now use the flex-price equilibrium values obtained in the first stage to solve the full-fledged equilibrium in this second stage:

$$p_t = p_t^{flex} - (1 - \theta) \left(\varepsilon_t^m - \varepsilon_t^s + \alpha \varepsilon_t^d + \lambda \varepsilon_t^c \right)$$
(A.14)

$$q_t = q_t^{flex} + \frac{(1+\lambda)(1-\theta)}{\eta + \sigma + \lambda} \left(\varepsilon_t^m - \varepsilon_t^s + \alpha \varepsilon_t^d + \lambda \varepsilon_t^c\right)$$
(A.15)

Brazilian Review of Econometrics 31(1) May 2011

$$s_t = s_t^{flex} + \frac{(1 - \eta - \sigma)}{(\eta + \sigma + \lambda)} (1 - \theta) \left(\varepsilon_t^m - \varepsilon_t^s + \alpha \varepsilon_t^d + \lambda \varepsilon_t^c\right)$$
(A.16)

$$y_t^d = y_t + \frac{(1+\lambda)(1-\theta)(\eta+\sigma)}{\eta+\sigma+\lambda} \left(\varepsilon_t^m - \varepsilon_t^s + \alpha\varepsilon_t^d + \lambda\varepsilon_t^c\right)$$
(A.17)

$$i_t = i_t^{flex} - \frac{(1 - \eta - \sigma)}{(\eta + \sigma + \lambda)} (1 - \theta) \left(\varepsilon_t^m - \varepsilon_t^s + \alpha \varepsilon_t^d + \lambda \varepsilon_t^c\right)$$
(A.18)

where $\alpha = \frac{\lambda \gamma}{(\eta + \sigma)(1 + \lambda)}$. The effect of a one-unit increase in the variable j innovation at date $t(\varepsilon_t^j)$ on the value of variable l at time t, holding all other innovations constant, is given by $\frac{\partial l}{\partial \varepsilon_t^j}$:

$$\begin{split} \frac{\partial y_t^d}{\partial \varepsilon_t^s} &= 1 - \frac{(1+\lambda)(1-\theta)(\eta+\sigma)}{\eta+\sigma+\lambda} \ge 0\\ \frac{\partial y_t^d}{\partial \varepsilon_t^d} &= \frac{(1+\lambda)(1-\theta)(\eta+\sigma)}{\eta+\sigma+\lambda} \alpha \ge 0\\ \frac{\partial y_t^d}{\partial \varepsilon_t^m} &= \frac{(1+\lambda)(1-\theta)(\eta+\sigma)}{\eta+\sigma+\lambda} \ge 0\\ \frac{\partial y_t^d}{\partial \varepsilon_t^c} &= \frac{(1+\lambda)(1-\theta)(\eta+\sigma)}{\eta+\sigma+\lambda} \lambda \ge 0\\ \frac{\partial p_t}{\partial \varepsilon_t^s} &= -\theta \le 0\\ \frac{\partial p_t}{\partial \varepsilon_t^m} &= \theta \ge 0\\ \frac{\partial p_t}{\partial \varepsilon_t^m} &= \theta \ge 0\\ \frac{\partial p_t}{\partial \varepsilon_t^c} &= \lambda \theta \ge 0\\ \frac{\partial q_t}{\partial \varepsilon_t^s} &= \frac{1}{\eta} - \frac{(1+\lambda)(1-\theta)}{\eta+\sigma+\lambda}\\ \frac{\partial q_t}{\partial \varepsilon_t^d} &= -\frac{1}{\eta} + \frac{\gamma\sigma}{\eta(\eta+\sigma)} + \frac{(1+\lambda)(1-\theta)}{\eta+\sigma+\lambda}\alpha \end{split}$$

[The conventional view is that $\frac{\partial q_t}{\partial \varepsilon_t^d} \leq 0$]

Brazilian Review of Econometrics 31(1) May 2011

$$\begin{aligned} \frac{\partial q_t}{\partial \varepsilon_t^m} &= \frac{(1+\lambda)(1-\theta)}{\eta+\sigma+\lambda} \ge 0\\ \frac{\partial q_t}{\partial \varepsilon_t^c} &= \frac{\lambda}{\eta} + \frac{(1+\lambda)(1-\theta)}{\eta+\sigma+\lambda}\lambda \ge 0\\ \frac{\partial i_t}{\partial \varepsilon_t^s} &= \frac{(1-\eta-\sigma)}{(\eta+\sigma+\lambda)}(1-\theta) \ge 0 \quad \text{if} \quad \eta+\sigma \le 1\\ \frac{\partial i_t}{\partial \varepsilon_t^d} &= \left(\frac{\gamma}{\eta+\sigma}-\alpha\right) - \frac{\alpha(1-\eta-\sigma)}{(\eta+\sigma+\lambda)}(1-\theta) \end{aligned}$$

[The conventional view is that $\frac{\partial i_t}{\partial \varepsilon_t^d} \ge 0$]

$$\begin{split} \frac{\partial i_t}{\partial \varepsilon_t^m} &= -\frac{(1-\eta-\sigma)}{(\eta+\sigma+\lambda)}(1-\theta) \leq 0 \quad \text{if} \quad \eta+\sigma \leq 1 \\ \frac{\partial i_t}{\partial \varepsilon_t^c} &= 1 - \frac{(1-\eta-\sigma)}{(\eta+\sigma+\lambda)}(1-\theta)\lambda \geq 0 \end{split}$$