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

First version : February, 2009. This version : October, 2011.

Abstract

We propose a new methodology for the analysis of impulse response functions in VAR or VARMA models. More precisely, we build our results on the non ambiguous notion of innovation of a stochastic process and we consider the impact of any kind of new information at a given date t on the future values of the process. This methodology allows to take into account qualitative or quantitative information, either on the innovation, or on future responses, or on filters, or on future paths of variables of interest. We show, among other results, that our approach encompasses several standard methodologies found in the literature, such as the orthogonalization of shocks (Sims (1980)), the "structural" identification of shocks (Blanchard and Quah (1989)), the "generalized" impulse responses (Pesaran and Shin (1998)), or the impulse vectors (Uhlig (2005)). Finally, working with a parsimonious Gaussian VAR(p) model estimated on U.S. quarterly data, we exploit the NIRF methodology to address two monetary policy issues. The first one concerns the shift (at the end of the 1970) to a more anti-inflationary monetary policy, while the second one focuses on the effects on macro variables of the FOMC stabilization announcement (at the end of 2008) of the future short rate around the zero lower bound.

Keywords: impulse response functions, innovation, new information, linear filter, monetary policy issues, 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:

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. A part of this literature is devoted to the notion of orthogonalized shocks while another important one, initiated by Blanchard and Watson (1986), Bernanke (1986) and Blanchard and Quah (1989), discusses the definition of "structural" shocks. However, both approaches rely on ad-hoc assumptions, such as the ordering of the variable in the VAR or expected effects that such shocks should have on a given variable. Unfortunately, these hypotheses are not always consensual. 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), Cochrane (1994)). Finally, in response to shortcomings of traditional orthogonalized and structural approaches, a third part of the literature proposes a statistical or "agnostic" approach, either in a bayesian way (Uhlig (2005)), or in a classical way (Pesaran and Shin (1998)).

In this paper we try to push as far as possible this statistical approach. The idea is to build our results on the notion of innovation ε_t (say) of a stochastic process, that is, the difference between the value of the process and its conditional expectation given its past. In contrast with "structural shocks", the statistical innovation is non ambiguous as it is defined in an unique way. Furthermore, the idea is to exploit the fact that we sometimes have at our disposal an information set on this innovation process which is larger, or smaller, than the simple information on the contemporaneous realization of one of its component. For instance, this "new information" can be related to the values, the signs, or more generally to a given possible interval, for one or several components of ε_t, or to the value of one or several average responses. It can also be related to a linear filter of ε_t, or even to the future path of some of its components. In this case, one may be interested in analysing the expected dynamical effects of such a "new information" on variables of the model. Therefore in

Fulvio.PEGORARO@banque-france.fr] and CREST, Laboratoire de Finance-Assurance [E-mail: pegoraro@ensae.fr].

We would like to thank Mike Chernov, Christian Gourieroux and Stephane Grégoir, as well as participants at the Banque de France October 2008 seminar, June 2009 Bank of France - Bundesbank Conference on "The Macroeconomy and Financial Systems in Normal Times and in Times of Stress" for helpful comments and remarks.

this paper we are interested in the responses to this notion of "new information", and we develop a general setting which is suitable to deal with these issues. This general methodology is named *New Information Response Function*.

More formally, we start with the important case where the new information only concerns the contemporaneous value ε_t of this innovation process, i.e. is of the form $a(\varepsilon_t) = \alpha$, where $a(.)$ is some given function and α some given vector of real numbers. This general setting contains many relevant particular cases. We first consider the "full new information" case where $a(.)$ is one-to-one. Here we have a unique value for the innovation and we show that the standard orthogonalized shocks, the impulse vectors introduced by Uhlig (2005) and the structural shocks can be viewed as particular cases of such full information. Second, we consider the case of "continuous limited new information" where $a(.)$ is not one-to-one and $a(\varepsilon_t)$ has a continuous probability distribution. This case includes the "generalized" impulse response function introduced by Pesaran and Shin (1998), based on the information on one innovation, but it also allows to take into account many other kinds of informations, including informations on several innovations and informations on responses. Third, we study the "discrete limited new information" case where the new information is based on discrete functions, like indicator functions and, in particular, sign functions, on either the innovation itself, or on an impulse vector, or on a response. Fourth, this general setting is used to consider new information on a linear filter of the vector of interest and responses of a linear filter.

Although the case where the new information only concerns the contemporaneous value ε_t of the innovation process is the more frequent, we also consider other important cases. First, we study the case where the information also depends on the observed past values $\underline{Y_{t-1}} = (Y'_{t-1}, Y'_{t-2}, \ldots)'$ of the process itself, that is the case where the new information can be written as $a(\varepsilon_t, \underline{Y_{t-1}}) = \alpha$. Second, we study the case where the new information also depends on future values of the innovation process (or, equivalently, on future values of the process itself), and in this case the new information can be written as $a(\underline{\varepsilon_{t:T}}, \underline{Y_{t-1}}) = \alpha$, where $\underline{\varepsilon_{t:T}} = (\varepsilon'_t, \varepsilon'_{t+1}, \ldots, \varepsilon'_T)'$.

Finally, in order to provide an empirical illustration of the NIRF methodology we address two monetary issues using a Gaussian VAR(p) model that links macroeconomic variables and interest rates. On the one hand, we test whether or not there is a shift toward a more aggressive anti-inflationary monetary policy in the U.S. at the end of the 1970s. For that purpose, we analyse the effects of a new information one the one-year ahead expected inflation, which can be expressed as a linear filter on the variables in the VAR (namely, short rate, one-year spread, GDP growth and inflation rate), before and after 1979:Q4. On the other hand, we investigate the effects on interest rates and economic variables of the communication by the Federal Open Market Committee (FOMC) in 2008:Q4, regarding the stability of future short term interest rate.

The paper is organized as follows. In Section 2 we define the new information response function. In Section 3 this concept is applied to the full new information case. Section 4 is devoted to the continuous limited new information case, while Section 5 deals with the discrete limited new information one. In Section 6 we show how these results can be used to analyze shocks on a linear filter and responses of a filter. Section 7 deals with path-dependent new information response functions, while in Section 8 we present the empirical applications. Finally, Section 9 concludes and proposes further developments, while Appendices gather Tables and further results.

2 Response to a new information on a function of a VAR innovation

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 $\tau = 0, ..., t$ 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:

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

(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 $\Theta_s = 0$ if $s < 0$, $\Phi_0 = I$, $\Phi_i = 0$ if $i > p$. Equation (5) provides a straightforward way to compute recursively the matrices Θ_τ.

Denoting $Y_{\underline{t}} = (Y'_t, Y'_{t-1}, ..., Y'_{t-p})'$, equation (3) implies:

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

so $\Theta_h \varepsilon_t$ measures the differential impact of the knowledge of ε_t on the prediction updating of Y_{t+h} between dates $t-1$ and t.

More generally, let us consider the differential impact on the prediction of Y_{t+h} of a new information

$a(\varepsilon_t) = \alpha$, where $a(.)$ is some function and α is given. Obvious examples of such functions are : $a(\varepsilon_t) = \varepsilon_t$, $a(\varepsilon_t) = b'\varepsilon_t$, $a(\varepsilon_t) = \mathbb{1}_{\mathbb{R}^+}(b'\varepsilon_t)$, where b is some vector, etc. This impact, also called new information response function (NIRF), is by definition:

$$E(Y_{t+h}|a(\varepsilon_t), \underline{Y_{t-1}}) - E(Y_{t+h}|\underline{Y_{t-1}})\,, \tag{7}$$

and a key result is the equivalent representation of (7) by:

$$\begin{aligned} E\left\{[E(Y_{t+h}|\varepsilon_t, \underline{Y_{t-1}}) - E(Y_{t+h}|\underline{Y_{t-1}})]|a(\varepsilon_t), \underline{Y_{t-1}}\right\} &= E[\Theta_h \varepsilon_t | a(\varepsilon_t), \underline{Y_{t-1}}] \\ &= \Theta_h \, E[\varepsilon_t | a(\varepsilon_t)]\,. \end{aligned} \tag{8}$$

Thus the average impact on Y_{t+h} of the new information $a(\varepsilon_t)$ at time t is the same as the one which would be implied by the new information $\varepsilon_t = \delta$ with $\delta = E[\varepsilon_t \,|\, a(\varepsilon_t)]$. Note that this impact $\Theta_h \, \delta$ can also be obtained from (2), with $\nu = 0$, by computing recursively $Y_t, \ldots, Y_{t+h}$, with $Y_s = 0$, $s < t$, $\varepsilon_t = \delta$ and $\varepsilon_s = 0$, $s > t$.

In Sections from 3 to 7 we will distinguish five important situations according to the properties of function $a(.)$:

i) the "full new information" case, when $a(.)$ is one-to-one.

ii) the "continuous limited new information" case, when $a(.)$ is not one-to-one and when the probability distribution of $a(\varepsilon_t)$ is continuous (i.e., absolutely continuous with respect to the Lebesgue measure).

iii) the "discrete limited new information" case, when the distribution of $a(\varepsilon_t)$ has a discrete component.

iv) the case of a new information on a linear filter $\widetilde{Y}_t = F(L)\, Y_t$, with innovation $\widetilde{\varepsilon}_t = F(0)\, \varepsilon_t$, defined by $\widetilde{a}[F(0)\, \varepsilon_t] = \alpha$.

v) the case of path-dependent new information in which $a(\varepsilon_t, \underline{Y_{t-1}}) = \alpha$ (”past path-dependent” NIRF) or $a(\underline{\varepsilon_{t:T}}, \underline{Y_{t-1}}) = \alpha$, with $\underline{\varepsilon_{t:T}} = (\varepsilon'_t, \varepsilon'_{t+1}, \ldots, \varepsilon'_T)'$ (”future path-dependent” NIRF).

3 Full new information

If $a(.)$ is one-to-one, the average impact on Y_{t+h} of the new information $a(\varepsilon_t) = \alpha$ is obviously $\Theta_h a^{-1}(\alpha)$. This simple situation contains the following well known cases: 1) the orthogonalized shocks; 2) the Uhlig (2005)'s impulse vectors and 3) the structural shocks.

3.1 Orthogonalized shocks

Let us consider the lower triangular matrix P defined by $\Sigma = PP'$ and the orthogonalized errors ξ_t defined by $\varepsilon_t = P\xi_t$. The distribution of ξ_t is obviously $N(0, I)$ and it is usual to consider a new information e_j on ξ_t, where e_j is the j^{th} column of the $n \times n$ identity matrix I. Such a new information is called a “shock” of 1 on ξ_{jt} or a “shock” of e_j on ξ_t. It is clear that the impact on Y_{t+h} of such a shock is the same as the shock $\delta = Pe_j$ on ε_t, namely $\Theta_h P e_j$, or $\Theta_h P^{(j)}$, where $P^{(j)}$ is the j^{th} column of P. In particular, the immediate impact on ε_t (or Y_t) is $P^{(j)}$, so there is no immediate impact on the component Y_{it} if $i < j$, and the immediate impact on Y_{jt} is P_{jj} (the (j, j) entry of P).

If we want an immediate impact on Y_{jt} equal to one, we can consider the lower triangular matrix $\widetilde{P} = PD^{-1}$, where D is the diagonal matrix $diag(P_{jj})$, and the vector ζ_t defined by $\zeta_t = D\xi_t$ or $\varepsilon_t = \widetilde{P}\zeta_t$. Now, a shock e_j on ζ_t has the impact $\overline{\delta} = \widetilde{P}^{(j)}$ on ε_t (or Y_t) and $\Theta_h \widetilde{P}^{(j)}$ on Y_{t+h}. Also note that (1) can be rewritten:

$$\widetilde{P}^{-1}\Phi(L)Y_t = \widetilde{P}^{-1}\nu + \zeta_t \tag{9}$$

and since $\widetilde{P}^{-1}$ is lower triangular with diagonal terms equal to 1, (9) is a recursive form of the VAR. So the average impact on Y_{t+h} of a shock e_j on ζ_t, could be obtained from (9) with $\nu = 0$, by computing recursively $Y_t, Y_{t+1}, ..., Y_{t+h}$ with $Y_s = 0,\ s < t,\ \zeta_t = e_j$ and $\zeta_s = 0,\ s > t$.

3.2 Uhlig (2005)'s impulse vectors

Uhlig (2005) defined an impulse vector $\gamma \in \mathbb{R}^n$ as a vector such that there exists a matrix A verifying $AA' = \Sigma$ and admitting γ as a column. The set of vectors satisfying this definition can be seen as all the possible new informations on ε_t implied by a shock of 1 on a component of a "fundamental" error η_t satisfying $\varepsilon_t = A\eta_t$ and $V(\eta_t) = I$.

It turns out [see Uhlig (2005)] that those vectors γ are characterized by $\gamma = P\beta$, where P is defined in Section 3.1, and β is a unit length vector of $\mathbb{R}^n$. Equivalently, these vectors belong to the set Γ defined by $\gamma'(P^{-1})'P^{-1}\gamma = 1$ or $\gamma'\Sigma^{-1}\gamma = 1$ and therefore, they generate an hyperellipsoïd.

An impulse vector γ is a particular full new information on ε_t whose impact on Y_{t+h} is $\Theta_h\gamma$ and the set of all possible impacts on Y_{t+h} coming from an impulse vector is $\Theta_h P\beta$, where β is of length one.

3.3 Structural shocks

A structural error is a vector η_t satisfying $\varepsilon_t = A\eta_t$, with $\Sigma = AA'$, and, therefore $V(\eta_t) = I$, like a "fundamental" vector considered in Section 3.2. Moreover, a structural error is uniquely defined by identification conditions which could be based on short run restrictions, imposing for instance that a shock e_j on η_t has no immediate impact on ε_{it}, i.e. $A_{ij} = 0$, or which could be based on long-run restrictions when Y_t is non-stationary and admits r cointegrating relationships. In the latter case, we can 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), and $\Lambda' Y_t$ a r-dimensional vector of cointegrating relationships, and such that W_t

has a stationary VAR representation of the form:

$$\Gamma(L)W_t = C\nu + C\varepsilon_t$$

where $C = \begin{pmatrix} I_{n-r} & 0 \\ & \Lambda' \end{pmatrix}$ is invertible.

The long run impact on the scalar component y_{it}, $i \leq n-r$, of a shock e_j on η_t is $[\Gamma^{-1}(1)CA^{(j)}]_i$ where $A^{(j)}$ is the j^{th} column of A, and imposing that such long run impacts are zero may imply identification [see Blanchard and Quah (1993) and Rubio-Ramirez, Waggoner and Zha (2008)]. In any case, a shock e_j on η_t is a full information $A^{(j)}$ on ε_t.

4 Continuous limited new information

Let us now consider the case where $a(.)$ is not one-to-one and $a(\varepsilon_t)$ has an absolutely continuous distribution. In this situation the new information $a(\varepsilon_t) = \alpha$ (say) does not define ε_t and we have to compute $\delta = E[\varepsilon_t | a(\varepsilon_t) = \alpha]$ in order to obtain the impact $\Theta_h \delta$ on Y_{t+h}. Since the event $a(\varepsilon_t) = \alpha$ has probability zero, we have to find the conditional expectation in a continuous distribution context and some examples are given below.

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

Pesaran and Shin (1998) considered the case where $a(\varepsilon_t) \equiv \varepsilon_{jt}$, that is the case where we have a new information only for a 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)].

4.2 New information on a set of individual innovations

If $a(\varepsilon_t) \equiv \varepsilon_t^K$, where ε_t^K is a K-dimensional subvector of ε_t containing any 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}.

4.3 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 $\delta = \Sigma \, \Theta_{h_1}^{K_1 \prime} (\Theta_{h_1}^{K_1} \Sigma \, \Theta_{h_1}^{K_1 \prime})^{-1} \alpha_1$ 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 4.2, 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 $M = (\Theta_{h_1}^{K_1 \prime}, \mathcal{S}_2')'$, we have:

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

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

4.4 New information on long run behavior

Assuming the same framework as in Section 3.3, we can consider a partial new information imposing that the long run impact on y_{it} is zero, that is:

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

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

4.5 Information defined as the set of impulse vectors Γ

A natural question is to identify the impact of the information imposing that ε_t belongs to the set Γ of the impulse vectors introduced by Uhlig (2005). As we have seen in Section 3.2, the set of impulse vectors is $\Gamma = \{\gamma \in \mathbb{R}^n : \gamma' \Sigma^{-1} \gamma = 1\}$, or equivalently $\Gamma = \{\gamma \in \mathbb{R}^n : \gamma = P\beta, \beta'\beta = 1\}$, where P is the lower triangular matrix satisfying $\Sigma = PP'$.

If the new information is $\varepsilon_t \in \Gamma$, i.e. $\varepsilon_t' \Sigma^{-1} \varepsilon_t = 1$, that is, if $a(\varepsilon_t) = \varepsilon_t' \Sigma^{-1} \varepsilon_t$ and $\alpha = 1$, we have to compute $E[\varepsilon_t | \varepsilon_t \in \Gamma]$. Since $\varepsilon_t = P\xi_t$, with $\xi_t \sim N(0, I)$ and $E[\varepsilon_t | \varepsilon_t \in \Gamma] = PE[\xi_t | \xi_t' \xi_t = 1]$, we have by symmetry $E[\varepsilon_t | \varepsilon_t \in \Gamma] = 0$. Therefore, the new information $\varepsilon_t \in \Gamma$ has no impact in average on Y_{t+h}. Additional constraints are considered in Section 5.5.

5 Discrete limited new information

5.1 Definition of the new information

Let us now consider the case where the distribution of $a(\varepsilon_t)$ has a discrete component. More precisely we assume that $a(.) = \begin{pmatrix} a_1(.) \\ a_2(.) \end{pmatrix}$, where $a_1(\varepsilon_t)$ has a continuous distribution and $a_2(\varepsilon_t)$ is valued in a finite set $\overline{\alpha}_2 = \{\alpha_{21}, ..., \alpha_{2L}\}$. In this case the conditional distribution of any component ε_{it} of ε_t given $a_1(\varepsilon_t) = \alpha_1$ and $a_2(\varepsilon_t) = \alpha_{2j} \in \overline{\alpha}_2$ is obtained by the conditional distribution of ε_{it} given $a_1(\varepsilon_t) = \alpha_1$ restricted to the set $a_2(\varepsilon_t) = \alpha_{2j}$. In other words, for any set S:

$$P(\varepsilon_{it} \in S \,|\, a_1(\varepsilon_t) = \alpha_1, a_2(\varepsilon_t) = \alpha_{2j}) = \frac{P(\varepsilon_{it} \in S, a_2(\varepsilon_t) = \alpha_{2j} | a_1(\varepsilon_t) = \alpha_1)}{P(a_2(\varepsilon_t) = \alpha_{2j} | a_1(\varepsilon_t) = \alpha_1)} .$$

Note that a simulation in this conditional distribution of ε_t given $a_1(\varepsilon_t) = \alpha_1$ and $a_2(\varepsilon_t) = \alpha_{2j}$ can be obtained by simulating independently a sequence in the conditional distribution of ε_t given $a_1(\varepsilon_t) = \alpha_1$ and keeping the first simulation $\widetilde{\varepsilon}_t$ satisfying $a_2(\widetilde{\varepsilon}_t) = \alpha_{2j}$. It is a simple rejection algorithm. The conditional expectation $E[g(\varepsilon_t) | a_1(\varepsilon_t) = \alpha_1 \text{ and } a_2(\varepsilon_t) = \alpha_{2j}]$, where g is some given function, can be approximated by the empirical mean of $g(\widetilde{\varepsilon}_t^s)$, $s = 1, ..., S$ and where $\widetilde{\varepsilon}_t^s$ are obtained by keeping the simulations satisfying $a_2(\widetilde{\varepsilon}_t) = \alpha_{2j}$ in a sequence of independent simulations in the conditional distribution of ε_t given $a_1(\varepsilon_t) = \alpha_1$. However, in some cases explicit forms of such conditional expectations are available.

5.2 Quantitative information and one interval information

Let us consider the case where $a_2(\varepsilon_t) = 1\!\!1_{]\,c,\,d\,[}(\varepsilon_{jt})$ and $a_1(\varepsilon_t) = \varepsilon_t^K$ with c and d real numbers $(c < d)$ and $K \subset \{1, ..., n\}$ such that $j \notin K$. Our purpose is 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.

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:

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

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:

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

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

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

5.3 Quantitative information and several interval informations

We still assume $a_1(\varepsilon_t) = \varepsilon_t^K$, but now $a_2(\varepsilon_t)$ is the set of functions $\left\{1\!\!1_{]\,c_j, d_j\,[}(\varepsilon_{jt}), j \in J\right\}$, 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

$$\begin{aligned} & E[\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, c_j < \varepsilon_{jt} < d_j, j \in J] \qquad (10) \\ = \quad & E[E(\varepsilon_{it} \,|\, \varepsilon_t^K = \alpha, \varepsilon_{jt}, j \in J) \,|\, \varepsilon_t^K = \alpha, c_j < \varepsilon_{jt} < d_j, j \in J]\,, \end{aligned}$$

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:

$$\mathcal{X} = \Phi^{-1}\{[\Phi(d) - \Phi(c)]\,\mathcal{U} + \Phi(c)\}\,, \qquad (11)$$

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

5.4 Quantitative information and interval informations on responses

The quantitative information is still $\varepsilon_t^K = \alpha$ but the interval informations are related to some responses at some horizons. More precisely the interval informations are:

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

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

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

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 $\Theta_h^{(j)\prime} \varepsilon_t$, $(j, h) \in S$, are components of the transformed random vector.

This method provides a way to find the new information on ε_t, which satisfies some constraints $\varepsilon_t^K = \alpha$ and some constraints on the impulse response function and which is the more in agreement with the stochastic behavior of ε_t.

5.5 Impulse vector and set information on responses: looking for structural shocks

Uhlig (2005) considered the case where the information is $\varepsilon_t \in \Gamma$, the set of impulse vectors, i.e. $\varepsilon_t'\Sigma^{-1}\varepsilon_t = 1$, and sign information on responses: $\Theta_h^{(j)\prime}\varepsilon_t > 0$, $(j,h) \in S$.
The conditional expectation

$$\delta_i = E[\varepsilon_{it}|\varepsilon_t'\Sigma^{-1}\varepsilon_t = 1, \Theta_h^{(j)\prime}\varepsilon_t > 0, (j,h) \in S],$$

can still be computed by a Monte-Carlo method. Indeed the conditional distribution of ε_t given $\varepsilon_t'\Sigma^{-1}\varepsilon_t = 1$ is the image by P of the conditional distribution of ξ_t given, $\xi_t'\xi_t = 1$, where $\xi_t \backsim N(0, I)$ which is the uniform distribution on the unit sphere. So, the method is a follows:

- draw ξ from $N(0, I)$
- compute $\widetilde{\xi} = \dfrac{\xi}{(\xi'\xi)^{1/2}}$
- compute $\widetilde{\varepsilon} = P\widetilde{\xi}$
- keep the simulation if $\Theta_h^{(j)\prime}\widetilde{\varepsilon} > 0$, $(j,h) \in S$.

The expectations are obtained from the empirical means of the retained simulations.
Uhlig (2005) used a bayesian approach requiring n_1 (say) drawings in the posterior of the VAR parameters, then for each drawing of the parameters, n_2 drawings of ξ uniformly on the unit sphere; then for each pair of drawings $(\Sigma, \widetilde{\xi})$, the computation of the Choleski matrix $\widetilde{P}$, the computation of the candidate impulse vector $\widetilde{\gamma} = \widetilde{P}\widetilde{\xi}$ and of the whole impulse response functions $\Theta_h^{(j)\prime}\widetilde{\gamma}$, $h \in \{1, \ldots, H\}$, and, finally, keeping the candidate $\widetilde{\gamma}$ if the IRF satisfies the sign constraints, this method provides a drawing in the posterior of the impulse vector of interest.
In our method, the conditional expectation $\delta = (\delta_1, \ldots, \delta_n)'$ obviously satisfies the condition $\delta'\Sigma^{-1}\delta = 1$ and is therefore an impulse vector which is easily computed, does not necessitate

any prior distribution and is nicely interpreted as the best predictor impulse vector satisfying the sign constraints. Moreover, our method is easily extended to the case where the information on responses is a set information imposing that the vector $[\Theta_h^{(j)\prime}\widetilde{\varepsilon}, (j,h) \in S]$ belongs to some set $\mathcal{E}$, for instance a product of intervals.

6 New information on a filter and responses of a filter

6.1 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 Section 8 and Jardet, Monfort, Pegoraro (2011)].

Let us consider a filter $\widetilde{Y}_t = F(L)Y_t$, where $F(L) = (F_1(L), ..., F_n(L))$ is a row vector of polynomials in the lag operator L. The innovation of $\widetilde{Y}_t$ at t is $\widetilde{\varepsilon}_t = F(0)\varepsilon_t$, and therefore an information on $\widetilde{\varepsilon}_t$, defined by $\widetilde{a}(\widetilde{\varepsilon}_t) = \alpha$, can be written as $\widetilde{a}[F(0)\varepsilon_t] = \alpha$ or $a(\varepsilon_t) = \alpha$ (say). This means that, an information 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 $\widetilde{\varepsilon}_t = 1$, the impact on Y_{t+h} is $\Theta_h\delta$, where

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

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 informations on several filters in a straightforward way, and extend the technique to interval informations.

Let us note that if we are interested in k filters $\widetilde{Y}_{1t}, ..., \widetilde{Y}_{kt}$, it is always possible to complete with $n-k$ components of Y_t and to apply the NIRF techniques to the dynamic model followed by the vector thus obtained Y_t^*. However this would be an awkward method since Y_t^* has a VARMA representation implying tedious computations.

6.2 Response of a filter

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 $s < 0$. It is clear that we can also impose interval constraints on some responses of a filter to a new information which, in turn, may involve this filter or other filters.

7 Path-Dependent New Information Response Functions

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 (8) the impact of $\underline{Y_{t-1}}$ cancels out. Let us now consider two situations in which the past values $\underline{Y_{t-1}}$ or the present and future values of some components of Y_t matter.

7.1 Past Path-Dependent NIRF

A relevant practical case in which the impact of the past values $\underline{Y_{t-1}}$ does not disappear is when the new information at t is no longer of the form $a(\varepsilon_t) = \alpha$, but is given by:

$$a(\varepsilon_t, \underline{Y_{t-1}}) = \alpha \,. \tag{12}$$

For instance, 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:

$$
\begin{aligned}
&\widehat{Y}_{t|t-1}^K(\underline{Y_{t-1}}) + \varepsilon_t^K = Y_{t-1}^K\,, \\
&\text{or}\ \ \varepsilon_t^K - Y_{t-1}^K + \widehat{Y}_{t|t-1}^K(\underline{Y_{t-1}}) = 0\,,
\end{aligned}
\tag{13}
$$

which is of the form $a(\varepsilon_t, \underline{Y_{t-1}}) = \alpha$. In this new setting, the NIRF becomes:

$$
\begin{aligned}
&E(Y_{t+h}|a(\varepsilon_t, \underline{Y_{t-1}}) = \alpha, \underline{Y_{t-1}} = \underline{y_{t-1}}) - E(Y_{t+h}|\underline{Y_{t-1}} = \underline{y_{t-1}}) \\
&\text{or}\ \ E(Y_{t+h}|a(\varepsilon_t, \underline{y_{t-1}}) = \alpha, \underline{Y_{t-1}} = \underline{y_{t-1}}) - E(Y_{t+h}|\underline{Y_{t-1}} = \underline{y_{t-1}})\,.
\end{aligned}
\tag{14}
$$

The first term of (14) can be written as:

$$
E[E(Y_{t+h}|\varepsilon_t, \underline{Y_{t-1}} = \underline{y_{t-1}})\,|\,a(\varepsilon_t, \underline{y_{t-1}}) = \alpha, \underline{Y_{t-1}} = \underline{y_{t-1}}]
\tag{15}
$$

and the NIRF becomes:

$$
\begin{aligned}
& E\left\{[E(Y_{t+h}|\varepsilon_t, \underline{Y_{t-1}} = \underline{y_{t-1}}) - E(Y_{t+h}\,|\,\underline{Y_{t-1}} = \underline{y_{t-1}})]\,|\,a(\varepsilon_t, \underline{y_{t-1}}) = \alpha, \underline{Y_{t-1}} = \underline{y_{t-1}}\right\} \\
= \;& E[\Theta_h \varepsilon_t\,|\,a(\varepsilon_t, \underline{y_{t-1}}) = \alpha, \underline{Y_{t-1}} = \underline{y_{t-1}}] \\
= \;& E[\Theta_h \varepsilon_t\,|\,a(\varepsilon_t, \underline{y_{t-1}}) = \alpha] \\
= \;& \Theta_h\, E[\varepsilon_t\,|\,a(\varepsilon_t, \underline{y_{t-1}}) = \alpha] \\
= \;& \Theta_h\, \delta_t\,,
\end{aligned}
\tag{16}
$$

with $\delta_t = E[\varepsilon_t\,|\,a(\varepsilon_t, \underline{y_{t-1}}) = \alpha]$. Therefore, the computation of the NIRF still boils down to the computation of a conditional expectation of ε_t given (now) some function of ε_t and $\underline{y_{t-1}}$. In example (13) we need to compute:

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

and we get:

$$\delta_t = \Sigma^K \, \Sigma_{K\,K}^{-1} \, [y_{t-1}^K - \widehat{y}_{t|t-1}^K(\underline{y_{t-1}})] \,, \tag{18}$$

where Σ^K and $\Sigma_{K\,K}$ are defined like in Section 4.2.

7.2 Future Path-Dependent NIRF

In some situations it is interesting to study the behavior of the future values of the endogeneous variables of a dynamic system, when the future path of one or several of them is imposed, and when simultaneously the present value of another set of variables is also fixed. For instance (see the application in the next section) we could impose the future values of the short rate and, at the same time, fix the present value of the GDP growth to its past value or the innovation of the GDP growth equal to zero. 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)], and we consider the possibility to add an information on the present value (or some isolated future values) of some other variables.

Let us partition Y_t into $(Y'_{1,t}, Y'_{2,t})$ where $Y_{1,t}$ is of size n_1 and $Y_{2,t}$ is of size $n_2 = n - n_1$. We assume, without loss of generality, that the values of $Y_{2,t}$ are imposed at all dates between t (today) and T, whereas the values of some components of $Y_{1,t}$ may be imposed at some present of future dates (but not all). More precisely, using the notation $Y_{2,t:T} = (Y'_{2,t}, ..., Y'_{2,T})'$, we impose the information:

$$\begin{aligned} &Y_{2,t:T} = \alpha_2 \,, \text{ and} \\ &C'_{t+h} Y_{1,t+h} = \alpha_{1,h} \,, \quad h = 0, ..., T-t \end{aligned}$$

where C_{t+h} is some selection vector choosing some components of $Y_{1,t+h}$ or more concisely:

$$\begin{aligned} Y_{2,t:T} &= \alpha_2 \\ C' Y_{1,t:T} &= \alpha_1 \end{aligned}$$

This information can also be written, with obvious notations:

$$a(\varepsilon_{t:T}, \underline{Y_{t-1}}) = \alpha$$

and the new feature is the presence of future innovations in function $a(.)$. The NIRF is given by:

$$E(Y_{t+h}|Y_{2,t:T}, C'Y_{1,t:T}, Y_{\underline{t-1}}) - E(Y_{t+h}|Y_{\underline{t-1}})\,. \tag{19}$$

Using the notation $\widetilde{Z}_t = (Y'_{1,t}, Y'_{1,t-1}, \ldots, Y'_{1,t-p+1})'$ and $X_t = (Y'_{2,t}, Y'_{2,t-1}, \ldots, Y'_{2,t-p+1})'$, the VAR($p$) process (1) defining the dynamics of Y_t can be written in the following block-recursive form:

$$\left\{\begin{array}{lcl} Y_{1,t} & = & \nu_1 + \widetilde{A}_{11}\, \widetilde{Z}_{t-1} + \widetilde{A}_{12}\, X_{t-1} + \varepsilon_{1,t} \\ \\ Y_{2,t} & = & \widetilde{\nu}_2 + A_{21} \begin{pmatrix} Y_{1,t} \\ \widetilde{Z}_{t-1} \end{pmatrix} + A_{22}\, X_{t-1} + \widetilde{\varepsilon}_{2,t}\,, \end{array}\right. \tag{20}$$

where $\varepsilon_{1,t}$ and $\widetilde{\varepsilon}_{2,t}$ are independent. Introducing the notation $Z_t = \begin{pmatrix} Y_{1,t} \\ \widetilde{Z}_{t-1} \end{pmatrix}$, the previous system can be written as:

$$\left\{\begin{array}{lcl} Z_t & = & \nu_1^* + A_{11}\, Z_{t-1} + A_{12}\, X_{t-1} + \varepsilon^*_{1,t} \\ Y_{2,t} & = & \widetilde{\nu}_2 + A_{21}\, Z_t + A_{22}\, X_{t-1} + \widetilde{\varepsilon}_{2,t}\,, \end{array}\right. \tag{21}$$

with $\nu_1^* = \begin{bmatrix} \nu_1 \\ 0 \end{bmatrix}$, $A_{11} = \begin{bmatrix} \widetilde{A}_{11} & 0 \\ I & 0 \end{bmatrix}$, $A_{12} = \begin{bmatrix} \widetilde{A}_{12} \\ 0 \end{bmatrix}$, $\varepsilon^*_{1t} = \begin{bmatrix} \varepsilon_{1t} \\ 0 \end{bmatrix}$.

Starting the previous system at t, the computation of $E(Y_{1,t+h}|Y_{2,t:T}, C'Y_{1,t:T}, \underline{Y_{t-1}})$ for $h = 0, ..., T-t$ can be viewed as the computation of the smoothed values of $Y_{1,t}, Y_{1,t+1}, ..., Y_{1,T}$ or equivalently $(Z_t, Z_{t+1}, ..., Z_T)$ in the linear Gaussian state-space system:

$$\begin{cases} Z_\tau & = & \nu_1^* + A_{11}Z_{\tau-1} + A_{12}X_{\tau-1} + \epsilon_{1,\tau}^* \\ Y_{2,\tau} & = & \widetilde{\nu}_2 + A_{21}Z_\tau + A_22X_{\tau-1} + \widetilde{\epsilon}_{2,\tau} \\ \alpha_{1,\tau-t} & = & \widetilde{C}_\tau' Z_\tau \ \tau \geq t \end{cases} \tag{22}$$

where $\widetilde{C}_\tau' = [C_\tau', 0_{(1\times n_1 p)}]$. The latent variable of this system is Z_τ. $(Y_{2,t}', \alpha_{1,\tau-t}')$ is the observed variable and $X_{\tau-1}$ is a function of the past values. Note that the dimension of the observed variable, and, therefore, the number of measurement equations may depend on time. The initial condition of this state-space model is just the degenerate distribution at the observed value of $Z_{t-\tau}$.

Thus, the Kalman filter and smoother provide recursive methods for the computation of $E(Y_{1,t+h} \mid Y_{2,t:T}, C'Y_{1,t:T}, \underline{Y_{t-1}})$, $h = 0, \ldots, T-t$, and we are able to compute the first set of components of the NIRF (19), for $h = 0, \ldots, T-t$:

$$E(Y_{1,t+h} \,|\, Y_{2,t:T} = y_{2,t:T}, C'Y_{1,t:T}, \underline{Y_{t-1}} = \underline{y_{t-1}}) - E(Y_{1,t+h} \,|\, \underline{Y_{t-1}} = \underline{y_{t-1}}) \,, \tag{23}$$

which measures the average differential impact on the future values of $Y_{1,t}$ of a new information on the present and future values of $Y_{2,t}$ and some future values of $Y_{1,t}$ compared to the average future values of $Y_{1,t}$ when we do not take into account this new information.

The second set of components of the new information response function (19), for $h = 0, \ldots, T-t$, is simply:

$$y_{2,t+h} - E(Y_{2,t+h} \,|\, \underline{Y_{t-1}} = \underline{y_{t-1}}) \,, \tag{24}$$

and gives the differences between the announced futures values of $Y_{2,t}$ and their expected values at $t-1$.

It is worth noting that the NIRF can also be computed for $h > T-t$. Indeed, we have to compute, for $j \geq 1$:

$$E(Y_{T+j} \,|\, Y_{2,t:T}, C'Y_{1,t:T}, \underline{Y_{t-1}}) - E(Y_{T+j} \,|\, \underline{Y_{t-1}}) \,. \tag{25}$$

The second term of this difference is just the standard prediction of Y_{T+j} at $t-1$, whereas the first term can be computed recursively using the VAR equations and the initial values $E(Y_T\,|Y_{2,t:T}, C'Y_{1,t:T}, \underline{Y_{t-1}}), \ldots, E(Y_{T-p+1}\,|\; Y_{2,t:T}, C'Y_{1,t:T}, \underline{Y_{t-1}})$ previously computed. In particular, by changing the value of T in (19) we can measure the impact on the NIRF of the horizon T up to which the values of $Y_{2,t}$ are guaranteed.

8 Applications to Monetary Policy

In this section we propose empirical illustrations of some results discussed in previous sections. Based on a parsimonious Gaussian VAR(p) model estimated with U.S. quarterly data, we address two monetary policy issues. First, in order to illustrate the impact of a new information on a filter, we investigate whether the effects on the short rate (i.e. the reaction of the central bank) of a new information on the one-year ahead expected inflation are stronger over the three last decades, confirming the shift at the end of the 1970 to a more anti-inflationary monetary policy [emphasized, for instance, in Brissimis and Magginas (2006), Castelnuovo and Surico (2010)]. Second, in order to illustrate the impact of a new information on "future path", we investigate how the variables of our model (and filters of theses variables) respond to the information of a stabilization of future short term interest rates around the zero lower bound, as it has been announced by the FOMC at the end of 2008.

8.1 Description of the Data

Our data set contains 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 zero-coupon yield 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 1964:Q1 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[4]. 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). Summary statistics about the short rate and the spread $S_t = R_t - r_t$ (expressed on a quarterly basis), the GDP growth and the inflation rate are presented in Table 1.

8.2 Model and Decompositions

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

$$Y_t = \nu + \sum_{j=1}^{p} \Phi_j Y_{t-j} + \varepsilon_t \,, \tag{26}$$

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 (and starting from a maximum lag of $p=4$), the lag length is selected to be $p=3$ (see Table 2), and the OLS estimation of the model is presented in Table 3.

It is well known that any H-year nominal yield $R_t(H)$ can be decomposed into the following two terms:

$$R_t(H) = EX_t(H) + TP_t(H)\,, \tag{27}$$

where

$$EX_t(H) = \frac{1}{H} E\left(\sum_{h=0}^{H-1} r_{t+h} | \Omega_t\right) \tag{28}$$

is the expectation part of $R_t(H)$, $TP_t(H) = R_t(H) - EX_t(H)$ is, by definition, the corresponding

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

term premium and Ω_t is the available information set at date t. In what follows we will consider two main cases for Ω_t, depending whether it includes or not an information about the future path of the short term interest rate (see sections 8.3 and 8.4). In addition, $EX_t(H)$ can be decomposed into two components:

$$EX_t(H) = \widetilde{EX}_t(H) + \Pi_t^e(H) \tag{29}$$

where:

$$\widetilde{EX}_t(H) = \frac{1}{H} E\left(\sum_{h=1}^{H} \widetilde{r}_{t+h} \mid \Omega_t\right) \tag{30}$$

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

$$\Pi_t^e(H) = \frac{1}{H} E\left(\sum_{h=1}^{H} \pi_{t+h} \mid \Omega_t\right) \tag{31}$$

is the inflation expectation over $(t, t+H)$. It is possible to show that these three components can be written as linear filter of $Y_t = (r_t, S_t, g_t, \pi_t)'$ and thus the NIRF approach can be adopted (see Appendix 2). In particular, in the following two sections, we will consider the case $H = 4$ quarters and we will apply the results presented in Section 6 and 7, respectively[5].

8.3 Responses to a new information on the 1-year ahead expected inflation

Transmission delay of monetary policy to the economy and, more particularly, to the inflation rate can lead central banks to adopt a pre-emptive strategy, responding to the forecasted value of inflation instead of its actual or past value. For that reason, some authors have included in VAR models variables supposed to reflect central banks expectations (Brissimis and Magginas (2006), Castelnuovo and Surico (2010)). In the following, we show that the NIRF methodology is well

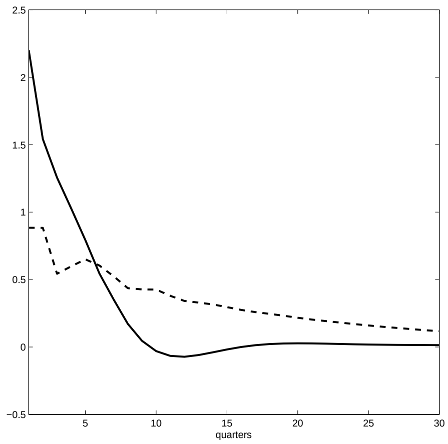
[5]Considering another maturity for H is possible, but requires the estimation of an affine term structure model.

adapted to investigate forward-looking monetary policy strategies conducted by central banks. More precisely, by means of the parsimonious VAR model presented above, we show how to analyze the reaction of central banks to an increase (decrease) in the inflation expectation, given that the latter can be expressed as a linear filter of the variables in the VAR (see Appendix 2) and, thus, the technique of Section 6 can be applied.

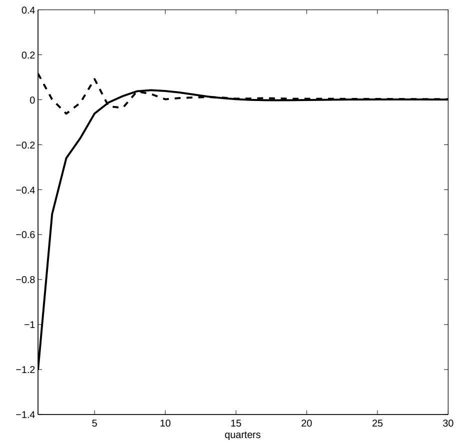
In what follows we focus on responses of the short term interest rate, interpreted as the reaction of the monetary policy, to an increase in the expectation of the one-year ahead inflation rate. This information includes the following elements. First, in order to isolate specific effects of this information, we assume that the instantaneous effect of the rise in the expectation of the 1-year ahead inflation is one-for-one on the 1-year interest rate. In other words, we assume that the instantaneous response of $TP_t + \widetilde{EX}_t$ is zero. Second, we assume, as it is usual in empirical literature, that the response of real GDP growth occurs with a one-quarter lag. In other words, the instantaneous response of GDP growth is zero. In figure 1 we report responses of the short term interest rate, of the one-year spread and of the expectation of one-year inflation before 1979:Q3 (dashed lines) and after 1979:Q4 (solid lines), when the increase in the 1-year ahead expected inflation is one.

In both sub-samples, responses of the short term interest rate are positive, which is in accordance with the conventional view of a monetary policy rule in which the central bank adjusts the policy rate in response to (expected) inflation. However, magnitude of this adjustment depends on the sample. Before 1979:Q3, the rise in the short rates is less than proportional to the increase in the expected inflation, leading to an increase in the spread (as the response of the 1-year interest rate is constructed to be one-for-one). In contrast, response of short term interest rate in the post-1979 period is twice the rise in expected inflation. The instantaneous response of the 1-year spread is negative accordingly. In addition we notice that the impact on the expected inflation reverts faster to zero in the post-1979 sample than in the pre-1979 period.

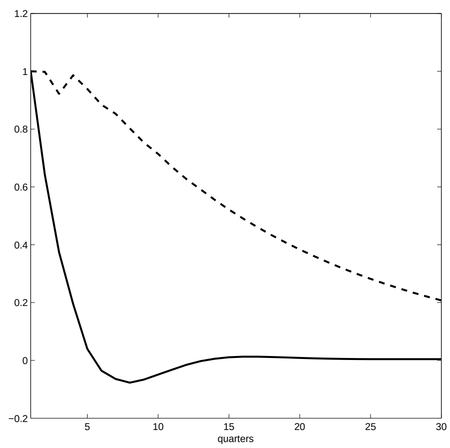
Evidences of a shift in the conduct of the U.S. monetary policy at the end of the 1970s have



(a) Short rate



(b) 1-year spread



(c) 1-year ahead expected inflation

Figure 1: Responses to a shock on the 1-year ahead expected inflation, before 1979:Q3 (dashed lines) and after 1979:Q4 (solid lines)

been emphasized in the literature [see Judd and Rudebusch (1998), Clarida, Gali and Gertler (2000), Boivin and Giannoni (2006), Lubik and Schorfheide (2004) among others]. This shift is associated with a significant move to an active anti-inflationary monetary policy stance after 1979. Our results also confirm these facts.

8.4 Responses to unconventional monetary policy: effect of forward policy guidance

Central banks are sometimes confronted to the key issue of how restoring good economic and financial conditions when the short term interest rate is near the zero lower bound. For instance, the U.S. Federal Reserve has been recently concerned with this problem. 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 continue to manage longer-term interest rates to a level consistent with a given objective of price stability and economic growth. Communication regarding the future path of short term interest rate is the key ingredient to achieve this goal[6].

Forward guidance on monetary policy has been recently 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 example is the August 2011 FOMC statement: *"The commitee 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"*.

What are the (expected) responses of economic variables to such information? Does this communication leads to reduce the medium and long term rates? Are these responses different from those obtained without taking into account this information about the future path of the short term rate?

Tools developed in section 7 allow us to address these questions. More precisely, we can estimate the expected response of $Y_t = (r_t, S_t, g_t, \pi_t)'$ and linear filters of Y_t to this new information about

[6] There are several examples of central banks using communication on the future path of short term interest rate, for instance New Zealand, Norway and Sweden by means of policy rate projections, or Canada and Japan by means of communication regarding the timing and conditions for rate moves.

the future path of short term interest rate[7].

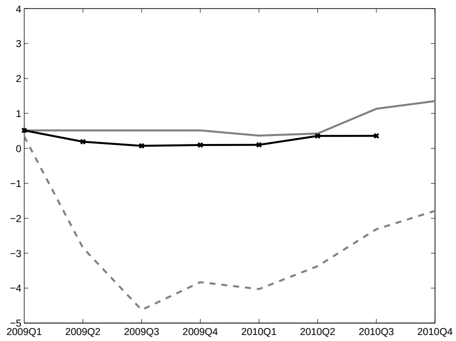
In what follows, we assume that the path of future interest rate is supposed to remain unchanged for a known given period. We report in figures 2 and 3 expected responses in 2009:Q1 (and for the 8 following quarters) of r_t, g_t,π_t, the 1-year interest rate and its components, namely Π_t^e, EX_t and TP_t, when agents take into account the fact that short rate will remain constant over the 4 following quarters[8] (for the case of 8 quarters, see Appendix 3). In order to deal with the monetary policy shift observed at the end of 1970s, and to perform real-time exercise, the VAR is estimated over the period 1979:Q4 to 2008:Q4. We also report in dashed lines the expected response obtained without taking into account future path of short term interest rates. Line with markers represent realized ex-post values of corresponding variables.

It is worth noting that taking into account this new information significantly improves forecasts of variables. Predictions of short term interest rate obtained from the VAR go to negative values (see figure 2(a)). Similarly, expectations of the 1-year interest rate are significantly improved (figure 2(b)) and remain positive. This is notably due to a better prediction of the mean of future short rates (see 2 (c)). Finally, we can note that forecasts of annual inflation is also improved (3(b)). Similar results are obtained in the case of a future path of 8 quarters (see figures 4 and 5 in Appendix 3).

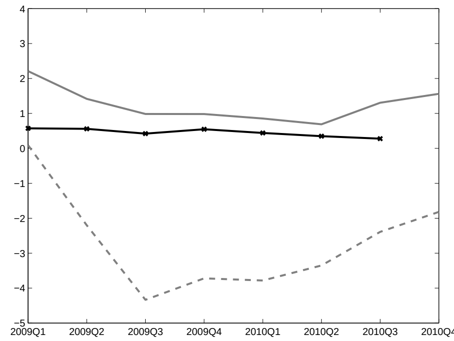
All in all, this empirical illustration stresses that considering information on future path of variables is a key element not only for forecasting purposes, but also for a precise anticipation of the future effects of a monetary policy intervention (i.e., a short rate path or scenario). New information responses function methodology provide a promising framework for that purpose.

[7]It has to be noted that the Federal Reserve does not always communicate about the time slot during which short rates will remain constant. In our application however this time is supposed to be known and, therefore, superior and inferior limits of expected response could be obtained considering a range of possible horizon. This is beyond the scope of the present illustrative exercise.

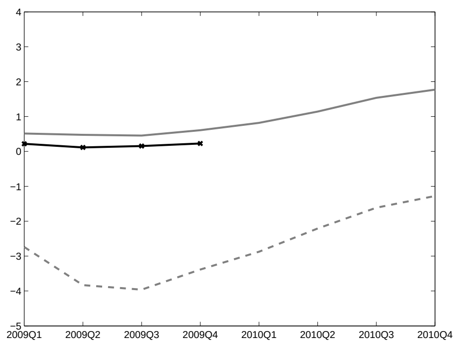
[8]This scenario is also suggested by market-based expectations of the future U.S. policy rate based on overnight index swap.



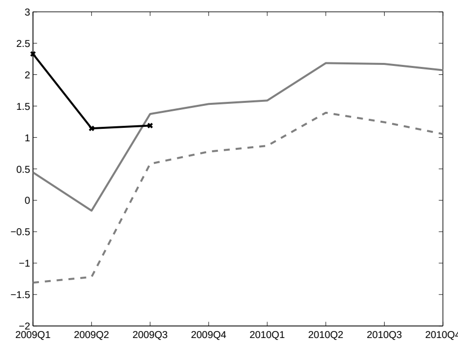
(a) Short term interest rate



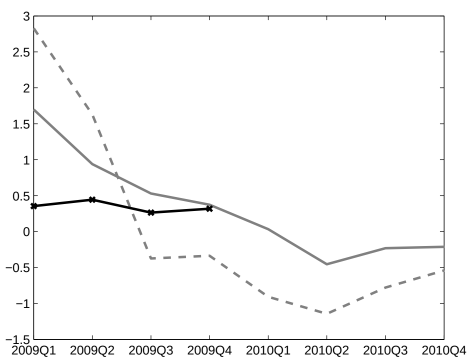
(b) 1-year interest rate



(c) 1-year ahead nominal expectation term

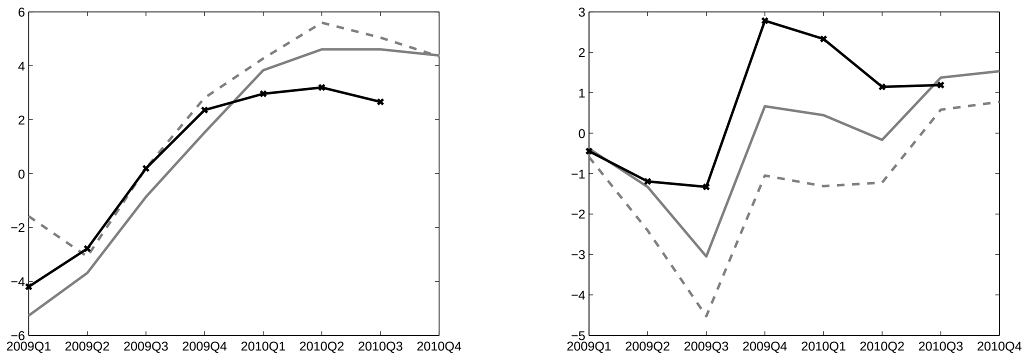


(d) 1-year ahead expected inflation



(e) 1-year nominal term premium

Figure 2: Responses to a new information in 2008:Q4 of a constant short term interest rate for the next 4 quarters. Future path-dependent NIRF-based expected responses (grey solid line), ex-post realization (black solid line with markers), VAR implied forecasts (grey dashed lines). The ex-post realizations at date t reported in figures (c), (d) and (e) are respectively $\frac{1}{4}\sum_{h=0}^{3} r_{t+h}$, $\frac{1}{4}\sum_{h=1}^{4} \pi_{t+h}$ and $R_t - \frac{1}{4}\sum_{h=0}^{3} r_{t+h}$ where R_t is the 1-year interest rate.



(a) Annual real output growth (b) Annual inflation rate

Figure 3: Responses to a new information in 2008:Q4 of a constant short term interest rate for the next 4 quarters. Future path-dependent NIRF-based expected responses (grey solid line), ex-post realization (black solid line with markers), VAR implied forecasts (grey dashed lines).

9 Conclusions and Further Developments

In this paper we propose a new 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 also show that this methodology is well suited to analyse the effects of a new information on the sign or on the average response of some variables, as well as on linear filters of the basic variables of our model. It is also well adapted to study the effects of an information on (possibly several) interval values associated to some innovation or some response at a given horizon, or on the past or future paths of these basic variables.

In the last section of the paper we provide two empirical illustrations of the NIRF methodology based on U.S. data. First, we focus on the impulse responses of the short-term term interest rate to a new information on the one-year ahead expectation of inflation. We show that the U.S. federal Reserve adjusts more significantly the short rate in the post-1979 period, confirming the shift to a more aggressive anti-inflationary policy at the early 1980s. Second, in order to illustrate the impact of a new information on a "future path", we investigate how variables of our VAR respond to the information released by the FOMC on December 2008, about the stabilization of future short term interest rate around the zero lower bound. We show that taking into account this information is critical and improves significantly forecasts of the macroeconomic and financial factors.

The results of this paper has 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 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 be also extended to VARMA(p,q). 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 also an interesting line of future research.

Appendix 1 : Tables.

Yields	r_t	S_t	g_t	π_t
Mean	0.0143	0.0005	0.0074	0.0104
Std. Dev.	0.0076	0.0014	0.0085	0.0083
Skewness	0.7738	-0.0143	1.5200	0.2431
Kurtosis	4.2884	7.4666	4.4473	7.4491
Minimum	0.0001	-0.0055	-0.0207	-0.0329
Maximum	0.0398	0.0068	0.0385	0.0407

Table 1: Summary Statistics on U.S. 1-quarter short rate (r_t), 4-quarters spread (S_t), 1-quarter GDP growth rate (g_t) and inflation rate (π_t) observed quarterly from 1964:Q1 to 2010:Q3 [Gurkaynak, Sack and Wright (2007) data base for the 1-quarter and 4-quarters yields; FRED data base for GDP growth rate (GDPC1) and inflation rate (CPIAUCSL)].

Lag p	LR	FPE	AIC	SIC	HQ
0	N.A.	3.64e-19	-31.10	-31.03	-31.07
1	528.51	2.23e-20	-33.89	-33.54*	-33.75
2	63.41	1.84e-20	-34.08	-33.45	-33.83*
3	33.57*	1.80e-20*	-34.11*	-33.19	-33.74
4	16.64	1.95e-20	-34.03	-32.84	-33.55

Table 2: Criteria for VAR order selection. Given a sample period of size T, and a n-dimensional Gaussian VAR(p) process with empirical white noise covariance matrix $\hat{\Omega}(p)$, $LR = (T-m)[log|\hat{\Omega}(p-1)| - log|\hat{\Omega}(p)|]$ denotes, for each lag p, the sequential modified [Sims (1980)] likelihood ratio (LR) test statistic, where m is the number of parameters per equation under the alternative. The modified LR statistics are compared to the 5% critical values. $FPE = [(T+np+1)/(T-np-1)]^n \det(\hat{\Omega}(p))$ denotes, for each lag p, the final prediction error criterion. If we denote by log-L $= -(Tn/2)\log(2\pi) + (T/2)\log(|\hat{\Omega}(p)^{-1}|) - (Tn/2)$ the maximum value of the log-likelihood function associated to the VAR(p) model, $AIC = -2\text{log-L}/T + 2pn^2/T$, $SIC = -2\text{log-L}/T + (\log(T)/T)pn^2$ and $HQ = -2\text{log-L}/T + (2\log(\log(T))/T)pn^2$ denote, respectively and for each lag p, the Akaike, Schwarz and Hannan-Quinn information criteria. For each criterion, and starting from a maximum lag of $p = 4$, (*) denotes the optimal number of lags.

	ν	Φ_1				Φ_2				Φ_3			
r_t	-0.0013	0.6919	0.2615	0.1031	-0.0319	0.0389	-0.1286	0.0215	0.1148	0.1447	-0.0598	0.0567	0.0731
	[-2.3975]	[7.5388]	[1.4478]	[4.0668]	[-0.9528]	[0.3555]	[-0.6882]	[0.7909]	[3.4329]	[1.5652]	[-0.3441]	[2.2213]	[2.1452]
S_t	0.0003	0.0040	0.3597	0.0093	0.0073	0.1276	0.1769	0.0092	-0.0395	-0.0979	0.0139	-0.0302	-0.0092
	[1.1686]	[0.0898]	[4.0648]	[0.7522]	[0.4450]	[2.3785]	[1.8766]	[0.6927]	[-2.4133]	[-2.1610]	[0.1636]	[-2.4150]	[-0.5555]
g_t	0.0047	-0.6772	0.2773	0.2388	-0.0702	0.3526	0.2542	0.2156	0.0178	0.3646	0.5419	-0.0176	-0.1255
	[2.9315]	[-2.4584]	[0.5116]	[3.1391]	[-0.6992]	[1.0726]	[0.4401]	[2.6383]	[0.1782]	[1.3133]	[1.0390]	[-0.2294]	[-1.2263]
π_t	0.0010	0.5982	-0.0505	0.1067	0.1962	-0.4890	-0.7074	0.0416	0.2460	-0.0170	0.2202	-0.0375	0.2872
	[0.8403]	[2.7956]	[-0.1200]	[1.8058]	[2.5137]	[-1.9145]	[-1.5763]	[0.6563]	[3.1554]	[-0.0789]	[0.5436]	[-0.6295]	[3.6106]

$\Omega \times 10^6$				Corr.				log-L	$\lvert \psi \rvert$
6.52	-1.63	1.42	4.91	ρ_{12}	ρ_{13}	ρ_{14}		3189.64	0.9653
[9.2466]	[-5.9387]	[0.9450]	[4.0162]	-0.5097	0.0724	0.3227			0.8494
.	1.57	-1.30	-1.05		ρ_{23}	ρ_{24}			0.6493$(^c)$
	[9.2466]	[-1.7577]	[-1.8260]		-0.1356	-0.1410			0.5191$(^c)$
.	.	58.7	2.17			ρ_{34}			0.4875
		[9.2466]	[0.6211]			0.0475			0.4847$(^c)$
.	.	.	35.5						0.3877
			[9.2466]						0.0794$(^c)$

Table 3: Parameter estimates of the state dynamics $X_t = \nu + \sum_{j=1}^{3} \Phi_j X_{t-j} + \varepsilon_t$, with $X_t = (r_t, S_t, g_t, \pi_t)'$ [Gurkaynak, Sack and Wright (2007) data base for the 1-quarter and 4-quarters yields; FRED data base for GDP growth rate (GDPC1) and inflation rate (CPIAUCSL); sample period : 1964:Q1 - 2010:Q3]. t-values are in brackets. ρ_{ij} denotes the (empirical) correlation between (ε_{it}) and (ε_{jt}). log-L denotes the maximum value of the log-Likelihood function. $\lvert \psi \rvert$ indicates the modulus of the roots of equation $|\widetilde{\Phi}(\psi)| = 0$, with $\widetilde{\Phi}(\psi) = (I_4 \psi^3 - \Phi_1 \psi^2 - \Phi_2 \psi - \Phi_3)$ denoting the characteristic polynomial; $(^c)$ indicates a pair of complex conjugate roots.

Appendix 2. 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:

$$Y_t = \nu + \sum_{j=1}^{p} \Phi_j Y_{t-j} + \varepsilon_t \tag{32}$$

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

$$Z_t = \widetilde{\nu} + \Phi\, Z_{t-1} + \widetilde{\varepsilon}_t \tag{33}$$

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

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

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 date t consists in the past and present values of Z_t, that is $\Omega_t = \underline{Z_t}$. In this case:

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

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

$$\begin{aligned} EX_t(H) &= d_0(H) + c_0(H)Z_t\,, \;\; \Pi^e_t(H) = d_1(H) + c_1(H)Z_t \\ TP_t(H) &= d_2(H) + c_2(H)Z_t\,, \;\; \widetilde{EX}_t(H) = d_3(H) + c_3(H)Z_t \end{aligned}$$

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 components of $R_t(H)$ can be expressed as linear filter of the variables in the VAR and, thus, the technique of Section 6 can be applied (see Section 8.3).

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

$$E(EX_{t+k}(H)|\Omega_t) - E(EX_{t+k}(H)|\underline{Y_{t-1}})\,. \tag{34}$$

The first component of this expression is:

$$
\begin{aligned}
E(EX_{t+k}(H)|\Omega_t) &= \frac{1}{H}E\left(\sum_{h=0}^{H-1}E(r_{t+k+h}|\Omega_{t+k})|\Omega_t\right) &(35)\\
&= \frac{1}{H}\sum_{h=0}^{H-1}E\left(r_{t+k+h}|\Omega_t\right) &(36)
\end{aligned}
$$

where $E\left(r_{t+k+h}|\Omega_t\right) = \overline{r}_{t+k+h}$ for $h+k \leq \overline{H}-1$.

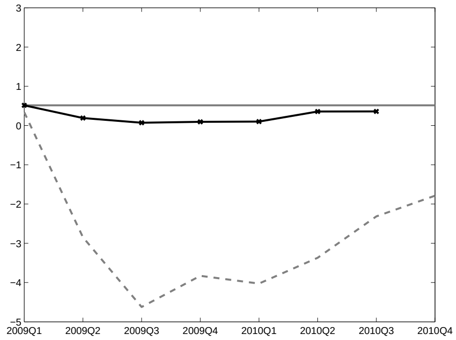
Similarly, the NIRF of $\Pi_t^e(H)$ at date t is given by:

$$
E(\Pi_{t+k}^e(H)|\Omega_t) - E(\Pi_{t+k}^e(H)|Y_{\underline{t-1}}) \,. \tag{37}
$$

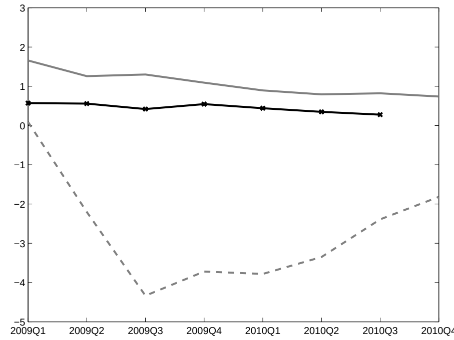
The first component of this expression is:

$$
\begin{aligned}
E(\Pi_{t+k}^e(H)|\Omega_t) &= \frac{1}{H}E\left(\sum_{h=1}^{H}E(\pi_{t+k+h}|\Omega_{t+k})|\Omega_t\right) &(38)\\
&= \frac{1}{H}\sum_{h=1}^{H}E\left(\pi_{t+k+h}|\Omega_t\right). &(39)
\end{aligned}
$$

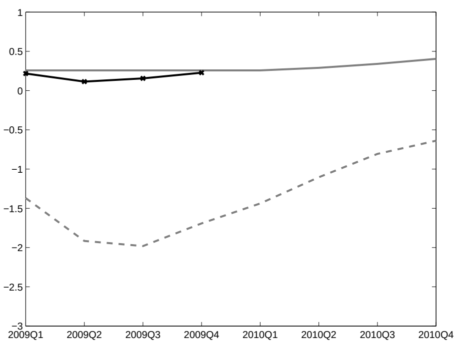
Appendix 3: The case of a short rate future path of 8-quarters.



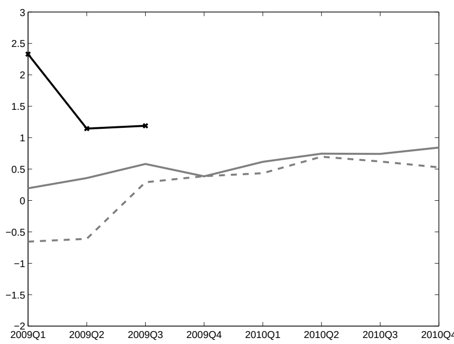
(a) Short term interest rate



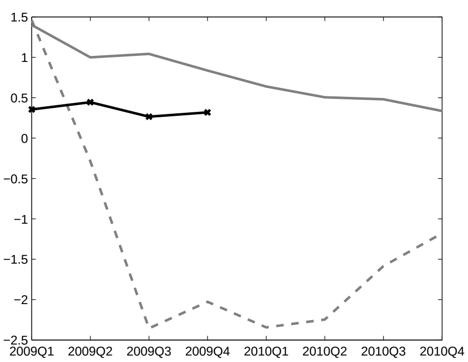
(b) 1-year interest rate



(c) 1-year ahead nominal expectation term

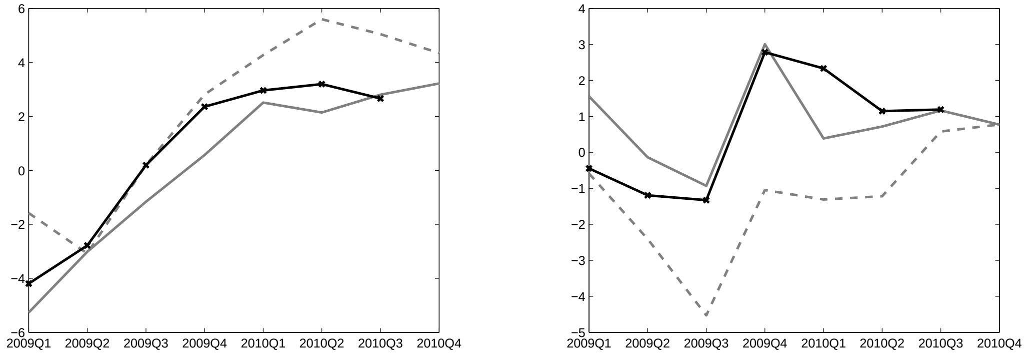


(d) 1-year ahead expected inflation



(e) 1-year nominal term premium

Figure 4: Responses to a new information in 2008:Q4 of a constant short term interest rate for the next 8 quarters. Future path-dependent NIRF-based expected responses (grey solid line), ex-post realization (black solid line with markers), VAR implied forecasts (grey dashed lines).



(a) Annual real output growth

(b) Annual inflation rate

Figure 5: Responses to a new information in 2008:Q4 of a constant short term interest rate for the next 8 quarters. Future path-dependent NIRF-based expected responses (grey solid line), ex-post realization (black solid line with markers), VAR implied forecasts (grey dashed lines).

References

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

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.

Jardet C., Monfort A., and F. Pegoraro, 2009, "No-arbitrage near-cointegrated VAR(p) term structure models, term premia and GDP growth", working paper Banque de France.

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.

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.

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

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

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.

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.

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.