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 : April, 2012.

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 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. It can also be related 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 two monetary policy issues. The first one evaluates the shift (at the end of the 1970s) to a more anti-inflationary monetary policy, while the second one compares 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, monetary policy issues, 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:

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 stochastic process (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

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 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.

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, 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). More precisely, this "new information" can be related 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 innovation of a linear filter 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 to 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 start with a new information that only concerns the contemporaneous value ε_t of this innovation process, i.e. of the form $a(\varepsilon_t) = \alpha$, where $a(.)$ is some given function and α some given vector of real numbers. This kind of new information contains not only the classical "full new information" case where $a(.)$ is one-to-one (and, thus, for any given α we have a unique value for the innovation)[4], but also other relevant cases. First, we consider the case of "continuous partial 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 only, but it also allows to take into account many other kinds of informations, including informations on a subset of innovations, informations on responses or a combination of both. Second, we study the "discrete partial new information" case

[4]For sake of completeness we show in Appendix 1 of the paper that the standard orthogonalized shocks, the impulse vectors introduced by Uhlig (2005) and the structural shocks (that are fully identified, either by short-run or by long-run restrictions) can be viewed as particular cases of such full information.

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. Third, 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 exploit the flexibility of the NIRF methodology to consider important generalizations. First, we study the case where the new 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 involves future values of the innovation process, that is future values of the process itself or of some associated linear filters, 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 usefulness of the NIRF methodology, we address two monetary policy issues using a Gaussian VAR(p) model that links macroeconomic variables and interest rates (namely, short rate, one-year spread, GDP growth and inflation rate). 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 analyze the effects of a new information on the one-year ahead expected inflation, which can be expressed as a linear filter on the VAR variables, before and after 1979:Q4. On the other hand, 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.

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 continuous partial new information case. Section 4 is devoted to the discrete partial new information case, while in Section 5 we show how these results can be used to analyze the responses to new informations on a linear filter and responses of a filter to new informations. Section 6 deals with past and future path-dependent new information response functions, while in Section 7 we present the two empirical applications. Finally, Section 8 concludes and proposes further developments. Proofs of propositions and some theoretical results are gathered in appendices.

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 $(y'_{-1}, y'_{-2}, ..., y'_{-p})' \equiv y_{\underline{-p}}$.

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:

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

(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 $Y_{\underline{t}} = (Y_t', Y_{t-1}', ..., Y_{t-p}')'$, 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 formalized in the following:

Proposition 1 (New Information Response Function): *Given a new information about* ε_t, *namely* $a(\varepsilon_t) = \alpha$, *with* $a(.)$ *some function and* α *a given constant, the New Information Response Function (NIRF) is the differential impact on the prediction of* Y_{t+h} *of the new information* $a(\varepsilon_t) = \alpha$:

$$
\begin{aligned}
NIRF_{t,t+h}(\varepsilon_t) &= E(Y_{t+h}|a(\varepsilon_t), \underline{Y_{t-1}}) - E(Y_{t+h}|\underline{Y_{t-1}}), \\
&= \Theta_h \, E[\varepsilon_t | a(\varepsilon_t)] \, .
\end{aligned}
\tag{7}
$$

[Proof: see Appendix 1].

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 $\varepsilon_s = 0$, $s > t$.

The standard case, when $a(.)$ is one-to-one, is the "full new information" case. For sake of completeness we show in Appendix 2 that it contains, as particular cases, the orthogonalized shocks, the Uhlig (2005) ' s impulse vectors and the structural shocks. Then, in Sections from 3 to 5 we distinguish three less standard cases:

i) the "continuous partial 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 a Lebesgue measure).

ii) the "discrete partial new information" case, when the distribution of $a(\varepsilon_t)$ has discrete components.

iii) 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 $\widetilde{a}[F(0)\, \varepsilon_t] = \alpha$.

In Section 6 we consider two generalizations of the previous setting where $a(.)$ is allowed to be not only function of ε_t, but also function of the observed past values of Y_t or function of the future values of the innovation. More precisely, we focus on the path-dependent new information setting in which the function $a(.)$ is specified as:

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

v) or as $a(\underline{\varepsilon_{t:T}}, \underline{Y_{t-1}}) = \alpha$, with $\underline{\varepsilon_{t:T}} = (\varepsilon_t', \varepsilon_{t+1}', \ldots, \varepsilon_T')'$.

The first case is named "past path-dependent" NIRF, while the second one is called "future path-dependent" NIRF.

3 Continuous partial 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.

3.1 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 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)].

3.2 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 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}.

3.3 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 3.2, 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).

3.4 New information on long-run behavior

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), and $\Lambda' Y_t$ a r-dimensional vector of cointegrating relationships, and such that W_t has a stationary VAR representation of the form:

$$\Gamma(L) W_t = C\nu + C\varepsilon_t \tag{8}$$

where $C = \begin{pmatrix} I_{n-r} & 0 \\ & \Lambda' \end{pmatrix}$ is invertible. We can consider 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 $\Gamma^{-1}(1)$.

3.5 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 it is shown in Section B) of Appendix 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 4.3.

4 Discrete partial new information

4.1 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 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_t given $a_1(\varepsilon_t) = \alpha_1$ restricted to the set $a_2(\varepsilon_t) = \alpha_{2j}$. In other words, for any set S:

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

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 $E[g(\varepsilon_t)|a_1(\varepsilon_t) = \alpha_1$ 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.

4.2 Quantitative information and 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.

Proposition 2: *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 $K \cap J = \varnothing$, is considered in Appendix 1 as well as the case where the interval informations are related to responses.

4.3 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: $\Theta_h^j \varepsilon_t > 0$, where Θ_h^j is the j^{th} row of Θ_h and $(j, h) \in S$. The conditional expectation

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

can 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 (defined in Section 3.5) of the conditional distribution of ξ_t given, $\xi_t' \xi_t = 1$, where $\xi_t \sim N(0, I)$ 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 $\Theta_h^j \widetilde{\varepsilon} > 0$, $(j, h) \in S$.

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 $(\widetilde{\Sigma}, \widetilde{\xi})$, 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 $\Theta_h^j \widetilde{\gamma}$, $h \in$

$\{1, \ldots, H\}$, and, finally, keeping the candidate $\widetilde{\gamma}$ if the impulse response function 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 prediction of the innovation among the impulse vectors 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 $[\Theta_h^j \widetilde{\varepsilon}, (j,h) \in S]$ belongs to some set $\mathcal{E}$, for instance a product of intervals.

5 New information on a filter and responses of a filter

5.1 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 7 and Jardet, Monfort, Pegoraro (2012)]. In that case, the NIRF framework still applies, as it is formalized in the following:

Proposition 3 (NIRF on a Linear Filter): *Given the linear 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, and its innovation $\widetilde{\varepsilon}_t = F(0)\varepsilon_t$, and given the new information on $\widetilde{\varepsilon}_t$, defined by $\widetilde{a}(\widetilde{\varepsilon}_t) = \alpha$, with $\widetilde{a}(.)$ some function and α a given constant, the differential impact on the prediction of Y_{t+h} is:*

$$\begin{aligned} &NIRF_{t,t+h}(\varepsilon_t) = \Theta_h \, E[\varepsilon_t \,|\, a(\varepsilon_t) = \alpha] \,, \\ &\textit{with} \;\; a(\varepsilon_t) := \widetilde{a}(F(0)\,\varepsilon_t) \,. \end{aligned} \tag{9}$$

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 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 would be possible to complete these filters 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.

5.2 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.

6 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 (7) 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.

6.1 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:

$$a(\varepsilon_t, \underline{Y_{t-1}}) = \alpha \,. \tag{10}$$

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:

$$\begin{aligned} &\widehat{Y}_{t|t-1}^K(\underline{Y_{t-1}}) + \varepsilon_t^K = Y_{t-1}^K \,, \\ &\text{or} \;\; \varepsilon_t^K - Y_{t-1}^K + \widehat{Y}_{t|t-1}^K(\underline{Y_{t-1}}) = 0 \,, \end{aligned} \tag{11}$$

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

Proposition 4 (Past Path-Dependent NIRF): *Given a new information $a(\varepsilon_t, \underline{Y_{t-1}}) = \alpha$, with $a(.)$ some function and α a given constant, the differential impact on the prediction of Y_{t+h} is:*

$$\begin{aligned} NIRF_{t,t+h}(\varepsilon_t, \underline{Y_{t-1}}) &= E(Y_{t+h} | a(\varepsilon_t, \underline{y_{t-1}}) = \alpha, \underline{Y_{t-1}} = \underline{y_{t-1}}) - E(Y_{t+h} | \underline{Y_{t-1}} = \underline{y_{t-1}}) \\ &= \Theta_h \, \delta_t \,, \;\; \textit{with} \;\; \delta_t = E[\varepsilon_t \,|\, a(\varepsilon_t, \underline{y_{t-1}}) = \alpha] \end{aligned} \tag{12}$$

[Proof: see Appendix 1].

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 (11) we need to compute:

$$\delta_t = E[\varepsilon_t \,|\, \varepsilon_t^K = y_{t-1}^K - \widehat{y}_{t|t-1}^K(\underline{y_{t-1}})] \tag{13}$$

and we get:

$$\delta_t = \Sigma^K \, \Sigma_{K\,K}^{-1} \, [y_{t-1}^K - \widehat{y}_{t|t-1}^K(\underline{y_{t-1}})] \,, \tag{14}$$

where Σ^K and $\Sigma_{K\,K}$ are defined like in Section 3.2.

6.2 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 linear filters of them 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 this section we show how to deal with this issue in the NIRF framework and, moreover, how to add an information on the future path of some linear filters of the variables.

Let us consider a new information on the future path of a $\widetilde{n}_t$-dimensional vector $\widetilde{Z}_t$ including some components of Y_t or/and some linear filters of Y_t. Formally, $\widetilde{Z}_t$ can be written as:

$$\widetilde{Z}_t = \gamma_t + M_{0,t}Y_t + M_{1,t}Y_{t-1} + ... + M_{J,t}Y_{t-J} \tag{15}$$

where γ_t and $\{M_{j,t}\}_{j=0,...,J}$ are respectively a $\widetilde{n}_t$-dimensional vector and $(\widetilde{n}_t, n)$ matrices, which are known and which may be time-varying. We assume that the values of $\widetilde{Z}_t$ are imposed at all dates between t (tomorrow) and T. More precisely, using the notation $\widetilde{Z}_{t:T} = (\widetilde{Z}'_t, ..., \widetilde{Z}'_T)'$, we impose the information:

$$\widetilde{Z}_{t:T} = \alpha \tag{16}$$

where $\alpha = (\alpha_t, ..., \alpha_T)'$ is a vector of known values. This information can also be written, with obvious notations:

$$a(\varepsilon_{t:T}, \underline{Y_{t-1}}) = \alpha \tag{17}$$

and the new feature is the presence of future innovations in function $a(.)$. In this case we have:

Proposition 5 (Future Path-Dependent NIRF): *Given a new information* $a(\varepsilon_{t:T}, \underline{Y_{t-1}}) = \alpha$, *with* $a(.)$ *some function and* $\alpha = (\alpha_t, ..., \alpha_T)'$, *given* $Z_t := \left(Y_t', ..., Y_{t-K}'\right)'$, *where* $K = max\{p, T-t, J\}$, *and assuming* $T-t > p$ *and* $T-t > J$ $(K = T-t)$, *the VAR(p) process defining the dynamics of* Y_t *and the restrictions* $\widetilde{Z}_{t+h} = \alpha_{t+h}$, *for* $h \in \{0, ..., T-t\}$, *can be written as the following linear Gaussian state-space system:*

$$\begin{cases} Z_t &= \nu_t + AZ_{t-1} + \varepsilon_t^* \\ \widetilde{Z}_t &= \gamma_t + M_t Z_t \end{cases} \tag{18}$$

with

$$\nu_t = \begin{pmatrix} \nu \\ 0 \end{pmatrix}, \quad \varepsilon_t^* = \begin{pmatrix} \varepsilon_t \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} \Phi_1 & \Phi_2 & \dots & \Phi_p & 0 \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \ddots & 0 & 0 \\ 0 & 0 & \dots & I & 0 \end{pmatrix}, \quad M_t = (M_{0,t}, ..., M_{J,t}) \,. \tag{19}$$

Then, the differential impact on the prediction of Y_{t+h}, *for* $h \in \{0, ..., T-t\}$, *is given by:*

$$NIRF_{t,t+h}(\varepsilon_{t:T}, \underline{Y_{t-1}}) = E\left(Y_{t+h} | \widetilde{Z}_{t:T} = \alpha, \underline{Y_{t-1}}\right) - E\left(Y_{t+h} | \underline{Y_{t-1}}\right), \tag{20}$$

and it is calculated in the following way. Starting the system at t, *the computation of* $E(Y_{t+h} | \widetilde{Z}_{t:T}, \underline{Y_{t-1}})$ *can be viewed as the computation of the* $(h \times n + 1)^{th}, (h \times n + 2)^{th}, ..., (h \times n + n)^{th}$ *components of the filtered values of* Z_T, *while the computation of* $E\left(Y_{t+h} | \underline{Y_{t-1}}\right)$ *is obtained from standard VAR predictions.*

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

$$E(Y_{T+j} \,|\, \widetilde{Z}_{t:T}, \underline{Y_{t-1}}) - E(Y_{T+j} \,|\, \underline{Y_{t-1}}) \,. \tag{21}$$

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 previously computed initial values $E(Y_T \,| \widetilde{Z}_{t:T}, \underline{Y_{t-1}})$, ...,$E(Y_{T-p+1} \,|\, \widetilde{Z}_{t:T}, \underline{Y_{t-1}})$.

7 Applications to Monetary Policy

In this section we propose two illustrations of the empirical relevance of the NIRF methodology. Using a 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. In this exercise we evaluate the shift, at the end of the 1970, to a more anti-inflationary monetary policy emphasized, for instance, by Brissimis and Magginas (2006) and Castelnuovo and Surico (2010). Second, in order to illustrate the impact of a new information on "future paths", 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.

7.1 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 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[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).

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

7.2 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:

$$Y_t = \nu + \sum_{j=1}^{p} \Phi_j Y_{t-j} + \varepsilon_t\,, \tag{22}$$

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 over the sample periods characterizing any of the proposed empirical exercises (the results are available upon requests from the authors).

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

$$R_\tau(H) = EX_\tau(H) + TP_\tau(H)\,, \tag{23}$$

where

$$EX_\tau(H) = \frac{1}{H} E\left(\sum_{h=0}^{H-1} r_{\tau+h} | \Omega_\tau \right) \tag{24}$$

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 what follows we will consider two main cases for Ω_τ, depending whether it includes or not an information about the future path of the short-term interest rate (see sections 7.3 and 7.4). In addition, $EX_\tau(H)$ can be decomposed into two components:

$$EX_\tau(H) = \widetilde{EX}_\tau(H) + \Pi^e_\tau(H) \tag{25}$$

where:

$$\widetilde{EX}_\tau(H) = \frac{1}{H} E\left(\sum_{h=1}^{H} \widetilde{r}_{\tau+h} \,|\, \Omega_\tau \right) \tag{26}$$

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

$$\Pi_\tau^e(H) = \frac{1}{H} E\left(\sum_{h=1}^{H} \pi_{\tau+h} \,|\, \Omega_\tau\right) \tag{27}$$

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 two sections we will consider the case $H = 4$ quarters and we will apply the results presented in Section 5 and 6, respectively[6].

7.3 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 values 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 adapted to investigate forward-looking monetary policy strategies conducted by central banks. More precisely, by means of the 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 3) and, thus, the technique of Section 5 applies.

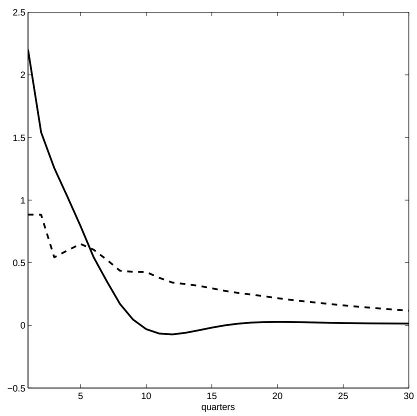
In what follows we focus on responses of the short-term interest rate, interpreted as the reaction of the monetary policy, to an increase in the expectation of the one-year ahead inflation rate. More precisely, the new information includes the following elements. First, in order to isolate its specific

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

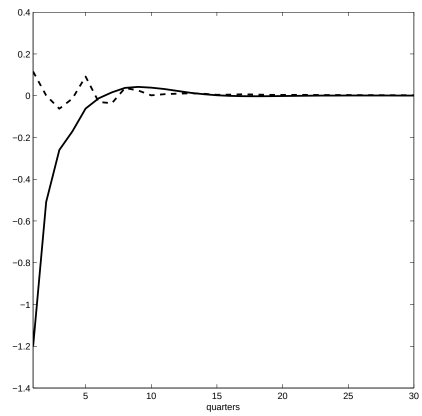
effect, we assume that the instantaneous effect of the rise in the expectation of the 1-year ahead inflation is one-for-one on the 1-year interest rate. In other words, we assume that the instantaneous response of $TP_t + \widetilde{EX}_t$ is zero. Second, we assume, as it is usual in the empirical literature, 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 the responses of the short-term interest rate, of the one-year spread and the one-year ahead expected inflation when the model is estimated from 1964:Q1 to 1979:Q3 (dashed lines) and from 1979:Q4 to 2010:Q3 (solid lines), and when the current increase of the expected inflation is equal to 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, the magnitude of this adjustment depends on the sample. Before 1979:Q3, the instantaneous rise in the short rate 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. Accordingly, the instantaneous response of the 1-year spread is negative. 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.

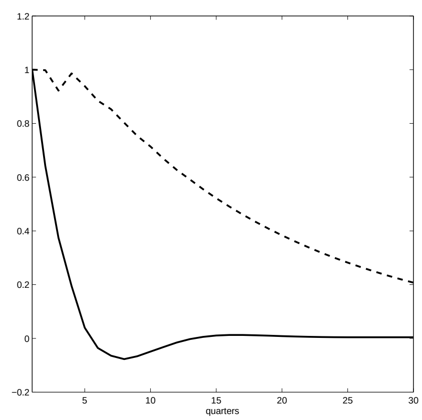
Evidence of a shift in the conduct of the U.S. monetary policy, at the end of the 1970s, have been already emphasized in the literature [see Judd and Rudebusch (1998), Clarida, Gali and Gertler (2000), Boivin and Giannoni (2006), Lubik and Schorfheide (2004) among others] and this shift is associated with a significant move to an active anti-inflationary monetary policy stance after 1979. Our results confirm and quantify these findings.



(a) Short-term interest 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).

7.4 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 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 (24)) 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):

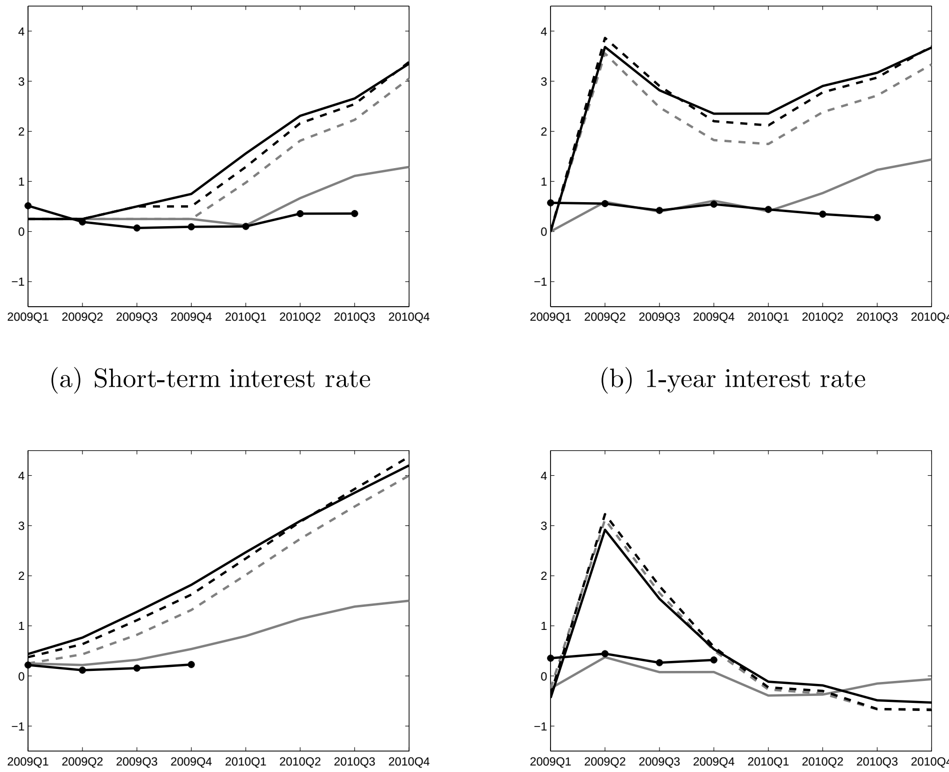
- Benchmark scenario: the short rate is fixed for the 4 quarters following the FOMC statement (i.e., $t-1 =$ 2008:Q4). Hence the new information is $\widetilde{Z}_{t,t+3} = \{r_t = \overline{r}, r_{t+1} = \overline{r}, r_{t+2} = \overline{r}, r_{t+3} = \overline{r}\}$;

- 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.



(a) Short-term interest rate (b) 1-year interest rate

(c) 1-year ahead nominal expectation term (d) 1-year nominal term premium

Figure 2: 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.

We report in figure 2 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 2 (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 $E\left(r_{t+3}\,|\,\Omega_{t-1}\right) = E\left(E_{t+1}(r_{t+3})\,|\,\Omega_{t-1}\right) = \overline{r}$.

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 2 (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 2 (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 2 (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.

8 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 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 also show that an important aspect of the NIRF is to be able to take into account an information about the past or future values of the variables of interest.

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 interest rate to a new information on the one-year ahead expected 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 usefulness of a new information on a "future path", 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 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:

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

we have:

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

and the result is proved.

Proof of Proposition 2:

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:

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

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 $\left\{\mathbb{1}_{]\,c_j, d_j\,[}(\varepsilon_{jt}), j \in J\right\}$, 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

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

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:

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

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

Quantitative information and interval informations on responses

Let us now consider the case where 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:

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

where the pair $(j, h) \in S \subset \{1, \ldots, n\} \times \{1, \ldots, H\}$ 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.

PROOF OF PROPOSITION 4: from Proposition 1 and given $a(\varepsilon_t, \underline{Y_{t-1}}) = \alpha$, the NIRF becomes:

$$\begin{aligned} & E(Y_{t+h}|a(\varepsilon_t, \underline{Y_{t-1}}) = \alpha, \underline{Y_{t-1}} = \underline{y_{t-1}}) - E(Y_{t+h}|\underline{Y_{t-1}} = \underline{y_{t-1}}) \\ \text{or } \; & E(Y_{t+h}|a(\varepsilon_t, \underline{y_{t-1}}) = \alpha, \underline{Y_{t-1}} = \underline{y_{t-1}}) - E(Y_{t+h}|\underline{Y_{t-1}} = \underline{y_{t-1}})\,. \end{aligned} \tag{A.5}$$

The first term of $(A.5)$ can be written as:

$$E[E(Y_{t+h}|\varepsilon_t, \underline{Y_{t-1}} = \underline{y_{t-1}}) \,|\, a(\varepsilon_t, \underline{y_{t-1}}) = \alpha, \underline{Y_{t-1}} = \underline{y_{t-1}}] \tag{A.6}$$

and the NIRF becomes:

$$\begin{aligned} & E\left\{[E(Y_{t+h}|\varepsilon_t, \underline{Y_{t-1}} = \underline{y_{t-1}}) - E(Y_{t+h}\,|\,\underline{Y_{t-1}} = \underline{y_{t-1}})] \,|\, a(\varepsilon_t, \underline{y_{t-1}}) = \alpha, \underline{Y_{t-1}} = \underline{y_{t-1}}\right\} \\ = \; & E[\Theta_h \varepsilon_t \,|\, a(\varepsilon_t, \underline{y_{t-1}}) = \alpha, \underline{Y_{t-1}} = \underline{y_{t-1}}] \\ = \; & E[\Theta_h \varepsilon_t \,|\, a(\varepsilon_t, \underline{y_{t-1}}) = \alpha] \\ = \; & \Theta_h \, E[\varepsilon_t \,|\, a(\varepsilon_t, \underline{y_{t-1}}) = \alpha] \\ = \; & \Theta_h \, \delta_t \,, \end{aligned} \tag{A.7}$$

with $\delta_t = E[\varepsilon_t \,|\, a(\varepsilon_t, \underline{y_{t-1}}) = \alpha]$.

Appendix 2: The full new information case

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: A) the orthogonalized shocks; B) the Uhlig (2005)'s impulse vectors and C) the structural shocks.

A) *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 $\Theta_h Pe_j$, 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:

$$\widetilde{P}^{-1}\Phi(L)Y_t = \widetilde{P}^{-1}\nu + \zeta_t \qquad (A.8)$$

and since $\widetilde{P}^{-1}$ is lower triangular with diagonal terms equal to 1, $(A.8)$ 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 $(A.8)$ 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 $\zeta_s = 0$, $s > t$.

B) *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 A), 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.

C) *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 B). 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. $A_{ij} = 0$. Such conditions could also be based on long-run restrictions when Y_t is non-stationary, admits r cointegrating relationships and has the following stationary VAR dynamics:

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

where $W_t = [\Delta\widetilde{Y}_t', (\Lambda' Y_t)']'$ and $C = \begin{pmatrix} I_{n-r} & 0 \\ & \Lambda' \end{pmatrix}$ 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); see also the survey by Kilian (2011)].

In any case, a shock structural e_j on η_t is a full information $A^{(j)}$ on ε_t. If we consider n shocks e_j, with $j = \{1, ..., n\}$, we can always rewrite any $\delta = E\left(\varepsilon_t | a(\varepsilon_t)\right)$ like a linear combination of the $A^{(j)}$ and thus, we may see the NIRF as a combination of structural shocks which are the most likely to generate that new information.

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:

$$Y_t = \nu + \sum_{j=1}^{p} \Phi_j Y_{t-j} + \varepsilon_t \qquad (A.9)$$

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

$$Z_t = \widetilde{\nu} + \Phi\, Z_{t-1} + \widetilde{\varepsilon}_t \qquad (A.10)$$

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:

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

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 5 can be applied (see Section 7.3).

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 $\Omega_{t-1} = \left\{ \widetilde{Z}_{t,t+\overline{H}-1}, Y_{\underline{t-1}} \right\}$. The NIRF of $EX_{t+k-1}(H)$, given Ω_{t-1}, is:

$$E(EX_{t+k-1}(H) | \Omega_{t-1}) - E(EX_{t+k-1}(H) | Y_{\underline{t-1}})\,, \tag{A.11}$$

and the first component in $(A.11)$ is:

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

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:

$$E(\Pi^e_{t+k-1}(H)|\Omega_{t-1}) - E(\Pi^e_{t+k-1}(H)|Y_{\underline{t-1}})\,, \tag{A.12}$$

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

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

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.

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, 2012, "No-arbitrage near-cointegrated VAR(p) term structure models, term premia and GDP growth", Working Paper available at http://www.crest.fr /pageperso/pegoraro/pegoraro.htm.

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.