New Information Response Functions

Caroline JARDET\(^{(1)}\) Alain MONFORT\(^{(2)}\)
Banque de France CNAM, CREST and Banque de France

Fulvio PEGORARO\(^{(3)}\)
Banque de France and CREST

First version : February, 2009. This version : July, 2009

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 the future responses, as well as informations on filters. 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)).

Keywords: impulse response functions, innovation, new information.

JEL Codes: C10, C32.

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

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

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

We would like to thank Mike Chernov, 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.
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. Finally, a third one uses 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 building our results on the non ambiguous notion of innovation $\varepsilon_t$ (say) of a stochastic process, that is to say, the difference between the value of the process and its conditional expectation given its past. We consider the impact of any kind of new information $a(\varepsilon_t)$ (say) at a given date $t$ on the future values of the process. The key remark is that such an impact is characterized by a shock on the innovation at $t$ defined by its conditional expectation given the new information.

We will study three important cases depending on the properties of function $a(.)$. 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 has a continuous probability distribution. This case includes the “generalized” impulse response function introduced by Pesaran and Shin (1998), the case of a set of impulse vectors, but also other informations on the subset of innovations. Third, we study the “discrete limited new information” case where the new information includes discrete functions, like sign functions, on either the innovation itself, or on an impulse vector or on a response. This general setting is then used to consider shocks on a linear filter of the vector of interest and responses of a linear filter.

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. Finally, Section 7 concludes and proposes further developments.
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:
\[
\Phi(L)Y_t = \nu + \varepsilon_t 
\]
where $\Phi(L) = I + \Phi_1 L + \ldots + \Phi_p L^p$, $L$ being the lag operator; $\varepsilon_t$ 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 $(y'_{-1}, y'_{-2}, \ldots, y'_{-p})' = y_{-p}$.

By considering the recursive equations:
\[
Y_{\tau} = \nu - \Phi_1 Y_{\tau-1} - \ldots - \Phi_p Y_{\tau-p} + \varepsilon_{\tau} 
\]
at $\tau = 0, \ldots, t$ and eliminating $Y_0, \ldots, Y_{t-1}$ we get a moving average representation of the form:
\[
Y_t = \mu_t + \sum_{\tau=0}^{t} \Theta_{\tau} \varepsilon_{t-\tau} 
\]
where $\mu_t$ is a function of $y_{-p}$ and the sequence $\Theta_{\tau}$ is such that:
\[
\left( \sum_{i=0}^{p} \Phi_i L^i \right) \left( \sum_{\tau=0}^{t} \Theta_{\tau} L^{\tau} \right) = I 
\]
which implies,
\[
\Theta_0 = I \quad \text{and} \quad \Theta_{\tau} = -\sum_{i=1}^{\tau} \Phi_i \Theta_{\tau-i}, \tau \geq 1, 
\]
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 $\Theta_{\tau}$.

Denoting $Y_L = (Y_t', Y_{t-1}', \ldots, Y_{t-p}')'$, equation (3) implies:
\[
E(Y_{t+h}|Y_L) - E(Y_{t+h}|Y_{t-1}) = \Theta_h \varepsilon_t 
\]
so $\Theta_h \varepsilon_t$ measures the differential impact of the knowledge of $\varepsilon_t$ on the updating of predictions 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 \( a(\cdot) \) is some function and \( \alpha \) is given. Obvious examples of such functions are: 
\[
a(\varepsilon_t) = \varepsilon_t, \quad a(\varepsilon_t) = b' \varepsilon_t, \quad a(\varepsilon_t) = \mathbb{1}_{\mathbb{R}^+}(b' \varepsilon_t)
\]
where \( b \) is some vector, etc. This impact, also called impulse response function, is:
\[
E(Y_{t+h}|a(\varepsilon_t), Y_{t-1}) - E(Y_{t+h}|Y_{t-1})
\]
\[
= E\left\{E(Y_{t+h}|\varepsilon_t, Y_{t-1}) - E(Y_{t+h}|Y_{t-1})|a(\varepsilon_t), Y_{t-1}\right\}
\]
\[
= E[\Theta_h \varepsilon_t|a(\varepsilon_t), Y_{t-1}]
\]
\[
= \Theta_h E[\varepsilon_t|a(\varepsilon_t)]. \tag{7}
\]
This means that the average impact on \( Y_{t+h} \) of a new information \( a(\varepsilon_t) \) at time \( t \) is the same as the one which would be implied by a shock \( \delta = E[\varepsilon_t|a(\varepsilon_t)] \) on the innovation \( \varepsilon_t \).

In the following, we will distinguish three important situations according to the properties of the function \( a(\cdot) \):

i) the “full new information” case, when \( a(\cdot) \) is one-to-one.

ii) the “continuous limited new information” case, when \( a(\cdot) \) 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.

3 Full new information

If \( a(\cdot) \) 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 \( \xi_t \) defined by \( \varepsilon_t = P\xi_t \). The distribution of \( \xi_t \) is obviously \( N(0, I) \) and it is usual to consider a shock
\(\epsilon_j\) on \(\xi_t\), where \(\epsilon_j\) is the \(j^{th}\) column of the \(n \times n\) identity matrix \(I\) (i.e., a shock of 1 on \(\xi_{jt}\) and of 0 on the other components). It is clear that the impact on \(Y_{t+h}\) of such a shock is the same as the shock \(\delta = Pe_j\) on \(\epsilon_t\), 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 \(\epsilon_t\) (or \(Y_t\)) is \(\Theta_h Pe_j\), or \(\Theta_h P^{(j)}\), where \(P^{(j)}\) is the \(j^{th}\) column of \(P\). In particular, the immediate impact on \(\epsilon_t\) (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 \(\tilde{P} = PD^{-1}\), where \(D\) is the diagonal matrix \(diag(P_{jj})\), and the vector \(\zeta_t\) defined by \(\zeta_t = D\xi_t\) or \(\epsilon_t = \tilde{P}\zeta_t\). Now, a shock \(e_j\) on \(\zeta_t\) has the impact \(\delta = \tilde{P}^{(j)}\) on \(\epsilon_t\) (or \(Y_t\)) and \(\Theta_h \tilde{P}^{(j)}\) on \(Y_{t+h}\). Also note that (1) can be rewritten:

\[
\tilde{P}^{-1}\Phi(L)Y_t = \tilde{P}^{-1}\nu + \zeta_t
\]

and since \(\tilde{P}^{-1}\) is lower triangular with diagonal terms equal to 1, (8) is the recursive form of the VAR. So the average impact on \(Y_{t+h}\) of a shock \(e_j\) on \(\zeta_t\), could be obtained recursively from (8) by computing \(Y_t, Y_{t+1}, ..., Y_{t+h}\) 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 the vector such that there exists a matrix \(A\) verifying \(AA' = \Sigma\) and admitting \(\gamma\) as a column. The set of vectors satisfying this definition can be seen as all the possible shocks on \(\epsilon_t\) implied by a shock \(e_j\) on a “fundamental” error \(\eta_t\) satisfying \(\epsilon_t = A\eta_t\) and \(V(\eta_t) = I\).

It turns out [see Uhlig (2005)] that those vectors \(\gamma\) are characterized by \(\gamma = P\beta\), where \(P\) is defined in Section 3.1, and \(\beta\) is a unit length vector of \(\mathbb{R}^n\). Equivalently, these vectors are such that \(\gamma'(P^{-1})'P^{-1}\gamma = 1\) or \(\gamma'S^{-1}\gamma = 1\) and therefore, they are an hyperellipsoid.

An impulse vector \(\gamma\) is a particular full new information on \(\epsilon_t\) 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 \(\beta\) is of length one.

### 3.3 Structural shocks

A structural error is defined as a vector \(\eta_t\) satisfying \(\epsilon_t = A\eta_t\), with \(\Sigma = AA'\), and, therefore \(V(\eta_t) = I\), like the “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 an impact $e_j$ on $\eta_t$ has no immediate impact on $\varepsilon_{it}$, 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:

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

where $\tilde{Y}_t$ is the subvector of $Y_t$ given by its first $(n - r)$ rows, 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 $C = \begin{pmatrix} I_{n-r} & 0 \\ \Lambda' \end{pmatrix}$.

The long run impact on the scalar components $y_{it}, i \leq n-r$, of a shock $e_j$ on $\eta_t$ 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, an information $e_j$ on $\eta_t$ is a full information $A^{(j)}$ on $\varepsilon_t$.

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 $\varepsilon_t$ 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.


Pesaran and Shin (1998) considered the case where $a(\varepsilon_t) \equiv \varepsilon_{jt}$. In the Gaussian case, the computation of $E[\varepsilon_t|a(\varepsilon_t) = \alpha]$ is straightforward and we get:

$$E [\varepsilon_{it}|\varepsilon_{jt} = \alpha] = \frac{\Sigma_{ij}}{\Sigma_{jj}} \alpha$$
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 $\Sigma$. 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 $a(\varepsilon_t) \equiv \varepsilon^K_t$, where $\varepsilon^K_t$ is a $K$-dimensional subvector of $\varepsilon_t$ containing any $\varepsilon_{jt}$ with $j \in K$ and $K \subset \{1, \ldots, n\}$, we have to compute $\delta = E[\varepsilon_t|\varepsilon^K_t = \alpha]$.

Again, in the Gaussian case we immediately get:

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

where $\Sigma^K$ is the matrix given by the columns $\Sigma^{(j)}$ of $\Sigma$ such that $j \in K$ and $\Sigma_{KK}$ is the variance-covariance matrix of $\varepsilon^K_t$.

For instance, if the new information is $\varepsilon_{jt} = 1$ and $\varepsilon_{kt} = 0$, the $i^{th}$ component of $\delta$ ($i \neq j$ and $i \neq k$) will be the coefficient of $\varepsilon_{jt}$ in the theoretical regression of $\varepsilon_{it}$ on $\varepsilon_{kt}$ and $\varepsilon_{jt}$.

4.3 Information defined as the set of impulse vectors

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 for instance 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 sign constraints will be 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 $\mathcal{S}_2 = \{\alpha_{21}, ..., \alpha_{2L}\}$. In this case the conditional distribution of any component $\varepsilon_{it}$ of $\varepsilon_t$ given $a_1(\varepsilon_t) = \alpha_1$ and $a_2(\varepsilon_t) = \alpha_{2j} \in \mathcal{S}_2$ is obtained by the conditional distribution of $\varepsilon_{it}$ given $a_1(\varepsilon_t) = \alpha_1$ restricted to the set $a_2(\varepsilon_t) = \alpha_{2j}$. In other words:

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

Note that simulations in this conditional distribution of $\varepsilon_t$ 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 $\varepsilon_t$ given $a_1(\varepsilon_t) = \alpha_1$ and keeping the first simulation $\tilde{\varepsilon}_t$ satisfying $a_2(\tilde{\varepsilon}_t) = \alpha_{2j}$. 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 $g(\tilde{\varepsilon}_t^s)$, $s = 1, ..., S$ and where $\tilde{\varepsilon}_t^s$ are obtained by keeping the simulation satisfying $a_2(\tilde{\varepsilon}_t) = \alpha_{2j}$ in a sequence of independent simulations in the conditional distribution of $\varepsilon_t$ given $a_1(\varepsilon_t) = \alpha_1$. However, in some cases explicit forms of such conditional expectations are available.

### 5.2 Quantitative informations and one sign information

Let us consider the case where $a_2(\varepsilon_t) = 1_{\mathbb{R}^+}(\varepsilon_{jt})$ and $a_1(\varepsilon_t) = \varepsilon^K_t$ with $K \subset \{1, ..., n\}$ such that $j \notin K$. Our purpose is to compute

$$E[\varepsilon_{jt} | \varepsilon^K_t = \alpha, \varepsilon_{jt} > 0]$$

and

$$E[\varepsilon_{it} | \varepsilon^K_t = \alpha, \varepsilon_{jt} > 0],$$

with $i \notin K$ and $i \neq j$. In both cases, explicit formulas are available.

i) Computation of $E[\varepsilon_{jt} | \varepsilon^K_t = \alpha, \varepsilon_{jt} > 0]$:

the conditional distribution of $\varepsilon_{jt}$ given $\varepsilon^K_t = \alpha$ is easily found; it is a Gaussian distribution with mean $\mu^K_{j\alpha}$ and variance $(\sigma^K_{j\alpha})^2$ (say) (where $\mu^K_{j\alpha}$ is a row vector). So $E[\varepsilon_{jt} | \varepsilon^K_t = \alpha, \varepsilon_{jt} > 0]$
is given by \( E[\mu_j^K \alpha + \sigma_j^K U | \mu_j^K \alpha + \sigma_j^K U > 0] \) where \( U \sim N(0, 1) \). We find

\[
E[\varepsilon_{jt} | \varepsilon_{jt} > 0] = \mu_j^K \alpha + \sigma_j^K \frac{\phi(\frac{\mu_j^K \alpha}{\sigma_j^K})}{\Phi(\frac{\mu_j^K \alpha}{\sigma_j^K})},
\]

where \( \phi \) and \( \Phi \) are, respectively, the p.d.f and the c.d.f of \( N(0, 1) \), and \( \lambda(x) = \frac{\phi(x)}{\Phi(x)} \) is the Mill’s ratio.

ii) Computation of \( E[\varepsilon_{it} | \varepsilon_{Kt} = \alpha, \varepsilon_{jt} > 0] \):
we first find the conditional expectation of \( \varepsilon_{it} \) given \( \varepsilon_{Kt} = \alpha \) and \( \varepsilon_{jt} > 0 \), which can be written as \( \mu_{ij}^K \alpha + \nu_{ij}^K \varepsilon_{jt} \) (say) and we get:

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

5.3 Quantitative informations and several sign informations

We still assume \( a_1(\varepsilon_t) = \varepsilon_t^K \), but now \( a_2(\varepsilon_t) \) is the set of functions \( \{ \mathbb{1}_{\mathbb{R}^+}(\varepsilon_{jt}), j \in J \} \), with \( J \subset \{1, \ldots, n\} \) and \( K \cap J = \emptyset \).

We have to compute

\[
E[\varepsilon_{it} | \varepsilon_{it}^K = \alpha, \varepsilon_{jt} > 0, j \in J], \ i \in J,
\]

and

\[
E[\varepsilon_{it} | \varepsilon_{it}^K = \alpha, \varepsilon_{jt} > 0, j \in J], \ i \notin K, i \notin J.
\]

i) Computation of \( E[\varepsilon_{it} | \varepsilon_{it}^K = \alpha, \varepsilon_{jt} > 0, j \in J], i \in J \):
the joint conditional distribution of \( \varepsilon_t^j \) given \( \varepsilon_t^K = \alpha \) is Gaussian with mean \( \mu_j^K \alpha \) and variance-covariance matrix \( \Sigma^{JK} \) (say) and we have to compute the mean of this normal distribution restricted to the orthant \( (\varepsilon_{jt} > 0, j \in J) \) (see below).
ii) Computation of $E[\varepsilon_{it}|\varepsilon_{Kt} = \alpha, \varepsilon_{jt} > 0, j \in J], i \notin K, i \notin J$:

$$E[\varepsilon_{it}|\varepsilon_{Kt} = \alpha, \varepsilon_{jt} > 0, j \in J] = E[E(\varepsilon_{it}|\varepsilon_{Kt} = \alpha, \varepsilon_{jt} > 0, j \in J)|\varepsilon_{Kt} = \alpha, \varepsilon_{jt} > 0, j \in J]$$

$$= \mu_{t}^{JK} + \nu_{t}^{JK}E[\varepsilon_{Jt}|\varepsilon_{Kt} = \alpha, \varepsilon_{jt} > 0, j \in J].$$

Again the joint conditional distribution of $\varepsilon_{Jt}$ given $\varepsilon_{Kt} = \alpha$ is $N(\mu_{t}^{JK}, \Sigma_{t}^{JK})$ and, as above, we have to compute the mean of this normal distribution restricted to the orthant ($\varepsilon_{jt} > 0, j \in J$).

The restriction of a $J$–variate normal distribution $N(m, Q)$ to the positive orthant is not easily 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 $y_{0} = (y_{01}, \ldots, y_{0J})$ 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 $R^{+}$, its simulation is straightforward. This algorithm is usually faster than the rejection algorithm.

### 5.4 Quantitative informations and sign informations on responses

The quantitative information is still $\varepsilon_{it}^{K} = \alpha$ but the sign information is related to some responses at some horizons. More precisely the sign information is:

$$\Theta_{h}^{(j)}\varepsilon_{t} > 0$$

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

$$E[\varepsilon_{it}|\varepsilon_{Kt} = \alpha, \Theta_{h}^{(j)}\varepsilon_{t} > 0, (j, h) \in S],$$

where $i \in K = \{1, \ldots, n\} - K$.

The conditional distribution of $\varepsilon_{it}^{K}$ given $\varepsilon_{it}^{K} = \alpha$ 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.
5.5 Impulse vector and sign information on responses

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

The conditional expectation

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

can still be computed by a Monte Carlo method. Indeed the conditional distribution of $\varepsilon_t$ given $\varepsilon_t' \Sigma^{-1} \varepsilon_t = 1$ is the image by $P$ of the conditional distribution of $\xi_t$ 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 as follows:

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

The expectation are obtained from the empirical means of the retained simulations.

6 Shocks on a filter and responses of a filter

6.1 Shocks 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 variable (see Jardet, Monfort, Pegoraro (2009)).

Let us consider a filter $\tilde{Y}_t = F(L)Y_t$, where $F(L) = (F_1(L), ..., F_n(L))$ is a row vector of polynomials in $L$. The innovation of $\tilde{Y}_t$ at $t$ is $\tilde{\varepsilon}_t = F(0)\varepsilon_t$, and therefore an information on $\varepsilon_t$, defined by $a(\varepsilon_t) = \alpha$, can be written as $\tilde{a}[F(0)\varepsilon_t] = \alpha$ or $a(\varepsilon_t) = \alpha$ (say). This means that, an information on $\varepsilon_t$ can be viewed as an information on $\varepsilon_t$ and it can be treated as in the previous framework. Let us consider some examples.

If the information is $\tilde{\varepsilon}_t = 1$ and $\varepsilon_{jt} = 0$, $j = 1, ..., n - 1$, the impact on $Y_{t+h}$ is $\Theta_h \delta$, where...
If the information is $\tilde{\varepsilon}_t = 1$ and $\varepsilon_{jt} = 0$, the impact on $Y_{t+h}$ is $\Theta_h\delta$ where the $i^{th}$ component $\delta_i$ is the coefficient of $\tilde{\varepsilon}_i$ in the theoretical regression of $\varepsilon_{it}$ on $\tilde{\varepsilon}_t$ and $\varepsilon_{jt}$ (in particular $\delta_j = 0$).

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 $\tilde{Y}_t = G(L)Y_t$, we can compute the impact on $\tilde{Y}_{t+h}$ of a new information $a(\varepsilon_t) = \alpha$ at $t$. Indeed, since the impact on $Y_{t+h}$ is $\Theta_hE[\varepsilon_t|a(\varepsilon_t) = \alpha]$, the impact on $\tilde{Y}_{t+h}$ is obviously $G(L)\Theta_hE[\varepsilon_t|a(\varepsilon_t) = \alpha]$ where the lag operator $L$ is operating on $h$ and where $\Theta_s = 0$ if $s < 0$.

7 Conclusions and Further Developments

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 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($p,q$) models $\Phi(L)Y_t = \mu + \Psi(L)\varepsilon_t$ by computing the $\Theta_h$ in the following way

$$\Theta_\tau = \Psi_\tau - \sum_{i=1}^{\tau} \Phi_i \Theta_{\tau-i}$$

with $\Phi_i = 0$ if $i > p$, $\Psi_\tau = 0$ if $\tau > q$ and $\Theta_s = 0$ if $s < 0$, ($\Phi(0) = I$, $\Psi(0) = I$).

The sign constraints could be replaced by more general information sets tackled by Monte Carlo methods, for instance imposing that some innovations belong to some intervals.

The extension to the nonlinear framework (see Gallant, Rossi, Tauchen (1993), Koop, Pesaran, Potter (1996), Gourieroux and Jasiak (2005)) is less obvious and could be the objective of further research.
References


