Testing for Sufficient Information in Structural VARs*

Mario Forni†
Università di Modena e Reggio Emilia
CEPR and RECent

Luca Gambetti‡
Universitat Autonoma de Barcelona

February 22, 2011

Abstract

We derive necessary and sufficient conditions under which a set of variables is *informationally sufficient*, i.e. it contains enough information to estimate the structural shocks with a VAR model. Based on such conditions, we suggest a procedure to test for informational sufficiency. Moreover, we show how to amend the VAR if informational sufficiency is rejected. We apply our procedure to a VAR including TFP, unemployment and per-capita hours worked. We find that the three variables are not informationally sufficient. When adding missing information, the effects of technology shocks change dramatically.

JEL classification: C32, E32, E62.

Keywords: Structural VAR, non-fundamentalness, information, FAVAR models, technology shocks.

*We thank Fabio Canova, Marco Lippi and Luca Sala for helpful discussion.

†Financial support from Fondazione Cassa di Risparmio di Modena is gratefully acknowledged. Contact: Dipartimento di Economia Politica, via Berengario 51, 41100, Modena, Italy. Tel. +39 0592056851; e-mail: mario.forni@unimore.it

‡The financial support from the Spanish Ministry of Science and Innovation through grant ECO2009-09847 and the Barcelona Graduate School Research Network is gratefully acknowledged. Contact: Office B3.1130 Departament d'Economia i Historia Economica, Edifici B, Universitat Autonoma de Barcelona, Bellaterra 08193, Barcelona, Spain. Tel (+34) 935814569; e-mail: luca.gambetti@uab.cat

1 Introduction

Since Sims (1980)'s seminal paper, Structural Vector Autoregression (SVAR) models have become extremely popular for structural and policy analysis. The idea behind these models is that structural economic shocks can be found as linear combinations of the residuals of the linear projection of a vector of variables onto their past values, i.e. are innovations with respect to the econometrician's information set. Therefore, an obvious requirement for the analysis to be meaningful is that such an information set conveys all of the relevant information. This is implicitly assumed in any VAR application.

But is this assumption always sensible? Unfortunately the answer is no. The basic problem is that, while agents typically have access to rich information, VAR techniques allow to handle a limited number of variables. If the econometrician's information set does not span that of the agents the structural shocks are non-fundamental and cannot be obtained from a VAR (Hansen and Sargent, 1991, Lippi and Reichlin, 1993, 1994, Chari, Kehoe and Mcgrattan, 2008). Fernandez-Villaverde *et al.* (2007) derives a simple condition to check whether the shocks of a DSGE model are recoverable from a VAR and shows theoretical cases in which VAR techniques fail. Fiscal foresight and news shocks are two examples, see Leeper, Walker and Yang, (2008) and Yang (2008). Forni and Gambetti (2010), Forni and Gambetti and Sala (2010) and Gambetti (2010).

At now there are no testing procedures to verify whether a specific VAR suffers from this informational problem. The contribution of this paper is twofold. First we theoretically characterize necessary and sufficient conditions under which a set of variables is *informationally sufficient* in a VAR, i.e. it contains enough information to estimate the structural shocks. Second, we propose a testing procedure based on such conditions. When informational sufficiency is rejected we propose a strategy to amend the VAR to fill the informational gap.

We derive two main results under the general assumption the economy admits a state space representation. First, we provide a necessary and sufficient condition for informational sufficiency. The condition requires that there are no state variables that Granger cause the variables included in the VAR.[1] The intuition is that the state

[1] The precise relation between our sufficient information condition and Condition 1 of Villaverde *et al.* (2007) is explained in Section 2.3. An essential difference is that our condition can be tested without resorting to any particular economic model.

variables contain all of the relevant information; therefore, if they do not help to predict a vector, such vector must contain the same information. Second, we show that, even if the VAR is not informational sufficient, still a single shock of interest can be correctly estimated. In order for this to be the case, the shock must be orthogonal to the past of the state variables.

Such conditions can be tested empirically. Based on the former result, we suggest the following testing procedure. First, we estimate the space spanned by the state variables of the economy by using the principal components of a large dataset, containing all available macroeconomic information. Second, we test whether the estimated principal components Granger cause the variables included in the VAR. The variables are informationally sufficient if and only if the null hypothesis of no Granger causality is not rejected.

The latter result can be used to verify whether, even if the VAR is not informationally sufficient, a particular shock of interest can still be estimated. The test works as follows. First, we identify and estimate the structural shock. Second, we perform a test of orthogonality between the estimated shock and the lags of the principal components. If the null of orthogonality is rejected, then the shock obtained from the VAR cannot be structural.

If a set of variables is not sufficient, we suggest to estimate either a structural factor model like Forni *et al.* (2009) or a VAR augmented by by the principal components, i.e. the FAVAR model proposed by Bernanke Boivin and Eliasz (2005), where number of principal components is determined by applying a sequence of sufficient information tests.

As an application we study technology shocks in the US. We test whether a small-scale VAR model, such as those typically used to study the effects of technology shocks, is informationally sufficient. Specifically, we use a VAR with total factor productivity, the unemployment rate and per-capita hours worked. We find that these three variables are Granger caused by the first two principal components of a large dataset of US macrcoeconomic variables. Therefore we add such principal components to the VAR and show that the remaining principal components do not Granger cause the augmented VAR, meaning that the information conveyed in the augmented VAR is sufficient. Finally, we identify the technology shock as the only one driving total factor productivity in the long run, in both the original and the augmented VAR. Differences in the results in the two models are dramatic. While in the original VAR technology shocks increase hours and reduce unemployment, in the augmented VAR results are reversed: hours reduce and unemployment increases. In the augmented model, investment and GDP react very sluggishly to the shock, prices fall and the real wage increases. Overall the result are hard to reconcile with the view that technology shocks are an important source of business cycle fluctuations.

The remainder of the paper is organized as follows. Section 2 presents theoretical results, as well as our proposed testing procedures. Section 3 discusses the application. Section 4 concludes.

2 Theory

2.1 The macroeconomy

Let us start from the following MA representation of the macroeconomy.

Assumption 1 (MA representation). *The n-dimensional vector x_t of stationary macroeconomic time series satisfies*

$$x_t = F(L)u_t, \tag{1}$$

where u_t is a q-dimensional, orthonormal white noise vector of structural macroeconomic shocks and $F(L)$ is an $n \times q$ matrix of impulse response functions, i.e. square-summable linear filters in the non-negative powers of the lag operator L, such that $\operatorname{rank}(F(z)) = q$ *for some complex number z.*

Representation (1) can be thought of as the representation of a macroeconomic equilibrium. Consider for instance the state-space representation studied in Villaverde, Rubio-Ramirez, Sargent and Watson (2007), i.e.

$$s_t = As_{t-1} + Bu_t \tag{2}$$

$$x_t = Cs_{t-1} + Du_t \tag{3}$$

where s_t is an r-dimensional vector of stationary "state" variables, $q \leq r \leq n$, A, B, C and D are conformable matrices of parameters, B has a left inverse B^{-1} such that $B^{-1}B = I_q$. Pre-multiplying (2) by B^{-1} we get $u_t = B^{-1}(I - AL)s_t$. Substituting this into (3) and rearranging gives

$$x_t = \left(DB^{-1} + (C - DB^{-1}A)L\right) s_t. \tag{4}$$

Stationarity of s_t ensures invertibility of (2), so that $s_t = (I - AL)^{-1} B u_t$. Combining this with (4) we get the MA representation

$$x_t = \left(DB^{-1} + (C - DB^{-1}A)L\right)(I - AL)^{-1} B u_t, \tag{5}$$

which is a special case of (1).

The assumption on the rank $(F(L))$ ensures that the representation is not redundant in the sense that there is another representation with a smaller number of shocks.

2.2 Sufficient information

The SVAR econometrician observes x_t, possibly with error. Precisely,

Assumption 2. (Econometrician's information set) *The econometrician information set $\mathcal{X}_t^*$ is given by the closed linear space spanned by present and past values of the variables in x_t^* (in symbols $\mathcal{X}_t^* = \overline{\text{span}}(x_{1t}^*, \ldots, x_{nt}^*)$), where*

$$x_t^* = x_t + \xi_t = F(L) u_t + \xi_t, \tag{6}$$

ξ_t being a (possibly zero) vector of measurement errors, orthogonal to u_{jt-k}, $j = 1, \ldots, q$, any k, and ξ_{t-k}, $k > 0$.

In practice the number of observable variables n is very large, so that the econometrician needs to reduce it in order to estimate a VAR. The VAR information set is then spanned by an s-dimensional sub-vector of x_t^*, or more, generally, an s-dimensional linear combination of x_t^*, say $z_t^* = W x_t^*$ (with s not necessarily equal to q).

Assumption 3 (VAR information set). *The information set of the VAR is $\mathcal{Z}_t^* = \overline{\text{span}}(z_{1t-k}^*, \ldots, z_{st-k}^*, k \geq 0)$, $z_t^* = W x_t^*$, W being $s \times n$.*

Now, consider the theoretical projection equation of z_t^* on its past history, i.e.

$$z_t^* = P(z_t^* | \mathcal{Z}_{t-1}^*) + \epsilon_t. \tag{7}$$

The SVAR methodology consists in (a) estimating a VAR to get ϵ_t; (b) attempting to get the structural shocks as linear combinations of the estimated entries of ϵ_t. Hence a key property of z_t^* and the related information set, is that the entries of ϵ_t span the structural shocks, i.e. the information in the history of z_t^* is sufficient to estimate the shocks. We call such property "sufficient information".

Definition 1 (Sufficient information). *We say that z_t^* and the related information set $\mathcal{Z}_t^*$ contain "sufficient information" if and only if there exist a matrix M such that $u_t = M\epsilon_t$.*

Let us stress that sufficiency, defined in this way, is related only to the variables in z_t^* and has nothing to do with the choice of a proper identification scheme. The correct identification of M is a further problem, which does make sense only if sufficiency holds true.

2.3 Sufficient information and fundamentalness

From (6) and the definition of z_t^* we get

$$z_t^* = WF(L)u_t + W\xi_t = z_t + W\xi_t. \tag{8}$$

Structuralness is related to "fundamentalness" of the MA representation in (8).[2] Let us first recall the concept of fundamentalness.

Definition 2 (Fundamentalness). *We say that u_t is fundamental for $w_t = Hx_t$, and the MA representation $w_t = HF(L)u_t$ is fundamental, if and only if $u_t \in \mathcal{W}_t = \overline{\text{span}}(w_{1t-k}, \ldots, w_{mt-k}, k \geq 0)$ (i.e. $\mathcal{U}_t = \overline{\text{span}}(u_{1t-k}, \ldots, u_{qt-k}, k \geq 0) = \mathcal{W}_t$).*

The following proposition holds:

Proposition 1. *The information in z_t^* is sufficient if and only if (a) $z_{jt} \in \mathcal{Z}_t^*$ for any j and (b) u_t is fundamental with respect to z_t.*

Proof. If (a) and (b) hold true, then $u_t \in \mathcal{Z}_t^* = \mathcal{E}_t = \overline{\text{span}}(\epsilon_{1t-k}, \ldots, \epsilon_{st-k}, k \geq 0)$. Being orthogonal to $\mathcal{E}_{t-1}$, u_t belongs to $\overline{\text{span}}(\epsilon_{1t}, \ldots, \epsilon_{st})$. On the other hand, let us assume that z_t^* is sufficient, i.e. $u_t = M\epsilon_t$. Then (a) holds, because $z_{jt} \in \mathcal{U}_t$ and $\mathcal{U}_t \subseteq \mathcal{Z}_t^*$. As for (b), let $\mathcal{S}_t = \overline{\text{span}}(z_{1t-k}, \ldots, z_{st-k}, W\xi_{t-k}, k \geq 0)$. Now, $ujt \in \mathcal{S}_{t-1}$), $j = 1, \ldots, q$, since it belongs to $\mathcal{Z}_t^*$ and $\mathcal{Z}_t^* \subseteq \mathcal{S}_t$. But u_{jt} is orthogonal to ξ_{t-k}, $k \geq 0$ by Assumption 2. Hence $u_{jt} \in \mathcal{Z}_t$, $j = 1, \ldots, q$. QED

Proposition 1 says that, for z_t^* being sufficient, there must be a linear transformation of z_t^* which is free of measurement errors and have a fundamental representation in the structural shocks.

[2]Some important references about fundamentalness are Hansen and Sargent (1991), Lippi and Reichlin (1993, 1994), Chari, Kehoe and McGrattan (2008), Fernandez-Villaverde *et al.* (2007).

To conclude this section, let us observe that, in the particular case of $F(L)$ being a matrix of rational functions, fundamentalness of u_t for w_t, along with fundamentalness of the associated MA representation $w_t = HF(L)u_t$ is equivalent to the following condition (see e.g. Rozanov, 1967, Ch. 2).

Condition R. *The rank of $HF(z)$ is q for all z such that $|z| < 1$.*

Considering equation (5) and the case $w_t = x_t$, condition R is satisfied if and only if D is invertible and the eigenvalues of $A - BD^{-1}C$ are strictly less than one in modulus, which is Condition 1 of Villaverde *et al.* (2007).

2.4 Testable implications of sufficient information

Proposition 2. *If x_t^* Granger causes z_t^*, then z_t^* is not informationally sufficient.*

Proof. Assume that z_t^* is sufficient, so that $u_t = M\epsilon_t$. Then $u_{jt-k} \in \mathcal{Z}_{t-1}^*$ for $k > 0$. It follows that $P(z_t^*|\mathcal{Z}_{t-1}^*) = W\left(F_1 u_{t-1} + F_2 u_{t-2} + \ldots\right)$ and $\epsilon_t = WF_0 u_t + W\xi_t$. Hence ϵ_t is orthogonal to both u_{t-k}, $k > 0$, and, by serial uncorrelation of ξ_t (Assumption 2), ξ_{t-k}, $k > 0$. Therefore $\epsilon_t \perp x_{t-k}^*$, $k > 0$ and x_t^* does not Granger cause z_t^* . QED

The intuition is that, if a set of variables is sufficient, than it contains all of the existing information, so that no other variable or set of variables can Granger cause it.

Proposition 2 can be of some usefulness in practice.[3] In particular, if the econometrician believes that a given variable in x_t^*, say v_t, conveys relevant information, he can check whether v_t Granger causes z_t^* as a vector. If v_t Granger causes z_t^*, the VAR with z_t^* is misspecified. Observe that, according to Proposition 2, identification is not required to perform the test, consistently with the fact that sufficient information, as observed above, is independent of the identification scheme.

On the other hand, Proposition 2 has an important limitation in that, being only a necessary condition, it can be used to reject sufficiency but not to validate it. Clearly, testing all of the variables in x_t^* would be close to a validation, but unfortunately this is not feasible, since in practice x_t^* is of high dimension. On the one hand, we cannot use all of the variables simultaneously; on the other hand, testing each one of them separately would yield, with very high probability, to reject sufficiency even if z_t^* is informationally sufficient, owing to Type I error.

[3] Proposition 2 is derived (within somewhat different settings) in Forni and Reichlin (1996) and Giannone and Reichlin (2006).

We can provide a sufficient condition by assuming the state space representation above, i.e. by replacing Assumption 1 with the more restrictive Assumption 1′:

Assumption 1′ (ABCD representation). *The vector x_t of macroeconomic time series satisfies equations (2) and (3).*

It is easily seen from equations (6) and (4) that x_t^* follows the static factor model

$$x_t^* = Gf_t + \xi_t, \tag{9}$$

where $G = \begin{pmatrix} DB^{-1} & C - DB^{-1}A \end{pmatrix}$ and $f_t = \begin{pmatrix} s_t' & s_{t-1}' \end{pmatrix}'$.

In addition, we need to assume that the history of the structural shocks helps predicting z_t^*, or, equivalently, that z_t^* is autocorrelated to some extent (since otherwise nothing can Granger cause it).

Assumption 4 (Autocorrelation of z_t^*). *There exists a summable sequence $\{c_k\}_{k=1}^{\infty}$ such that $R = W\sum_{k=1}^{\infty} c_k F_k$ has rank q.*

The following proposition establishes a necessary and sufficient condition for informational sufficiency.

Proposition 3. *Let K be any non-singular $p \times p$ matrix, p being the dimension of f_t. z_t^* is informationally sufficient if and only if $g_t = Kf_t$ does not Granger cause z_t^*.*

Proof. Let us assume that z_t^* is sufficient, i.e. $u_t = M\epsilon_t$. Then ϵ_t is orthogonal to u_{t-k}, $k > 0$ and therefore to g_{t-k}, $k > 0$. Hence $P(z_t^*|\mathcal{Z}_{t-1}^*) = P(z_t^*|z_{jt-k}^*, g_{it-k}, j = 1, \ldots, s, i = 1, \ldots, p, k > 0)$, so that g_t does not Granger cause z_t^*. Regarding the opposite implication, let us assume that g_t does not Granger cause z_t^*. We have $P(z_t^*|\mathcal{Z}_{t-1}^*) = P(z_t^*|z_{jt-k}^*, g_{it-k}, j = 1, \ldots, s, i = 1, \ldots, p, k > 0)$. But the latter projection is equal to $P(z_t^*|u_{jt-k}, j = 1, \ldots, q, k > 0) = W\sum_{k=1}^{\infty} F_k u_{t-k} = \zeta_t$, since ζ_t belongs to $\overline{\text{span}}(z_{jt-k}^*, g_{it-k}, j = 1, \ldots, s, i = 1, \ldots, p, k > 0)$ and $z_t^* - \zeta_t$ is orthogonal to such space because of Assumption 2. On the other hand, $\zeta_t = P(z_t^*|\mathcal{Z}_{t-1}^*) = \sum_{k=1}^{\infty} A_k \epsilon_{t-k}$. Projecting both sums on $\overline{\text{span}}(\epsilon_{it-k}, u_{it-k}, i = 1, \ldots, s, j = 1, \ldots, r)$ we get $WF_k u_{t-k} = A_k \epsilon_{t-k}$ for all k, so that $WF_k u_t = A_k \epsilon_t$ for all k and $R = \left(W\sum_{k=1}^{\infty} c_k F_k\right) u_t = \left(\sum_{k=1}^{\infty} A_k\right) \epsilon_t$. Assumption 4 ensures that R has a left inverse, so that $u_t = R^{-1}\left(\sum_{k=1}^{\infty} A_k\right)\epsilon_t$. QED

The intuition for sufficiency is that, under Assumption 1', the factors contain all of the information available in the system; therefore they Granger cause every predictable vector, unless such vector contain the same information.

Proposition 3 is useful in that, besides providing a sufficient condition, allows us to summarize the signals in the large dimensional vector x_t into a relatively small number of factors (the entries of g_t). Such factors are unobservable, but, under suitable assumptions, can be consistently estimated by the principal components $\hat{g}_t$, as both the number of variables and the number of time observations go to infinity (Stock and Watson, 2002; Forni, Giannone, Lippi and Reichlin, 2009).

2.5 Testing for sufficient information

Proposition 3 provides the theoretical basis for the following testing procedure.

1. Take a large data set x_t^* capturing all of the relevant macroeconomic information.
2. Set a maximum number of factors P and compute the first P principal components of x_t^*.
3. Perform Granger causation tests to see whether the first h principal components, $h = 1, \ldots, P$, Granger cause z_t^*. If the null of no Granger causality is never rejected, z_t^* is informationally sufficient. Otherwise, sufficiency is rejected.

If informational sufficiency is rejected, we cannot use the VAR for global identification. However, partial identification could still provide correct results, as shown in the following subsection.

2.6 Structuralness of a single shock

Even if informational sufficiency is rejected, z_t^* could be sufficient to get a single shock of interest, say u_{1t}, or a subset of shocks $u_{1t}, \ldots, u_{jt}$, $j < q$. This is important in that for many applications the econometrician is interested in identifying just a single shock.

To see this, consider the following example

$$
\begin{aligned}
z_{1t}^* &= u_{1t} + u_{2t-1} \\
z_{2t}^* &= u_{1t} - u_{2t-1}
\end{aligned}
$$

In this case z_t^* is not sufficient for u_t by Proposition 1. In fact, since the determinant of the MA filter has a zero in zero, the MA representations non fundamental by

Condition R. Indeed, it is easily seen that u_{2t} cannot be recovered from the present and the past of z_t^*. Nevertheless, z_t^* is sufficient for u_{1t}, since $z_{1t}^* + z_{2t}^* = 2u_{1t}$.

The following proposition is an immediate consequence of Assumption 1.

Proposition 4. *The structural shock u_{jt}, $j = 1, \dots, q$ is orthogonal to x_{t-k}^*, $k > 0$, and the lagged factors f_{t-k}, $k > 0$.*

Proposition 4 essentially states that a structural shock is unpredictable. After having identified the shock of interest, we can verify whether it can be a structural shock by testing for orthogonality with respect to the past of the principal components.[4]

If orthogonality is not rejected, the econometrician could rely on the estimated shock. Let us stress however that orthogonality is only a necessary condition for structuralness. Hence even if it is not rejected, it is safer to enlarge the VAR information set as suggested below.

2.7 A solution for insufficient information

What should the econometrician do if sufficient information is rejected? A possibility is to estimate a factor model along the lines of Forni *et al.* (2009).

An alternative solution to fill the informational gap is to add the principal components $\hat{g}_t$ to the VAR information set and estimate a FAVAR with $w_t = (z_t^{*\prime} \;\; \hat{g}_t^{\prime})^{\prime}$.

By looking at equation (9) it is seen that the x's are linear combinations of the factors in f_t and therefore, asymptotically, are linear combinations of the entries of w_t, say $x_t = Q\,(z_t^{*\prime} \;\; g_t^{\prime})^{\prime}$. An immediate consequence is that we can estimate the impulse response functions of all of the x's simply as $\hat{Q}\hat{B}(L)$, where the entries of $\hat{Q}$ are the coefficients of the OLS projection of x_t^* on w_t and the entries of $\hat{B}(L)$ are the estimated impulse response functions of the VAR with w_t.

This is interesting in that it enables us to study the effects of our shock of interest on many variables. In addition, a key implication is that the shocks of interest can be identified by imposing restrictions on variables which are not included in the VAR. This is very useful since restrictions on the principal components would be very difficult to interpret.

[4] Ramey (2009) applies a version of this test to check whether the fiscal policy shock obtained with a SVAR *á la* Perotti (2007) is structural. She however does not use the principal components, but the forecast of public expenditure from the survey of professional forecasters.

A crucial problem is to establish how many principal components to retain. A first possibility is to rely on existing information criteria.[5] An alternative is to use again Proposition 3 as follows.

1. Take $w_t^h = (z_t^{*\prime} \;\; \hat{g}_{1t} \;\; \cdots \;\; \hat{g}_{ht})'$ and test for sufficiency of w_t^h as explained above, for $h = 1, \ldots, P$.

2. Retain p principal components if w_t^p is informationally sufficient whereas $w_t^1, \ldots, w_t^{p-1}$ are not.

   Such a procedure is the one we follow in the empirical application below.

3 An Application to Technology Shocks

3.1 Technology shocks and the business cycle

Do technology shock explain aggregate fluctuations? Despite the huge amount of works that have addressed this question over the last years, no consensus has been reached. The empirical evidence is mixed. In his seminal paper, Gali (1999) finds a very modest role for technology shocks as a source of economic fluctuations. The result echoes the finding in Blanchard and Quah (1989) that aggregate supply shocks are not important for the business cycle. On the contrary other authors, see for instance Christiano, Eichenbaum and Vigfusson (2003) and Beaudry and Portier (2006), provide evidence that technology shocks are capable of generating sizable fluctuations in macroeconomic aggregates.

Most of the existing evidence about the effects of technology shocks is obtained using small-scale VAR models. In many cases only two or three variables are used. Here, as an application of our testing procedure, we investigate whether a small scale model conveys enough information to identify the shocks, in particular the technology shock.

We consider the vector z_t^* including the growth rate of total factor productivity (TFP_t), the unemployment rate (u_t) and the logs of per capita hours worked (h_t). The space spanned by the state variables of the economy is estimated by using the principal components of a large dataset of US macroeconomic variables.[6]

[5] See for instance the criteria in Bai and Ng (2002) and Onatski (2010).

[6] See the Appendix for the precise definition and the treatment of the variables used in the dataset.

3.2 Testing for informational sufficiency

We apply our testing procedure to this VAR. We use the Gelper and Croux (2007) multivariate extension of the out-of-sample Granger causality test proposed by Harvey *et al.*(1998).

Table 1 shows the results. The first column of panel A shows the p-value of the test of the null hypothesis that the first principal component does not Granger cause z_t^*. The hypothesis is strongly rejected suggesting that the three variables do not contain sufficient information to correctly recovering the structural shocks. The second column of A shows the p-values of the test of the null hypothesis that the VAR augmented by the first principal component, i.e. $w_t^1 = (z_t' \;\; \hat{g}_{1t})'$, is not Granger caused by the remaining principal components from the second to the j-th, $j = 2, \dots, P$. For instance the third element of the column, i.e. 0.405, is the p-value obtained by testing that $(\hat{g}_{2t} \;\; \hat{g}_{3t})'$ does not Granger cause w_t^1. We reject that the principal components from the second up to the eleventh do not Granger cause w_t^1 at the 5% level, suggesting that not even w_t^1 is informationally sufficient. However we can not reject that w_t^2 is informationally sufficient since it is never Granger caused by the remaining principal components. Augmenting z_t^* with the first two principal components is sufficient to obtain the structural shocks, including the technology shock.

3.3 Testing for structuralness of the technology shock

As observed in subsection 2.6, even if the VAR is not informationally sufficient, still it could be possible to identify the technology shock. To check whether this is the case, we identify the technology shock, following Beaudry and Portier (2006), as the only one affecting total factor productivity in the long run. Then we test whether the shock is orthogonal to the past of the estimated principal components. Precisely, we run a regression of the estimated shock on the lagged principal components and perform an F-test of the null hypothesis that the coefficients are jointly zero. The first column of B in Table 1 displays the p-value of the test when only the first principal component is included as a regressor. The hypothesis is strongly rejected suggesting that the shock obtained from the original VAR is not structural.

Then we implement the same identification in the VARs for w_t^1 and w_t^2 and run the same orthogonality test. The second column reports the p-values for w_t^1. The null that the second principal component does not predict the shock is rejected at the 10% but not the 5% level. The hypothesis that the shock is orthogonal to the principal components from the second up to the eighth is strongly rejected. Finally, orthogonality is never rejected for the w_t^2 specification, consistently with the results of panel A.

3.4 Information and impulse response functions

Next we study the consequences of insufficient information in terms of impulse response functions. In particular, we investigate to what extent the effects of technology shocks change by augmenting the original VAR with the principal components. According to the results of the test, impulse response functions are expected to change when adding the first two principal components, but should remain essentially unchanged when adding further components.

Figure 1 shows the impulse response functions. The left column plots the impulse response functions for the three varables, total factor productivity, unemployment and per capita hours, for all the sixteen specifications $z_t^*, w_t^1, \ldots, w_t^{15}$. The solid line with dots represents the impulse response functions estimated with z_t^*. The line with crosses represents the impulse response functions estimated with w_t^2. The remaining lines are the estimated responses of the other models. The effects are expressed in percentage terms. The right column displays for the three variables the impact effect (dots), the effect at 1 year (crosses), 2 years (circles) and in the long run (diamonds). The horizontal axis displays the number of principal components included in the VAR.

The VAR without principal components predicts that the technology shock increases per-capita hours worked and reduces unemployment. Such results are in line with the theoretical predictions of standard RBC models and the empirical findings of Christiano, Eichenbaum and Vigfusson (2003) and Beaudry and Portier (2006). Total factor productivity reacts positively on impact and stays roughly constant afterward, with no delay in the diffusion process.

The picture changes dramatically when adding the principal components. The effects on both unemployment and hours change sign. Now, unemployment increases and hours reduce so that technology becomes contractionary. Moreover, the impact effect of productivity reduces substantially while the long run effect is roughly unchanged so that the diffusion process is substantially slower in line with the S-shape view and the recent news shocks literature (Beaudry and Portier, 2006, and Schmitt-Grohe and

Uribe, 2008).

Notice that, consistently with the results of the test, models including more than two principal components, all deliver the same impulse response functions. This can also be seen from the right panels of Figure 1. Impulse response functions change radically by adding the first principal component, and to a lesser extent by adding the second one, but are roughly constant from that point onward.

Figure 2 plots the impulse response functions of some variables of interest for the specification w_t^2 . The solid line represents the point estimate while the dotted lines are the 68% confidence bands. Investment and GDP do not react significantly on impact and start to increase significantly only after a few quarters, reaching their maximal level after about two years. The shape of the response of consumption is similar to that of investment and GDP (although the impact effect is slightly negative). The GDP deflator reduces immediately while real wages immediately increase.

Overall the picture that emerges is hard to reconcile with the view that technology shocks are an important source of business cycle fluctuations.

4 Conclusions

This paper derives necessary and sufficient conditions under which a set of variables is *informationally sufficient*, i.e. contains enough information to estimate the structural shocks with a l VAR model. Based on such conditions, a procedure to test for informational sufficiency is proposed. Moreover, a test is provided to verify whether a single shock obtained with partial identification is a structural shock. Finally, the paper shows how to amend the model if informational sufficiency and structuralness are rejected.

Our testing procedures are applied to a three-variable VAR including TFP, unemployment and per-capita hours worked. It is found that the VAR is not informationally sufficient, and the technology shock, identified as the only one affecting TFP in the long run, is not a structural shock. When amending the model by adding missing information, informational sufficiency and structuralness cannot be rejected. Results in terms of impulse response functions change dramatically: the reaction of both unemployment and hours worked changes sign, so that a positive shock becomes contractionary, and the response of TFP becomes S-shaped, in accordance with the recent "news" shock literature.

Appendix: Data

Transformations: 1=levels, 2= first differences of the original series, 4 = logs of the original series, 5= first differences of the logs of the original series .

no.series	Transf.	Mnemonic	Long Label
1	5	GDPC1	Real Gross Domestic Product, 1 Decimal
2	5	GNPC96	Real Gross National Product
3	5	NICUR/GDPDEF	National Income/GDPDEF
4	5	DPIC96	Real Disposable Personal Income
5	5	OUTNFB	Nonfarm Business Sector: Output
6	5	FINSLC1	Real Final Sales of Domestic Product, 1 Decimal
7	5	FPIC1	Real Private Fixed Investment, 1 Decimal
8	5	PRFIC1	Real Private Residential Fixed Investment, 1 Decimal
9	5	PNFIC1	Real Private Nonresidential Fixed Investment, 1 Decimal
10	5	GPDIC1	Real Gross Private Domestic Investment, 1 Decimal
11	5	PCECC96	Real Personal Consumption Expenditures
12	5	PCNDGC96	Real Personal Consumption Expenditures: Nondurable Goods
13	5	PCDGCC96	Real Personal Consumption Expenditures: Durable Goods
14	5	PCESVC96	Real Personal Consumption Expenditures: Services
15	5	GPSAVE/GDPDEF	Gross Private Saving/GDP Deflator
16	5	FGCEC1	Real Federal Consumption Expenditures & Gross Investment, 1 Decimal
17	5	FGEXPND/GDPDEF	Federal Government: Current Expenditures/ GDP deflator
18	5	FGRECPT/GDPDEF	Federal Government Current Receipts/ GDP deflator
19	2	FGDEF	Federal Real Expend-Real Receipts
20	1	CBIC1	Real Change in Private Inventories, 1 Decimal
21	5	EXPGSC1	Real Exports of Goods & Services, 1 Decimal
22	5	IMPGSC1	Real Imports of Goods & Services, 1 Decimal
23	5	CP/GDPDEF	Corporate Profits After Tax/GDP deflator
24	5	NFCPATAX/GDPDEF	Nonfinancial Corporate Business: Profits After Tax/GDP deflator
25	5	CNCF/GDPDEF	Corporate Net Cash Flow/GDP deflator
26	5	DIVIDEND/GDPDEF	Net Corporate Dividends/GDP deflator
27	5	HOANBS	Nonfarm Business Sector: Hours of All Persons
28	5	OPHNFB	Nonfarm Business Sector: Output Per Hour of All Persons
29	5	UNLPNBS	Nonfarm Business Sector: Unit Nonlabor Payments
30	5	ULCNFB	Nonfarm Business Sector: Unit Labor Cost
31	5	WASCUR/CPI	Compensation of Employees: Wages & Salary Accruals/CPI
32	1	COMPNFB	Nonfarm Business Sector: Compensation Per Hour
33	5	COMPRNFB	Nonfarm Business Sector: Real Compensation Per Hour
34	1	GDPCTPI	Gross Domestic Product: Chain-type Price Index
35	1	GNPCTPI	Gross National Product: Chain-type Price Index
36	1	GDPDEF	Gross Domestic Product: Implicit Price Deflator
37	1	GNPDEF	Gross National Product: Implicit Price Deflator

no.series	Transf.	Mnemonic	Long Label
38	5	INDPRO	Industrial Production Index
39	5	IPBUSEQ	Industrial Production: Business Equipment
40	5	IPCONGD	Industrial Production: Consumer Goods
41	5	IPDCONGD	Industrial Production: Durable Consumer Goods
42	5	IPFINAL	Industrial Production: Final Products (Market Group)
43	5	IPMAT	Industrial Production: Materials
44	5	IPNCONGD	Industrial Production: Nondurable Consumer Goods
45	1	AWHMAN	Average Weekly Hours: Manufacturing
46	1	AWOTMAN	Average Weekly Hours: Overtime: Manufacturing
47	2	CIVPART	Civilian Participation Rate
48	5	CLF16OV	Civilian Labor Force
49	5	CE16OV	Civilian Employment
50	5	USPRIV	All Employees: Total Private Industries
51	5	USGOOD	All Employees: Goods-Producing Industries
52	5	SRVPRD	All Employees: Service-Providing Industries
53	5	UNEMPLOY	Unemployed
54	1	UEMPMEAN	Average (Mean) Duration of Unemployment
55	1	UNRATE	Civilian Unemployment Rate
56	5	HOUST	Housing Starts: Total: New Privately Owned Housing Units Started
57	1	FEDFUNDS	Effective Federal Funds Rate
58	1	TB3MS	3-Month Treasury Bill: Secondary Market Rate
59	1	GS1	1-Year Treasury Constant Maturity Rate
60	1	GS10	10-Year Treasury Constant Maturity Rate
61	1	AAA	Moody's Seasoned Aaa Corporate Bond Yield
62	1	BAA	Moody's Seasoned Baa Corporate Bond Yield
63	1	MPRIME	Bank Prime Loan Rate
64	5	BOGNONBR	Non-Borrowed Reserves of Depository Institutions
65	5	TRARR	Board of Governors Total Reserves, Adjusted for Changes in Reserve
66	5	BOGAMBSL	Board of Governors Monetary Base, Adjusted for Changes in Reserve
67	5	M1SL	M1 Money Stock
68	5	M2MSL	M2 Minus
69	5	M2SL	M2 Money Stock
70	5	BUSLOANS	Commercial and Industrial Loans at All Commercial Banks
71	5	CONSUMER	Consumer (Individual) Loans at All Commercial Banks
72	5	LOANINV	Total Loans and Investments at All Commercial Banks
73	5	REALLN	Real Estate Loans at All Commercial Banks
74	5	TOTALSL	Total Consumer Credit Outstanding
75	5	CPIAUCSL	Consumer Price Index For All Urban Consumers: All Items
76	5	CPIULFSL	Consumer Price Index for All Urban Consumers: All Items Less Food
77	5	CPILEGSL	Consumer Price Index for All Urban Consumers: All Items Less Energy

no.series	Transf.	Mnemonic	Long Label
78	5	CPILFESL	Consumer Price Index for All Urban Consumers: All Items Less Food & Energy
79	5	CPIENGSL	Consumer Price Index for All Urban Consumers: Energy
80	5	CPIUFDSL	Consumer Price Index for All Urban Consumers: Food
81	5	PPICPE	Producer Price Index Finished Goods: Capital Equipment
82	5	PPICRM	Producer Price Index: Crude Materials for Further Processing
83	5	PPIFCG	Producer Price Index: Finished Consumer Goods
84	5	PPIFGS	Producer Price Index: Finished Goods
85	5	OILPRICE	Spot Oil Price: West Texas Intermediate
86	5	USSHRPRCF	US Dow Jones Industrials Share Price Index (EP) NADJ
87	5	US500STK	US Standard & Poor's Index if 500 Common Stocks
88	5	USI62...F	US Share Price Index NADJ
89	5	USNOIDN.D	US Manufacturers New Orders for Non Defense Capital Goods (BCI 27)
90	5	USCNORCGD	US New Orders of Consumer Goods & Materials (BCI 8) CONA
91	1	USNAPMNO	US ISM Manufacturers Survey: New Orders Index SADJ
92	5	USVACTOTO	US Index of Help Wanted Advertising VOLA
93	5	USCYLEAD	US The Conference Board Leading Economic Indicators Index SADJ
94	5	USECRIWLH	US Economic Cycle Research Institute Weekly Leading Index
95	1	GS10-FEDFUNDS	
96	1	GS1-FEDFUNDS	
97	1	BAA-FEDFUNDS	
98	5	GEXPND/GDPDEF	Government Current Expenditures/ GDP deflator
99	5	GRECPT/GDPDEF	Government Current Receipts/ GDP deflator
100	2	GDEF	Governnent Real Expend-Real Receipts
101	5	GCEC1	Real Government Consumption Expenditures & Gross Investment, 1 Decimal
102	1		Fernald's TFP growth CU adjusted
103	1		Fernald's TFP growth
104	5		DOW JOONES/GDP DEFL
105	5		S&P/GDP DEFL
106	1		Fernald's TFP growth - Investment
107	1		Fernald's TFP growth - Consumption
108	1		Fernald's TFP growth CU - Investment
109	1		Fernald's TFP growth CU - Consumption
110	1		Personal Finance Current
111	1		Personal Finance Expected
112	1		Business Condition 12 Months
113	1		Business Condition 5 Years
114	1		Buying Conditions
115	1		Consumer's sentiment: Current Index
116	1		Consumer's sentiment: Expected Index
117	4		Per-capita hours worked (HOANBS/Civilian Polulation 16 and over)

References

[1] Bai, J., and S. Ng (2002). Determining the number of factors in approximate factor models, Econometrica 70, 191-221.

[2] Beaudry, P. and F. Portier (2006). Stock Prices, News, and Economic Fluctuations American Economic Review, 96(4): 1293-1307.

[3] Bernanke, B. S., J. Boivin and P. Eliasz (2005). Measuring Monetary Policy: A Factor Augmented Autoregressive (FAVAR) Approach, The Quarterly Journal of Economics 120, 387-422.

[4] Blanchard, O. J., and D. Quah (1989). The Dynamic Effects of Aggregate Demand and Supply Disturbances. American Economic Review, 79(4): 655-73

[5] Chari, V.V., Kehoe, P.J. and E.R. McGrattan (2008). Are structural VARs with long-run restrictions useful in developing business cycle theory? Journal of Monetary Economics 55: 1337-1352.

[6] Christiano, L., M. Eichenbaum and R. Vigfusson (2003), "What Happens After a Technology Shock?", NBER Working Papers 9819.

[7] Fernandez-Villaverde J., J.F. Rubio-Ramirez, T.J. Sargent and M.W. Watson (2007). ABCs (and Ds) of Understanding VARs. American Economic Review, American Economic Association, 97(3):1021-1026.

[8] Forni, M. and L. Gambetti (2010). Fiscal Foresight and the Effects of Government Spending, CEPR DP series no. 7840.

[9] Forni, Gambetti, L. and L. Sala (2010). News shocks do not drive the business cycle. Mimeo

[10] Forni, M., D. Giannone, M. Lippi and L. Reichlin (2009). Opening the Black Box: Structural Factor Models with Large Cross-Sections, Econometric Theory 25, 1319-1347.

[11] Forni M. and L., Reichlin (1996). Dynamic Common Factors in Large Cross-Sections, Empirical Economics 21, 27-42.

[12] Galí, J. (1999), "Technology, Employment and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?", *American Economic Review*, 89(1): 249-271.

[13] Gambetti, L. (2010). Fiscal Policy, Foresight and the Trade Balance in the U.S. UFAE and IAE Working Papers 852.10.

[14] Gelper, S. and C. Croux (2007), Multivariate out-of-sample tests for Granger causality, Computational Statistics & Data Analysis 51, 3319-3329.

[15] Giannone, D., and L. Reichlin (2006). Does Information Help Recovering Structural Shocks from Past Observations? Journal of the European Economic Association, 4(2-3), 455465.

[16] Hansen, L.P., and T.J. Sargent (1991). Two problems in interpreting vector autoregressions. In Rational Expectations Econometrics, L.P. Hansen and T.J. Sargent, eds. Boulder: Westview, pp.77-119.

[17] Harvey, D. I., Leybourne, S. J. and P. Newbold, (1998) Tests for Forecast Encompassing, Journal of Business & Economic Statistics 16, 254-259.

[18] Leeper, E.M., Walker, T.B. and S.S. Yang (2008). Fiscal Foresight: Analytics and Econometrics, NBER Working Paper No. 14028.

[19] Leeper, E.M., Walker, T.B. (2009). Information Flows and News Driven Business Cycles, mimeo Indiana University

[20] Lippi, M. and L. Reichlin (1993). The Dynamic Effects of Aggregate Demand and Supply Disturbances: Comment, American Economic Review 83, 644-652.

[21] Lippi, M. and L. Reichlin (1994). VAR analysis, non fundamental representation, Blaschke matrices, Journal of Econometrics 63, 307-325.

[22] Onatski, A. (2010). Determining the Number of Factors from Empirical Distribution of Eigenvalues, The Review of Economics and Statistics 92, 1004-1016.

[23] Rozanov, Yu. (1967). Stationary Random processes. San Francisco: Holden Day

[24] Schmitt-Grohe, S and M., Uribe (2008). What's News in Business Cycles, NBER WP 14215

[25] Sims, C. A. (1980). Macroeconomics and Reality. Econometrica, 48(1), 148.

[26] Stock, J.H. and M.W. Watson (2002a). Macroeconomic Forecasting Using Diffusion Indexes, Journal of Business and Economic Statistics 20, 147-162.

[27] Stock, J.H. and M.W. Watson (2002b). Forecasting Using Principal Components from a Large Number of Predictors, Journal of the American Statistical Association 97, 1167-1179.

[28] Yang, S.S. (2008). Quantifying tax effects under policy foresight. Journal of Monetary Economics, 52(8):1557-1568.

Tables

	A			B		
j	z_t^*	w_t^1	w_t^2	z_t^*	w_t^1	w_t^2
1	0.000	–	–	0.005	–	–
2	–	0.480	–	–	0.055	–
3		0.405	0.475	–	0.113	0.977
4	–	0.620	0.375	–	0.091	0.452
5	–	0.125	0.250	–	0.115	0.581
6	–	0.105	0.500	–	0.142	0.641
7	–	0.125	0.545	–	0.126	0.186
8	–	0.285	0.785	–	0.027	0.197
9	–	0.125	0.705	–	–	0.216
10	–	0.085	0.450	–	–	0.207
11	–	0.050	0.660	–	–	0.148
12	–	–	0.355	–	–	0.186
13	–	–	0.395	–	–	0.239
14	–	–	0.560	–	–	0.279
15	–	–	0.720	–	–	0.337

Table 1: p-values A: Test for informational sufficiency B: Test for structuralness of the technology shock.

Figures

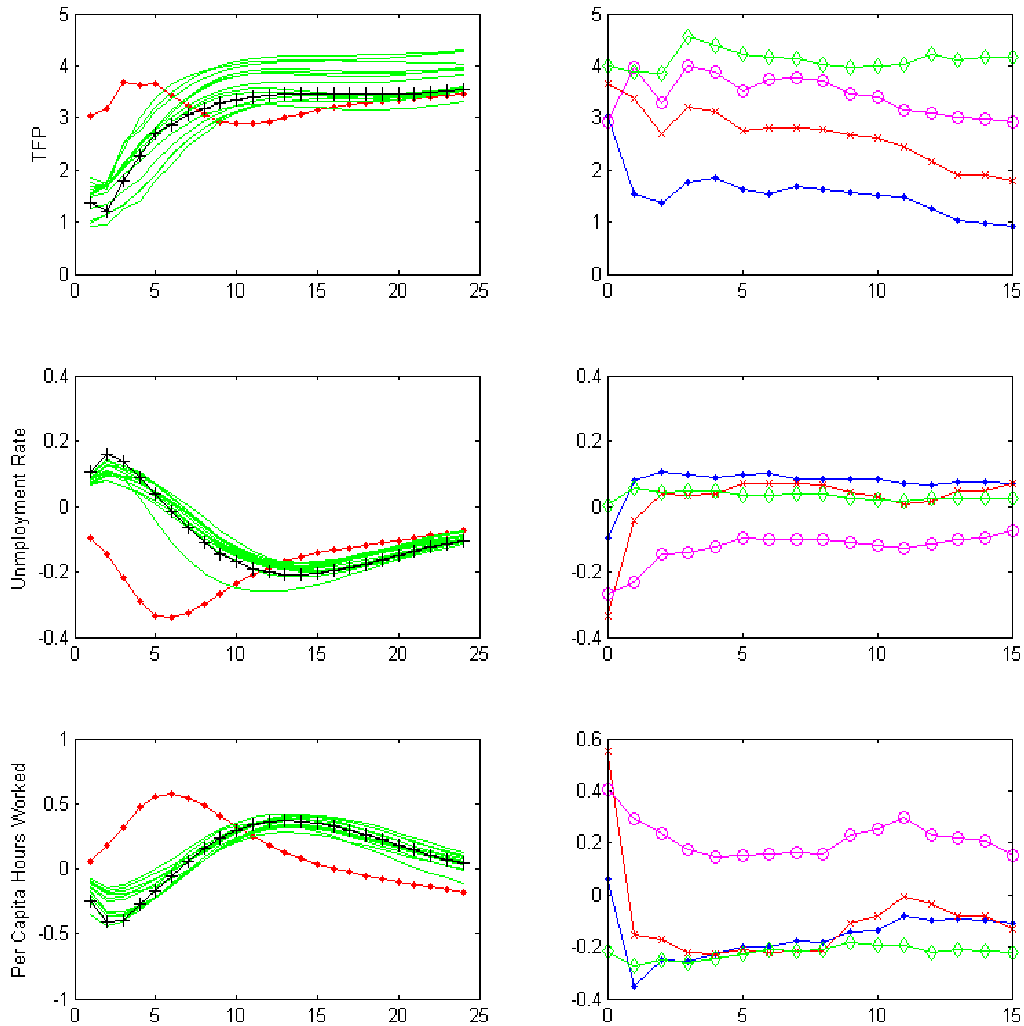

Figure 1: Impulse response functions

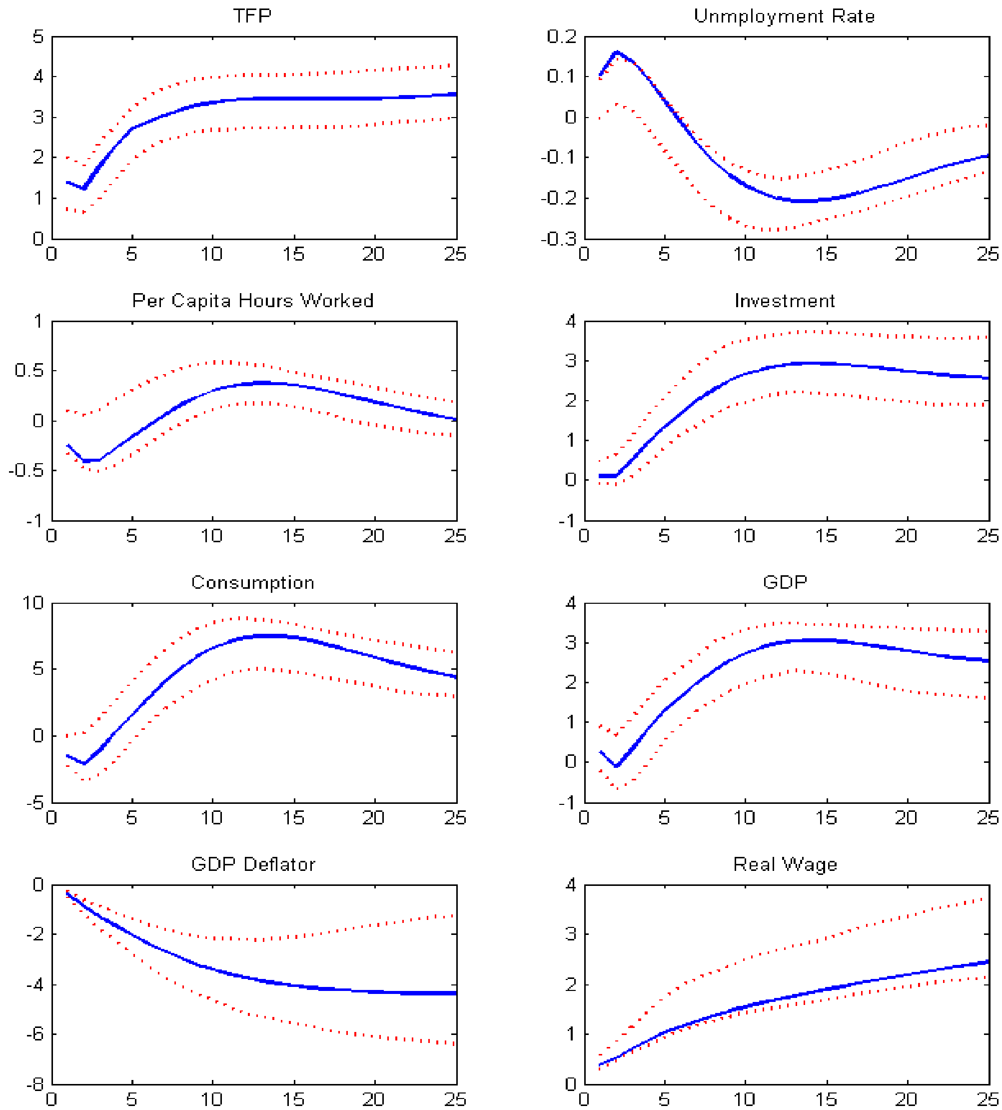

Figure 2: Impulse response functions.