AUCO Czech Economic Review 4 (2010) 139–167
Acta Universitatis Carolinae Oeconomica

Monetary Policy Transmission in Italy: A BVAR Analysis with Sign Restriction

Carlo Migliardo[*]

Received 10 October 2008; Accepted 21 January 2010

Abstract In this paper, we propose a Bayesian VAR model to examine the short term effects of monetary policy shocks on the Italian economy. Firstly, our BVAR model uses the Cholesky decomposition to identify four kinds of macroeconomic shocks, namely, supply, demand, interest rate and monetary shocks. Then, from the theoretical model, we derive and impose a minimum set of robust sign restrictions to identify the transmission mechanism of monetary tightening. The outcomes from the sign identification confirm the micro evidence on inflation persistence. Moreover, our results show a greater persistence of inflation to monetary restriction than Cholesky identification presents. Overall, we find that a monetary innovation brings a decline of 30 basis point of GDP, this result is almost invariant across both prior and identification technique.

Keywords Bayesian VAR methods, conjugate prior, Litterman prior, Markov chain Monte Carlo, monetary policy, regime switching, sign restriction identification
JEL classification C11, C32, E12, E32, E58

1. Introduction

A lot of research has been conducted to provide the stylized facts on the effect of monetary policy transmission mechanism on nominal and real economic variables. While economic analysis of monetary transmission through an atheoretical approach has a long history, it is only recently that interest in Bayesian techniques has been growing among academics and in political economic research as well. As Cogley and Sargent (2005) pointed out: "*...Fed uses Bayesian methods to update estimates of three models of the Phillips curve: a Samuelson-Solow model, a Solow-Tobin model, and a Lucas model. Every period, the central bank updates the probability that it assigns to each of these three models, and then determines its first-period decision by solving a 'Bayesian linear regulator problem'. Although by the mid 1970s the U.S. data induce the Fed to assign very high probability to the Lucas model, the government refrains from adopting its low-inflation policy recommendation because that policy has very bad consequences under one of the other low (but not zero) probability models. The statistical model is thus able to explain the puzzling delay in the Fed's decision to deflate after learning the natural rate hypothesis*." On the other hand, the Euro-

[*] University of Messina, Department of Economics and Social Sciences (SEFISAST), Via dei Verdi 75, 98122 Messina, Italy. Phone +39-090-6764514, E-mail: cmigliardo@unime.it.

pean Central Bank (hereinafter ECB) has set up a research group that uses a Bayesian approach to implement monetary policy strategy.[1]

At the same time, Vector Autoregressive models (hereinafter VAR) undoubtedly have extensive use in the empirical analysis of money effects.[2] De Arcangelis and Di Giorgio (1999) and Angeloni et al. (2003) developed a structural VAR to identify monetary policy shocks testing both the "money view" and the "credit view". They have shown that interest rate should be a preferred rather than a reserve aggregate as a policy measure. They argue that, in the 90s, the Bank of Italy implemented the monetary policy mainly to counteract inflation. Smets (1997) provides a small open macroeconometric model for European Monetary Union (hereinafter EMU) countries to overcome the typical exchange rate identification problem in the VAR analysis.

Bayesian inference, in VAR models (hereinafter BVAR), was introduced by Doan et al. (1984) and Litterman (1983, 1986). The considerable empirical literature on the use of BVARs showed their ease of use and satisfactory forecasting performance. Besides such empirical studies, there are currently very few existing Bayesian VARs. Amisano et al. (1997b) build a BVAR model for Italy using monthly and quarterly data. Ciccarelli and Rebucci (2006) applied a Bayesian VAR methodology to investigate the evolution of the transmission mechanism of European monetary policy at cross-country level in the run-up to EMU.

In line with previous empirical works based on the Bayesian framework, this paper aims to provide an econometric model for economic policy analysis of the Italian economy, with a special focus on assessment of the dynamic responses of the Italian economy to monetary policy innovations.

One of the goals of this paper focuses on the short term effects of monetary policy in Italy on a Bayesian VAR study. This analysis introduces some innovations with regard to previous empirical literature on the BVAR for Italy. Above all, the contribution of the paper with respect to the existing literature is twofold. Firstly, since the findings of previous literature far from unequivocally establish the impact of a monetary policy on output and prices,[3] this paper provides new evidence about the impact of monetary policy on the Italian economy over the last 26 years by applying sign restriction as in Canova and De Nicolò (2002). This technique allows us to identify monetary policy shocks as well as to rule out eventual puzzles.[4]

Even though this paper is an application of existing methodology, the Bayesian

[1] See Sims (2002) for other evidence on the use of the Bayesian method by Central Banks.

[2] See the references reported, for instance Sims (1980), Christiano and Eichenbaum (1992), Bernanke and Blinder (1992), Amisano and Giannini (1997a), Bagliano and Favero (1999), among many others. Additional excellent surveys for Italy are in Smets (1997), Bagliano and Favero (1999), De Arcangelis and Di Giorgio (1999), Angeloni et al. (2003), Mojon and Peersman (2003).

[3] The results comparing the effect of interest rate shock have not shown a high degree of consistency across studies, Rafiq and Mallick (2008) found a positive impact on output of a contractionary monetary policy; on the contrary, Mojon and Peersman (2003) show an output declining in response to an increase in interest rate.

[4] Other examples of identification techniques are found in Blanchard and Quah (1989), who impose a zero impulse response at infinity; Lippi and Reichlin (1994), who impose a particular shape on the impulse response. In our paper, as in Canova and Gambetti (2009), we impose sign restrictions on the cross-correlations of variables in response to shocks. Uhlig (2005) uses an identification procedure that selects the "correct" orthogonal decomposition according to the economic theory.

methods in general and the sign restriction in particular are rather novel and not very common in data on countries other than the U.S. (Canova and De Nicolò, 2002). Although previous research on Italy used BVAR, this study introduces an identification achieved without imposing zero constraints on impact responses of variables to structural shocks.[5] Secondly, in contrast with the U.S. case, for Italy most empirical study has not previously been conducted over a long sample period due to the shift in the monetary policy regime as well as to the several structural breaks that have occurred in the Italian economy. However, the use of this new methodology achieves a robust identification of the shocks despite these data problems. Furthermore, in order to avoid other puzzles[6] in our analysis, the VAR exercise has been estimated in a closed economy framework. The sign identification also allows us to deal with the persistence on inflation, output and money.

The analysis is structured in a two-step approach. In the first part, the study analyzes the outcomes derived from standard Bayesian VAR with Cholesky identification. Several problems occur with Cholesky factorization, i.e. the results tend to depend on the specific set of identifying restrictions. This implies that the effects of shocks to economic variables are dependent on their position in the orthogonalization order. In practice, the order in which the variables are inserted into the model is quite relevant and this has a fundamental implication for the outcomes. Thus, a variation in the order in which variables are inserted dramatically changes the response and puzzles are not infrequent. In particular, the Bayesian VAR applied to real data shows that we can usefully interpret Cholesky decomposed data on IS and Phillips shocks, but we have to be careful about interpreting those on a contractionary monetary policy (central bank raises interest rate). Furthermore, the prior sensitivity analysis[7] reveals a liquidity "conundrum", especially in the Conjugate and Diffuse prior, even though the error bands are not so well defined.

In the second step, the research evaluates the persistence of a negative monetary policy impulse through a sign restriction identification methodology. In our model, we introduce some restrictions that are derived from economic theory and are consistent with the New Keynesian sticky price model as well as with Real Business Cycle (hereinafter RBC) models. We show that effects of monetary policy on prices and on output are usually qualitatively similar to the ones obtained by Amisano et al. (1997). The outcomes from the sign identification confirm the micro evidence on inflation persistence.[8] Moreover, our results show a greater persistence of inflation to monetary

[5] The contemporaneous zero restrictions conventionally used do not have economic insight, since they are absent in the Dynamic Stochastic General Equilibrium (hereinafter DSGE) models that economists use to evaluate VAR results.

[6] The contemporaneous relation between interest rate and exchange rate innovations causes new empirical conundrums, such as that of an impact depreciation of the nominal exchange rate following a contractionary monetary policy shock in the domestic country, this issue be may due to the simultaneous relation between exchange rate and interest rate, see Eichenbaum and Evans (1995), Grilli and Rubini (1996), or Bagliano et al. (1999).

[7] The prior sensitivity analysis checks the robustness of the estimated effects on changes in the prior.

[8] See Benigno and Lopez-Salido (2006) and Galì et al. (2003) for EURO area. For Italy see Fabiani et al. (2004). The evidence provided by micro analysis of price dynamics shows that inflation dynamics in Italy last at least 5–6 quarters.

restriction than Cholesky identification presents.

The paper is organized as follows: Section 2 describes data and the techniques to make them stationary. Section 3 shows the estimation procedure. Section 4 presents and analyzes the empirical results. Section 5 concludes.

2. The data

This section contains a description of the dataset. The sample is restricted to 1981 onwards, when the Bank of Italy was separated from the Treasury, becoming truly independent. The model thus spans from 1981:1 to 2006:4 at quarterly frequency. All time series used were obtained from *DataStream*. The variables used in the BVAR model are defined as follows:

GDP: Real gross domestic product, quarterly seasonally adjusted at constant price of 2000.
Price: consumer price index, seasonally adjusted.
Interest rate:[9] Short term interest rate (3 month T-bill rate) percent annum.
Money: monetary aggregate (M3, which should be the most stable).

All data is in logs except the interest rate and inflation, thus, the coefficients can be interpreted as elasticity. The series are plotted in Figure 1. The data has been annualized and expressed in percentages.

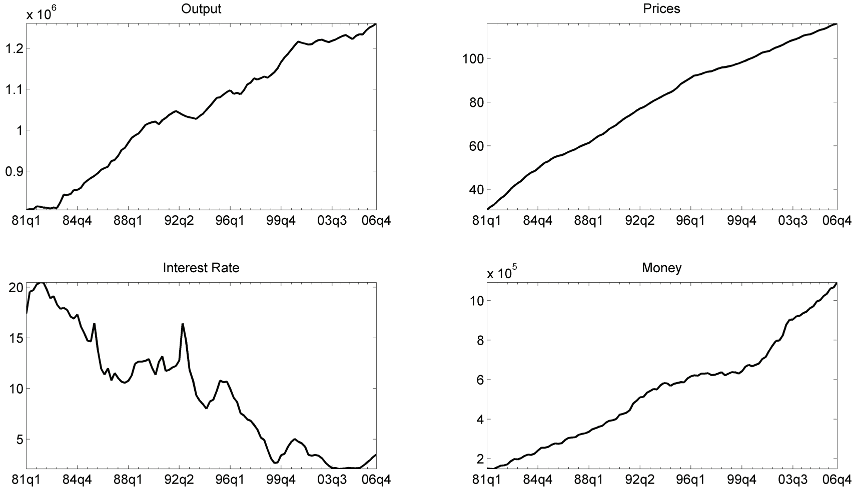


Figure 1. The dynamic of variables in level

[9] We follow the approach suggested by Bernanke and Blinder (1992), also confirmed for Italy by De Arcangelis and Di Giorgio (1998), to adopt as monetary policy instrument the interest rate than the reserve aggregate.

Visual inspection of Real data denotes a deterministic trend in all series. As expected, real GDP and Price levels show a deterministic pattern. It is interesting to note that interest rate and inflation show non-stationary behavior, perhaps because of Italy's run-up to EMU.

The data has been filtered attenuating the deterministic component in all series by applying the Hodrick-Prescott (HP) filter:[10]

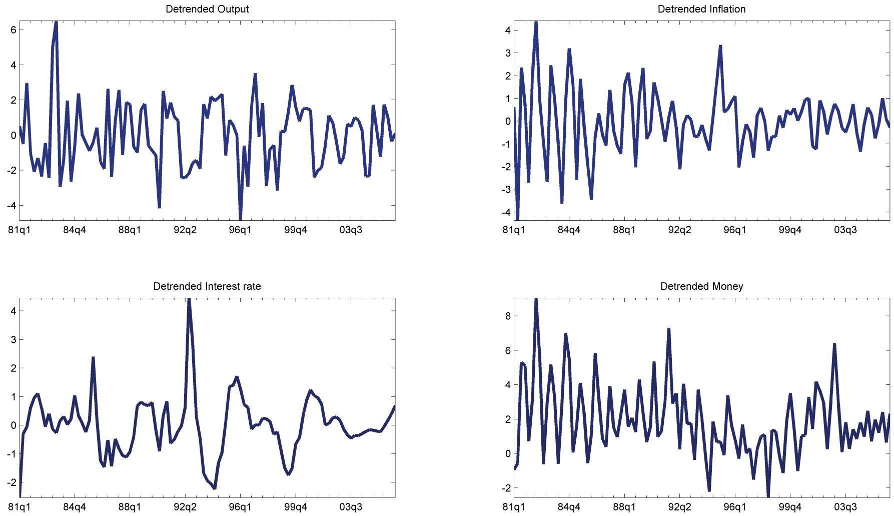


Figure 2. Filtered data applying the HP filter

Hence, our BVAR model utilizes these four variables: y_t is the linearly detrended log of the GDP, π_t is the first difference of the log price level, and m_t is the linearly detrended log of money stock (M3).

3. Econometric Methodology

This part describes the estimation procedure. Essentially, our research adopts several different techniques, though all based on the qualitative evaluation of impulse response of the VAR. Although the system considered may contain integrated or even cointegrated variables, this paper focuses on the short term dynamics with stationary data set around a deterministic trend. Moreover, for some cointegration constraints, it may be hard to give a single economic interpretation, given that several structural[11] shocks

[10] We also estimate the model using the inflation and GDP deviations from the German data, but find similar results.

[11] Such as the European Monetary System (EMS) crisis, the entrance into the European Monetary Union. In order to test for structural stability. We apply the Quandt Likelihood Ratio (QLR) test statistics. Even

occurred during the sample period. Therefore, tests for integration and cointegration among variables are likely to have low efficiency and this could affect economic inference (Sims et al. 1990; Campos et al. 1996 found that imposition of cointegrating relations may reduce the efficiency of estimation). Moreover, the analysis focuses on short-run dynamic response of the system, since we assume that monetary is neutral in the long-run.

The lag length has been chosen using the Hannan-Quinn information criterion. At that lag length, normality of residuals cannot be rejected when we run an appropriate test that compares the third and fourth moments of the residuals with the values 0 and 3, respectively (the results are available upon request), given the detrended time series for our BVAR model, the Schwarz criterion and the Hannan and Queen test identify the optimal dynamics described by a BVAR in (1).

3.1 Identification and estimation

The structural form of a Vector Autoregression is:

$$Ay_t = B(L)y_{t-1} + u_t, \quad u \sim N(0, \Sigma_u), \tag{1}$$

where y_t is a vector of endogenous variables (GDP, Inflation, Interest rate, Money), u is a $n \times 1$ vector of structural error terms. A is the matrix constituted by the contemporaneous effect among the endogenous variables. $B(L) = B_0 + B_1 L + B_2 L^2 + \ldots + B_p L^p$ is the lag equation. Σ_u is the variance covariance matrix and the lag length is p.

In this model each variable of the system is a linear function of the previous values of the other endogenous variables. In order to estimate the coefficient and to identify shocks of the VAR model, we need to express the structural form in the reduced version:

$$y_t = \Phi(L)y_{t-1} + \varepsilon, \tag{2}$$

where $\Phi(L) = A^{-1}B(L)$ and $\varepsilon_t = A^{-1}u_t$, the variance-covariance matrix of the residual Σ_ε can be estimated by OLS methods. However, in the reduced form, the impact of structural innovations to the economic variables cannot be directly computed. Moreover, in the standard VAR representation, innovations are not contemporaneously independent of each other. Therefore, in order to estimate the dynamics response of the variables and consequently to identify the shocks inside the Vector Autoregression model, we will adopt and compare both the Cholesky decomposition and the Sign restriction.

The former utilizes the property of Cholesky factorization. Since any real positive symmetric matrix Ω can be decomposed as $\Omega = ADA' = \hat{A}\hat{A}'$ where $\hat{A} \equiv AD^{1/2}$ where A is a lower triangular matrix with 1s in the main diagonal and D is a diagonal matrix.

Applying the Wold representation form (MA(∞)) to the (2) and contemporaneously uncorrelated shocks, it can be represented by:

$$Y_t = \mu + \Psi(L)\varepsilon_t, \tag{3}$$

if the QLR test revels that maximum value in the 1992:4, the null hypothesis can be rejected at the 10% significance level. Moreover we re-estimated the model on a shorter sample beginning in 1999:1. As is demonstrated in the Appendix, switching to the shorter model does not change the picture significantly.

where $\Psi(L)$ is a matrix polynomial in the lag operator: $\Psi(L) = \sum_{i=0}^{\infty} \Psi_i L^i$; $\Psi_0 = I$; $\Psi_1 = A_1, \ldots, \Psi_k = A_1 \Psi_{k-1}$; $\mu = (I - A_1)^{-1} A_0$, $\varepsilon_t = A^{-1} u_t$. Given that $E(u_t u_t') \equiv \Sigma$, Σ_u is the variance-covariance matrix and $E(\varepsilon_t \varepsilon_t') = A^{-1} E(u_t u_t')(A^{-1})' = A^{-1} \Sigma (A^{-1})'$ applying the Cholesky decomposition to $\Sigma \to \Sigma = ADA'$ (where D is a diagonal matrix with the variance of u_t) the outcome is: $E(\varepsilon_t \varepsilon_t') = A^{-1} ADA'(A^{-1})' = D$. Therefore the error terms ε_t are mutually uncorrelated. In this way, we have orthogonalized the variance-covariance matrix of innovations, allowing the identification of the shocks. Moreover, this approach has the advantage that shocks in the VAR system can be identified as shocks to the endogenous variables.

Several problems arise from the Cholesky factorization, e.g. the results tend to depend on the specific set of identifying restrictions. This implies that the effects of shocks to economic variables are dependent on their position in the orthogonalization ordering. Since a variation in the order in which variables are inserted dramatically changes the response, puzzles are not infrequent. Furthermore, we have to drop some contemporaneous effects, imposing some elements of the variance-covariance matrix at zero. Cooley and Dwyer (1998) have shown contemporaneously recursive structure is hard to obtain in a general equilibrium model. Moreover, Canova and Pina (1999) have showed that in DSGE this kind of identification creates substantial misspecification.

In the construction of the empirical model econometricians, in the face of several problems, use some prior information, e.g. the variables to insert into the VAR, the correct selection of lag length, identification restriction, as well as problems in the quality of the sample or in the length of the time series. Nowadays, in this context, the Bayesian approach is the natural solution. Born for the purpose of forecasting, Bayesian VARs are now estimated for several other goals.

As Sims (1993), Canova[12] and Sala (2006) and Robertson and Tallman (1999), among others, have shown, VARs with Bayesian restriction give greater precision and better forecasts than Autoregressive Integrated Moving Average (hereinafter ARIMA) models and traditional multivariate simultaneous equations. In particular, the former reduce the dimensionality of the problem imposing probability distributions on the coefficients of the VAR.

In order to detect lack of identification a sequence of priors[13] has been adopted, using the economic theory for parameterization. Above all, since the results of Bayesian analysis are sensitive to the priors, we used three popular priors, Minnesota, Conjugate and Diffuse, to check the sensitivity of the findings. The first prior was formulated by Litterman[14](1986). In VAR models, Litterman introduced prior distributions that induced random walk prior mean for coefficients with a set of hyper-parameters, which drive their variance. In more details, Litterman's prior holds that variables behave like a random walk with an unknown deterministic component, hence the systematic variation in the data is relatively small compared with the random variation. Doan et al.

[12] Canova (2007) showed, on average, BVARs perform better than other econometric models (Auto Regressive, ARIMA and Unrestricted VARs), both in Mean Square Error tests and in turning point predictions.

[13] Bayesian estimations try to improve the estimates by introducing prior information (i.e. information drawn from theory or the visual inspection of the sample) into the analysis.

[14] This prior is also called Minnesota, since Litterman and others developed it at the University of Minnesota.

(1984) used this kind of prior to make conditional forecasts. According to these authors, the Minnesota prior may reveal empirical regularities unknown using standard procedures. Furthermore, as Marimon and Scott (2001) pointed out, the Litterman prior prevents the misspecification of coefficients (time varying instead of constant coefficient). At the same time, it corrects the possible presence of serial correlation within the error terms due to over parametrization, isolating the systemic component of variation inside the series.

For these reasons, we suppose that the Minnesota prior is the natural candidate for inflation. However, other kinds of prior will be tested, and in the end we will select the one that minimizes the mean square error.

3.2 Sign restriction identification

In contrast to the conventional identification, sign restriction does not drop any contemporaneous effects (the variance-covariance matrix is full); Given that an orthonormal matrix Σ can be orthogonal decomposed in the product of matrix of its eigenvalues times the matrix of its eigenvector as $\Sigma = \Pi\Xi\Pi'$ where Π is a diagonal matrix of eigenvectors of Σ, Ξ is a diagonal matrix of eigenvalues of Σ_ε and setting $\tilde{P}$ as $\tilde{P} = \Pi\Xi^{1/2}$ to simplify the notation then $\Sigma = \tilde{P}\tilde{P}'$. While the Choleky decomposition imposes the upper tringular part of the innovation matrix to zero, the sign identification uses a full variance-covariance matrix. Since, there exist infinite combinations of possible matrices $\tilde{P}$, our identification selects, over the space of the orthogonalisations, only the matrices with the sign required by theory, i.e. demand shocks raise price, liquidity effect and so on.

First step: Orthogonalized error term through the orthogonal decomposition for the variance-covariance matrix of the VAR in the (1) $\Sigma_\varepsilon = \Pi\Xi\Pi'$ and setting Π as $\Pi = P_{m,n}Q_{m,n}(\theta_i)$, where $Q_{m,n}(\theta_i)$ are orthonormal rotation matrices equal to:

$$Q_{m,n}(\theta_i) = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \cos(\theta_i) & \dots & -\sin(\theta_i) & 0 \\ \vdots & \vdots & \vdots & 1 & \vdots & \vdots \\ 0 & 0 & \sin(\theta_i) & \dots & \cos(\theta_i) & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix},$$

where θ is the rotation angle $(0 < \theta \leq \pi)$ for the row m and n.

Since the BVAR model contains four variables, there are six $(N(N-1)/2)$ bivariate rotations admissible $(i = 1, \dots, 6)$:

$$Q_{m,n}(\theta_i) = \begin{vmatrix} \cos(\theta_1) & -\sin(\theta_1) & 0 & 0 \\ \sin(\theta_1) & \cos(\theta_1) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos(\theta_2) & -\sin(\theta_2) \\ 0 & 0 & \sin(\theta_2) & \cos(\theta_2) \end{vmatrix}$$

$$\begin{vmatrix} \cos(\theta_3) & 0 & -\sin(\theta_3) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(\theta_3) & 0 & \cos(\theta_3) & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} \cos(\theta_4) & 0 & 0 & -\sin(\theta_4) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sin(\theta_4) & 0 & 0 & \cos(\theta_4) \end{vmatrix}$$

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_5) & -\sin(\theta_5) & 0 \\ 0 & \sin(\theta_5) & \cos(\theta_5) & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} & 0 & & 0 \\ & \cos(\theta_6) & -\sin(\theta_6) & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \sin(\theta_6) & \cos(\theta_6) & 1 \end{vmatrix}$$

Then, we can set $\tilde{P} = P \quad \Xi^{1/2} Q_{m,n}(\theta_i)$ given that Q is orthonormal, we can decompose the variance-covariance matrix Σ of the BVAR: $\Sigma = \tilde{P}\tilde{P}' = PQ_{m,n}(\theta)\,Q_{m,n}(\theta)'\,P' = PD^{0.5}Q_{m,n}(\theta)\,Q_{m,n}(\theta)'\,D^{0.5}P'$. Therefore, starting from this decomposition, we can decouple it in one direction, for each value of θ. We grid the interval into M points and derive 6M points for the orthogonal decomposition of Σ. Even though, the variance-covariance matrix of the innovations has been orthogonalized, the number of admissible P is very large, thus in the second step we select only the orthonormal matrices that agree with theoretical economic signs.

The MATlab routine works as follow: First it builds matrices Π and D. It checks if the sign of the shocks match the expected responses. If this does not occur, the algorithm assembles new matrices $\Pi_{m,n}(\theta_i) = \Pi \times Q_{m,n}(\theta_i)$ using several values of m, n and θ_i. The vector of identified shocks is a particular transformation of sine and cosine functions, the rotation matrix Q defines the selected identification and Q is an explicit function of the sines and cosines of an angle. Therefore, the sign restriction achieves identification by restricting the signs of structural responses and eliminates any kind of possible puzzle by construction. For more details see Canova and De Nicolò (2002).

4. Empirical results

This section deals with model responses to several different shocks, such as IS, cost push shock and monetary policy innovation. The innovations have been identified by Cholesky techniques and sign restriction as well. As the first step, we apply a conventional identification tool to compute the impulse responses functions.

The Cholesky identification is imposed on the model with the following causality order: $y_t = [GDP_t, \pi_t, i_t, M_t]$. This specification follows the principle that monetary variables should be ordered last, since they are expected to respond faster to the real economy rather than vice versa (Favero 2001). Therefore, the Central Bank through its policy instrument (the interest rate) reacts immediately to disturbances in the macroeconomic environment, while the non-policy variables (such as output and prices) react with a lag to monetary policy impulse. GDP is placed at the top of the ordering, assuming that it will only be affected by exogenous fiscal policy contemporaneously. The price variable is ordered in the middle given that prices are supposed to react immediately to a GDP innovation as well as reacting to monetary variables only with a lag. Therefore Cholesky decomposition drops some contemporaneous effects by construction and a possible solution for this issue is to use a full matrix with a minimal set of sign restrictions that are robust across the theoretical models.[15] Specifically, we draw attention to monetary policy shock persistence on GDP, Inflation and Money Stock.

4.1 Impulse response function

Let us consider a theoretical point of view, what we would expect is that: An IS shock (i.e. a rise in government spending) directly raises output gap and inflation. Following the Taylor rule, the Central Bank responds to this by raising the interest rate. Over time, all three variables should gradually return to their steady state level. A supply/Phillips shock (i.e. unions raising their wage claims) raises inflation and (as real wages are also higher) make firms hire less labor, so output falls; over time, the higher inflation offsets part of the impact the nominal-wage rise had on real wages and as real wages fall again, so the output gap variable rises as well. Two opposite effects (through the Taylor rule) affect the interest rate: on the one hand, higher inflation tends to raise it; on the other hand, the lower output gap tends to lower it; overall, firstly the former effect should dominate, therefore raising the interest rate, whereas later on the latter effect dominates.

The Taylor shock (central bank raises interest rate): the higher price of money makes people hold less of it and borrow less for consumption and investment, so both inflation and the output gap decline sharply and recover only gradually afterwards, just as the interest rate returns to zero as well.

4.2 Impulse response functions from the data

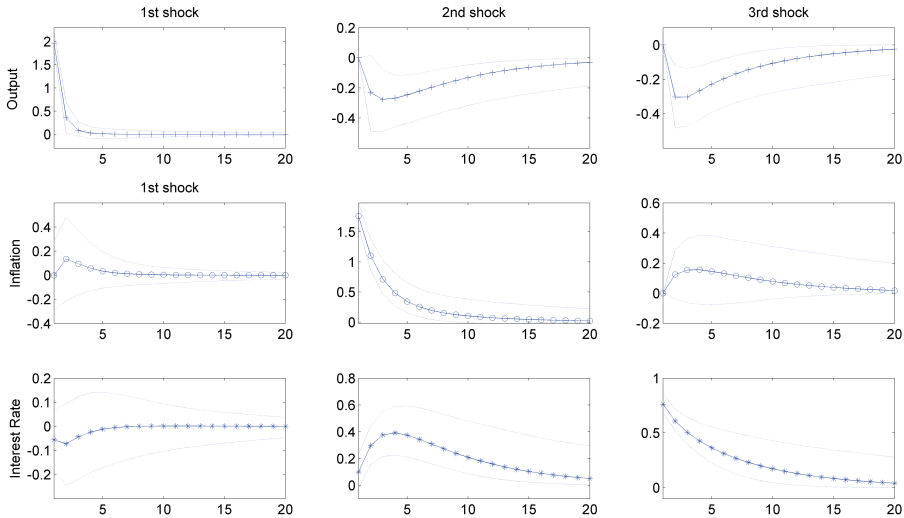
The first estimation comes from a VAR with three variables: output-gap y_t, inflation π_t, the nominal interest rate i_t. Then, we introduce the money stock m_t.

Firstly, we estimate Bayesian VAR with Conjugate prior information, in which the likelihood is a Normal-inverse Wishart (NIW). Therefore, assuming an NIW prior implies that the posterior is also NIW and posterior parameters can be computed analytically. Also, we calibrate an uninformative prior,[16] by setting prior precision to zero. Then, we make 1000 draws from the posterior distribution of the VAR coefficients. Above all, we employed the Cholesky decomposition procedure to isolate the impact of the three different shocks.

As we can see in the Figure 3, confidence intervals are wide. The estimated effects of demand and supply shocks on output and prices are qualitatively in line with previous empirical findings, such as those in Angeloni et al. (2003). In more detail, the IS shock (i.e. a rise in government spending) directly raises inflation, in particular for Litterman's prior (see Appendix A).

[15] For a recent example of such models, see Walsh (2003) or Woodford (2003).

[16] Bayes' theorem requires the specification of a prior that represents the belief of the researcher. When the prior is non-informative it minimizes the influence of the prior on the posterior. A non-informative prior distribution is basically equivalent to the Bayesian analysis with prior distributions that have very large variances; see the appendix on Bayesian inference for more details.



Note: One standard deviation in size. Median and the 95% confidence interval.

Figure 3. Impulse responses to an exogenous shock (main diagonal) in real GDP, prices and interest rate

The supply/Phillips shock (i.e. unions raising their wage claims) raises inflation and (as real wages are also higher) make firms hire less labor, so GDP falls; over time, the higher inflation offsets part of the impact the nominal-wage rise had on real wages and as real wages fall again, so the output gap variable rises as well. This is nevertheless not the case for the third (monetary policy) shock: as expected, GDP declines,[17] but inflation hardly responds at all in the median, and both exhibit wide confidence intervals. This absence of a clear output gap and inflation response may be due to the use of triangular orthogonalization. Above all, the impulse response functions denote that the Cholesky Identifying Restriction is not justified for the interest rate shock. On the other hand, the graph may represent an example of the counter-intuitive price puzzle[18], where an interest rate rise makes inflation rise rather than fall because it raises people's inflation expectations, but not significantly, since the 95% confidence interval also includes negative values.

What we would expect in a Taylor shock (central bank raises interest rate) is that the higher price of money makes people hold less of it and borrow less for consumption and investment, so both inflation and the output gap decline sharply and recover only gradually afterwards, just as the interest rate returns to zero as well. For all three variables, the median and 5th and 95th percentiles are almost entirely congruent, i.e.

[17] Angeloni et al. (2003) found that the fall in GDP following a contractionary monetary policy shock is concentrated in investment. They argue that the credit channel is a significant mechanism through which the effects of monetary policy operate in Italy.
[18] This anomaly first noted by Sims (1992) and labeled "the price puzzle" by Eichenbaum (1992).

the 95% confidence interval is very narrow, resulting from our large sample size.

Overall, this VAR exercise shows that we can usefully interpret Cholesky decomposed data on IS and Phillips shocks, but we have to be careful about interpreting those on Taylor shocks. The responses to the IS shock are as expected and are indeed where the model fits the data best. The only difference is that the interest rate response is more muted. A possible explanation might be that the present Monetary Authority, in our case the Italian one, has a stronger preference for interest rate smoothing than suggested by the DSGE models.

The supply (hereinafter AS) shock to inflation dies out much more quickly than in the DSGE models, or in other words, inflation is less persistent. Furthermore, the inflation hike fails to cause a drop in output,[19] which might be taken to suggest that while prices are less sticky, real variables are more so: even when real wages are higher, firms do not want to or cannot (due to firing regulations, sunk costs, etc.) respond by immediately cutting production. Above, we have shown that the Cholesky decomposition seems to be appropriate in this case; this observation may validly be understood to be a good representation of reality. When studying the effects of a contractionary monetary policy shock, we must remember our finding from above that the Cholesky identification of Taylor shock responses is not entirely reliable. Keeping this qualification in mind, we find firstly that the path of the interest rate itself is less convex than in the standard literature on New Keynesian models. Secondly, a tightening monetary policy shock causes the negative response of the output gap, as we would have expected. Thirdly and most strikingly, we see a positive response of inflation, which is now significant at the 95% confidence level. A possible explanation for this price puzzle is that the rate rise raises people's inflation expectations; because they understand the rate rise as implying that the central bank expects higher future inflation.[20]

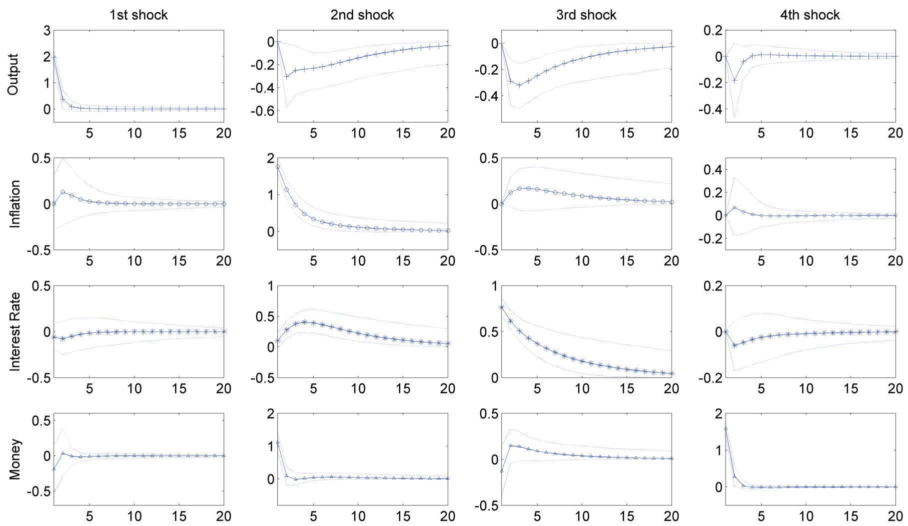
However, it is interesting to note that the Central Bank appears to adopt a strict inflation-targeting rule that responds only to cost push shock. Since, the Bank of Italy, notoriously less inflation averse than the ECB, conducted monetary policy until Italy's entry into EMU, it is not straightforward to interpret this dynamic behavior. Furthermore, Bernanke (2004) argues that in specific historical periods[21] changes in inflation expectations are essential for identifying truly structural shocks. In support of this Castelnuovo and Surico (2006) stressed that when the policy framework does not mandate sufficient response to inflation pressures, the behavior of private sector expectations give rise to perverse dynamics, like inflation persistence and the price puzzle, that are not necessarily intrinsic characteristics of the economy.

Now let us introduce a narrow money growth (M3) into our model and we get the following figures:

[19] This result is prior sensitive. In particular, when we use Litterman prior information, GDP reacts positively. On the contrary, we get a negative response by using other alternative Bayesian priors, such as Diffuse or Conjugate.

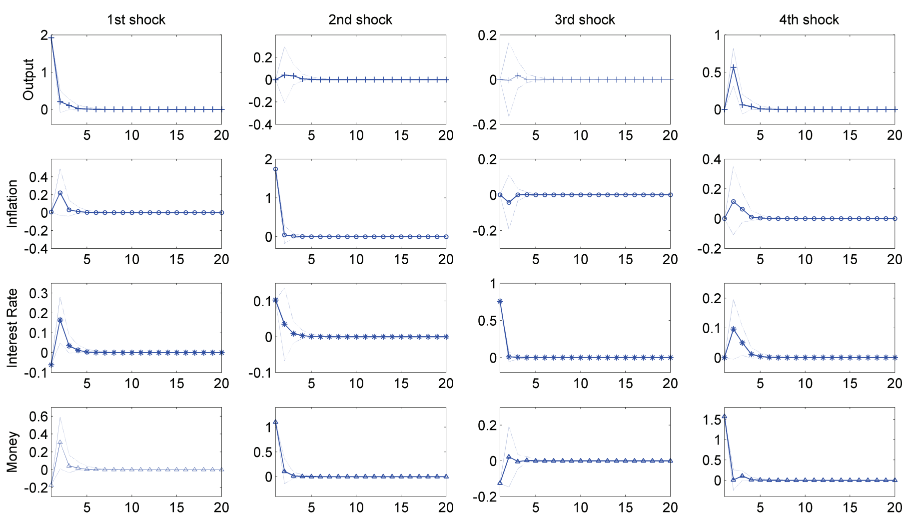
[20] An alternative explanation about the price puzzle is given by Post Keynesian point (e.g. Lavoie) that considers the interest rate to be a cost in the production process and as such its increase drives up prices.

[21] Castelnuovo and Surico (2006) find that the positive response of prices to a monetary policy shock is historically limited to the subsamples associated with a *weak* Central Bank response to inflation. For instance, these subsamples correspond to the pre-Volcker period for the United States and the period prior to the introduction of the inflation targeting rule for the United Kingdom.



Note: One standard deviation in size. Median and the 95% confidence interval.

Figure 4. Impulse responses to an exogenous shock (main diagonal) in real GDP, prices, interest rate and money stock



Note: One standard deviation in size. Median and the 95% confidence interval.

Figure 5. Impulse responses to an exogenous shock (main diagonal) in real GDP, prices, interest rate and money stock using the Litterman Prior

The impulse response for money shock is prior sensitive. Indeed, both conjugate and diffuse priors indicate that money stock has no real effects. Furthermore, a money shock decreases the interest rate, as expected. On the contrary, the Litterman prior reveals a real money effect one to one on GDP.[22] In support of this result Angeloni et al. (2003) stress the importance of cash flow effects finding that money plays a central role in the transmission mechanism of monetary policy in Italy.

Summarizing our finding with standard identification we may state that the impact of a money shock has a very short effect, confirming the monetary neutrality in the long-run. The impulse responses are sensitive to changes in the assumed prior, confirming that the Litterman prior performs better than conjugate and diffuse priors. In particular, for the first two shocks, the Minnesota prior shows the smallest error band and no puzzles either. This outcome validates the Random walk assumption to fit economic time series. Furthermore, it is hard to identify the policy reaction function because the dataset includes the switch of monetary policy from the Bank of Italy to the ECB. This finding is qualitatively consistent with previous empirical works (De Arcangelis and Di Giorgio 1999), based on the pre-EMU period, revealing that Policy Makers follow an inflation target rule.

4.3 Identification: Cholesky vs. sign restrictions

Since the response impulse function problems are more concentrated on the monetary restrictions, a sign restriction method has been implemented to identify a monetary policy innovation. We do not impose any restrictions on shocks 1, 2 and 4. Shock 3 is labeled as a monetary policy shock if it simultaneously lowers output, inflation and the money stock, while increasing the nominal interest rate. Hence liquidity and price puzzles are ruled out by construction, at least in the short run. The horizon over which the sign restriction is binding is set to three quarters.

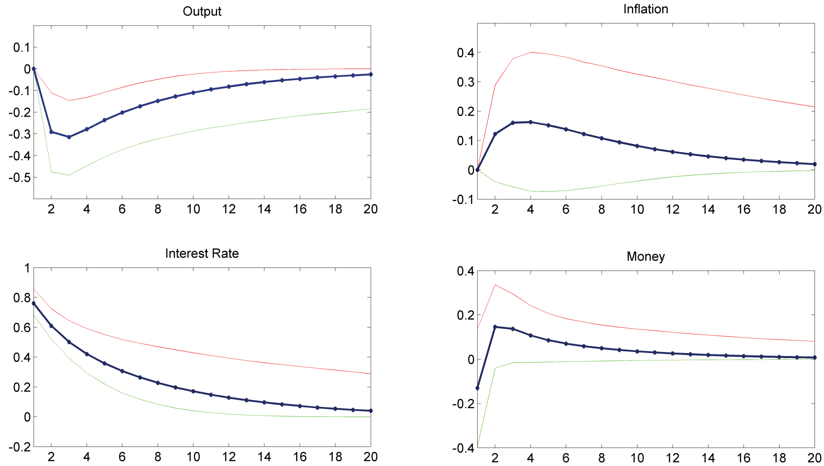
Table 1. Expected theoretical responses from New Keynesian models

	Output	Inflation	Money	Interest Rate
Supply	+	−	−	−
Demand	+	+	+	+
Money	+	+	+	−
Monetary policy	−	−	−	+

We then make another random selection from the posterior distribution of the VAR coefficients and premultiply by an orthogonal rotation matrix to get another decomposition. (Effectively we make a joint draw from the posterior of the VAR coefficients and from the space of rotation matrices). We check the sign restrictions again. We keep

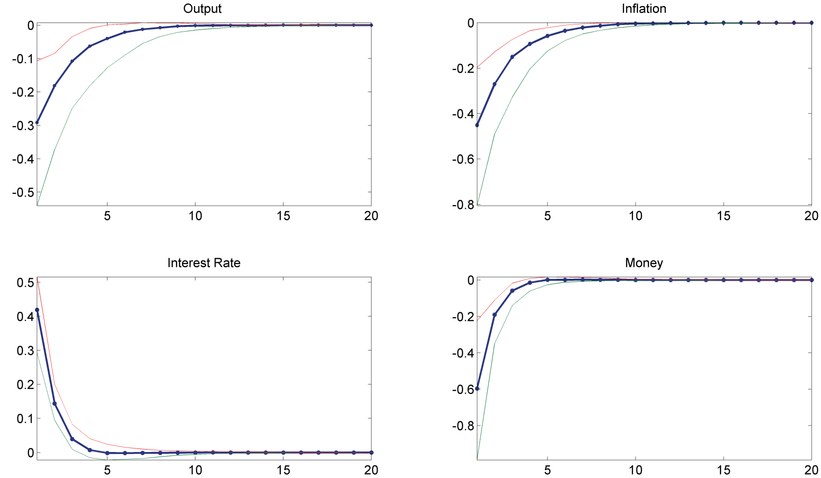
[22] Walsh (2003), in Chapter 1, supplies a wide review of the empirical literature on the relationship between money, prices and output. The Litterman prior confirms the consensus on the short term effects of money. Above all, a money innovation produces hump-shaped movements in real economic activity with a correlation of 1.

the impulse responses that satisfy the restrictions and discard the ones that do not. This process is repeated until a sufficient number of impulse responses have been achieved.



Note: One standard deviation in size. Median and the 95% confidence interval.

Figure 6. Impulse responses to a contractionary monetary policy shock using the Cholesky decomposition



Note: One standard deviation in size. Median and the 95% confidence interval.

Figure 7. Impulse responses to a contractionary monetary policy innovation using the sign restriction approach

The impulse responses from the Cholesky identification scheme exhibit what is known as Liquidity Puzzle: the money stock increases after a tightening of monetary policy. Rafiq and Mallick(2008) found a liquidity puzzle using M1 instead of M3. Furthermore, there is a clear and significant Price Puzzle: the inflation rate rises together with the interest rate. The drop in output however is consistent with the previous VAR literature on the mechanism transmission (e.g. Bagliano and Favero 1999).

Using the sign restrictions (Figure 7), we see that the drop in output is persistent far beyond the horizon over which we impose our sign restriction. It is notable that this identification technique allowed us to compile well-behaved responses of output and of prices to a monetary policy shock. Moreover, these impulse responses are qualitatively consistent with the consensus view on the transmission mechanism that monetary policy has real short-term effects. However, some additional features are worth nothing. The greatest impact, through the liquidity effect, is on money stock. Inflation exhibits more persistence than output[23] as well as appearing less persistent than we found with Cholesky identification; in particular, prices return to their steady state level only after 5 quarters.[24] Money stock and output show similar dynamics, at horizons above 6-8 quarters; the signs of the responses of the interest rate and the money stock are indeterminate. Above all, the sign restriction reduces bias, given that the confidence bands for impulse responses are tight to the median value.

Overall, these results are closest to those reported by previous studies on monetary transmission using VAR methodology. We find that a monetary innovation brings a decline of 30 basis point of GDP,[25] this result is almost invariant across both prior and identification technique; conversely Rafiq and Mallick (2008) found a rise in output following a 1% rise in interest rate.

4.4 Robustness check: alternative estimates

The sample used in our analysis is supposed to be affected by regime changes. These structural breaks may be source of the possibility of structural instability of the estimated model. The findings, obtained from full sample estimation, may fail to uncover the truth identification of shocks; consequently shortening the period under investigation may produce different dynamics.

Among the most important changes were the announcements of two systematic monetary policy regimes. In September 1992, the domestic currency (Italian lira) left the exchange rate mechanism of the European Monetary System. Furthermore, on 1 January 1999 there has been a major change in the way monetary policy is conducted,

[23] This finding is consistent with those found by Rafiq and Mallick (2008), however the results contrast with those found by Mojon and Peersman(2003) who found prices less responsive to tight monetary policy than output.

[24] Ascari and Vaona (2007) using sectoral and regional data for Italy, find significant regional disparities in inflation persistence. They argue that heterogeneous levels of inflation persistence are associated with different local degrees of competitiveness in the retail sector.

[25] There exists a large body of empirical evidence on the impact of monetary policy shocks in Italy. The previous studies found similar effects of a temporary 100 basis points increase in the interest rate on output in Italy for example Mojon and Peersman (2003) −0.12%, BIS (1995) −0.44%, Smets (1997) −0.31%, Barran et al. (1996) −0.30%, Ehrmann et al. (2003) −0.42%, Dedola and Lippi (2005) −1.07%.

the European Central Bank assumed responsibility for monetary policy in the euro area. Both dates can be considered as significant turning points in preferences of monetary policy and in its behavior.

Since Lucas' "critique", it has been recognized that regimes changes impact on the expectations formation mechanisms and hence on the economy's response to policy. For these reasons, we re-estimated the model on a shorter sample beginning in 1999:1.

The findings reported in appendix show that switching to the shorter sample changes significantly the dynamics for the BVAR with Cholesky identification. In particular, the results seem to confirm that there was indeed an important structural break during the nineties, and it might have been the regime change of monetary policy. If we restrict our dataset to contain only observations from 1999 on, the posterior distribution became more volatile and less concentrated around the median (compare Figures 8, 9, 10 with Figures 11, 12 and 13). However, the identification of monetary policy impulses became much easier in the sense that only a small set of possible shock vectors require the sign restrictions.

From the point of view of monetary transmission mechanism, the most important change is the reaction of price level. Indeed, the immediate response to monetary tightening of typical magnitude is negative, and it starts to decline at the end of the first quarter. The pace of the decrease is very fast, the greatest effect (0.15–0.25%, depending on prior imposition) can be observed during the fourth-fifth quarter after the shock. This is in sharp contrast with full sample estimates, where an initial rise in prices was followed by a drop above zero. Because the latter phenomenon occurred only for the full sample with standard identification, we can attribute the bulk of price puzzle to the data prior to 1999.

Furthermore, the response of the output seems to be sensitive across sample choices. The short sample model produces the less smooth decline (it drops immediately to the minimum value of -0.20%) as well as the size of the recession and the beginning of the recovery are quite different: the level of output decreases by roughly 0.2–0.5% within the first two periods after the shock and became positive at the end of the second quarter.

It is worth noting that remarkable difference between estimates on different time span is present only when we adopt the Cholesky identification strategy. Indeed, on data starting in 1999 the sign restriction sets produced almost the same picture that fits the full sample findings. We can conclude, therefore, that the sign identification strategies are a good characterization of monetary policy shocks, also when they are applied on a sample with several structural breaks.

5. Conclusions

This paper presents an empirical analysis of the dynamic behavior over the last twenty years of the Italian economy. More specifically, we derive the monetary policy effects in Italy within a BVAR with sign restriction techniques. This approach has the advantage of overcoming the Cholesky identification. At the same time, sign restrictions eliminate possible puzzles. Moreover, it avoids an *a priori* imposition on the contemporaneous effects, both within the real and nominal variables.

Overall, the Bayesian methodology reduces bias and the standard errors are not wide using the Minnesota prior. However, the BVAR exercise shows that this result is not statistically significant using the Conjugate or Diffuse prior. Furthermore, our study reveals that the coefficient on expected inflation is upwardly biased. One possibility is that the Taylor rule does not accurately describe the way the Monetary Authority conducts monetary policy and that the Central Bank reacts differently to different economic perturbations. Another possible explanation is given by the Cholesky identification, which influences the empirical results.

Policy parameters have qualitative and quantitative implications for the relationship between macro dynamics and structural shocks. When the Central Bank in Italy reacts less strongly to deviations of expected inflation from its target, two different effects take place: on the one hand, inflation returns faster to the target in response to an AS and IS shock; on the other hand, the economy enters a longer recession in response to a cost-push innovation. Furthermore, this analysis highlights the dynamics of the main economic variables in Italy. Our results reveal some information about interaction between monetary policy and the other variables. For instance, we find that prices and output decline after monetary tightening with a bottom response occurring after about one period. The dynamics are quite similar to those discussed by the literature.

Several conclusions can be drawn from this study. First, the results of our model are consistent with the consensus view on the transmission mechanisms. We confirm that the qualitative effects of monetary policy are quite similar to the ones described in the previous empirical studies (see De Arcangelis and Di Giorgio 1999) on short run monetary effects on output, since an expected increase in the interest rate leads to a temporary fall in GDP that peaks around three quarters (in De Arcangelis and Di Giorgio 1999, four quarters with an impact of 30 basis points) after the shock. On the contrary, the heterogeneous results in terms of magnitude and persistence for the dynamic responses of inflation across the identifications and priors denote that sign restriction is the only instrument for identifying truly structural shocks. Thus, this paper, through the sign restriction, overcomes the identification issue associated with the shift in monetary policy that Bernanke (2004) relates to the specific historical period changes in inflation expectations.

It is hard to identify univocally the monetary policy reaction function performed by the Monetary Authority, due to the structural break in the switch of monetary policy regime (from the Bank of Italy to the European Central Bank). Nonetheless, the primary objective of the ECB is to promote price stability throughout the Euro Area and to design and implement monetary policy that is consistent with this objective. In fact at its meeting on 13 October 1998 the Governing Council of the ECB announced: "*... the main elements of the stability-oriented monetary policy strategy of the ESCB. These elements concern: the quantitative definition of the primary objective of the single monetary policy, price stability ... As mandated by the Treaty establishing the European Community, the maintenance of price stability will be the primary objective of the ESCB. Therefore, the ESCB's monetary policy strategy will focus strictly on this* *objective.”*[26] Since the degree of inflation aversion is greater than the weight given to output volatility, our analysis confirms that firstly the Bank of Italy, and then the ECB, have followed the Bundesbank policy of remaining anchored to low inflation.[27]

Another important finding from this study is that, through sign restriction identification, the estimated effect of monetary policy shock on output, prices and money is less persistent than in the standard Cholesky identification. Finally, the dynamic responses of the sign identification seem to be the most robust result across sample choices although it is affected by structural breaks.

In future research, it would be worth testing the robustness of our model for the purpose of forecasting. Another extension would be to estimate the BVAR within an open macroeconomic framework. From an econometric perspective, additional research efforts are needed to provide a macro dynamic model that displays richer and more realistic dynamics, such as adding time varying coefficient into the VAR model in order to explore eventual structural changes.

Acknowledgment I wish to warmly thank professors Fabio Canova and Marco Mazzoli, for their excellent supervision and constant encouragement during this research. I also wish to thank Laura Barbieri, Maurizio Baussola, Marteen Doschee and Vivien Lewis, Martin Uribe, Mario Veneziani, for helpful comments and suggestions. I am also indebted to Nicholas Whithorn for his peerless checking of the text. Finally, I would like to thank the anonymous referees and the editor Roman Horvath for their invaluable corrections, comments and suggestions. I am grateful to the “Bonino-Pulejo Foundation” and the Catholic University for their financial support, which made this research possible. Part of this research has been developed during my visiting research period at Universitat Pompeu Fabra-Barcelona, Spain.

References

Amisano, G. and Giannini, C. (1997a). *Topic in structural VAR Econometrics.* Berlin, Springer Verlag.

Amisano, G., Serrati, M. and Giannini, C. (1997b). Tecniche BVAR per la costruzione di modelli previsivi mensili e trimestrali. Rome, Bank of Italy, Temi di discussione del Servizio Studi Banca d’Italia No. 302.

Angeloni, I., Kashyap, A. K., Mojon, B. and Terlizzese, D. (2003). The Output Composition Puzzle: A Difference in the Monetary Transmission Mechanism in the Euro Area and United States. *Journal of Money Credit and Banking,* 35(6), 1265–1306.

Ascari, G. and Vaona, A. (2007). Regional Patterns of Inflation Persistence in Italy. Pavia, University of Pavia, Quaderni di Dipartimento No. 192(2).

[26] In this context, the Governing Council of the ECB has adopted the following definition: “Price stability shall be defined as a year-on-year increase in the Harmonised Index of Consumer Prices (HICP) for the euro area of below 2%.”

[27] De Arcangelis and Di Giorgio (1999) find that the Bank of Italy’s monetary policy has been based on inflation target. Rotondi and Vaciago (2007) have verified the hypothesis of an ECB with ’Bundesbank preferences’.

Bagliano, F. C., and Favero, C. A. (1999a). Information from Financial Markets and VAR Measures of Monetary Policy. *European Economic Review*, 43(6), 825–837.

Bagliano, F. C., Favero, C. A. and Franco, F. (1999b) Measuring Monetary Policy in Open Economies. London, C.E.P.R., Discussion Paper No. 2079.

Barran, F., Coudert, V. and Mojon B. (1996). The Transmission of Monetary Policy in the European Countries. Paris, CEPII Research Center, Working Paper No. 3.

Benigno, P. and Lopez-Salido, J. D. (2006). Inflation Persistence and Optimal Monetary Policy in the Euro Area. *Journal of Money, Credit and Banking*, 38(3), 587–614.

Bernanke, B. S. and Blinder, A. (1992). The Federal Funds Rate and the Channels of Monetary Transmission. *American Economic Review*, 82(4), 901–21.

BIS (1995). *Financial Structure and the Monetary Policy Transmission Mechanism*. Basle, Bank for International Settlements.

Blanchard, O. J. and Quah, D. (1989). The Dynamic Effects of Aggregate Demand and Supply Disturbances. *American Economic Review*, 79(4), 655–673.

Campos, J., Ericsson, N. R. and Hendry, D. F. (1996). Cointegration Tests in the Presence of Structural Breaks. *Journal of Econometrics*, 70, 187–220.

Canova, F. (2007). *Methods for Applied Macroeconomic Research.* Princeton, Princeton University Press.

Canova, F. and De Nicolò, G. (2002). Monetary Disturbances Matter for Business Fluctuations in the G-7. *Journal of Monetary Economics*, 49(6), 1131–59.

Canova, F. and Gambetti, L. (2009). Structural Changes in the U.S. Economy: Bad Luck or Bad Policy?" *Journal of Economic Dynamics and Control*, 33(2), 477–490.

Canova, F. and Pina, J. P. (1999). Monetary Policy Misspecification in VAR models. London, C.E.P.R., Working Paper No. 2333.

Canova, F. and Sala, L. (2006). Back to Square One: Identification Issues in DSGE Models. Frankfurt am Main, European Central Bank, Working Paper No. 583.

Castelnuovo, E. and Surico, P. (2006). The Price Puzzle: Fact or Artefact? London, Bank of England, Working Paper No. 288.

Christiano, L. J., Eichenbaum, M. and Evans, C. L. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy,* 113(1), 1–45.

Ciccarelli, M. and Rebucci, A. (2006). Has the Transmission Mechanism of European Monetary Policy Changed in the Run-Up to EMU? *European Economic Review*, 50(3), 737–776.

Clarida, R., Gali, J. and Gertler, M. (1999). The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature,* 37(4), 1661–707.

Cogley, T. and Sargent, T. (2005). The Conquest of U.S. Inflation: Learning and Robustness to Model Uncertainty. *Review of Economic Dynamics*, 8(2), 528–563.

Cooley, T. F. and Dwyer, M. (1998). Business Cycle Analysis without Much Theory: A Look at Structural VARs. *Journal of Econometrics*, 83(1–2), 57–88.

De Arcangelis, G. and Di Giorgio, G. (1999). *Monetary Policy Shocks and Transmission in Italy: A VAR Analysis*. Rome, Luiss University Press.

Dedola, L. and Lippi, F. (2005). The Monetary Transmission Mechanism: Evidence from the Industries of Five OECD Countries. *European Economic Review*, 49(6), 1543–1569.

Doan, T., Litterman, R. and Sims, C. (1984). Forecasting and Conditional Projection Using Realistic Prior Distributions. *Econometric Reviews*, 3(1), 1–100.

Eichenbaum, M. (1992). Comments on "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy". *European Economic Review*, 36(5), 1001–1011.

Eichenbaum, M. and Evans, C. (1995). Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates. *Quarterly Journal of Economics*, 110(4), 975–1009.

Ehrmann, M., Gambacorta, L., Martinez-Pagès, J., Sevestre, P. and Worms, A. (2003). The Effects of Monetary Policy in the Euro Area. *Oxford Review of Economic Policy*, 19(1), 58–72.

Fabiani, S., Gattulli, A. and Sabbatini, R. (2004). The Pricing Behaviour of Italian Firms: New Survey Evidence on Price Stickiness. Frankfurt am Main, European Central Bank, Working Paper No. 3.

Galì, J. and Gertler, M. (1999). Inflation Dynamics: A Structural Econometric Analysis. *Journal of Monetary Economics,* 44(2), 195–222.

Galì, J., Gertler, M. and Lopez-Salido, J. D. (2003). European Inflation Dynamics. *European Economic Review*, 47(4), 759–60.

Grilli, V. and Roubini, N. (1995). Liquidity and Exchange Rates: Puzzling Evidence from the G-7 Countries. New York, New York University, Working Paper No. 95-17.

Hamilton, J. (2004). *Time Series Analysis*. Princeton, Princeton University Press.

Lippi, M. and Reichlin, L. (1994). VAR Analysis, Non-fundamental Representations, Blaschke Matrices. *Journal of Econometrics*, 63(1), 307–325.

Litterman, R. B. (1983). A Random Walk, Markov Model for the Distribution of Time Series. *Journal of Business and Economic Statistics*, 1(2), 169–173.

Litterman, R. B. (1986). Forecasting with Bayesian Vector Autoregressions: 5 Years of Experience. *Journal of Business and Economic Statistics*, 4(1), 25–38.

Marimon, R. and Scott, A. (2001). *Computational Methods for the Study of Dynamic Economies.* Oxford, Oxford University Press.

Mojon, B. and Peersman, G. (2003). A VAR Description of the Effects of Monetary Policy in the Individual Countries of the Euro Area. In Angeloni, I., Kashyap, A. K. and Mojon, B. (eds.), *Monetary Policy Transmission in the Euro Area*. Cambridge, Cambridge University Press, 56–74.

Rafiq, M. S. and Mallick, S. K. (2008). The Effect of Monetary Policy on Output in EMU: A Sign Restriction Approach. *Journal of Macroeconomics*, 30(4), 1756–1791.

Robertson, J. C. and Tallman, E. W. (1999). Prior Parameter Uncertainty: Some Implications for Forecasting and Policy Analysis with VAR models. Atlanta, Federal Reserve Bank of Atlanta, Working Paper No. 99-13.

Rotondi, Z. and Vaciago, G. (2007). Lessons from the ECB Experience: Frankfurt Still Matters! *Economic Notes*, 36(2), 147–170.

Smets, F. (1997). Measuring Monetary Policy Shocks in France, Germany and Italy: The Role of The Exchange Rate. *Swiss Journal of Economics and Statistics*, 133(III), 597–616.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, 48(1), 1–48.

Sims, C. A. (1992). Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy. *European Economic Review*, 36(5), 975–1000.

Sims, C. A. (1993). A Nine Variable Probabilistic Macroeconomic Forecasting Model. Minneapolis, Federal Reserve Bank of Minneapolis, Institute for Empirical Macroeconomics, Discussion Paper No. 14.

Sims, C. A. (2002). The Role of Models and Probabilities in the Monetary Policy Process. *Brookings Papers on Economic Activity*, 2, 1–62.

Uhlig, H. (2005). What Are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure. *Journal of Monetary Economics*, 52(2), 381–419.

Walsh, C. E. (2003). *Monetary Theory and Policy.* Cambridge MA, The MIT Press.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, Princeton University Press.

Appendix A: Sensitivity Analysis

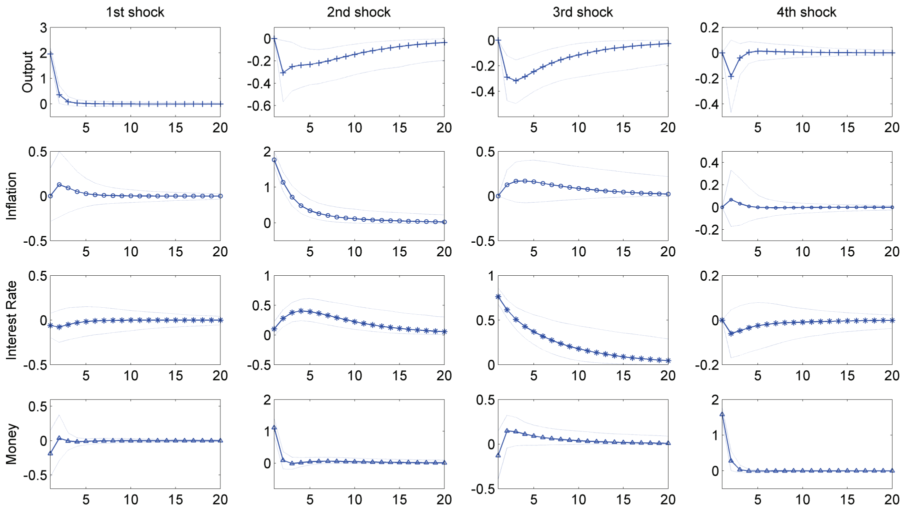


Figure 8. Impulse responses to an exogenous shock in real GDP, prices, interest rate and money stock: Conjugate prior and full sample

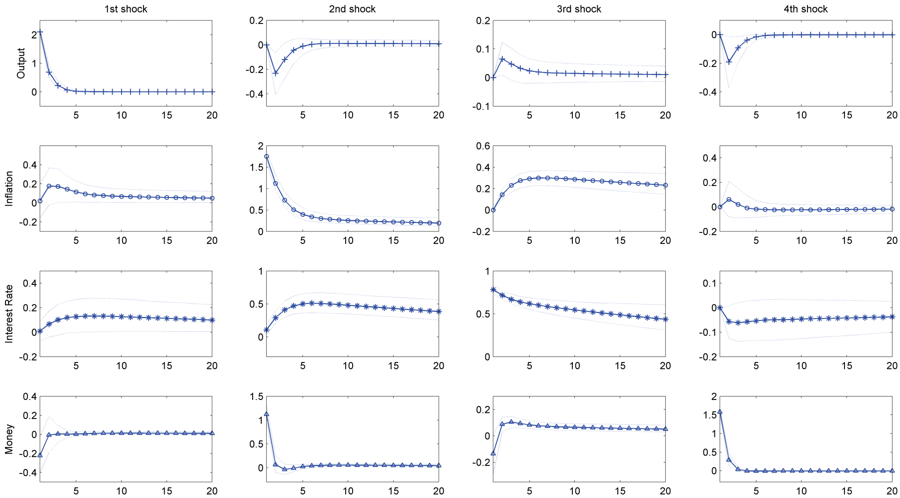


Figure 9. Impulse responses to an exogenous shock in real GDP, prices, interest rate and money stock: Diffuse prior and full sample

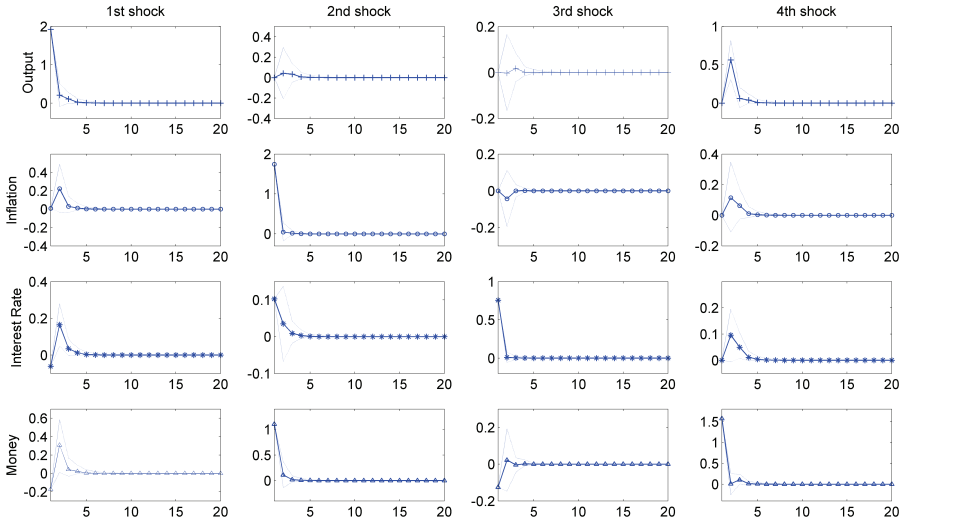


Figure 10. Impulse responses to an exogenous shock in real GDP, prices, interest rate and money stock: Litterman prior and full sample

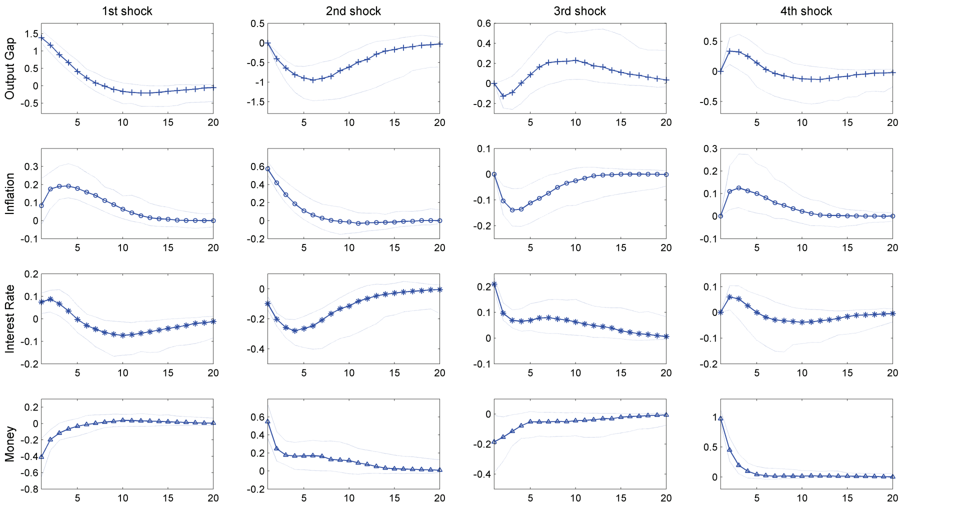


Figure 11. Impulse responses to an exogenous shock in real GDP, prices, interest rate and money stock: Conjugate prior and short sample (1999 onwards)

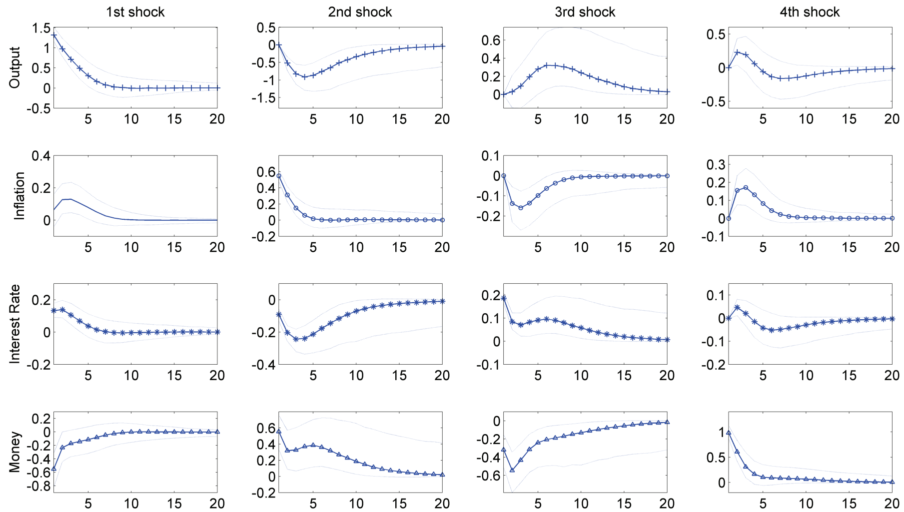


Figure 12. Impulse responses to an exogenous shock in real GDP, prices, interest rate and money stock: Diffuse prior and short sample (1999 onwards)

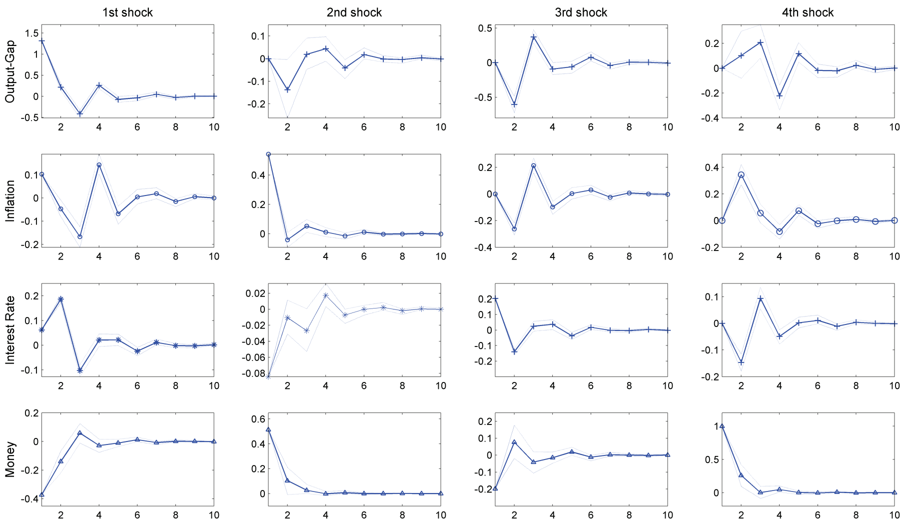


Figure 13. Impulse responses to an exogenous shock in real GDP, prices, interest rate and money stock: Litterman prior and short sample (1999 onwards)

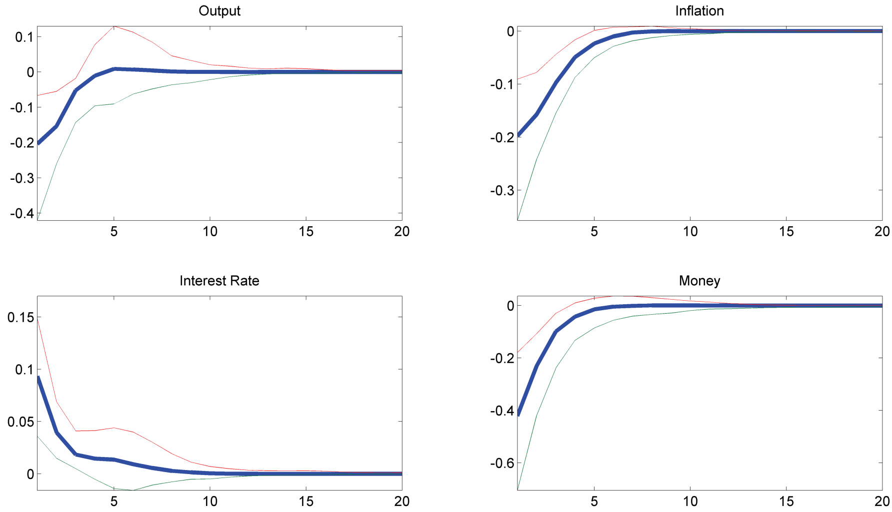


Figure 14. Impulse responses to a contractionary Monetary Policy innovation: the sign restriction approach and short sample (1999 onwards)

Appendix B: Bayesian inference

Given that the Vector Autoregression model: $y = (X \otimes I_k)\,\alpha + u$. The standard routine applies OLS to estimate the unknown coefficients α and Σ. The BVAR tries to improve the estimation of α, by introducing into the analysis, prior information i.e. on the statistical or economic nature of the sample. The information on the parameter α is represented by the prior density function $g(\alpha,\Sigma)$. In this context, the sample likelihood is given by the conditional density function $g(Y,X|\alpha,\Sigma)$, applying Bayes' rule:

$$\begin{aligned} g(\alpha,\Sigma|Y) &= \frac{g(Y|\alpha,\Sigma)\,g(\alpha,\Sigma)}{g(Y)} \\ &= \frac{l(Y|\alpha\,,\Sigma)\,g(\alpha,\Sigma)}{g(Y)} \rightarrow g(\alpha,\Sigma|Y) \propto l(Y|\alpha\,,\Sigma)\,g(\alpha,\Sigma) \end{aligned}$$

Therefore, the posterior is proportional to the prior times the likelihood, where the likelihood function is: $l(Y|\alpha,\Sigma) \propto |\Sigma|^{-T\frac{1}{2}} \exp\left\{-\frac{1}{2}\sum_t \Sigma^{-1}(Y - X\alpha)'(Y - X\alpha)\right\}$

The marginal posterior $g(\alpha|Y)$ and g$(\Sigma|Y)$ can be obtained by integrating out α and Σ from $g(Y,X|\alpha,\Sigma)$. Once, we have assumed that u is normally distributed, we need only to choose a prior, represented by the prior density function $g(\alpha,\Sigma)$.

Minnesota Prior

The Minnesota prior is a particular case when it is assumed that α has a normal distribution: $g(\alpha) \sim N\left(\overline{\alpha},\overline{\Sigma}_a\right)$, where $\overline{\alpha}$ and $\overline{\Sigma}_a$are the prior mean and the prior variance-covariance matrix of α.

Given that the assumption of gauss distribution for α, the probability density function of the data conditional on the parameters of the model (the likelihood) is:
$l(Y|\alpha,\Sigma_e) \propto |\Sigma_e|^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}\Sigma_e^{-1}(Y - X\alpha)'(Y - X\alpha)\right\}$, the residual variance-covariance matrix (Σ_e) is assumed fixed and diagonal. Let the prior be $\alpha = \overline{\alpha} + v_a$ with $v_a \sim N(0,\overline{\Sigma}_a)$, where α and $\overline{\Sigma}_a^{-1}$ are function of a small number of hyper-parameters. Set $\overline{\alpha} = 0$ except for $\overline{\alpha}_{i1} = 1$, Σ_a the variance-covariance matrix is a priori diagonal with element $\sigma_{ij,l}$ is the lag l of variable j in equation i is:

$$\begin{cases} \dfrac{\varphi_0}{h(l)} & \text{if } i = j \\ \varphi_0 \times \dfrac{\varphi_1}{h(l)} \times \left(\dfrac{\sigma_j}{\sigma_i}\right)^2 & \text{if } i \neq j,\ j \text{ endogenous variable} \\ \varphi_0 \times \varphi_2 & \text{if } i \neq j,\ j \text{ exogenous variable} \end{cases}$$

Therefore, the variance-covariance matrix is a-priori selected to be diagonal.

Let φ_i be the hyper-parameters measuring the tightness of the prior. If $\varphi_i \rightarrow \infty$, the prior becomes less important and the posterior estimates converge to the OLS estimates. On the contrary a small value for φ_i represents greater confidence in the prior information and will force the parameter estimates to be closer to the values predicted by the random walk assumption. In particular, φ_0 is the tightness on the variance of the first lag, φ_1 represents the tightness of other variables, φ_2 is the relative tightness of the exogenous variables, $h(l)$ is the tightness of the variance of lags other than the first one. The scaling factors $(\sigma_j/\sigma_i)^2$ describe the difference in scale and variability of the data; we set the scaling factors equal to the variance of the residuals from a univariate autoregressive model for the variables. Since the variance-covariance matrix is diagonal, there is no relationship among the various coefficients of the VAR system. The posterior density of α, $g(\alpha|y)$ is:

$$g(\alpha|Y) \propto \exp\left\{-0.5\left[\left(\alpha' - \Sigma_a\alpha\right)'\tilde{\Sigma}_a^{-1}\left(\alpha' - \Sigma_a\alpha\right)\right]\right\},$$

where $\alpha = \Sigma_a\left(\overline{\Sigma}_a^{-1} + \Sigma_e^{-1}X'Y\right)$ and $\Sigma_a = \left(\overline{\Sigma}_a^{-1} + \Sigma_e^{-1}X'X\right)^{-1}$. The calibration for theta has been achieved following the default values suggested by RATS manual:

$\theta_0 = 0.2$	controls the relative importance of sample and prior information
$\theta_1 = 1$	sets the relative importance for lags of other variables
$\theta_2 = 1000$	regulates the relative importance of the information contained in the exogenous variables

This calibration implies a relatively lose prior and an informative prior for the exogenous variables. Note that the Minnesota prior requires seasonally adjusted data.

The Conjugate prior (Normal-Wishart)

The Conjugate prior is very common within the Bayesian applied econometrics, because it has two main advantages: it is easy to interpret and to calculate, since, the posterior distribution follows the same parametric form as the prior distribution and therefore the prior information can be interpreted in the same way as likelihood function information. Moreover the Conjugate prior allows us to overcome the assumption of a fixed and diagonal variance-covariance matrix of the error terms solving the two main weaknesses in the Litterman prior: the posterior independence between equations and the fixed residual variance-covariance matrix. The Conjugate prior assumes a normal Wishart distribution: $g(\alpha|\Sigma) = N\left(\overline{\alpha}, \left|\overline{\Sigma}\otimes\overline{\Omega}\right.\right)$, and $g(\Sigma) = iW\left(\overline{\Sigma}, \alpha\right)$. The unconditional prior distribution of α is $\alpha = N\left(\overline{\alpha}, (m-n-1)^{-1}\overline{\Sigma}\otimes\overline{\Omega}\right)$, where m denotes the degrees of freedom of the inverse-Wishart, with $m > n+1$. The posterior distribution is given by:

$$g(\alpha|\Sigma, Y) = N\left(\tilde{\alpha}, \Sigma\otimes\tilde{\Omega}\right) \text{ and } g(\Sigma|Y) = iW\left(\tilde{\Sigma}, T+m\right),$$

where $\tilde{\Omega} = \left(\overline{\Omega}^{-1} + X'X\right)^{-1}$, $\tilde{\alpha} = \tilde{\Omega}\left(\overline{\Omega}^{-1}\overline{\alpha} + X'X\alpha_{OLS}\right)$ and $\tilde{\Sigma} = \alpha'_{OLS}X'X\alpha_{OLS} + \alpha'\overline{\Omega}^{-1}\overline{\alpha} + \overline{\Sigma} + (Y - X\alpha_{OLS})'(Y - X\alpha_{OLS}) - \tilde{\alpha}'\left(\overline{\Omega}^{-1} + X'X\right)\tilde{\alpha}$.

The likelihood function is:

$$g(\alpha, \Sigma_e) \propto |\Sigma_e \otimes I_T|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}[y-(I_m \otimes X)\alpha]' \left(\Sigma_e^{-1} \otimes I_T\right)[y-(I_m \otimes X)\alpha]\right\}$$

As in Canova (2007), a useful decomposition is given:

$$g(\alpha, \Sigma_e) \propto N(\alpha \,|\, \alpha_{OLS}, \Sigma_e, X, y) \times W\left(\Sigma_e^{-1} \,|\, y, X, \alpha_{OLS}, T-k-m-1\right)$$

Therefore, the Conjugate prior involves a normal distribution for α conditional on α_{OLS} and Σ_e, and a Wishart distribution for Σ_e^{-1} the normal inverse Wishart distribution is extensively used as it estimates a covariance matrix assuming it is a multivariate-normal covariance matrix.

The Normal Diffuse or Jeffrey prior:

The Normal Diffuse prior was introduced by Zellner (1971). As for the Conjugate prior, the Jeffrey prior has the advantage of allowing a non diagonal residual variance-covariance matrix, furthermore it avoids the Normal-Wishart class of restrictions on the variance-covariance matrix of α.

The Jeffrey prior assumes a prior independence between α and Σ, thus for α it is $g(\alpha)$=N$\left(\overline{\alpha}, \overline{\Omega}\right)$ and for Σ it is $g(\Sigma) = |\Sigma|^{-\frac{1}{2}(n+1)}$ where n is the number of variables inserted into the system. The posterior is the combination of the prior with the data and thus:

$$g(\alpha, \Sigma \,|\, Y) = g(\alpha \,|\, \Sigma, Y)\, g(\Sigma \,|\, Y),$$

where $g(\alpha \,|\, \Sigma, Y) = N\left(\tilde{\alpha}, \tilde{\Omega}\right)$, $g(\Sigma \,|\, \alpha, Y) \sim W\left(\tilde{\Sigma}^{-1}, T\right)$, $\tilde{\Omega} = \left(\overline{\Omega}^{-1} + \left(\Sigma^{-1} \otimes X'X\right)\right)^{-1}$, $\tilde{\Sigma} = (Y - X\alpha_{OLS})'(Y - X\alpha_{OLS}) + (\alpha - \alpha_{OLS})' X'X(\alpha - \alpha_{OLS})$, $\tilde{\alpha} = \tilde{\Omega}(\overline{\Omega}^{-1}\overline{\alpha} + (\Sigma^{-1} \otimes X'X)\alpha_{OLS})$ with $W\left(\tilde{\Sigma}^{-1}, T\right)$ representing the inverse of the inverted Wishart distribution. This prior is quite useful for single-parameter problems but can be seriously lacking in multi-parameter settings.