New Information Response Functions and Applications to Monetary Policy

Caroline JARDET [(1)]
Banque de France

Alain MONFORT [(2)]
CREST, Banque de France
and Maastricht University

Fulvio PEGORARO[(3)]
Banque de France, CREST
and HEC Lausanne

This version : April, 2013.

Abstract

We propose a general methodology, built on the non ambiguous notion of statistical innovation of a VAR or VARMA model, in order to study the dynamic effects of a "new information", on a set of variables of interest, in a more flexible way than traditional impulse response functions. This new information can be related not only to the values, the signs, the range of one or several components of the innovation, or to the value of one or several average responses, but also to a linear filter of this innovation, or to the future paths of some of its components or associated linear filters or both. The methodology is called New Information Response Functions (NIRF) and is shown to encompass standard methodologies typically adopted in the literature. In order to illustrate the usefulness of the NIRF methodology we estimate a Gaussian VAR(p) model estimated on U.S. quarterly data and we address a current and relevant monetary policy issue: we compare the effects on the future long-rate as well as on its expectation term and term premium components, of alternative kinds of FOMC stabilization announcements (like the one at the end of 2008) involving the future path of the short rate and its expectations around the zero lower bound.

Keywords: impulse response functions, innovation, new information, linear filter, future path, forward policy guidance, stabilization announcements, zero lower bound.

JEL Codes: C10, C32, E52, E58.

---

[1]Banque de France, Financial Economics Research Service [DGEI-DEMFI-RECFIN; E-mail: Caroline.JARDET@banque-france.fr].

[2]CNAM and CREST, Laboratoire de Finance-Assurance, Banque de France, Financial Economics Research Service [E-mail: monfort@ensae.fr].

[3]Banque de France, Financial Economics Research Service [DGEI-DEMFI-RECFIN; E-mail: Fulvio.PEGORARO@banque-france.fr] and CREST, Laboratoire de Finance-Assurance [E-mail: pegoraro@ensae.fr].

We would like to thank Mike Chernov, Hans Dewatcher, Domenico Giannone, Christian Gourieroux, Stephane

1 Introduction

The pioneering paper by Sims (1980) has triggered a large literature on the definition of shocks and impulse response functions in VAR or VARMA models. Since then, this literature has continuously evolved notably in response to recurrent issues related to the identification of structural shocks (see the survey by Kilian (2011)). A first part of this literature is devoted to the identification of structural VAR models by short-run restrictions (see Blanchard and Watson (1986), Bernanke (1986), Bernanke and Mihov (1998), among others) while, a second part promotes identification by long-run restrictions or by a mix of short- and long-run restrictions (see Blanchard and Quah (1989), Gali (1992)). Both approaches rely on assumptions, mainly based on economic theory, such as the ordering of the variable in the VAR or, more generally, the expected short-run or long-run effects that a given shock should have on a given variable. However, these assumptions are not always consensual and this lack of general agreement leads to different response functions which make difficult to bring out a clear economic message (see for instance Lütkepohl (1991) and Cochrane (1994)). The shortcomings of the traditional identifying assumptions have spurred the development of statistical (sometimes called "agnostic") approaches. Notably, some authors have proposed to identify structural shocks by means of sign restrictions (see Uhlig (2005), Mountford and Uhlig (2009), Peersman and Straub (2009), Inoue and Kilian (2011) among others). Alternatively, some other authors have focused on the notion of statistical innovation of a given variable in the VAR model (see Pesaran and Shin (1998)).

In this paper we propose a general methodology, built on the non ambiguous notion of statistical innovation ε_t (say) of a VAR (or VARMA) stochastic process Y_t (that is, the difference between the value of the process and its conditional expectation given its past), giving the possibility to study impulse responses in a more flexible way than traditional impulse response functions (IRFs). We aim at exploiting the fact that we may naturally have at our disposal, for a given analysis,

Grégoir, Chris Otrok and Glenn Rudebusch, as well as participants at the June 2009 Bank of France - Bundesbank Conference on "The Macroeconomy and Financial Systems in Normal Times and in Times of Stress", the NASM 2011 Conference in St. Louis, the Banque de France October 2011 seminar, the October 2011 FUNDP Namur workshop "Financial markets and macroeconomic policies in the aftermath of crisis", for helpful comments and remarks.

an information on this innovation process which could be different from an information on the contemporaneous realization of one of its components, or different from an information providing an exact identification (structural shocks). Indeed, this "new information" can be related not only to the values, the signs, the ranges of one or several components of ε_t, or to the value of one or several average responses to a new information. It can also be related to the innovations of *linear filters* on the variables of interest, or to the *future path of some of these innovations*. In addition, these future paths can be effective or hypothetical (i.e., counterfactuals or scenarios): for instance one may want to analyze and compare forecast scenarios associated with alternative possible future paths of some of the innovations or of some of their linear filters. In all of these situations, one may be interested in analyzing the expected dynamical effects of such a "new information" on any of the variables of interest. In this paper we develop a statistical setting which is suitable to deal with these issues. This general methodology is named *New Information Response Function*.

More formally, we assume that, given the observed values $\underline{Y_{t-1}}$ (say) of our process of interest, a genuine or hypothetical new information on the future path $\underline{Y_{t,T}} = (Y_t', Y_{t+1}', \ldots, Y_T')'$ becomes available. This new information, denoted I_t, is defined by some function $a(\underline{Y_{t,T}})$ of the future path and it can also be written as $I_t = \widetilde{a}(\underline{\varepsilon_{t,T}}, \underline{Y_{t-1}})$, where $\underline{\varepsilon_{t,T}}$ denotes the corresponding sequence of future innovations. For computation purposes, we then distinguish the case $T = t$, where the new information I_t involves only $\underline{Y_t}$, namely ε_t and $\underline{Y_{t-1}}$, from the general case $T > t$ where I_t concerns the future path of the process and, possibly, the future values of some linear filters of that process or the values of other linear filters belonging to some interval. In the former case, the response function is obtained calculating a conditional expectation depending on the moving average representation of the process and on the information we have about ε_t. In order to detail the flexibility of the NIRF approach, we provide many examples of both kind of elements. In the latter case, we show that the response function can be calculated through a Kalman filter iteration, even if the new information contains future values of some linear filters of the process. In addition, if this new information concerns the future values of other linear filters belonging to a given interval, we provide a computationally fast simulation procedure based on a useful property of the Gaussian vectors and on Kalman filter iterations [Durbin and Koopman (2002)]. Indeed, we show that the VAR model equipped with the new information $\widetilde{a}(\underline{\varepsilon_{t,T}}, \underline{Y_{t-1}})$ can still be represented in a State Space form.

Finally, in order to provide an empirical illustration of the usefulness of the NIRF methodology, we address a current and relevant monetary policy issue using a Gaussian VAR(p) model that links macroeconomic variables and interest rates (namely, short rate, one-year spread, GDP growth and inflation rate). More precisely, in order to illustrate the impact of a new information on "future paths", we consider the stabilization announcement by the Federal Open Market Committee (FOMC) in 2008:Q4 to *warrant exceptionally low levels of the federal fund rate for some times*, and we investigate how the long rate, as well as its expectation component and term premium, respond to alternative kinds of scenarios involving not only the future path of the short-term rate but, possibly, also the future path of its expectation component. In other words, we show that the NIRF can be adopted to suggest a preferred monetary policy communication strategy. Our empirical exercise highlight that, if the purpose of the monetary authority is to keep the long rate stable, an *accurate and unambiguous* commitment on the future path of the short-term interest rates seems to be the successful choice.

The paper is organized as follows. In Section 2 we provide the definition and the computation of the new information response function for the case $T = t$ and the general case $T > t$. In Section 3 we present several cases of NIRFs depending on the assumption we make about the function $\widetilde{a}(.)$. Starting with the case $T = t$ and $I_t = \widetilde{a}(\varepsilon_t)$, we first mention the so-called "full new information" case ($\widetilde{a}(\varepsilon_t)$ is one-to-one and it contains standard orthogonalized shocks, the impulse vectors introduced by Uhlig (2005) and the structural shocks as particular cases), and we present the "continuous partial new information" case that includes the "generalized" impulse response function introduced by Pesaran and Shin (1998), or other kinds of information, like information on a subset of innovations, information on responses or a combination of both (Section 3.1). Second, we describe the "discrete partial new information" case, where the new information $\widetilde{a}(\varepsilon_t)$ is based on discrete functions, like indicator functions and, in particular, sign functions (Section 3.2). Third, we present the case of a new information on a linear filter where the new information can still be represented by $\widetilde{a}(\varepsilon_t)$ (Section 3.3). Finally, in Section 3.4 we mention the general case $T > t$ and $I_t = \widetilde{a}(\underline{\varepsilon_{t,T}}, \underline{Y_{t-1}})$ that is adopted in Section 4 to develop the empirical application to forward policy guidance. Section 5 concludes and proposes further developments, while proofs of propositions and some theoretical results are gathered in appendices.

2 Definition and Computation of the New Information Response Function (NIRF)

2.1 The Model

Let us consider a n-dimensional VAR(p) process Y_t satisfying:

$$\Phi(L)Y_t = \nu + \varepsilon_t \, , \tag{1}$$

where $\Phi(L) = I + \Phi_1 L + ... + \Phi_p L^p$, L being the lag operator; ε_t is the n-dimensional Gaussian innovation process of Y_t with distribution $N(0, \Sigma)$. We do not necessarily assume that Y_t is stationary, so we have to assume some starting mechanism, defined by the initial values $(y'_{-1}, y'_{-2}, ..., y'_{-p})' \equiv \underline{y_{-p}}$. By considering the recursive equations:

$$Y_\tau = \nu - \Phi_1 Y_{\tau-1} - \ldots - \Phi_p Y_{\tau-p} + \varepsilon_\tau \, , \tag{2}$$

at $\tau \in \{0, ..., t\}$ and eliminating $Y_0, ..., Y_{t-1}$ we get a moving average representation of the form:

$$Y_t = \mu_t + \sum_{\tau=0}^{t} \Theta_\tau \varepsilon_{t-\tau} \, , \tag{3}$$

where μ_t is a function of t and $\underline{y_{-p}}$, and the sequence Θ_τ is such that:

$$\left[\left(\sum_{i=0}^{p} \Phi_i L^i \right) \left(\sum_{\tau=0}^{t} \Theta_\tau L^\tau \right) \right]_t = I \, , \tag{4}$$

(with $\Phi_0 = I$ and where $[\,\centerdot\,]_t$ is a notation for the polynomial obtained by retaining only the terms of degree smaller than or equal to t from the polynomial between brackets) which implies,

$$\begin{aligned} \Theta_0 &= I \text{ and} \\ \Theta_\tau &= -\sum_{i=1}^{\tau} \Phi_i \Theta_{\tau-i}\,, \tau \geq 1\,, \end{aligned} \tag{5}$$

with $\Phi_i = 0$ if $i > p$. Equation (5) provides a straightforward way to compute recursively the matrices Θ_τ. Denoting $\underline{Y_t} = (Y_t', Y_{t-1}', ..., Y_{t-p}')'$, equation (3) implies:

$$E(Y_{t+h}\,|\,\underline{Y_t}) - E(Y_{t+h}\,|\,\underline{Y_{t-1}}) = \Theta_h\,\varepsilon_t\,; \tag{6}$$

so $\Theta_h\,\varepsilon_t$ measures the differential impact of the knowledge of ε_t, or Y_t, on the prediction updating of Y_{t+h} when $\underline{Y_{t-1}}$ is known.

2.2 Definition of the NIRF

Let us now assume that $\underline{Y_{t-1}}$ has been observed and a genuine or hypothetical new information on the future path $\underline{Y_{t,T}} = (Y_t', Y_{t+1}', \ldots, Y_T')'$ is considered. This new information, denoted I_t, is defined by some function $a(\underline{Y_{t,T}})$ of the future path and it can also be written as $I_t = \widetilde{a}(\underline{\varepsilon_{t,T}}, \underline{Y_{t-1}})$, where $\underline{\varepsilon_{t,T}} = (\varepsilon_t', \ldots, \varepsilon_T')'$ is the sequence of the future innovations. The *New Information Response Function* (NIRF) is defined by:

$$\begin{aligned} NIRF_{t,t+h}(I_t) &= E(Y_{t+h}\,|\,I_t, \underline{Y_{t-1}}) - E(Y_{t+h}\,|\,\underline{Y_{t-1}}) \\ &= E(Y_{t+h}\,|\,\widetilde{a}(\underline{\varepsilon_{t,T}}, \underline{Y_{t-1}}), \underline{Y_{t-1}}) - E(Y_{t+h}\,|\,\underline{Y_{t-1}})\,, \quad h \in \{1, \ldots, H\}\,, \end{aligned} \tag{7}$$

and thus $NIRF_{t,t+h}(I_t)$ measures the differential impact on the prediction of Y_{t+h} of the additional knowledge of I_t, when $\underline{Y_{t-1}}$ has been observed.

2.3 Computation of the NIRF

We distinguish the case $T = t$, where the new information I_t involves only $\underline{Y_t}$, namely ε_t and $\underline{Y_{t-1}}$, and the general case $T > t$ where $I_t = a(\underline{Y_{t,T}})$ involves the future path of the process and, possibly, of future filters of that process.

i) <u>CASE $T = t$</u>

In this case we have $I_t = a(\underline{Y_t}) = \widetilde{a}(\varepsilon_t, \underline{Y_{t-1}})$, and the $NIRF_{t,t+h}(I_t)$ is easily computed in the following way.

PROPOSITION 1: *If the new information is $I_t = \widetilde{a}(\varepsilon_t, \underline{Y_{t-1}})$, we have $NIRF_{t,t+h}(I_t) = \Theta_h \, \delta_t$, where $\delta_t = E(\varepsilon_t \,|\, \widetilde{a}(\varepsilon_t, \underline{Y_{t-1}}))$. Proof: see Appendix 1.*

An frequent particular case is when the new information I_t does not involve the observed past values $\underline{Y_{t-1}}$ and is given by $I_t = \widetilde{a}(\varepsilon_t)$. In this case $NIRF_{t,t+h}(I_t) = \Theta_h \, \delta$, where $\delta = E(\varepsilon_t \,|\, \widetilde{a}(\varepsilon_t))$.

ii) <u>CASE $T > t$</u>

Let us assume now that the new information $I_t = a(\underline{Y_{t,T}})$ is given by:

$$I_t = a(\underline{Y_{t,T}}) = [a_\tau(Y_{t-J,\tau}), \tau \in D_t \subset \{t, \ldots, T\}] \tag{8}$$

where $a_\tau(\underline{Y_{\tau-J,\tau}})$, for any $\tau \in D_t$, denotes a date-τ *filter of order* J of Y_τ, i.e. a function of $Y_\tau, Y_{\tau-1}, \ldots, Y_{\tau-J}$. Thus, the new information consists in the knowledge of *future filters* of the variables $Y_t, \ldots, Y_T$. Let us introduce the notations $K = \max(p, J, T-t)$ and $Z_\tau = (Y'_\tau, \ldots, Y'_{\tau-K})'$. For any $\tau \in \{t, \ldots, T\}$, $a_\tau(\underline{Y_{\tau-J,\tau}})$ can be rewritten as $a^*_\tau(Z_\tau)$ and model (1) can be equivalently represented as:

$$Z_t \;\; = \;\; \nu^* + \Phi^* \, Z_{t-1} + \varepsilon^*_t \tag{9}$$

with

$$\nu^* = \begin{pmatrix} \nu \\ 0 \end{pmatrix}, \ \varepsilon_t^* = \begin{pmatrix} \varepsilon_t \\ 0 \end{pmatrix}, \ \Phi^* = \begin{pmatrix} \Phi_1 & \Phi_2 & \dots & \Phi_p & 0 \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \ddots & 0 & 0 \\ 0 & 0 & \dots & I & 0 \end{pmatrix}. \tag{10}$$

For $h \in \{1, \dots, T-t\}$, the $NIRF_{t,t+h}(I_t)$ can be computed in the following way.

Proposition 2: *Let us consider the State Space model defined by the transition and measurement equations:*

$$\begin{cases} Z_\tau &= \nu^* + \Phi^* Z_{\tau-1} + \varepsilon_\tau^* \\ X_\tau &= a_\tau^*(Z_\tau), \ \ \tau \in D_t, \end{cases} \tag{11}$$

where X_τ is the known value of $a_\tau^*(Z_\tau)$ and where the initial conditions Z_{t-1} are the observed values. The value of $E(Y_{t+h} \,|\, I_t, \underline{Y_{t-1}})$ is the component of the filtered value $\widehat{Z}_{T|T}$ (say) of Z_T corresponding to Y_{t+h} and $NIRF_{t,t+h}(I_t)$ is obtained by subtracting the standard VAR prediction $E(Y_{t+h} \,|\, \underline{Y_{t-1}})$. Proof: see Appendix 1.

Note that the dimension of $a_\tau^*(Z_\tau)$ may vary with τ and that, for some future date, we may have no measurement equation at all. The computation of $NIRF_{t,t+h}(I_t)$, for $h > T-t$, is easily seen to be equal to the difference of two standard VAR predictions based on model (1). Indeed, we have to compute, for $j \geq 1$:

$$E(Y_{T+j} \,|\, I_t, \underline{Y_{t-1}}) - E(Y_{T+j} \,|\, \underline{Y_{t-1}}). \tag{12}$$

The second term of this difference is just the standard VAR prediction of Y_{T+j} made at time $t-1$ and using $(Y_{t-1}, \dots, Y_{t-p})$ as initial conditions, whereas the first term is a standard VAR prediction made at time T and using the previously computed initial values $E(Y_{T-i} \,|\, I_t, \underline{Y_{t-1}})$, $i \in \{0, \dots, p-1\}$.

The computation of the $NIRF_{t,t+h}(I_t)$ for any $h \in \{1, \dots, T-t\}$, based on Proposition 2, is particularly easy when the a_τ^* functions are linear, that is to say when the new information is the knowledge of *linear filters of future values* of Y_t. Indeed, in this case, the value of $\widehat{Z}_{T|T}$ is obtained by the standard Kalman filter.

If, moreover, the information I_t is made not only of the exact knowledge of the values X_τ of some linear filters $a^*_\tau(Z_\tau)$, but also of the knowledge that some other linear filter $\widetilde{a}^*_\tau(Z_\tau)$ belong to some interval (for instance, the positive or negative half real line), it is possible to take into account this additional information in the following way. First, simulate a path $Z^{(s)}_\tau$, $\tau \in \{t, \dots, T\}$, in the conditional distribution of $(Z'_t, \dots, Z'_T)'$ given $(a^*_\tau(Z_\tau) = X_\tau, \tau \in D_t)$. Second, discard this simulated path if it does not satisfy the constraints on the filters $\widetilde{a}^*_\tau(Z_\tau)$. Third, repeat the previous steps S times and compute the empirical means of the retained simulated values of $Y^{(s)}_{t+h}$ in order to get a Monte Carlo approximation of $E(Y_{t+h} \,|\, I_t, \underline{Y_{t-1}})$. In this procedure, the simulation in the first step is key and easily done thanks to the following nice property of Gaussian vectors (compare with Durbin and Koopman (2002)).

LEMMA: *If $U = (U'_1, U'_2)'$ is Gaussian, the conditional distribution of U_1 given $U_2 = u_2$, is identical to the unconditional distribution of $U_1 - E(U_1 \,|\, U_2) + E(U_1 \,|\, U_2 = u_2)$. Proof: see Appendix 1.*

The previous result can be used to easily simulate in the conditional distribution of $U_1 = (Z'_t, \dots, Z'_T)'$, given that $U_2 = (a^*_\tau(Z_\tau), \tau \in D_t)$ is equal to $u_2 = (X_\tau, \tau \in D_t)$. Indeed, to simulate $(U^{(s)'}_1, U^{(s)'}_2)'$ in the unconditional distribution of $(U'_1, U'_2)'$ reduces here to straightforwardly simulate in the unconditional distribution of $U_1 = (Z'_t, \dots, Z'_T)'$ (since U_2 is function of U_1). Moreover, the computation of $E(U_1 \,|\, U_2 = u_2)$ and $E(U_1 \,|\, U_2 = U^{(s)}_2)$ are easily obtained from the Kalman filter.

3 Examples of New Information Response Function

The New Information Response Function framework introduced in the previous section clearly contains many particular cases depending on the assumption we make about the function $\widetilde{a}(.)$. For instance, the standard case (that we call "full new information" case), obtained when $T = t$ and $\widetilde{a}(\varepsilon_t)$ is one-to-one, contains as particular cases the orthogonalized shocks, the Uhlig (2005)'s impulse vectors and the structural shocks (see Jardet, Monfort and Pegoraro (2009) for details).

The purpose of this section is to detail the flexibility of the NIRF approach through a list of examples based on alternative and less standard specifications of $\widetilde{a}(.)$. More precisely, we will consider in Sections 3.1 and 3.2 the case $I_t = \widetilde{a}(\varepsilon_t)$ where $\widetilde{a}(.)$ is not one-to-one or it is valued in a finite set of real numbers, while Section 3.3 presents the case of a new information on a date-t linear filter, namely $I_t = a_t(\underline{Y_{t-J,t}})$. Section 3.4 focus on the past-path dependent NIRF case where the impact of $\underline{Y_{t-1}}$ on the response does not cancel out. Finally, Section 3.5 moves to the general future path-dependent case, with $T > t$ and $I_t = \widetilde{a}(\underline{\varepsilon_{t,T}}, \underline{Y_{t-1}})$, that is the general case where for instance both future values of basic variables and associated linear filters characterize the new information. This case will be the one adapted to treat the forward policy guidance empirical exercise presented in Section 4.

3.1 Continuous partial new information

Let us now consider the case, named "continuous partial new information" case, where $\widetilde{a}(.)$ only depends on ε_t (i.e. $T = t$), is not one-to-one and $\widetilde{a}(\varepsilon_t)$ has an absolutely continuous distribution.

3.1.1 Pesaran-Shin (1998) "generalized" impulse response functions

Pesaran and Shin (1998) considered the case where $\widetilde{a}(\varepsilon_t) \equiv \varepsilon_{jt}$, that is the case where we have a new information only about one component of ε_t, namely $\varepsilon_{jt} = \alpha$. In the Gaussian case, the computation of $E[\varepsilon_t|\varepsilon_{jt} = \alpha]$ is straightforward and we get:

$$E\left[\varepsilon_{it} \,|\, \varepsilon_{jt} = \alpha\right] = \frac{\Sigma_{ij}}{\Sigma_{jj}}\alpha\,.$$

In particular if $\alpha = 1$, the immediate impact $\delta = E[\varepsilon_t|\varepsilon_{jt} = 1]$ is $\Sigma^{(j)}\Sigma_{jj}^{-1}$ where $\Sigma^{(j)}$ is the j^{th}column of Σ. It is easily seen that this impact is different from the one obtained by an orthogonalized shock with immediate impact on Y_{jt} equal to one, except if $j = 1$ [see Pesaran and Shin (1998)].

3.1.2 New information on a set of individual innovations

If $\widetilde{a}(\varepsilon_t) \equiv \varepsilon_t^K$, where ε_t^K is a K-dimensional sub-vector of ε_t containing ε_{jt} with $j \in K$ and $K \subset \{1, \ldots, n\}$, we have to compute $\delta = E[\varepsilon_t \,|\, \varepsilon_t^K = \alpha]$ where α is now a vector. Again, in the

Gaussian case we immediately get:

$$\delta = \Sigma^K \, \Sigma_{KK}^{-1} \alpha$$

where Σ^K is the matrix given by the columns $\Sigma^{(j)}$ of Σ such that $j \in K$ and Σ_{KK} is the variance-covariance matrix of ε_t^K. For instance, if the new information is $\varepsilon_{jt} = 1$ and $\varepsilon_{kt} = 0$, the i^{th} component of δ ($i \neq j$ and $i \neq k$) will be the coefficient of ε_{jt} in the theoretical regression of ε_{it} on ε_{kt} and ε_{jt}.

3.1.3 New information on responses

We know from equation (6) that the expected response of Y_{t+h_1} (for a given h_1) to a value of ε_t is $\Theta_{h_1} \varepsilon_t$. We may want to impose that some components of this response are given, that is $\Theta_{h_1}^{K_1} \varepsilon_t = \alpha_1$, where $\Theta_{h_1}^{K_1}$ is the set of rows of Θ_{h_1} corresponding to the components of interest. If this new information is the only one, the NIRF has to be computed as $\Theta_h \, \delta$ with $\delta = E(\varepsilon_t \,|\, \Theta_{h_1}^{K_1} \varepsilon_t = \alpha_1)$, that is $\delta = \Sigma \, \Theta_{h_1}^{K_1'} (\Theta_{h_1}^{K_1} \Sigma \, \Theta_{h_1}^{K_1'})^{-1} \alpha_1$ in the Gaussian case. This new information can be combined with another one, for instance a new information on a set of individual innovations as in Section 3.1.2, i.e. $\varepsilon_t^{K_2} = \alpha_2$. In this case we have to take:

$$\delta = E(\varepsilon_t \,|\, \Theta_{h_1}^{K_1} \varepsilon_t = \alpha_1 \, , \varepsilon_t^{K_2} = \alpha_2) \, ,$$

which can be easily computed as soon as $K_1 + K_2 \leq n$. In the Gaussian case, if we denote by $\mathcal{S}_2$ the selection matrix such that $\varepsilon_t^{K_2} = \mathcal{S}_2 \, \varepsilon_t$, and given $M = (\Theta_{h_1}^{K_1'}, \mathcal{S}_2')'$, we have:

$$\delta = \Sigma \, M'(M \, \Sigma \, M')^{-1} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$$

(assuming that $M \, \Sigma \, M'$ is invertible).

3.1.4 New information on long-run behaviors

Let us assume that Y_t is non-stationary and admits r cointegrating relationships, and let us construct a vector W_t such that:

$$W_t = \begin{pmatrix} \Delta\widetilde{Y}_t \\ \Lambda' Y_t \end{pmatrix} ,$$

where $\widetilde{Y}_t$ is the subvector of Y_t given by its first $(n-r)$ rows (possibly after a reordering of the components of Y_t), $\Lambda' Y_t$ is a r-dimensional vector of cointegrating relationships, and W_t has a stationary VAR representation of the form:

$$\Gamma(L) W_t = C\nu + C\varepsilon_t \tag{13}$$

where $C = \begin{pmatrix} I_{n-r} & 0 \\ & \Lambda' \end{pmatrix}$ is invertible. We can consider, for instance, a partial new information imposing that the long-run impact on Y_{it} (with $i \leq n-r$) is zero, that is:

$$\Gamma_i^{-1}(1)\, C\, \varepsilon_t = 0\,,$$

where $\Gamma_i^{-1}(1)$ is the i^{th} row of $\Gamma^{-1}(1)$.

3.2 Discrete partial new information

3.2.1 Definition of the new information

Let us now consider the case where the distribution of $\widetilde{a}(\varepsilon_t)$ has a discrete component. More precisely, in this case, called "discrete partial new information" case, we assume that $\widetilde{a}(.) = \begin{pmatrix} \widetilde{a}_1(.) \\ \widetilde{a}_2(.) \end{pmatrix}$, where $\widetilde{a}_1(\varepsilon_t)$ has a continuous distribution and $\widetilde{a}_2(\varepsilon_t)$ is valued in a finite set $\overline{\alpha}_2 = \{\alpha_{21}, ..., \alpha_{2L}\}$. In this case the conditional distribution of ε_t given $\widetilde{a}_1(\varepsilon_t) = \alpha_1$ and $\widetilde{a}_2(\varepsilon_t) = \alpha_{2j} \in \overline{\alpha}_2$ is obtained by the conditional distribution of ε_t given $\widetilde{a}_1(\varepsilon_t) = \alpha_1$ restricted to the set $\widetilde{a}_2(\varepsilon_t) = \alpha_{2j}$.

3.2.2 Quantitative information and interval information

Let us consider the case where $\widetilde{a}_2(\varepsilon_t) = 1\!\!1_{]\,c,\,d\,[}(\varepsilon_{jt})$ and $\widetilde{a}_1(\varepsilon_t) = \varepsilon_t^K$ with c and d real numbers ($c < d$) and $K \subset \{1, ..., n\}$ such that $j \notin K$. Our purpose is thus to compute $E[\varepsilon_{jt} \,|\, \varepsilon_t^K = \alpha, c < \varepsilon_{jt} < d]$ and $E[\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, c < \varepsilon_{jt} < d]$, with $i \notin K$ and $i \neq j$. In both cases, explicit formulas are available.

Proposition 3: *For $c, d \in \mathbb{R}$ ($c < d$) and $K \subset \{1, ..., n\}$ such that $j \notin K$, we have:*

$$
\begin{aligned}
&E[\varepsilon_{jt} \,|\, \varepsilon_t^K = \alpha, c < \varepsilon_{jt} < d] = \mu_j^K \alpha + \sigma_j^K \,\frac{\varphi(c_j^K) - \varphi(d_j^K)}{\Phi(d_j^K) - \Phi(c_j^K)}\,, \\
&c_j^K := \frac{c - \mu_j^K \alpha}{\sigma_j^K}\,, \quad d_j^K := \frac{d - \mu_j^K \alpha}{\sigma_j^K}\,, \\
&\mu_j^K := E(\varepsilon_{jt} \,|\, \varepsilon_t^K = \alpha)'\,, \quad (\sigma_j^K)^2 := Var(\varepsilon_{jt} \,|\, \varepsilon_t^K = \alpha)\,,
\end{aligned}
$$

and for $i \notin K$, $i \neq j$ we have:

$$
\begin{aligned}
&E[\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, c < \varepsilon_{jt} < d] = \mu_{ij}^K \alpha + \nu_{ij}^K \left[\mu_j^K \alpha + \sigma_j^K \left(\frac{\varphi(c_j^K) - \varphi(d_j^K)}{\Phi(d_j^K) - \Phi(c_j^K)} \right) \right]\,, \\
&\text{with } \mu_{ij}^K \alpha + \nu_{ij}^K \varepsilon_{jt} := E(\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, \varepsilon_{jt})'\,,
\end{aligned}
$$

where φ and Φ are, respectively, the p.d.f and the c.d.f of $N(0,1)$. In particular, if $c = 0$ and $d = +\infty$ we get:

$$
E[\varepsilon_{jt} \,|\, \varepsilon_t^K = \alpha, 0 < \varepsilon_{jt}] \;=\; \mu_j^K \alpha + \sigma_j^K \,\lambda \left(\frac{\mu_j^K \alpha}{\sigma_j^K} \right)\,,
$$

where $\lambda(x) = \dfrac{\varphi(x)}{\Phi(x)}$ is the inverse Mill's ratio. Proof: see Appendix 1.

The case of interval information on several components ε_{jt}, $j \in J$, $J \subset \{1, \ldots, n\}$ and $K \cap J = \varnothing$, as well as the case where the interval information are related to responses, are presented in Appendix 1.

3.3 New information on a filter

In some situations, the relevant information is on a linear filter of the basic variables. For instance, in macro-finance models of the yield curve, this filter may be a term premium, or an expectation component [see Jardet, Monfort, Pegoraro (2013)]. In that case, the NIRF framework still applies, as it is formalized in the following:

PROPOSITION 4: *Given the linear filter* $\widetilde{Y}_t = a_t(\underline{Y_{t-J,t}}) = F(L)\,Y_t$, *where* $F(L) = (F_1(L), ..., F_n(L))$ *is a row vector of polynomials in the lag operator* L, *its innovation is* $\widetilde{\varepsilon}_t = F(0)\,\varepsilon_t$, *and given the new information on* $\widetilde{\varepsilon}_t$, *defined by* $\widetilde{a}(\widetilde{\varepsilon}_t) = \alpha$, *the differential impact on the prediction of* Y_{t+h} *is:*

$$\begin{aligned} NIRF_{t,t+h}(\varepsilon_t) &= \Theta_h\, E[\varepsilon_t \,|\, a(\varepsilon_t) = \alpha]\,, \\ \text{with } a(\varepsilon_t) &:= \widetilde{a}(F(0)\,\varepsilon_t)\,. \end{aligned} \tag{14}$$

Proof: see Appendix 1.

This means that, a new information $\widetilde{a}(\widetilde{\varepsilon}_t) = \alpha$ on $\widetilde{\varepsilon}_t$ can be viewed as an information on ε_t and it can be treated as in the previous framework. Let us consider some examples. If the information is $\widetilde{\varepsilon}_t = 1$ and $\varepsilon_{jt} = 0$, $j = 1, ..., n-1$, the impact on Y_{t+h} is $\Theta_h\delta$, where $\delta = E[\varepsilon_t|\widetilde{\varepsilon}_t = 1, \varepsilon_{jt} = 0, j = 1, ..., n-1]$ is equal to $(0, ..., 0, 1/F_n(0))$. If the information is just $\widetilde{\varepsilon}_t = 1$, the impact on Y_{t+h} is $\Theta_h\delta$, where

$$\delta \;=\; \frac{cov(\varepsilon_t, \widetilde{\varepsilon}_t)}{V(\widetilde{\varepsilon}_t)} = \frac{\Sigma F'(0)}{F(0)\Sigma F'(0)}$$

If the information is $\widetilde{\varepsilon}_t = 1$ and $\varepsilon_{jt} = 0$, the impact on Y_{t+h} is $\Theta_h\delta$ where the i^{th} component δ_i of δ is the coefficient of $\widetilde{\varepsilon}_t$ in the theoretical regression of ε_{it} on $\widetilde{\varepsilon}_t$ and ε_{jt} (in particular $\delta_j = 0$). We could also impose point information on several filters in a straightforward way, and extend the technique to interval information [4].

It is important to highlight that, if we are interested in studying the impact of a new information about k linear filters $\widetilde{Y}_{1t}, ..., \widetilde{Y}_{kt}$ on the basic variables forming Y_t, it would be possible to complete these filters with $n-k$ components of Y_t ($Y_t^{(2)}$, say) and to apply the NIRF techniques to the MA representation of the VARMA model followed by the vector thus obtained Y_t^* (see Appendix 2 for details). However, this would be a tedious method, compared to the one we suggest. Moreover,

[4] Similarly, we might be interested in the response of a linear filter to some new information. If we consider the univariate filter $\widetilde{Y}_t = G(L)Y_t$, we can compute the impact on $\widetilde{Y}_{t+h}$ of a new information $a(\varepsilon_t) = \alpha$ at t. Indeed, since the impact on Y_{t+h} is $\Theta_h E[\varepsilon_t|a(\varepsilon_t) = \alpha]$, the impact on $\widetilde{Y}_{t+h}$ is obviously $G(L)\Theta_h E[\varepsilon_t|a(\varepsilon_t) = \alpha]$ where the lag operator L is operating on h and where $\Theta_s = 0$ if $s < 0$.

by construction, the specification of Y_t^*, that is the selection of $Y_t^{(2)}$, implies that the impact on $Y_t^{(2)}$ only can be assessed and, therefore, in order to provide a comprehensive analysis, several specifications of Y_t^* have to be considered in order to determine the impact on *any* variable in Y_t. In sharp contrast, the NIRF approach to linear filters, formalized in Proposition 4, easily and immediately delivers through formula (14) the dynamic relationship between any linear filter and any basic variable.

3.4 Past path-dependent NIRF

In all situations considered above the values of the Y_t's actually observed do not play any role in the computation of the NIRF, since in equation (7) the impact of $\underline{Y_{t-1}}$ cancels out. Let us now consider an example in which the past values $\underline{Y_{t-1}}$ matter.

If we want to impose that a subset of components Y_t^K does not move between $t-1$ and t we have to impose $Y_t^K = Y_{t-1}^K$ or, denoting by $\widehat{Y}_{t|t-1}^K(\underline{Y_{t-1}})$ the prediction of Y_t^K made at $t-1$ (a linear function of $\underline{Y_{t-1}}$), we have to impose:

$$\begin{aligned} &\widehat{Y}_{t|t-1}^K(\underline{Y_{t-1}}) + \varepsilon_t^K = Y_{t-1}^K \,, \\ &\text{or} \;\; \varepsilon_t^K - Y_{t-1}^K + \widehat{Y}_{t|t-1}^K(\underline{Y_{t-1}}) = 0 \,, \end{aligned} \tag{15}$$

which is of the form $\widetilde{a}(\varepsilon_t, \underline{Y_{t-1}}) = \alpha$. In this case the NIRF is $\Theta_h \, \delta_t$ with:

$$\delta_t = E[\varepsilon_t \,|\, \varepsilon_t^K = y_{t-1}^K - \widehat{y}_{t|t-1}^K(\underline{y_{t-1}})] \tag{16}$$

and we get:

$$\delta_t = \Sigma^K \, \Sigma_{K\,K}^{-1} \, [y_{t-1}^K - \widehat{y}_{t|t-1}^K(\underline{y_{t-1}})] \,, \tag{17}$$

where Σ^K and $\Sigma_{K\,K}$ are defined like in Section 3.1.2.

3.5 Future path-dependent NIRF

In some situations it is interesting to study the behavior of the future values of the endogenous variables of a dynamic system, when the future path of one or several of them or/and when the future path of some filters are known or fixed for scenarios analysis. The first kind of information has already been imposed in the conditional prediction literature [see e.g. Waggoner and Zha (1999), Clarida and Coyle (1984), Doan, Litterman and Sims (1986), Jarocinski (2010), Banbura, Giannone, Lenza (2010)]. In Proposition 2 we have shown how to deal with this issue in the NIRF framework and, in particular, how to assess the impact, on a set of variables of interest, of a general new information $I_t = a(\underline{Y_{t,T}})$ on the future path of some filters of the variables.

4 Applications to Forward Policy Guidance *(preliminary)*

In this section we propose an illustration of the empirical relevance of the NIRF methodology. Using a Gaussian VAR(p) model estimated with U.S. quarterly data, we address a current and relevant monetary policy issue. More precisely, in order to illustrate the impact of a new information on "future paths" of a given variable *and* associated linear filter, we consider the short rate stabilization announcement by the FOMC at the end of 2008 and we investigate how the long-rate, as well as its expectation and term premium components (i.e., linear filters of the long-rate), respond to alternative kinds of scenarios involving the future paths of the short-term interest rate *and* of its expectations.

4.1 Description of the Data

Our data are quarterly observations of the U.S. short-term zero-coupon bond yield r_t, i.e the one-quarter yield, the spread between the one-year and the one-quarter yield S_t, the one-quarter inflation rate π_t and the growth rate of real Gross Domestic Product (GDP) g_t, for the period from 1979:Q4 to 2010:Q3. The quarterly inflation rate π_t, from $t-1$ to t, is given by $\pi_t = \log(P_t/P_{t-1})$, where P_t is the price index level observed the last month of the quarter. The GDP growth over the period $(t-1, t)$ is given by $g_t = \log(G_t/G_{t-1})$, where G_t is the real GDP level at quarter t.

The interest rate data are obtained from the Gurkaynak, Sack, and Wright (2007) [GSW (2007), hereafter] data base[5]. The price index and the real GDP data are obtained from the FRED database: P_t is the seasonally adjusted consumer price index for all urban consumer (all items, CPIAUCSL); G_t is the seasonally adjusted real GDP level, in billions of chained 2005 dollars (GDPC1).

4.2 Model and Decompositions

We collect these variables in the four-dimensional vector $Y_t = (r_t, S_t, g_t, \pi_t)'$ and we describe the joint dynamics of Y_t by the following Gaussian VAR(p) process:

$$Y_t = \nu + \sum_{j=1}^{p} \Phi_j Y_{t-j} + \varepsilon_t \,, \tag{18}$$

where ε_t is a 4-dimensional Gaussian white noise with $\mathcal{N}(0, \Sigma)$ distribution [Σ denotes the (4×4) conditional variance-covariance matrix]; Φ_j, for $j \in \{1, \ldots, p\}$, are (4×4) matrices, while ν is a 4-dimensional vector. On the basis of several lag order selection criteria, the lag length is selected to be $p = 3$ and the model is estimated by OLS.

It is well known that any H-year date-τ nominal yield $R_\tau(H)$ can be decomposed into the following two terms:

$$R_\tau(H) = EX_\tau(H) + TP_\tau(H) \,, \tag{19}$$

where

$$EX_\tau(H) = \frac{1}{H} E\left(\sum_{h=0}^{H-1} r_{\tau+h} | \Omega_\tau \right) \tag{20}$$

is the expectation part of $R_\tau(H)$, $TP_\tau(H) = R_\tau(H) - EX_\tau(H)$ is, by definition, the corresponding term premium and Ω_τ is the available information set at date τ. In addition, $EX_\tau(H)$ can be decomposed into two components:

$$EX_\tau(H) = \widetilde{EX}_\tau(H) + \Pi^e_\tau(H) \tag{21}$$

[5]Each observation in our sample is given by the daily value observed at the end of each quarter.

where:

$$\widetilde{EX}_\tau(H) = \frac{1}{H} E\left(\sum_{h=1}^{H} \widetilde{r}_{\tau+h} \,|\, \Omega_\tau\right) \tag{22}$$

is the expectation term of the real yield of residual maturity H, $\widetilde{r}_{\tau+1} = r_\tau - \pi_{\tau+1}$ is the one-quarter real (ex-post) interest rate, while

$$\Pi^e_\tau(H) = \frac{1}{H} E\left(\sum_{h=1}^{H} \pi_{\tau+h} \,|\, \Omega_\tau\right) \tag{23}$$

is the inflation expectation over $(\tau, \tau + H)$. These three components can be written as linear filters of $Y_t = (r_t, S_t, g_t, \pi_t)'$ and thus the NIRF approach can be adopted (see Appendix 3). In particular, in the following section we will consider the case $H = 4$ quarters[6] and we will apply the general result presented in Proposition 2.

4.3 Responses to unconventional monetary policy: effects of forward policy guidance

Central banks are sometimes confronted with the key issue of how restoring good economic and financial conditions when the short-term interest rate is near the zero lower bound. Among the set of measures proposed to handle this issue, known as unconventional monetary policy measures, one is the forward policy guidance. The idea is that if a central bank can credibly commit to future policy actions, it can manage longer-term interest rates to a level consistent with a given objective of price stability and economic growth. There are several examples of central banks using communications on the future path of the short-term interest rate like, for instance, New Zealand, Norway and Sweden by means of policy rate projections, or Canada and Japan by means of communications regarding the timing and conditions for rate moves. Forward guidance on monetary policy has also been implemented by the U.S. Federal Reserve. In its statement released in December 16, 2008, the FOMC announced "*that (anticipated) weak economic conditions are likely to warrant exceptionally low levels of the federal funds rate for some time*". A more recent

[6] Considering another maturity for H is possible, but requires the estimation of an affine term structure model.

example is the August 2011 FOMC statement: *"The committee currently anticipates that economic conditions - including low rates of resources utilization and a subdued outlook for inflation over the medium-run - are likely to warrant exceptionally low levels for the federal funds rate at least through mid-2013"*.

The communication regarding the future path of short-term interest rate is the key ingredient of such a policy. Does this communication reach its goal of reducing the medium- or long-term rate $R_t(H)$? In particular, since the aim is to manage as accurately as possible the expectation component $EX_t(H)$ (see (20)) of $R_t(H)$, the final effects of such communication should depend on the period of time during which short-term rates are known: the longer this period, the better the management of $EX_t(H)$ and, thus, of $R_t(H)$. When the central bank can commit only for a period of time (ℓ, say) shorter than the time-to-maturity of the long rate (H, namely), the response of $EX_t(H)$ will also depend in general on the path of short rate expectations over the non-committed remaining period ($H - \ell$). In this case, what are the effects of this communication on $R_t(H)$?

The NIRF approach and, in particular, the tools developed in Section 6 provide a rigorous way to address these kinds of questions. Actually, we can estimate the expected response of $Y_t = (r_t, S_t, g_t, \pi_t)'$ and linear filters of Y_t to a new information about the future paths of short-term interest rate and/or about the future paths of the expected short-term rate. More precisely, we focus on the expected responses of the short rate r_t, of the 1-year interest rate and of the linear filters EX_t and TP_t, in 2009:Q1 and the 7 following quarters, after the FOMC statement in 2008:Q4 to "*warrant exceptionally low levels of the federal fund rate for some time*". In a benchmark scenario, we assume that this statement is interpreted by the agents as the new information that the short-term interest rates will remain constant, over the four following quarters, at the level $\overline{r} = 0.25\%$ (annual basis). Then, in order to assess the usefulness of this possible announcement for the prediction of the future realizations of the variables of interest, we compare its effects on the long rate, EX_t and TP_t, with those provided by three alternative scenarios, in which the levels of the short rate are assumed to be fixed for only the first two quarters (2009:Q1 and 2009:Q2), and then it is its expected value which is fixed for the two following ones (2009:Q3 and 2009:Q4; see the scenarios below). We have also assumed in any scenario that, at the starting date $t =$

2009:Q1, the model-implied 1-year yield R_t is positive and, in line with previous application, that the instantaneous response of GDP growth in 2009:Q1 is zero. Note that, given the very low level of the short-term interest rate, only scenarios implying an increase in (expected) short rates will be considered. More precisely, the four scenarios are the following (for $t =$2009:Q1 and assuming annualized rates):

- Benchmark scenario: the short rate is fixed for the 4 quarters following the FOMC statement (i.e., $t-1=$ 2008:Q4). Hence the new information is $\widetilde{Z}_{t,t+3}=\{r_t=\overline{r}, r_{t+1}=\overline{r}, r_{t+2}=\overline{r}, r_{t+3}=\overline{r}\}$;

- Scenario 1: the short rate is constant in the first 2 quarters and then it is expected (at date $t+1$) to remain unchanged in the two following ones. The new information is $\widetilde{Z}_{t,t+3}=\{r_t=\overline{r}, r_{t+1}=\overline{r}, E_{t+1}(r_{t+2})=\overline{r}, E_{t+1}(r_{t+3})=\overline{r}\}$;

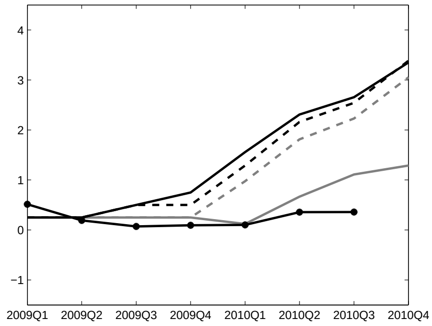
- Scenario 2: the short rate is constant in the first 2 quarters and it is expected (at date $t+1$) to increase by 25 basis points in the third quarter and, then, to remain unchanged. The new information is $\widetilde{Z}_{t,t+3}=\{r_t=\overline{r}, r_{t+1}=\overline{r}, E_{t+1}(r_{t+2})=\overline{r}+\frac{0.25}{100}, E_{t+1}(r_{t+3})=\overline{r}+\frac{0.25}{100}\}$;

- Scenario 3: the short rate is constant in the first 2 quarters and at date $t+1$ it is expected to increase by 25 basis points in the two following ones. The new information is $\widetilde{Z}_{t,t+3}=\{r_t=\overline{r}, r_{t+1}=\overline{r}, E_{t+1}(r_{t+2})=\overline{r}+\frac{0.25}{100}, E_{t+1}(r_{t+3})=\overline{r}+\frac{0.5}{100}\}$.

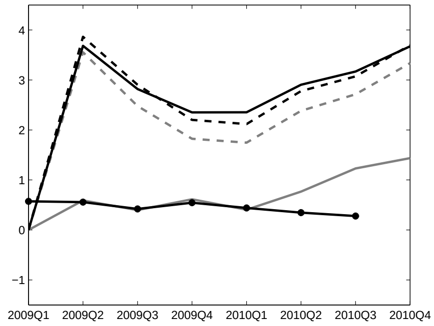
In order to deal with the monetary policy shift observed at the end of 1970s and to perform a real-time exercise, we estimate the VAR over the period 1979:Q4 to 2008:Q4.

We report in figure 1 the expected responses of r_t, of the 1-year interest rate and of the associated expectation term and term premium components for the four above mentioned scenarios. We observe that the benchmark scenario generates responses which are, compared with the alternative ones, closer to the ex-post realized values of these variables. In particular, we note that the short rate responses obtained with the benchmark scenario tend to remain at lower levels than those obtained with the other scenarios (see figure 1 (a)). Notably, although the short rates obtained with the benchmark and the first scenario are (by construction) identical over the first four quarters[7], the

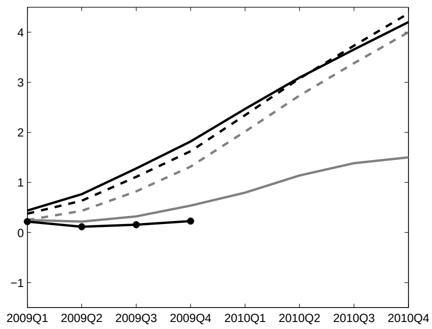
[7]With the scenario 1, $E\left(r_{t+2}\,|\,\Omega_{t-1}\right)=E\left(E_{t+1}(r_{t+2})\,|\,\Omega_{t-1}\right)=\overline{r}$ and $E\left(r_{t+3}\,|\,\Omega_{t-1}\right)=E\left(E_{t+1}(r_{t+3})\,|\,\Omega_{t-1}\right)=\overline{r}$.



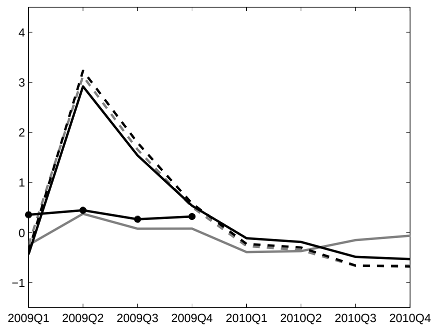
(a) Short-term interest rate



(b) 1-year interest rate



(c) 1-year ahead nominal expectation term



(d) 1-year nominal term premium

Figure 1: Responses (in annual percentage) to the four scenarios. Grey solid lines: benchmark scenario. Grey dashed lines: scenario 1. Black dashed lines: scenario 2. Black solid lines: scenario 3. Black lines with markers: ex-post realization of the variables.

short rate of scenario 1 reaches, after 8 quarters, a level that is more than twice higher than the one obtained with the benchmark scenario and the expectation component in the latter case remains at levels lower than the other ones (see figure 1 (c)). Similarly, the 1-year interest rate in the the benchmark scenario is, on average, lower than those obtained with the alternative scenarios and quite close to its realized values (see figure 1 (b)). Accordingly, the 1-year nominal term premium provided by the benchmark scenario is stable, lower than the alternative ones and closer to its realized values (during the first four quarters; see figure 1 (d)). These results strongly suggest that the central bank should benefit from an *accurate and unambiguous* commitment on the future path of the short-term interest rates.

All in all, this empirical illustration stresses that considering information on future paths of relevant variables is a key element not only for forecasting purposes, but also for a precise anticipation of the future effects of a monetary policy intervention (like the short rate path in our exercise) and the NIRF methodology provides a flexible and promising framework to handle these kinds of economically relevant issues.

5 Conclusions and Further Developments

In this paper we propose a general statistical methodology, the NIRF methodology, for the analysis of impulse response functions in VAR models, which encompasses several standard approaches, such that orthogonalization of shocks (Sims (1982)), the "generalized" impulse responses of Pesaran and Shin (1998), or the impulse vectors of Uhlig (2005). We show that an important aspect of the NIRF methodology is to be able to take into account an information about the past or future values of the variables of interest. In particular, this information may concerns the future path of the process and, possibly, the future values of some linear filters of that process or the values of other linear filters belonging to some interval. We also show that this methodology is well suited to analyze the effects of a new quantitative or qualitative information on one or several innovations and/or on the response of one or several variables to such an information, as well as a new information on linear filters of the basic variables of our model.

We provide an empirical illustration of the NIRF methodology based on U.S. data. More precisely, in order to illustrate the usefulness of a new information on a "future path" of variables and linear filters, we investigate how the long-term rate in our VAR model, and the associated expectation part and term premium component, respond to alternative kinds of communications that may be released by the FOMC (on December 2008) about the stabilization of the future short-term interest rate *and* its expectations around the zero lower bound. We show that taking into account this information is critical and significantly improves interest rates forecasts. In addition, by means of several scenarios based on possible paths of the short rate and its expectation component, we show how the NIRF can be used to anticipate future effects of monetary policy decisions and, thus, it can be adopted to suggest a preferred monetary policy communication strategy.

The results of this paper have been derived in the Gaussian case. If the distribution is no longer Gaussian and if function $a(.)$ is linear, the results are still valid if we replace the notion of conditional expectation by the notion of linear regression. If $a(.)$ is non linear, the conditional expectation $E[\varepsilon_t | a(\varepsilon_t) = \alpha]$ might be approximated by Monte Carlo and kernel techniques. The results could also be extended to VARMA(p,q) processes. The interval constraints could be replaced by more general set information tackled by Monte Carlo methods. Finally, the extension to the nonlinear framework [see Gallant, Rossi, Tauchen (1993), Koop, Pesaran, Potter (1996), Gourieroux and Jasiak (2005)] could be an interesting line of future research.

Appendix 1: Proofs of propositions.

Proof of Proposition 1: from the definition of new information response function:

$$NIRF_{t,t+h}(I_t) = E(Y_{t+h} \,|\, \widetilde{a}(\varepsilon_t, \underline{Y_{t-1}}), \underline{Y_{t-1}}) - E(Y_{t+h} \,|\, \underline{Y_{t-1}})\,, \tag{A.1}$$

we have:

$$\begin{aligned} NIRF_{t,t+h}(I_t) &= E\left\{[E(Y_{t+h} \,|\, \varepsilon_t, \underline{Y_{t-1}}) - E(Y_{t+h} \,|\, \underline{Y_{t-1}})] \,|\, \widetilde{a}(\varepsilon_t, \underline{Y_{t-1}}), \underline{Y_{t-1}}\right\} \\ &= E[\Theta_h\, \varepsilon_t \,|\, \widetilde{a}(\varepsilon_t, \underline{Y_{t-1}}), \underline{Y_{t-1}}] \\ &= \Theta_h\, E[\varepsilon_t \,|\, \widetilde{a}(\varepsilon_t, \underline{Y_{t-1}})]\,, \end{aligned} \tag{A.2}$$

and the result is proved.

Proof of Proposition 2: given $(I_t, \underline{Y_{t-1}})$ and for any $\tau \in \{t, \ldots, T\}$, the X_τ are observed whereas the Z_τ are latent and thus we get the state space model (11). The filtered value $\widehat{Z}_{T|T}$ is the conditional expectation of $Z_T = (Y'_T, \ldots, Y'_{T-K})'$ given $X_T = I_t$ and $\underline{Y_{t-1}}$. Since $K \geq T - t$, $E(Y_{t+h} \,|\, I_t, \underline{Y_{t-1}})$ is a component of $\widehat{Z}_{T|T}$ for any $h \in \{1, \ldots, T-t\}$.

Proof of Lemma: if $(U'_1, U'_2)'$ is Gaussian, we know that the conditional distribution of U_1 given $U_2 = u_2$ is Gaussian as well as the unconditional distribution of $W := U_1 - E(U_1 \,|\, U_2) + E(U_1 \,|\, U_2 = u_2)$. This means that we have just to check that both distributions have the same expectation and the same variance-covariance matrix. Now, we obviously have $E(W) = E(U_1 \,|\, U_2 = u_2)$ and the variance-covariance matrix $V(W)$ of W is the sum of $V[E(W \,|\, U_2)]$, which is equal to zero, and of $E[V(W \,|\, U_2)]$, which is equal to $V(W \,|\, U_2)$ (since $V(W \,|\, U_2)$ is non-random in the Gaussian case) and also equal to $V(U_1 \,|\, U_2)$.

Proof of Proposition 3:

i) Computation of $E[\varepsilon_{jt} \,|\, \varepsilon_t^K = \alpha, c < \varepsilon_{jt} < d]$:

the conditional distribution of ε_{jt} given $\varepsilon_t^K = \alpha$ is easily found; it is a Gaussian distribution with mean $\mu_j^K \alpha$ and variance $(\sigma_j^K)^2$ (say) (where μ_j^K is a row vector). So $E[\varepsilon_{jt} \,|\, \varepsilon_t^K = \alpha, c <$

$\varepsilon_{jt} < d]$ is given by $E[\mu_j^K \alpha + \sigma_j^K U \,|\, c < \mu_j^K \alpha + \sigma_j^K U < d]$ where $U \backsim N(0,1)$. We find

$$E[\varepsilon_{jt} \,|\, \varepsilon_t^K = \alpha, c < \varepsilon_{jt} < d] = \mu_j^K \alpha + \sigma_j^K \, E \left(U \,|\, \tfrac{c - \mu_j^K \alpha}{\sigma_j^K} < U < \tfrac{d - \mu_j^K \alpha}{\sigma_j^K} \right) .$$

Using the notations $c_j^K = \frac{c-\mu_j^K \alpha}{\sigma_j^K}$ and $d_j^K = \frac{d-\mu_j^K \alpha}{\sigma_j^K}$, we find:

$$E[\varepsilon_{jt} \,|\, \varepsilon_t^K = \alpha, c < \varepsilon_{jt} < d] \;=\; \mu_j^K \alpha + \sigma_j^K \frac{\varphi(c_j^K) - \varphi(d_j^K)}{\Phi(d_j^K) - \Phi(c_j^K)} ,$$

where φ and Φ are, respectively, the p.d.f and the c.d.f of $N(0,1)$. In particular, if $c = 0$ and $d = +\infty$, we find:

$$E[\varepsilon_{jt} \,|\, \varepsilon_t^K = \alpha, 0 < \varepsilon_{jt}] \;=\; \mu_j^K \alpha + \sigma_j^K \, \lambda \left(\frac{\mu_j^K \alpha}{\sigma_j^K} \right) ,$$

where $\lambda(x) = \dfrac{\varphi(x)}{\Phi(x)}$ is the inverse Mill's ratio.

ii) Computation of $E[\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, c < \varepsilon_{jt} < d]$:

we first find the conditional expectation of ε_{it} given $\varepsilon_t^K = \alpha$ and ε_{jt}, which can be written as $\mu_{ij}^K \alpha + \nu_{ij}^K \varepsilon_{jt}$ (say) and we get:

$$\begin{aligned} E[\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, c < \varepsilon_{jt} < d] &= E\left[E\left(\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, \varepsilon_{jt} \right) \,|\, \varepsilon_t^K = \alpha, c < \varepsilon_{jt} < d \right] \\ &= \mu_{ij}^K \alpha + \nu_{ij}^K E[\varepsilon_{jt} \,|\, \varepsilon_t^K = \alpha, c < \varepsilon_{jt} < d] \\ &= \mu_{ij}^K \alpha + \nu_{ij}^K \left[\mu_j^K \alpha + \sigma_j^K \left(\frac{\varphi(c_j^K) - \varphi(d_j^K)}{\Phi(d_j^K) - \Phi(c_j^K)} \right) \right] . \end{aligned}$$

In the particular case $c = 0, d = +\infty$, we find:

$$E[\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, 0 < \varepsilon_{jt}] = \mu_{ij}^K \alpha + \nu_{ij}^K \left[\mu_j^K \alpha + \sigma_j^K \lambda \left(\frac{\mu_j^K \alpha}{\sigma_j^K} \right) \right] .$$

Proof of Proposition 4: given the linear filter $\widetilde{Y}_t = F(L)\, Y_t$, its innovation is by definition:

$$\begin{aligned}
\widetilde{\varepsilon}_t &= \widetilde{Y}_t - E_t(\widetilde{Y}_t \,|\, \underline{Y_{t-1}}) \\
&= \sum_{i=1}^{n} \Big[F_{i0} Y_{i,t} + F_{i1} Y_{i,t-1} + \ldots + F_{iJ} Y_{i,t-J} - F_{i0} E(Y_{i,t} \,|\, \underline{Y_{t-1}}) - F_{i1} Y_{i,t-1} - \ldots - F_{iJ} Y_{i,t-J} \Big] \\
&= \sum_{i=1}^{n} F_{i0} \widetilde{\varepsilon}_{it} = F(0) \varepsilon_t
\end{aligned}$$

and, thus, formula (14) immediately follows.

Quantitative information and several interval information

We still assume $a_1(\varepsilon_t) = \varepsilon_t^K$, but now $a_2(\varepsilon_t)$ is the set of functions $\big\{ 1\!\!1_{]c_j, d_j[}(\varepsilon_{jt}), j \in J \big\}$, with c_j and d_j real numbers ($c_j < d_j$) for any $j \in J$, $J \subset \{1, \ldots, n\}$ and $K \cap J = \varnothing$. We have to compute:

$$E[\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, c_j < \varepsilon_{jt} < d_j, j \in J], \; i \in J\,,$$

and

$$E[\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, c_j < \varepsilon_{jt} < d_j, j \in J], \; i \notin K, i \notin J\,.$$

i) Computation of $E[\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, c_j < \varepsilon_{jt} < d_j, j \in J]$, $i \in J$:

the joint conditional distribution of ε_t^J given $\varepsilon_t^K = \alpha$ is Gaussian with mean $\mu^{JK}\alpha$ and variance-covariance matrix Σ^{JK} (say) and we have to compute the mean of this normal distribution restricted to $(c_j < \varepsilon_{jt} < d_j, j \in J)$ (see below).

ii) Computation of $E[\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, c_j < \varepsilon_{jt} < d_j, j \in J]$, $i \notin K$, $i \notin J$:

given that

$$\begin{aligned}
& E[\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, c_j < \varepsilon_{jt} < d_j, j \in J] \qquad (A.3) \\
= \; & E[E(\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, \varepsilon_{jt}, j \in J) \,|\, \varepsilon_t^K = \alpha, c_j < \varepsilon_{jt} < d_j, j \in J]\,,
\end{aligned}$$

and denoting by $\mu_i^{JK}\alpha + \nu_i^{JK}\, \varepsilon_t^J$, the conditional expectation of ε_{it} given $\varepsilon_t^K = \alpha$ and ε_t^J, we get:

$$E[\varepsilon_{it}\,|\,\varepsilon_t^K=\alpha, c_j<\varepsilon_{jt}<d_j, j\in J]=\mu_i^{JK}\alpha+\nu_i^{JK}E[\varepsilon_t^J\,|\,\varepsilon_t^K=\alpha, c_j<\varepsilon_{jt}<d_j, j\in J]\,.$$

Again, the joint conditional distribution of ε_t^J, given $\varepsilon_t^K=\alpha$, is $N(\mu^{JK}\alpha, \Sigma^{JK})$ and, as above, we have to compute the mean of this normal distribution restricted to the set $(c_j<\varepsilon_{jt}<d_j, j\in J)$.

The restriction of a $J-$variate normal distribution to a product of intervals is in general not analytically tractable, but it can be simulated either by the rejection algorithm mentioned above, or by using the Gibbs algorithm, and therefore its mean can be computed by a Monte-Carlo method. The principle of the Gibbs algorithm is to start from an initial value $y_0=(y_{01},...,y_{0J})$ and to successively draw a new component in its conditional distribution given the other components fixed at their more recent values. Since the conditional distribution of a component given the others is a univariate normal distribution restricted to an interval, its simulation is straightforward. Indeed, the simulation can be done by using a rejection method, or by using the fact a random variable $\mathcal{X}$ following the standard normal distribution restricted to an interval $]\,c,\,d\,[$ is deduced from a random variable $\mathcal{U}$ following the uniform distribution on $[0,1]$, by the formula:

$$\mathcal{X}=\Phi^{-1}\{[\Phi(d)-\Phi(c)]\mathcal{U}+\Phi(c)\}\,, \tag{A.4}$$

since $P(\mathcal{X}<x)=P(\Phi(\mathcal{X})<\Phi(x))=P\left[\mathcal{U}<\dfrac{\Phi(x)-\Phi(c)}{\Phi(d)-\Phi(c)}\right]=\dfrac{\Phi(x)-\Phi(c)}{\Phi(d)-\Phi(c)}$. This algorithm is usually faster than the rejection algorithm.

Quantitative information and interval information on responses

Let us now consider the case where the quantitative information is still $\varepsilon_t^K=\alpha$ but the interval information are related to some responses at some horizons. More precisely the interval information are:

$$c_{jh}<\Theta_h^j\varepsilon_t<d_{jh}$$

where the pair $(j,h) \in S \subset \{1,\dots,n\} \times \{1,\dots,H\}$ and Θ_h^j is the j^{th} row of Θ_h. In this case, we have to compute:

$$E[\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, c_{jh} < \Theta_h^j \varepsilon_t < d_{jh}, (j,h) \in S] \,,$$

where $i \in \overline{K} = \{1,...,n\} - K$. The conditional distribution of $\varepsilon_t^{\overline{K}}$ given $\varepsilon_t^K = \alpha$ is Gaussian and the previous expectation can be computed by a Monte Carlo method based on the rejection principle, that is, by using simulations in this distribution and keeping them if they satisfy the inequality constraints. If $card(S) \leq n$, the Gibbs algorithm can also be used, provided that a linear transformation is first done on ε_t in such a way that the $\Theta_h^j \varepsilon_t$, $(j,h) \in S$, are components of the transformed random vector.

Appendix 2: VARMA-distributed linear filters and associated MA representation

Let us represent the k-dimensional linear filter $\widetilde{Y}_t$ as follows:

$$\widetilde{Y}_t = F(L)\, Y_t = F_1(L)\, Y_t^{(1)} + F_2(L)\, Y_t^{(2)} \,, \qquad (A.5)$$

where $Y_t^{(1)}$ denotes the first k components of Y_t and $Y_t^{(2)}$ collects the last $(n-k)$ ones. The process $Y_t^* := \left(\widetilde{Y}_t', Y_t^{(2)\prime}\right)'$ can be written as:

$$Y_t^* = \begin{pmatrix} F_1(L) & F_2(L) \\ 0 & I \end{pmatrix} Y_t = F^*(L)\, Y_t \qquad (A.6)$$

and therefore

$$Y_t = \begin{pmatrix} F_1^{-1}(L) & -F_1^{-1}(L)\, F_2(L) \\ 0 & I \end{pmatrix} Y_t^* = [F^*(L)]^{-1}\, Y_t^* \,, \qquad (A.7)$$

thus implying the following VARMA representation for Y_t^*:

$$\Phi(L)\, F_{adj}^*(L)\, Y_t^* = \det\left(F_1(1)\right) \nu + \det\left(F_1(L)\right) \varepsilon_t \,, \qquad (A.8)$$

given that $[F^*(L)]^{-1} = [\det\left(F^*(L)\right)]^{-1} F_{adj}^*(L) = [\det\left(F_1(L)\right)]^{-1} F_{adj}^*(L)$, where $F_{adj}^*(L)$ is the adjoint matrix of $F^*(L)$. The MA representation, relevant for response function analysis, is given by:

$$Y_t^* = F^*(1)\,\Theta(1)\,\nu + \widetilde{\Theta}(L)\,\eta_t\,, \tag{A.9}$$

where $\widetilde{\Theta}(L) := F^*(L)\,\Theta(L)\,(F^*(0))^{-1}$, $\Theta(L) = \Phi^{-1}(L)$ and $\eta_t = F^*(0)\,\varepsilon_t$ is the innovation of Y_t^*.

Appendix 3: Decomposition of the long-term interest rate and application of the NIRF methodology

The joint dynamics of $Y_t = (r_t, S_t, g_t, \pi_t)'$ is described by the following Gaussian VAR(p) process:

$$Y_t = \nu + \sum_{j=1}^{p} \Phi_j Y_{t-j} + \varepsilon_t \tag{A.10}$$

that can be rewritten in a VAR(1) form:

$$Z_t = \widetilde{\nu} + \Phi\, Z_{t-1} + \widetilde{\varepsilon}_t \tag{A.11}$$

where $Z_t = (Y_t', Y_{t-1}', ..., Y_{t-p+1}')'$, $\widetilde{\nu} = (\nu', 0, ..., 0)'$ and

$$\Phi = \begin{pmatrix} \Phi_1 & \dots & \dots & \Phi_p \\ I_{(p\times p)} & 0_{(p\times p)} & \dots & 0_{(p\times p)} \\ 0_{(p\times p)} & I_{(p\times p)} & \dots & 0_{(p\times p)} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{(p\times p)} & \dots & I_{(p\times p)} & 0_{(p\times p)} \end{pmatrix}$$

where $I_{(p\times p)}$ is the $(p \times p)$ identity matrix and $0_{(p\times p)}$ the $(p \times p)$ matrix of zeros. Let us first assume that the set of information available to agents at the present date $t-1$ consists in the past and present values of Z_{t-1}, that is $\Omega_{t-1} = \underline{Z_{t-1}}$. In this case:

$$\begin{aligned} E\,(r_{t+h-1}\,|\,\Omega_{t-1}) &= e_1'((I-\Phi)^{-1}(I-\Phi^h)\widetilde{\nu} + \Phi^h Z_{t-1}) \\ E\,(\pi_{t+h-1}\,|\,\Omega_{t-1}) &= e_4'((I-\Phi)^{-1}(I-\Phi^h)\widetilde{\nu} + \Phi^h Z_{t-1}) \end{aligned}$$

where e_i is the i^{th} column of the $(4p \times 4p)$ identity matrix. Therefore, we have:

$$\begin{aligned}
EX_{t-1}(H) &= d_0(H) + c_0(H)Z_{t-1}\,, \;\; \Pi^e_{t-1}(H) = d_1(H) + c_1(H)Z_{t-1} \\
TP_{t-1}(H) &= d_2(H) + c_2(H)Z_{t-1}\,, \;\; \widetilde{EX}_{t-1}(H) = d_3(H) + c_3(H)Z_{t-1}
\end{aligned}$$

where

$$\begin{aligned}
d_0(H) &= \frac{1}{H}e_1'(I-\Phi)^{-1}\left(\sum_{h=0}^{H-1}(I-\Phi^h)\widetilde{\nu}\right) \\
c_0(H) &= \frac{1}{H}e_1'\left(\sum_{h=0}^{H-1}\Phi^h\right) \\
d_1(H) &= \frac{1}{H}e_4'(I-\Phi)^{-1}\left(\sum_{h=1}^{H}(I-\Phi^h)\widetilde{\nu}\right) \\
c_1(H) &= \frac{1}{H}e_4'\left(\sum_{h=1}^{H}\Phi^h\right) \\
d_2(H) &= -d_0(H)\,, \;\; c_2(H) = e_1' + e_2' - c_0(H) \\
d_3(H) &= d_0(H) - d_1(H)\,, \;\; c_3(H) = c_0(H) - c_1(H)
\end{aligned}$$

Hence, the components of $R_{t-1}(H)$ can be expressed as linear filter of the variables in the VAR and, thus, the technique of Section 3.3 applies.

Let us now assume that the set of information available at $t-1$ also includes some information about the future path of one variable in the VAR like the short rate. For instance, let us assume that future values of the short rate are known until date $t+\overline{H}-1$. We denote by $\widetilde{Z}_{t,t+\overline{H}-1} = \left\{\overline{r}_t, \overline{r}_{t+1}, ..., \overline{r}_{t+\overline{H}-1}\right\}$ these $\overline{H}$ known values of the short rate. Hence $\Omega_{t-1} = \left\{\widetilde{Z}_{t,t+\overline{H}-1}, \underline{Y_{t-1}}\right\}$. The NIRF of $EX_{t+k-1}(H)$, given Ω_{t-1}, is:

$$E(EX_{t+k-1}(H)|\Omega_{t-1}) - E(EX_{t+k-1}(H)|\underline{Y_{t-1}})\,, \qquad (A.12)$$

and the first component in $(A.12)$ is:

$$E(EX_{t+k-1}(H)|\Omega_{t-1}) = \frac{1}{H}E\left(\sum_{h=0}^{H-1}E(r_{t+k+h-1}\,|\,\Omega_{t+k-1})\,|\,\Omega_{t-1}\right)$$
$$= \frac{1}{H}\sum_{h=0}^{H-1}E\left(r_{t+k+h-1}\,|\,\Omega_{t-1}\right)\,,$$

where $\Omega_{t+k-1} = \{\widetilde{Z}_{t+k,t+\overline{H}-1}, Y_{\underline{t+k-1}}\}$ if $k \leq \overline{H}-1$ and $\Omega_{t+k-1} = Y_{\underline{t+k-1}}$ otherwise, and with $E\left(r_{t+k+h-1}|\Omega_{t-1}\right) = \overline{r}_{t+k+h-1}$ for $k+h \leq \overline{H}$. Similarly, the NIRF of $\Pi^e_{t-1}(H)$ is given by:

$$E(\Pi^e_{t+k-1}(H)|\Omega_{t-1}) - E(\Pi^e_{t+k-1}(H)|Y_{\underline{t-1}})\,, \qquad (A.13)$$

and the first component in $(A.13)$ is:

$$E(\Pi^e_{t+k-1}(H)|\Omega_{t-1}) = \frac{1}{H}E\left(\sum_{h=1}^{H}E(\pi_{t+k+h-1}|\Omega_{t+k-1})|\Omega_{t-1}\right)$$
$$= \frac{1}{H}\sum_{h=1}^{H}E\left(\pi_{t+k+h-1}|\Omega_{t-1}\right).$$

References

Banbura M., Giannone D. and M. Lenza, 2010, "Scenario analysis using the Kalman Filter: an application to large Euro Area VAR", mimeo.

Bernanke B., 1986, "Alternative explanation of the money-income correlation", Carnegie Rochester conference series on public policy, Vol. 15, pp. 49-100.

Bernanke B. and I. Mihov, 1998, "Measuring monetary policy", Quarterly Journal of Economics, Vol.113, pp.869-902.

Blanchard O., and D. Quah, 1989, "The Dynamic Effects of Aggregate Demand and Supply Disturbances", American Economic Review, Vol. 79, pp. 655-673.

Blanchard, O. and M. Watson, 1986, "Are Business Cycles All Alike?" In Robert Gordon (ed.), Continuity and Change in the American Business Cycle. NBER and the University of Chicago Press.

Boivin, J. and M. Giannoni, 2006, "Has monetary policy become more effective?"', Review of Economic and Statistics, Vol. 88, pp 445-462.

Brissimis, S. and S. Magginas, 2006, "Forward-looking information in VAR models and the price puzzle", Journal of Monetary Economics, Vol. 53, pp 1225-1234.

Castelnuovo, E. and P. Surico, "Monetary policy, Inflation expectations and the price puzzle", 2010, The Economic Journal, Vol. 120, pp 1262-1283.

Clarida, R., Gali, J. and M. Gertler, 2000, "Monetary policy rules and macroeconomic stability: evidence and some theory", Quarterly Journal of Economics, Vol. 115, pp 147-180.

Cochrane J., 1994, "Shocks", Meltzer, A., Plosser, C. (Eds.), Carnegie-Rochester Conference Series on Public Policy, vol. 41. North-Holland, Amsterdam, pp. 295-364.

Durbin, J., and S. J. Koopman, 2002, "A simple and efficient simulation smoother for state-space time series analysis", Biometrika, Vol. 89, pp 603-615.

Inoue A. and L. Kilian, 2011, "Inference on impulse response functions in structural VAR models", mimeo, University of Michigan.

Jardet C., Monfort A., and F. Pegoraro, 2009, "New Information Response Functions", Banque de France, Document de travail N. 235.

Jardet C., Monfort A., and F. Pegoraro, 2013, "No-arbitrage near-cointegrated VAR(p) term structure models, term premia and GDP growth", The Journal of Banking and Finance, 37, 2, 389-402.

Jarocinski M., 2010, "Conditional forecast and uncertainty about forecast revisions in vector autoregressions", Economic Letters, Vol. 108, pp.257-259.

Judd, J.P. and G. Rudebusch, 1998, "Taylor's rule and the Fed: 1970-1997", Federal reserve Bank of San Francisco Economic Review 98-03, pp 3-16.

Gali J. 1992, "How does the IS-LM model fit the post-war US data ?", Quarterly Journal of Economics, Vol. 107, pp. 709-738.

Gallant A., Rossi P., and G. Tauchen, 1993, "Nonlinear dynamic structures", Econometrica, Vol. 61(4), pp. 971-907.

Gourieroux C., and J. Jasiak, 2005, "Nonlinear innovations and impulse responses with application to VaR sensitivity", Annales d'Economie et de Statistique, 78, pp. 1-33.

Kilian L., 2011, "Structural Vector Auroregression", mimeo, University of Michigan.

Koop G., Pesaran H., and S. Potter, 1996, "Impulse response analysis in nonlinear multivariate models", Journal of Econometrics, Vol. 74, pp.119-148.

Lubik, T.A. and F. Schorfheide, 2004, "Testing for inderterminacy an application to U.S. monetary policy"', American economic review, Vol. 94, pp 190-217.

Lütkepohl H., 1991, "Introduction to multiple time series analysis", Springer-Verlag, Berlin.

Mountford A. and H Uligh, 2009, "What are the effects of fiscal policy shocks?", Journal of Applied Econometrics, Vol. 24, pp. 960-992.

Peersman G. and R. Straub, 2009, "Technology shocks and robust sign restrictions in a Euro Area SVAR", International Economic Review, Vol. 50, pp. 727-750.

Pesaran H., and Y. Shin, 1998, "Generalized Impulse Response Analysis in Linear Multivariate Models", Economics Letters, Vol.58, pp.17-29.

Rubio-Ramirez J., Waggoner D., and T. Zha, 2008, "Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference", Federal Reserve Bank of Atlanta, working paper 2008-18.

Sims C., 1980, "Macroeconomics and reality", Econometrica, Vol. 48, pp. 11-48.

Uhlig, H., 2005, "What are the effects of monetary policy on output? Results from an agnostic identification procedure" Journal of Monetary Economics, vol. 52(2), pp. 381-419.