Structural VARs and Non-invertible Macroeconomic Models

Mario Forni*
Università di Modena e Reggio Emilia,
CEPR and RECent

Luca Gambetti†
Collegio Carlo Alberto,
Università di Torino and BGSE

Luca Sala‡
Università Bocconi,
IGIER and Baffi Carefin

Abstract

We resume the line of research pioneered by Sims and Zha (2006) and make two novel contributions. First, we provide a formal treatment of partial fundamentalness, i.e. the idea that a structural VAR can recover, either exactly or with good approximation, a single shock or a subset of shocks, even when the underlying model is non-fundamental. In particular, we extend the measure of partial fundamentalness proposed by Sims and Zha to the finite-order case and study the implications of partial fundamentalness for impulse-response and variance-decomposition analysis. Second, we present an application where we validate a theory of news shocks and find it to be in line with the empirical evidence.

JEL classification: C32, E32.

Keywords: invertibility, non-fundamentalness, news shocks, DSGE model validation, structural VAR.

*Contact: Dipartimento di Economia Politica, Università di Modena e Reggio Emilia, via Berengario 51, 41100, Modena, Italy. Tel. +39 0592056851; e-mail: mario.forni@unimore.it. The financial support from FAR 2014, Università di Modena e Reggio Emilia is gratefully acknowledged.

†Corresponding author. Contact: Collegio Carlo Alberto, Piazza Vincenzo Arbarello, 8, 10122 Torino, Italy. Luca Gambetti acknowledges the financial support of the Spanish Ministry of Economy and Competitiveness through grant ECO2015-67602-P and through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2015-0563), and the Barcelona Graduate School Research Network.

‡Contact: Department of Economics, Università Bocconi, via Roentgen 1, 20136, Milan, Italy. Tel. +39 0258363062; e-mail: luca.sala@unibocconi.it.

1 Introduction

In recent years several works have challenged the validity of structural VARs as a tool for estimating and studying the transmission mechanisms of structural shocks. Actually, in many of the DSGE models recently studied in the literature, the solution is incompatible with a VAR representation. The basic problem has become known as "non-fundamentalness" or "non-invertibility" of the moving average representation implied by the DSGE. When such representation is not invertible, a VAR representation in terms of all of the structural shocks does not exist (Lippi and Reichlin, 1993).[1] In an important paper, Fernandez-Villaverde, Rubio-Ramirez, Sargent and Watson, 2007 derive a condition for the validity of VAR methods, related to the state-space representation of the economy. The condition, known as the "Poor Man's Condition" (PMC henceforth), implies fundamentalness of the corresponding moving average representation[2] and the possibility of recovering all of the structural shocks from a VAR.

Non-fundamentalness is best seen as an informational deficiency problem: the variables in the VAR do not convey enough information to recover the structural shocks. Informational deficiency is endemic in two relevant situations. First, when the number of shocks in the theoretical model is larger than the number of variables included in the VAR. This case is particularly interesting since the use of VAR models with few variables is widespread in the literature but modern DSGE models feature a relatively large number of shocks (see Smets and Wouters, 2007). Second, when the DSGE features the so-called "anticipated" shocks, that is, shocks with a delayed effects on some variables. Anticipated shocks have recently become a standard feature in DSGE models since they can match important empirical facts. Examples of anticipated shocks are fiscal shocks (Leeper, Walker and Yang, 2013, Mertens and Ravn, 2010, Forni and Gambetti, 2016), news shocks to TFP (Sims, 2012, Schmitt-Grohe and Uribe, 2012) and risk shocks (Christiano, Motto and Rostagno, 2014). In presence of anticipated shocks, VAR deficiency stems from the fact that, being the effects of the shock delayed, the current value of the series does not convey information about the current shocks.[3]

When one looks at the VAR literature of the last twenty years it emerges clearly that the

[1]See also Hansen and Sargent, 1991 and Lippi and Reichlin, 1994a, 1994b. A partial list of recent papers about non-fundamentlness includes Giannone, Reichlin and Sala, 2006, Giannone and Reichlin, 2006, Ravenna, 2007, Yang, 2008, Forni, Giannone, Lippi and Reichlin, 2009, Sims, 2012, Leeper, Walker and Yang, 2013, Giacomini, 2013, , Forni, Gambetti and Sala, 2014, Forni and Gambetti, 2014, Beaudry and Portier, 2015, Chen, Choi and Escanciano, 2015, Forni, Gambetti, Lippi and Sala, 2017a, 2017b, Canova and Sahneh, 2017.

[2]See Franchi and Paruolo, 2015, about the converse implication.

[3]Non-fundamentlness also occurs when agents do not observe one or more shocks, but only noisy signals on these shocks. A partial list of references on "noise" models includes Lorenzoni, 2009, Barsky and Sims, 2012, Blanchard et al., 2013, Forni et al., 2017a, 2017b, Chaurour and Jurado, 2016.

common practice is to identify just one shock or a subset of shocks of interest (the so-called *partial identification*). Partial identification characterizes for instance the extensive literature on monetary policy shocks, government spending shocks and technology shocks. As far as partial identification is concerned, the relevant issue is not to establish whether the VAR is fundamental or not, but whether it can correctly estimate the shock of interest and its impulse response functions.

Sims and Zha (2006) investigates the connection between DSGE and VAR and shows a very important result: some of the shocks, more specifically the monetary policy shock, can be recovered with good approximation using a SVAR model even when the number of shocks in the economy is larger than the number of variables in the VAR. This means that it is possible to recover a subset of shocks even when the underlying MA representation of the variables included in the VAR is non-fundamental. In this sense the non-fundamentalness problem is not as serious as was initially thought. Sims, 2012, reaches a similar conclusion showing, by means of Monte Carlo simulations based on a DSGE model, that a VAR may perform well in recovering the impulse response functions, even if the structural model is not invertible. Papers in the same line of research are Fève and Jidoud, 2012, Beaudry and Portier, 2014, and Beaudry *et al.*, 2015.

Our paper resumes this line of research and makes two contributions. First, we provide a formal treatment of those ideas, more specifically the idea that a VAR can recover with good approximation a single shock or a subset of shocks even when the underlying model is non-fundamental. Second, we present a new application related to news shocks.

We argue that fundamentalness or invertibility is unnecessarily restrictive and the appropriate condition is a shock-specific condition, that we call *informational sufficiency*. The concept of informational sufficiency is simple: a VAR is sufficient for a shock if such shock is a linear combination of the current and past values of the variables included in the VAR. The relation with fundamentalness is straightforward: fundamentalness holds if and only if the VAR is informationally sufficient for all of the structural shocks. As we will show below, partial sufficiency, far from being a statistical curiosity, holds true in economically relevant contexts such as a standard monetary policy model.

Moreover we show that, in order for partial identification to be successful, sufficiency has not to hold exactly. What is required is the parameters of the "true" model to be close to the sufficiency region. The reason is that the impulse response coefficients implied by the VAR are continuous functions of the true coefficients. Thus when the true coefficients are on the border between the sufficiency and the deficiency regions, small movements toward the interior of the deficiency region imply small changes of the impulse response functions.

A very simple and effective way to measure deficiency is to take the unexplained variance of the population orthogonal projection of the shock of interest, call it u_{it}, onto the VAR residuals. This measure, which we call δ_i, is the same measure used in Sims and Zha (2006), and can easily be computed for any calibrated or estimated DSGE model. It takes on values between zero and one: $\delta_i = 0$ implies sufficient information, so that the VAR specification is good; $\delta_i = 1$ means no information, so that the VAR specification is dramatically bad. If δ_i is close to zero the VAR is approximately sufficient and performs fairly well, as shown in the examples of Section 3 as well as the empirical application.

We show that, when $\delta_i = 0$ and the i-th structural shock is correctly identified, then the corresponding impulse response functions are consistently estimated irrespective of global fundamentalness. In addition, we show that if the VAR is informationally sufficient for one shock of interest, but is not sufficient for all of the structural shocks, then standard variance decomposition is downward biased for short-run horizons. By contrast, the variance decomposition obtained by integrating the spectral densities over a frequency band is not affected by this bias. We also introduce the finite-order deficiency measure $\delta_i(K)$, which is an extension of the deficiency measure to the case of finite-order VARs, highlighting the role of lag truncation, and study its properties and its relation with δ_i.

The deficiency measure can be regarded as a generalization of the PMC, in the sense that it essentially reduces to PMC when we require zero deficiency for all i. Unlike PMC, δ_i (a) is shock-specific, (b) provides information about the "degree" of non-fundamentalness, (c) can be computed even if the number of variables in the VAR is smaller than the number of structural shocks.

In the empirical application we illustrate the relevance of this generalization. We focus on a New-Keynesian DSGE, similar to the one used by Blanchard, Lorenzoni and L'Huillier (2013, BLL henceforth). The model features seven exogenous sources of fluctuations, including a news permanent shock and a surprise temporary shock in technology. A few parameters are calibrated by using BLL estimates, whereas the remaining ones are estimated by means of Bayesian techniques. We then use such parameter values to compute δ_i, all i, for ten VAR specifications, nine of them including less than seven variables. We find that fundamentalness (and hence PMC) does not hold in any of the specifications (including the seven-variable square specification). Despite this result, most specifications are sufficient, or almost sufficient, for at least one shock and most shocks can be perfectly, or almost perfectly, recovered from at least one specification. The news shock, in particular, exhibits $\delta < 0.01$ in four specifications.

VAR deficiency measures can be used for DSGE validation purposes as follows. Consider a calibrated or estimated DSGE model and consider a VAR specification including a subset of variables included in the DSGE. Compute δ_i under the null that the model is true. If δ_i is close to zero, the theoretical VAR and the DSGE are compatible. Therefore we can assess whether the impulse response functions derived from DSGE are in line with those obtained from the VAR estimated with actual data.[4] In the empirical application, we use this procedure to test for the transmission mechanisms of the news shocks in the DSGE model described above. We choose a relatively small VAR (a four-variable VAR) with $\delta_i = 0.006$. We identify the news shock consistently with the model and find that the theoretical response functions lie within the VAR confidence bands for almost all variables.[5]

A recent paper by Chahrour and Jurado, 2017, argues that what really matters for structural macroeconomic analysis is not invertibility, i.e. that the structural shocks can be obtained from the present and the past of the observables, but "recoverability" , i.e. that the structural shocks can be obtained from the present, past and future of the observables.

It is well-known that we can in principle use dynamic transfomations involving future values of the VAR residuals (the so-called Blaschke factors) and therefore of the observables, to perform the identification step (Lippi and Reichlin, 1993). In fact, this has been done for specific models (Mertens and Ravn, 2010, Forni et al, 2017a, 2017b).[6] However, the identification step is already a difficult task in standard structural VAR analysis; if we consider dynamic transformations instead of just static transformations, identification becomes much more demanding. While in special cases economic theory may be able to provide the correct identification restrictions, in general it is not.[7] Moreover, when the structural model has more shocks than variables we do not have recoverability; despite this, as anticipated above, standard structural VAR methods can provide good results for specific shocks. For both reasons, we think that in most cases approximate partial sufficiency, rather than recoverability, is what is needed in practice.[8]

The remainder of the paper is organized as follows. Section 2 presents and discusses the main ideas. Section 3 presents formal time series results and more technical material. Section 4 shows a couple of examples and the related simulations. Section 5 discusses an application of the theoretical framework to DSGE models and validates a theory of news shocks. Section

[4] See Canova, 2002, 2007, and Canova and Paustian, 2011, for DSGE validation through VARs.

[5] Structural VAR techniques can usefully be employed even when we do not have a well-defined theoretical model of reference. In this case, however, δ_i cannot be computed. In order to verify whether a given VAR is informationally sufficient for the shocks of interest, we can use the testing procedure suggested in Forni and Gambetti, 2014.

[6] The empirical illustration of Chahrour and Jurado, 2017, is essentially a bivariate version of Forni et. al. (2017a).

[7] It is worth mentioning that, even if we are able to identify and estimate a non-fundamental model, the end-of-sample values of the structural shocks cannot be recovered, since of course, at the end of the sample, future values of the observables are not available.

[8] Of course, it is also possible to define a concept of partial recoverability and partial approximate recoverability along the lines of the present paper. This is left for future research.

6 concludes. The Appendix reports the details of the DSGE model used in Section 5.

2 Theory: the main ideas

In the present Section we illustrate our main theoretical arguments; more technical details and formal results are postponed to the following Section.

2.1 The macroeconomy

We assume that the macroeconomic variables in the model have a Moving Average (MA) representation, possibly derived from a state-space representation, where the structural shocks propagate through linear impulse response functions. As a consequence, our results hold true for any theoretical model (not necessarily DSGE) which can be cast in MA form.[9]

Let us focus on the section of the macroeconomic model corresponding to the variables used in the VAR, i.e. the entries of the n-dimensional vector x_t. We assume that x_t has a non singular spectral-density matrix and can be represented as

$$x_t = \sum_{k=0}^{\infty} A_k u_{t-k} = A(L)u_t, \tag{1}$$

where $u_t = (u_{1,t} \ \cdots \ u_{q,t})'$ is a q-dimensional white noise vector of mutually orthogonal macroeconomic shocks, and $A(L) = \sum_{k=0}^{\infty} A_k L^k$ is an $n \times q$ matrix of rational impulse response functions.

Representation (1) is "structural" in the sense that the vector u_t includes all of the exogenous shocks driving x_t. However, we do not assume that all of the shocks in u_t have a structural economic interpretation: some of them may be statistical residuals, devoid of economic interest, arising from measurement errors. This enables us to evaluate VAR deficiency – and therefore its performance – with respect to the shocks of interest when actual variables are affected by measurement errors.[10]

We do not assume that the number of variables n is equal to the number of shocks q. In other words, representation (1) is not necessarily square. In particular, it can be "short", with more shocks than variables ($q > n$). Short systems are relevant for applied work for two reasons. First, several empirical analyses are based on small-scale VARs, with just two or three variables. If the economy is driven by a larger number of shocks, the above MA

[9]As observed above, a relevant problem for the validation of DSGE models through VARs is that most DSGE models are non-linear. The theoretical impulse response functions result from a linear approximation which in principle may be inaccurate. Since our focus is on informational deficiency, we do not address this issue here.

[10]Giannone, Reichlin and Sala, 2006, study the the impact of measurement errors in VAR estimation.

system will be short. Second, most variables are in practice affected by measurement errors and/or small shocks of limited economic interest, so that, even if we have as many variables as major structural shocks (or even more variables than shocks) the system may be short because measurement errors are included in the vector u_t.[11]

2.2 VAR deficiency

Given model (1), we want to evaluate whether a VAR in x_t conveys the information needed to recover the shocks of interest and the corresponding impulse response functions. In practice, the impulse response function obtained with a VAR are affected by estimation errors arising from the finiteness of the sample size. Such finiteness requires specification of low-order VARs, which might be affected by truncation bias. Since our main focus here is on the non-fundamentalness bias, we abstract from estimation errors and lag truncation, by replacing finite-sample, finite-order VARs with population orthogonal projections on infinite-dimensional information spaces.

Within this conceptual framework, the structural VAR procedure consists in performing the orthogonal decomposition

$$x_t = P(x_t|H^x_{t-1}) + \epsilon_t, \tag{2}$$

where H^x_{t-1} is the closed linear space $\overline{\text{span}}(x_{1,t-k}, \ldots, x_{n,t-k}, k = 1, \ldots, \infty)$ and ϵ_t is the Wold innovation, and capturing u_{it} (by means of suitable identification restrictions) as a linear combination of the entries of ϵ_t.[12]

The best possible result is the projection of u_{it} onto the entries of ϵ_t. We therefore consider the projection equation

$$u_{it} = M\epsilon_t + e_{it}. \tag{3}$$

We define informational deficiency as the fraction of unexplained variance in the above regression:[13]

$$\delta_i = \sigma^2_{e_i}/\sigma^2_{u_i}.$$

This measure has first been proposed by Sims and Zha, 2006. It can be computed from the theoretical model (1), that is from $A(L)$, according to the formulas provided in Section 3.

We say that x_t is informationally sufficient for u_{it} if and only if u_{it} is an exact linear combination of the entries of ϵ_t, i.e. $u_{it} = M\epsilon_t$ and $\delta_i = 0$. In Section 3 we show that projecting u_{it} onto the entries of ϵ_t is equivalent to projecting it onto the VAR information

[11] "Tall" systems, i.e. systems with more variables than shocks, are also interesting from a theoretical point of view, but are unlikely to occur in practice, because of measurement errors. We shall not consider them further in the present work.

[12] If the Wold representation of x_t is invertible, equation (2) can be written in the standard (possibly infinite) VAR form. We use here the projection notation in order to include the case of non-invertibility.

[13] For simplicity of notation we do not make explicit the dependence of δ_i on x_t.

set H_t^x. It follows that we have sufficiency for u_{it} if and only if $u_{it} \in H_t^x$; in words, if the structural shock is a linear combination of the present and past values of the variables included in the VAR. By the very definition of informational sufficiency, if the vector M is obtained (by imposing suitable identification restrictions derived from economic theory), then an informationally sufficient VAR for u_{it} delivers u_{it} without error, whereas an informationally deficient VAR for u_{it} produces an approximation, whose error is measured by δ_i.

However, the goal of many VAR exercises is to recover the impulse response functions. In Section 3 we show that a VAR, which is sufficient for u_{it} (but possibly deficient for the other structural shocks) — and correctly identifies it as $M\epsilon_t$ — delivers the correct impulse response functions. In addition, the impulse response coefficients implied by the VAR are continuous functions of the true ones. Hence, when the true coefficients change, moving from the border of the sufficiency region towards the interior of the deficiency region (so that δ_i becomes positive) δ_i provides a meaningful indication about the performance of the theoretical VAR in approximating the impulse response functions.

Of course, in practical situations we do not have theoretical VARs, but only finite-sample, finite-order VARs, which are affected by estimation and lag-truncation errors. Such real world VARs provide estimates whose asymptotic bias is measured by δ_i, as the sample size and the truncation lag increase at appropriate rates.

Although the truncation bias is not our focus here, the deficiency measure can be naturally extended to the case of finite-order VARs. Deficiency of a VAR(K) with respect to u_{it}, denoted by $\delta_i(K)$, is given by the fraction of unexplained variance of the population projection of u_{it} onto the truncated VAR information space spanned by present and past values of the x's, until the maximum lag K. By its very definition, the sequence $\delta_i(K)$ is nonincreasing in K. The finite-order deficiency $\delta_i(K)$ measures the total asymptotic bias due to deficiency *plus* lag truncation, provided that the VAR residuals are serially uncorrelated (see Section 3 for details).

2.3 Beyond the ABCs (and Ds)

How does our deficiency measure relate to fundamentalness and existing fundamentalness conditions? We have fundamentalness when all of the shocks in u_t belong to the econometrician information set, i.e. $u_{it} \in H_t^x$, for all i. Hence we have fundamentalness if, and only if, the VAR is informationally sufficient for all shocks, that is $\delta_i = 0$ for all i. Sufficient information is then a notion of "partial fundamentalness", a straightforward shock-specific generalization of the fundamentalness concept.

Note that short systems are never fundamental (see Section 3, Proposition 1). This is quite intuitive: if we have just n variables we cannot estimate consistently more than n orthogonal shocks. By contrast, square systems can be either fundamental or not, depending on the roots of the determinant of $A(L)$: we have fundamentalness if there are no roots smaller than 1 in modulus.

Fernandez-Villaverde et al., 2007, propose a fundamentalness condition, the PMC, based on the state-space representation of the economy. Consider the following linear equilibrium representation of a DSGE model

$$s_t = As_{t-1} + Bu_t \tag{4}$$
$$x_t = Cs_{t-1} + Du_t \tag{5}$$

where s_t is an m-dimensional vector of stationary "state" variables, x_t is the n-dimensional vector of variables observed by the econometrician, and u_t is the q-dimensional vector of shocks with $q \leq m$. A, B, C and D are conformable matrices of parameters, B has a left inverse B^{-1} such that $B^{-1}B = I_q$. Representation (4)-(5) can always be cast in form (1). If the matrix D is square (this implies that the system is square, $q = n$) and invertible, the matrix $A(L)$ appearing in representation (1) can be written as

$$A(L) = DB^{-1}\left[I - (A - BD^{-1}C)L\right](I - AL)^{-1}B.$$

The PMC is that all the eigenvalues of the matrix $A - BD^{-1}C$ are strictly less than one in modulus. It is easily seen that, if the PMC holds, the MA representation of x_t is invertible and u_t can be represented as a linear combination of the present and past values of x_t. In other words, the PMC implies fundamentalness, i.e. sufficient information for all shocks.[14] Hence, if the system is square, and the PMC holds, then $\delta_i = 0$ for all i.

Summing up, the deficiency measure can be regarded as a generalization of the PMC (as well as other existing fundamentalness conditions), both because it is shock specific and because it provides information about the "degree" of non-fundamentalness. In addition, the PMC is only defined for square systems, whereas deficiency does not require $q = n$.

This generalization is very important for applied work. As anticipated above, short systems are never fundamental. Adding variables to the VAR in such a way to get a square system does not necessarily solve the problem. Let us stress again that non-fundamental structural MA representations, far from being an oddity, are quite common in modern macroeconomic models. We show below that, despite non-fundamentalness, we may have small deficiency for a single shock of interest even when $q > n$. This implies a relevant fact, which we believe is important to stress: small-scale VARs can in principle be successfully employed even when the number of shocks driving the economy is large.[15]

[14] About the converse implication see Franchi and Paruolo (2015).

[15] This does not mean that small scale VAR are *always* successful. For instance Sims (1992) and Caldara and Herbst (2016) show cases where the omission of important variables leads to misleading results.

Sims (2012) makes the point that a VAR may perform reasonably well even if fundamentalness does not hold. With his words, non-fundamentalness "should not be thought of as an "either/or" proposition – even if the model has a non-invertibility, the wedge between VAR innovations and economic shocks may be small, and structural VARs may nonetheless perform reliably" (Sims, 2012, abstract).[16] The deficiency measure δ_i formalizes and clarifies the notion of "wedge" discussed and analyzed with Monte Carlo experiments in Sims' paper.

3 Theory: formal results

This Section presents formal time series results, the related proofs and some technical discussion. It can be skipped by the non-technical reader without loosing the general message of the paper and without threatening the comprehension of the following sections.

3.1 Fundamentalness: definition and standard results

Let us begin by reviewing the definition of fundamentalness and a few related results.

Definition 1 (Fundamentalness). *We say that u_t is fundamental for x_t, and the MA representation $x_t = A(L)u_t$ is fundamental, if and only if $u_{it} \in H_t^x$, $i = 1, \dots, q$, where $H_t^x = \overline{\text{span}}(x_{1,t-k}, \dots, x_{n,t-k}, k = 0, \dots, \infty)$.*

Now, consider the theoretical projection equation of x_t on its past history, i.e. equation (2). The Wold representation of x_t is

$$x_t = B(L)\epsilon_t, \tag{6}$$

where $B(0) = I_n$ and ϵ_t is white noise.

The following result is standard in time series theory.

Proposition 1. *u_t is fundamental for x_t if and only if there exist a nonsingular matrix Q such that $u_t = Q\epsilon_t$.*

Proof. If $u_t = Q\epsilon_t$, fundamentalness of u_t is implied by equation (2). On the other hand, assuming fundamentalness of u_t, we have $[A(L) - A(0)]u_t \in H_{t-1}^x$. Moreover, equation (1) and serial uncorrelation of u_t imply $u_t \perp H_{t-1}^x$. Hence we have $\mathrm{P}(x_t|H_{t-1}^x) = [A(L) - A(0)]u_t$ and $A(0)u_t$ is the residual of the projection, i.e. $A(0)u_t = \epsilon_t$. Non-singularity of $A(0)$ is implied by the assumption that x_t has a non-singular spectral-density matrix, which in turn implies that ϵ_t has a non-singular variance-covariance matrix. It follows that $u_t = Q\epsilon_t$ with $Q = A(0)^{-1}$. QED

[16] Beaudry and Portier (2015) and this work provide further evidence about this fact.

It is apparent from the above condition that fundamentalness cannot hold if the system is short. In this case, a matrix Q satisfying Proposition 1 does not exist, since for any $q \times n$ matrix Q, with $q > n$, the entries of $Q\epsilon_t$ are linearly dependent, whereas the entries of u_t are mutually orthogonal.

By contrast, in the square case $n = q$ fundamentalness clearly holds if the impulse response function matrix $A(L)$ is invertible. In this case, we can write the structural model (1) as

$$A(L)^{-1}x_t = u_t,$$

so that the condition defining fundamentalness is fulfilled.

Fundamentalness of u_t for x_t is equivalent to the following condition (see e.g. Rozanov, 1967, Ch. 2).

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

When $A(L)$ is a square matrix the above condition reduces to the well known condition that the determinant of $A(z)$ has no roots smaller than one in modulus. Fundamentalness is therefore slightly different from invertibility, since invertibility rules out also roots with modulus equal to 1. Hence invertibility implies fundamentalness, whereas the converse is not true.[17]

3.2 Sufficient information

For simplicity we shall assume here that the target of VAR estimation is the single shock of interest u_{it} (along with the corresponding impulse response functions). The generalization to any subvector v_t of u_t, including $s \leq q$ shocks, is straightforward.

Let us go back to the VAR representation of x_t, i.e. equation (2), and the projection equation (3). The following proposition says that the structural VAR strategy, i.e. approximating u_{it} by means of the VAR residuals, is optimal in the sense that it provides the best linear approximation, given the VAR information set.

Proposition 2 (Optimality of the structural VAR procedure). *The projection of u_{it} onto the entries of ϵ_t, i.e. $M\epsilon_t$, is equal to the projection of u_{it} onto H_t^x.*

Proof. From (3) it is seen that H_t^x is the direct sum of the two orthogonal spaces H_{t-1}^x and $\text{span}(\epsilon_{jt}, j = 1, \ldots, n)$. Hence $P(u_{it}|H_t^x) = P(u_{it}|\epsilon_{jt}, j = 1, \ldots, n) + P(u_{it}|H_{t-1}^x)$. Since u_{it} is orthogonal to the past values of the x's, the latter projection is zero. Hence $P(u_{it}|H_t^x) = P(u_{it}|\epsilon_{jt}, j = 1, \ldots, n) = M\epsilon_t$. QED

[17] The unit root case is economically interesting in that, if $x_t = \Delta X_t$ and the determinant of $A(z)$ vanishes for $z = 1$, then the entries of X_t are cointegrated. Non-invertibility implies that x_t does not have a VAR representation and VAR estimates do not have good properties. However this problem can be solved by estimating an ECM or a VAR in the levels X_t.

Proposition 2 motivates the following definitions.

Definition 2 (VAR deficiency and sufficient information). *The informational deficiency of x_t (and the related VAR information set H_t^x) with respect to u_{it} is*

$$\delta_i = \mathrm{var}[u_{it} - P(u_{it}|H_t^x)]/\sigma_{u_i}^2 = \sigma_{e_i}^2/\sigma_{u_i}^2.$$

We say that x_t is informationally sufficient for u_{it} if and only if $\delta_i = 0$, i.e. $u_{it} \in H_t^x$, or, equivalently, $u_{it} = M\epsilon_t$.

As an immediate consequence of Definitions 1 and 2, we have the following result.

Proposition 3. u_t is fundamental for x_t if and only if x_t is informationally sufficient for u_{it}, $i = 1, \ldots, q$, i.e. $\delta_i = 0$ for all i.

3.3 Partial sufficiency: IRFs

Until now we have focused on the conditions under which the VAR is able to recover the shock u_{it}. However, the ultimate goal of the VAR validation procedure are the impulse response functions, rather than the shock itself. Hence a basic question in our framework is the following. Let the VAR be sufficient for u_{it} and the identification restrictions be correct. Are the impulse response functions obtained from the VAR equal to the theoretical ones? Standard results in VAR identification theory guarantee a positive answer when the MA representation is (globally) fundamental. But what happens if this is not the case?

The structural VAR procedure consists in inverting the VAR representation to estimate the Wold representation (6) and choosing identification restrictions delivering an "identification matrix", say Q.[18] Having the matrix Q, the structural shocks are obtained as $v_t = Q\epsilon_t$ and the corresponding impulse response functions as $A^*(L) = B(L)Q^{-1}$. If the matrix Q is the "correct" one, then identification is successful, i.e. $v_t = u_t$ and $A^*(L) = A(L)$. On the contrary, if the identification scheme is wrong and therefore Q is not the correct identification matrix, then $v_t \neq u_t$ and $A^*(L) \neq A(L)$.

If we are only interested in identifying the single shock u_{it}, the only one relevant feature of the matrix Q is that its i-th row, say Q_i, is equal to the vector M. Hence we have several correct identification matrices, all fulfilling the equality $Q_i = M$.

In order to state the main result of this subsection we need a formal definition of a correct identification scheme.

[18]For instance, with the recursive identification scheme we impose that $A(0)$ is lower triangular and that $\Sigma_u = I$. These restrictions deliver the identification matrix Q such that Q is lower triangular and $QQ' = \Sigma_\epsilon$, Σ_ϵ being the variance-covariance matrix of the Wold residuals.

Definition 4 (Correct identification). *An identification matrix is a nonsingular $n \times n$ matrix Q such that $Q\Sigma_\epsilon Q'$ is diagonal, i.e. the entries of $v_t = Q\epsilon_t$ are orthogonal. An identification matrix is correct for u_{it} if and only if $v_t = Q\epsilon_t$ is such that $v_{it} = u_{it}$, i.e., denoting with Q_i the i-th row of Q, $Q_i = M$. An identification scheme is a set of restrictions on $A(L)$ and Σ_u defining an identification matrix. An identification scheme is correct for u_{it} if and only if the related identification matrix is correct for u_{it}.*

A correct identification matrix is fully characterized by the vector M and the orthogonality conditions for the shocks. Knowing the identification vector M is equivalent to knowing a correct identification matrix: M can be obtained from Q by the relation $M = Q_i$; conversely, a correct identification matrix Q can be obtained from M by imposing $Q_i = M$ and completing the matrix in such a way that $Q\Sigma_\epsilon Q'$ is diagonal.[19]

If Q is a correct identification matrix, we can write the impulse response function representation derived from the VAR as

$$x_t = A^*(L)v_t = A_i^*(L)v_{it} + A_{-i}^*(L)z_t = A_i^*(L)u_{it} + A_{-i}^*(L)z_t, \tag{7}$$

where $A_i^*(L)$ is the i-th column of $A^*(L)$, $A_{-i}^*(L)$ is the $n \times n-1$ matrix obtained by eliminating the i-th column from $A^*(L)$ and $z_t = (v_{1t} \ \cdots v_{i-1,t} \ v_{i+1,t} \ \cdots v_{nt})'$. $A_i^*(L)$ is the vector of impulse response functions derived from the VAR – let us say the "empirical" impulse response functions, even if, of course, we are speaking of the population VAR (with infinite sample size and infinite number of lags).

Now let $A_i(L)$ be the i-th column of $A(L)$ and $A_{-i}(L)$ be the $n \times q-1$ matrix obtained by eliminating the i-th column from $A(L)$. The structural MA representation can then be written as

$$x_t = A(L)u_t = A_i(L)u_{it} + A_{-i}(L)w_t, \tag{8}$$

where $w_t = (u_{1t} \ \cdots u_{i-1,t} \ u_{i+1,t} \ \cdots u_{qt})'$. $A_i(L)$ is the vector of the "true" impulse response functions.

Proposition 4. *Let x_t and the related VAR be informationally sufficient, and the identification scheme be correct for u_{it}. Then the empirical impulse response functions are equal to the true impulse response functions, i.e. $A_i^*(L) = A_i(L)$.*

Proof. Let us first observe that the entries of v_t are orthogonal at all leads and lags, since $v_t = Q\epsilon_t$ is a vector white noise and Q is an identification matrix. It follows that u_{it} is

[19] Starting from M, we can find a correct matrix Q as follows. Let D be any $(n-1) \times n$ matrix whose rows are mutually orthogonal and span the null space of of MC^{-1}, C being the Cholesky matrix of Σ_ϵ, i.e. the matrix such that $CC' = \Sigma_\epsilon$. The remaining $n-1$ rows of Q are the rows of the matrix DC.

orthogonal to the entries of z_t at all leads and lags. Moreover, by the assumptions of model (1), u_{it} is also orthogonal to u_{jt}, $j \neq i$, and therefore to the entries of w_t, at all leads and lags. From (7) and (8) we get $A_i^*(L)u_{it} + A_{-i}^*(L)z_t = A_i(L)u_{it} + A_{-i}(L)w_t$. Projecting both sides onto $u_{i,t-k}, k \geq 0$ we get $A_i^*(L)u_{it} = A_i(L)u_{it}$, which implies the result. QED

Let us remark that the equality result in Proposition 4 translates into a consistency result for real-world, finite-sample VARs, provided that the parameters of the population VAR are estimated consistently, and the truncation lag increases with the sample size, following a consistent information criterion.

Of course, if the identification scheme is not correct, then $v_{it} \neq u_{it}$ and, in general, $A_i^*(L)$ will be different from $A_i(L)$. This is true for any structural VAR model. From this point of view, partial fundamentalness does not make any difference with respect to global fundamentalness.

Proposition 4 refers to exact identification schemes, i.e. schemes which uniquely define the shock of interest and the related impulse response functions. Schemes based on sign restrictions are not exact, in that they only define a set of identification matrices (possibly including both correct and incorrect matrices) and therefore a set of possible IRFs. If we impose inequality restrictions delivering a set of IRFs, the relevant question is therefore whether such set includes the true IRFs or not. A thorough treatment of this issue is beyond the scope of the present paper. Notice however that, if the sign restrictions deliver a set of identification matrices including a correct identification matrix, then by Proposition 4 the set of impulse response functions obtained with the sign restrictions will necessarily include the true ones.

3.4 Partial sufficiency: variance decomposition

Under sufficiency the "empirical" IRFs of u_{it} are equal to the true IRFs. But what about the forecast error variance decomposition when sufficiency does not hold for all shocks? It turns out that it is downward biased, as shown by the following argument.

It is well known that $H_t^\epsilon \subseteq H_t^u$ and, if u_t is non-fundamental for x_t, $H_t^\epsilon \subset H_t^u$. As a consequence, in the non-fundamental case, the prediction error of v_t is larger than the one of u_t[20]. Precisely, for any horizon $s \geq 0$, we have

$$\text{var}\,[P(x_{i,t+s}|H_{t-1}^u) - x_{i,t+s}] \leq \text{var}\,[P(x_{i,t+s}|H_{t-1}^v) - x_{i,t+s}],$$

and the inequality is strict at least for $s = 0$ if u_t is non-fundamental for x_t. Hence if x_t is informationally sufficient for u_{it}, but not for all shocks, then the total forecast error variance is overestimated by the VAR model at short horizons. On the other hand, Proposition 4 implies that the impulse response functions of u_{it}, and therefore the variance of the forecast errors, is

[20]See, among others, Sims, 2012 and Fernández-Villaverde and Rubio-Ramírez, 2006

estimated consistently. Putting things together, the fraction of total variance accounted for by u_{it}, derived from the VAR, is downward biased, since the numerator is unbiased, whereas the denominator is upward biased.[21]

An alternative, though not equivalent, variance decomposition, which is not affected by this bias, is obtained by using integrals of the spectral densities over suitable frequency bands (see e.g. Forni, Gambetti and Sala, 2014).

Let $A_{i,j}(L)$ and $A^*_{h,j}(L)$, be the j-th elements of the matrices $A_i(L)$, $A^*_h(L)$, respectively. As is well known, the variance of the component of x_{jt} which is attributable to v_{ht} can be computed as $\sigma^2_{v_h} \int_0^\pi A_{-i,j}(e^{-i\theta})A_{-i,j}(e^{i\theta})d\theta/\pi$. If we are interested for instance in the variance of waves of business cycle periodicity, say between 8 and 32 quarters, the corresponding angular frequencies (with quarterly data) are $\theta_1 = \pi/4$ and $\theta_2 = \pi/16$ and the corresponding variance is $\sigma^2_{v_h} \int_{\theta_1}^{\theta^2} A_{-i,j}(e^{-i\theta})A_{-i,j}(e^{i\theta})d\theta/\pi$. On the other hand, the total "cyclical" variance of x_{jt} is given by $\int_{\theta_1}^{\theta^2} S_{x_j}(\theta)d\theta/\pi$, where $S_{x_j}(\theta)$ denotes the spectral density of x_{jt}. Hence the contribution of v_{ht} to the cyclical variance of x_{jt} is given by

$$\frac{\sigma^2_{v_h} \int_{\theta_1}^{\theta^2} A^*_{h,j}(e^{-i\theta})A^*_{h,j}(e^{i\theta})d\theta}{\int_{\theta_1}^{\theta^2} S_{x_j}(\theta)d\theta}. \tag{9}$$

Similarly, the contribution of u_{it} is given by

$$\frac{\sigma^2_{u_i} \int_{\theta_1}^{\theta^2} A_{i,j}(e^{-i\theta})A_{i,j}(e^{i\theta})d\theta}{\int_{\theta_1}^{\theta^2} S_{x_j}(\theta)d\theta}. \tag{10}$$

Under the assumptions of Proposition 4, the numerators are equal, so that the ratios are equal. Therefore the spectral variance decomposition proposed here is preferable to the standard forecast error variance decomposition in that it is not biased in the case of partial fundamentalness. It turns out that the spectral variance decomposition evaluated at $\theta_1 = 0$ and $\theta_2 = \pi$ is equivalent to the forecast error variance decomposition at horizon infinity.

3.5 Finite-order deficiency

In this subsection we consider a quite natural extension of the deficiency measure to the case of finite-order VARs. Let us denote the VAR(K) information set as $H^x_t(K) = \text{span}(x_{j,t-k}, j = 1, \dots, n, k = 0, \dots, K)$ and consider the orthogonal decompositions

$$\begin{aligned} x_t &= P(x_t|H^x_{t-1}(K)) + \epsilon^K_t \\ u_{it} &= M^K \epsilon^K_t + e^K_{it}. \end{aligned}$$

[21] If $x_{it} = \Delta X_{it}$ a similar result holds for the decomposition of the forecast error variance of the level $X_{i,t+s}$. This explains the large estimation error, at horizon 0, reported in Table (1) for the variable r_t.

Proposition 2 still holds for the finite-order VAR.

Proposition 2′ *The projection of u_{it} onto the entries of ϵ_t^K, i.e. $M^K\epsilon_t^K$, is equal to the projection of u_{it} onto $H_t^x(K)$.*

The proof is the same as that of Proposition 2, with ϵ_t^K in place of ϵ_t and $H_\tau^x(K)$ in place of H_τ^x, $\tau = t, t-1$. The VAR(K) deficiency can then be defined as

$$\delta_i(K) = \mathrm{var}[u_{it} - P(u_{it}|H_t^x(K))]/\sigma_{u_i}^2 = \sigma_{e_i^K}^2/\sigma_{u_i}^2. \tag{11}$$

Correspondingly, we can say that $H_t^x(K)$ *is informationally sufficient for* u_{it} if and only if $\delta_i(K) = 0$, i.e. $u_{it} \in H_t^x(K)$, or, equivalently, $u_{it} = M^K\epsilon_t^K$.

As K increases, the spaces $H_t^x(K)$ are nested, so that the sequence $\delta_i(K)$ is non-increasing in K and $\delta_i \leq \delta_i(K)$ for any K. The difference $\delta_i(K) - \delta_i$ provides information about the additional effect of lag truncation on estimation of the shock u_{it}.

By inverting the finite-order VAR, we get the representation

$$x_t = B^K(L)\epsilon_t^K,$$

and, by imposing the identification constraints, we get the "shocks" $v_t^K = Q\epsilon_t^K$ and the corresponding impulse response functions

$$A^K(L) = B^K(L)Q^{-1}. \tag{12}$$

Proposition 4, as it is, does not hold for the finite-order VAR. It might be the case that the K-order VAR is sufficient for u_{it}, but the corresponding impulse response functions are biased. This is because the VAR residuals in ϵ_t^K might be serially correlated, owing to lag truncation. To get unbiasedness of these response functions we need to reinforce the assumptions of Proposition 4 with the additional condition that ϵ_t^K is a vector white noise (i.e. x_t has an exact VAR(K) representation).

Proposition 4′. *Let a VAR(K) be informationally sufficient, and the identification matrix be correct for u_{it}. Assume further that the VAR residual ϵ_t^K is a vector white noise. Then the empirical impulse response functions are equal to the true impulse response functions, i.e. there is a column h of $A^K(L)$ such that $A_h^K(L) = A_i(L)$.*

We omit the proof, which is essentially the same as that of Proposition 4. By comparing $\delta_i(K)$ with δ_i we can get an idea of the additional bias due to lag truncation.

3.6 Computing deficiency

In this subsection we present an algorithm to compute VAR deficiency from our reference macroeconomic model (1). Let us write the projection equation of u_{it} onto $H_t^x(K)$ as

$$u_{it} = P(u_{it}|H_t^x(K)) + e_{it}^K = Fy_t + e_{it}^K,$$

where $y_t = (x_t' \ \cdots \ x_{t-K}')'$ and $F = E(u_{it}y_t')\Sigma_y^{-1}$, Σ_y being the variance covariance matrix of y_t. From (11) and the above equation we get

$$\delta_i(K) = 1 - F\Sigma_y F'/\sigma_{u_i}^2 = 1 - E(u_{it}y_t')\Sigma_y^{-1}E(u_{it}y_t')/\sigma_{u_i}^2.$$

Using (8) it is easily seen that $E(u_{it}x_t') = A_i(0)'$ and $E(u_{it}x_{t-k}') = 0$ for all $k > 0$, so that $E(u_{it}y_t') = (A_i(0)' \ 0 \ \cdots \ 0)$.

Hence

$$\delta_i(K) = 1 - A_i(0)'GA_i(0)/\sigma_{u_i}^2, \tag{13}$$

where G is the $n \times n$ upper-left submatrix of Σ_y^{-1}, Σ_y being

$$\Sigma_y = \begin{pmatrix} \Gamma_0 & \Gamma_1 & \cdots & \Gamma_{K-1} \\ \Gamma_{-1} & \Gamma_0 & \cdots & \Gamma_{K-2} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{-K+1} & \Gamma_{-K+2} & \cdots & \Gamma_0 \end{pmatrix}, \tag{14}$$

where $\Gamma_k = E(x_t x_{t-k}')$, $k = 0, \ldots, K-1$. The covariance matrices of x_t can be computed from the MA representation (1) by using the formulas

$$\Gamma_k = \sum_{j=0}^{\infty} A_{j+k}\Sigma_u A_j'. \tag{15}$$

As for δ_i, it can simply be approximated with any desired precision by using a suitably large K.

A step by step algorithm to compute $\delta_i(K)$ and δ_i is then the following.

1. Compute the covariance matrices Γ_k, $k = 0, \ldots, K-1$, according to (15). If $A(L)$ is infinite, approximate Γ_k, $k = 0, \ldots, K-1$ by truncating the sum appearing in (15) at some lag (e.g. lag 100).

2. Compute the matrix Σ_y according to (14).

3. Compute the inverse Σ_y^{-1} and its $n \times n$ upper-left submatrix G.

4. Compute $\delta_i(K)$ according to formula (13).

5. Approximate δ_i with $\delta_i(K)$ with a suitably large K (in the application below we set $\delta_i \approx \delta_i(1000)$.[22]

[22] An exact formula for δ_i is

$$\delta_i = 1 - \sigma_{u_i}^2 A_i(0)'\Sigma_\epsilon^{-1}A_i(0).$$

3.7 Near sufficiency and the empirical IRFs

What happens if δ_i is not exactly zero, but close to zero? The following continuity argument shows that the empirical IRFs are close to the true IRFs.

We have assumed that the true IRFs are rational functions, i.e. functions of the form $A_{ij}(L) = N_{ij}(L)/D_{ij}(L)$ with maximum lag l_1 at the numerator and l_2 at the denominator, so that each entry is characterized (under the normalization $D_{ij}(0) = 1$) by $l = l_1 + l_2 - 1$ parameters. Moreover, without loss of generality, let us normalize the structural shocks in such a way that $\Sigma_u = I_q$. Hence the parameters of the macroeconomic model can be assembled into the finite-dimensional vector $\theta \in \mathbb{R}^m$, with $m = nq(l+1)$. Now, the point is that the empirical IRFs are continuous functions of θ.

To see this, consider the companion form of the VAR(K), i.e.

$$y_t = Ry_{t-1} + s_t,$$

where $y_t = (x_t' \ \cdots \ x_{t-K+1})'$, $s_t = [\epsilon_t^{K'} \ 0 \ \cdots \ 0]'$ and

$$R = \begin{pmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{K-1} & \Phi_K \\ I_n & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_n & 0 \end{pmatrix} \tag{16}$$

$$\begin{pmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_K \end{pmatrix} = \begin{pmatrix} \Gamma_1 & \Gamma_2 & \cdots & \Gamma_K \end{pmatrix} \Sigma_y^{-1}, \tag{17}$$

Σ_y being defined as in (14). The empirical impulse response functions $A^K(L)$ are given by equation (12), where $B^K(L)$ is the $n \times n$ upper-left submatrix of

$$I_s + RL + R^2L^2 + \cdots, \quad s = nK. \tag{18}$$

The matrices Γ_k, $k = 1, \ldots, K$ are convolutions of the true IRF (matrix) coefficients, as implied by (15), so that they are continuous functions of θ. Moreover R, and ultimately the empirical

---

This formula is obtained from (3) by observing that $M = E(u_{it}\epsilon_t{}')\Sigma_\epsilon^{-1}$ and noting that, by (2) and (1), $E(\epsilon_t u_{it}) = E(x_t u_{it}) = A_i(0)\sigma_{u_i}^2$. The problem with this formula is computation of Σ_ϵ. However, if the model can be written in the state-space form

$$\begin{aligned} s_t &= As_{t-1} + Bu_t \\ x_t &= Hs_t, \end{aligned}$$

the matrix Σ_ϵ can be obtained from the Wold representation

$$x_t = \left\{I_n + H(I_m - AL)^{-1}SL\right\}\epsilon_t$$

where S is the steady-state Kalman gain. Σ_ϵ is given by HPH', where P is the steady-state variance-covariance of the states.

IRFs in $A^K(L) = A_0^K + A_1^K L + A_2^K L^2 + \cdots + A_k^K L^K + \cdots$, are continuous functions of the covariances Γ_k, $k = 1, \ldots, K$, as implied by (16), (17) and (18). Hence $A_{k,j,i}^K$, the j, i entry of A_k^K, is a continuous function of θ and so is the bias $\beta_{k,j,i}^K = |A_{k,j,i}^K - A_{k,j,i}|$. Similarly, $\delta_i(K)$ is a continuous function of θ (see equation (13)), denoted as $\delta_i(K, \theta)$.

Now, let the assumptions of Proposition $4'$ be fulfilled within some sufficiency region and let $\mathcal{B}_i \subset \mathbb{R}^m$ be the boundary of such region. If $\delta_i(K, \theta)$ is small, then θ must be close to some $\bar{\theta} \in \mathcal{B}_i$ where, by Proposition $4'$, $\beta_{k,j,i}^K(\bar{\theta}) = 0$ for any j and k. By continuity, $\beta_{k,j,i}^K(\theta)$ must be small as well, for any variable and lag.

4 Examples and simulations

In this Section we show two simple examples illustrating the concepts of partial sufficiency and approximate partial sufficiency. In addition, we study the VAR implications of a fully fledged DSGE model featuring news shocks. A few simulations are performed to show the connection between the deficiency measure δ_i and the VAR performance.

4.1 Example 1: Exact partial sufficiency in a square system

Let us assume that output deviates from its potential value because of a demand shock d_t inducing temporary fluctuations, and reacts negatively to the interest rate r_t, expressed in mean deviation, with a one-period delay. Precisely, the output gap y_t is given by

$$y_t = (1 + \alpha L)d_t - \beta r_{t-1},$$

where α and β are positive. The central bank aims at stabilizing output by responding to output gap deviations, so that the interest rate follows the rule

$$r_t = \gamma y_t + v_t,$$

where v_t is a discretionary monetary policy shock and $\gamma > 0$.[23] The structural MA representation for the output gap and the interest rate is then

$$\begin{pmatrix} y_t \\ r_t \end{pmatrix} = \frac{1}{1 + \gamma \beta L} \begin{pmatrix} 1 + \alpha L & -\beta L \\ \gamma(1 + \alpha L) & 1 \end{pmatrix} \begin{pmatrix} d_t \\ v_t \end{pmatrix}. \tag{19}$$

Here the determinant of the MA matrix is $(1 + \alpha L)$, which vanishes for $L = -1/\alpha$, so that the representation is non-fundamental if $|\alpha| > 1$. From the policy rule we see that $v_t = r_t - \gamma y_t$,

[23] The example can be easily generalized to the case of a more realistic rule where the interest rate also reacts to inflation.

so that the monetary policy shock can be recovered from the present values of the variables included in the VAR, irrespective of α (of course, d_t cannot be found from the x's if $|\alpha| > 1$).

What happens when the above model is non-fundamental ($\alpha > 1$) and the econometrician tries to estimate the monetary policy shock and the related impulse-response functions? To answer this question we generated 1000 artificial data sets with 200 time observations from (19), with $\alpha = 3$, $\gamma = 0.4$, $\beta = 1$ and standard normal shocks. We then estimated for each data set a VAR with 4 lags and identified both shocks by imposing a standard Cholesky, lower triangular impact effect matrix, consistently with the model.

Figure 1 displays the true impulse response functions (red solid lines) along with the median (black dashed lines), the 5-th and the 95-th percentiles (grey area) of the distribution of the estimated impulse-response function to the demand shock d_t (first column) and the monetary policy shock v_t (second column). In the lower panels we report the distributions of the correlation coefficients between the estimated shocks and the true shocks.

The figure shows clearly that the impulse response functions are very poorly estimated for d_t, but very precisely for v_t. A similar result holds for the shocks themselves: the distribution of the correlation coefficients is very close to 1 for v_t and far from 1 for d_t.

Let us now have a look to the true and estimated variance decomposition. Table 1 shows the fraction of the forecast error variance of y_t and r_t accounted for by the monetary policy shock. The contribution of the monetary policy shock to total variance is severely underestimated on impact, slightly underestimated at horizon 1 and well estimated at longer horizons. This is in line with the theoretical discussion in Section 3.4. Table 2 reports the variance decomposition over frequency bands for the true model, equation (10), and the estimated VAR, equation (9), over three frequency bands, corresponding to periodicities below 2 years, between 2 and 8 years and beyond 8 years. The estimated decomposition is in line with the true decomposition, as argued in Section 3.4.

Table 3 shows the values of $\delta(K)$, $K = 1, 4, 1000$ for d_t and v_t. The VAR is dramatically deficient for the demand shock, consistently with Figure 1, but exhibits perfect information for the second shock, the monetary policy shock. Note that x_t must be deficient for the demand shock, since the MA representation is non-fundamental.

4.2 Example 2: Near sufficiency in a short system

Partial approximate sufficiency, far from being a statistical curiosity, is relevant in practice. As already noticed, most observed variables are likely affected by small macroeconomic shocks and/or measurement errors. Owing to these minor shocks, the applied researcher is usually faced with short systems, which are necessarily non-fundamental, but may be approximately sufficient for the shocks of interest.

As an example, consider the following news shock model, similar to the one used in Forni, Gambetti and Sala (2014). Total factor productivity, a_t, follows the slow diffusion process

$$a_t = a_{t-1} + \alpha\varepsilon_t + \varepsilon_{t-1} \tag{20}$$

where $0 \leq \alpha < 1$.

The representative consumer maximizes

$$E_t \sum_{t=0}^{\infty} \beta^t c_t,$$

where E_t denotes expectation at time t, c_t is consumption and β is a discount factor, subject to the constraint $c_t + \bar{p}_t n_{t+1} = (\bar{p}_t + a_t)n_t$, where $\bar{p}_t$ is the price of a share, n_t is the number of shares and $(\bar{p}_t + a_t)n_t$ is the total amount of resources available at time t. The equilibrium value for asset prices is given by:

$$\bar{p}_t = \sum_{j=1}^{\infty} \beta^j E_t a_{t+j}.$$

Using (20), we see that $E_t a_{t+k} = a_t + \varepsilon_t$ for all $k > 0$. Hence, $\bar{p}_t = (a_t + \varepsilon_t)\beta/(1-\beta)$ and $\Delta\bar{p}_t = b(1+\alpha)\varepsilon_t$, where $b = \beta/(1-\beta)$. Let us assume further that actual prices p_t are subject to a temporary deviation from the equilibrium, driven by the shock d_t, so that $p_t = \bar{p}_t + \gamma d_t$. In addition, let us add an orthogonal measurement error e_t to the technology variable a_t^*, observed by the econometrician. The structural MA representation of Δa_t^* and Δp_t is short, because we have three shocks and just two variables:

$$\begin{pmatrix} \Delta a_t^* \\ \Delta p_t \end{pmatrix} = \begin{pmatrix} \alpha + L & 0 & \theta(1-L) \\ b(1+\alpha) & \gamma(1-L) & 0 \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ d_t \\ e_t \end{pmatrix}. \tag{21}$$

We assume unit variance shocks, so we add a scaling factor θ to the impulse response function of e_t to control for the size of the measurement error. We set $\beta = 0.99$, $\alpha = 0.5$ and $\gamma = 20$. Moreover, we set $\theta = 0.5$, so that the measurement error is large (it explains more than 25% of the total variance of Δa_t^*).

Figure 2 shows the estimation results obtained by a Monte Carlo exercise with $T = 200$, i.i.d. unit variance Gaussian shocks and 1000 artificial data sets. The VAR is estimated with 4 lags and is identified by assuming that ε_t is the only shock affecting a_t^* in the long run, consistently with the model. The estimates of the technology shock ε_t and the related impulse response functions are fairly good, even if a small distortion is visible. On the contrary, the temporary shock to stock prices d_t and the associated responses are poorly estimated. Correspondingly, the deficiency measure is 0.03 for ε_t and 0.97 for d_t (see Table 4).

4.3 A fully fledged model: computing the δ's

Let us now consider a fully fledged economic model. The model is a New-Keynesian DSGE, similar to the one used in Blanchard, Lorenzoni and L'Huillier (2013, BLL henceforth). It features several frictions, such as internal habit formation in consumption, adjustment costs in investment, variable capital utilization, Calvo price and wage stickiness. The model also features seven exogenous sources of fluctuations; namely, a news shock and a surprise shock in technology, an investment-specific shock, a monetary policy shock, a shock to price markups, a shock to wage markups and a shock to government expenditures.

In particular we assume that the logarithm of technology follows the process

$$\begin{aligned} a_t &= a_{t-1} + \varepsilon_{t-4} + (1-L)T_t \\ T_t &= \rho T_{t-1} + v_t \end{aligned}$$

where ε_t is a news shock, which is observed by agents at time t, but will be reflected in a_t at time $t+4$. The component T_t is a temporary component driven by the surprise technology shock v_t. The news shock ε_t is the main focus of our analysis.

As for the parameters, some of them are calibrated using the posterior mean values estimated by BLL (see Table 5). The remaining ones are estimated using Bayesian techniques; estimation results are reported in Table 6. A complete description of the model and the estimation details are reported in Appendix A.

Having the parameters, we can now compute VAR deficiency for different VAR specifications. We examine ten specifications:

S1 : TFP, consumption.

S2 : TFP, investment.

S3 : TFP, consumption, hours worked, interest rate.

S4 : TFP, GDP, consumption, hours, interest rate.

S5 : TFP, GDP, consumption, hours, inflation.

S6 : TFP, GDP, investment, hours, interest rate.

S7 : TFP, GDP, investment, hours, inflation.

S8 : TFP, GDP, investment, hours, interest rate and inflation.

S9 : TFP, GDP, consumption, investment, hours, interest rate.

S10 : TFP, GDP, consumption, investment, hours, interest rate and inflation.

The informational deficiency measure is computed on the variables transformed to obtain stationarity (see Appendix A). The structural shocks cannot be fundamental for specifications 1-9, since they are short. Specification S10 is square, so that in principle we might have fundamentalness. For each specification, we compute the value of δ^K, $K = 1, 2, 4, 12, 1000$. The value of $\delta(1000)$ is taken as our approximation of δ. Even if we are primarily interested in the news shock, we compute δ for all shocks for illustrative purposes.

Table 7 shows the results. Few observations are in order. First, the structural shocks are non-fundamental for the variables in S10, since for some i we have $\delta_i > 0$. Of course the PMC is not satisfied, the largest eigenvalue being 1.13. Hence all of the specifications considered have a non-fundamental representation in the structural shocks.

Second, despite non-fundamentalness, several specifications exhibit very low deficiency for a few specific shocks. For instance, in the 4-variable specification S3, $\delta_i < 0.01$ for both the news and the surprise technology shocks (columns 1 and 2, respectively).

Third, a specification may be highly deficient for some shocks and sufficient for other shocks. For instance specification S9 is sufficient, or almost sufficient, for four shocks out of seven, but has a value of δ as high as 0.78 for the third shock, which is the price markup shock. This is far from surprising, since the inflation rate is not included in S9.

Focusing on our shock of interest (first column), we see that there are four specifications, namely S3, S4, S9 and S10, exhibiting a value of δ smaller than 0.05. Table 8 shows the values of $\delta(K)$ for the news shocks, $K = 1, 2, 4, 12, 1000$, for these four specifications. As already observed, these numbers should be interpreted as measuring the total bias due to non-fundamentalness and lag truncation and therefore can provide some guidance for the number of lags to use in the VAR. We see from the Table that including just one lag is inappropriate for all specifications. By contrast, for $K \geq 2$ the value of δ is smaller than 0.05. The most interesting specifications are S3, which is parsimonious, and S10, which enables the researcher to consider more variables.

4.4 A fully fledged model: simulations

Now let us focus on our preferred specifications S3 and S10. Chari et al., 2008, highlights that the VAR may be affected by large truncation and estimation bias. To evaluate the different sources of bias involved in VAR estimation, we generate artificial data for the variables in levels from the model and estimate the VAR on artificial data in order to see whether the VAR is able to reproduce the true impulse response functions. We use 500 Monte Carlo replications. For each artificial data set we estimate a Bayesian VAR with diffuse priors and take the average of the posterior impulse responses over 50 draws from the posterior distribution of the VAR parameters.

Following Beaudry and Portier (2015), the news shock is identified by imposing that (i) no shocks other than the technology surprise shock affect TFP on impact; (ii) surprise and news shocks are the only ones affecting TFP at a given horizon.[24] Restrictions (i) and (ii) are just identifying and, of course, are consistent with the theory (see the above equation). This identification scheme is the same as the one used in Forni, Gambetti and Sala (2014), where condition (ii) is replaced by the equivalent condition that the effect of news on TFP at the given horizon is maximized.

Figures 3 report the results for the four-variable specification S3. The blue dashed lines are the theoretical impulse response functions. The red thick solid lines are obtained with 5000 time observations and 12 lags. We take them as showing the asymptotic deficiency bias. The green dotted lines are obtained with 5000 time observations and 4 lags. We interpret them as showing the total asymptotic bias due to deficiency *and* lag truncation. The black thin solid lines are obtained with 243 observations[25] and 4 lags. They show the total bias due to deficiency, truncation and small sample estimation. The dark gray and light gray areas are the 68% and 90% posterior probability intervals obtained with 4 lags and 243 observations.

We see from the figure that (a) the deficiency bias is negligible, in that the red lines are almost identical to the blue dashed lines; (b) the truncation bias is fairly small, even if it is clearly visible in the TFP and the hours-worked panels (green-dotted versus red-solid lines); (c) the small sample bias is sizable (black-solid versus green-dotted lines). The total bias is relevant, particularly for consumption and hours worked. However, the theoretical impulse response functions lie within the narrower bands, except for the corner of the TFP response at lag 4.

Figure 4 reports results for the seven-variable specification. The conclusions are similar to the previous ones, even if in this case the the total bias is somewhat less pronounced at medium-long horizons.

5 Application: Validating a theory of news shocks

VAR analysis has been widely used in the literature for validation purposes or to discriminate between competing models. A prominent example is the technology-hours debate. The RBC model predicts that hours should increase while the New-Keynesian model predicts that hours

[24]Here we use the five year horizon (lag 20). Other choices in the range 12-28 lags produce very similar results. Longer horizons, such as the ten-year horizon, produce a larger estimation bias in the present case.

[25]This is in order to replicate the sample size in US data.

should fall in response to a contemporaneous and permanent technology shock. Gali, 1999, estimates a bivariate VAR in labor productivity and hours and identifies the technology shock as the only one shock driving labor productivity in the long run. He then checks whether hours respond positively, as implied by the RBC model, or negatively, as implied by sticky prices models.

In this Section, we perform a validation exercise in the spirit of the above literature. The reference model is the DSGE model described in the previous Section. As before, we focus on the news shock.

5.1 Choosing the VAR specification

In order for the results obtained from VAR analysis to be meaningful, it is important to check that the deficiency measure associated to the VAR specification used is close to zero. Only in that case the VAR evidence can support or reject the theory. If, on the other side, the measure is large, the VAR specification is inappropriate.

Validation should therefore be performed through the following steps.

1. Consider a calibrated/estimated DSGE model and its linear equilibrium representation as in (4)-(5) or (1). Select the shock of interest u_{it}.

2. Explore different VAR specifications (including variables represented in the model) by computing δ_i. Choose a specification whose δ_i is smaller than a pre-specified threshold (e.g. 0.05).

3. Estimate the VAR and verify whether the theoretical impulse response function lie within the VAR confidence bands. If they do, the model is validated. If they do not, there is something wrong either in the values of the parameters or in the model itself.

We have already explored in the previous Section the values of δ_i for the news shock for several VAR specifications. As we have seen, specifications S3 and S10 are good and therefore we use them for our validation exercise.

We use US quarterly data on Total Factor Productivity, real per-capita GDP, real per-capita consumption of nondurables and services, real per-capita investment, per-capita hours worked, the federal funds rate and the inflation rate. The sample span is 1954Q3-2015Q2. Further details about the data and their treatment are provided in Appendix B.

All VAR estimates are Bayesian estimates with diffuse prior. Data are taken in levels. The number of lags is 4. Point estimates of the impulse response functions are obtained as averages of the posterior distribution across 500 draws. Identification of the news shock is obtained as explained in subsection 4.4.

Let us anticipate that neither S3 nor S10 reject the model. Before considering S3 and S10 in more detail, however, let us see what happens if, ignoring deficiency, we use instead S1 or S2, which have a large δ_i. Figure 5 shows the VAR results for the news shock obtained with real data. The top panels refer to S1 while the bottom panels refer to S2. The black solid lines are the empirical impulse response functions to unit-variance shocks. The blue dashed lines are the theoretical impulse response functions to unit-variance shocks. According to S1, the theoretical model should be rejected, because the effects of news shocks are largely overstated. According to S2, the theoretical model should again be rejected, but for the opposite reason: the effect of news is understated. Both conclusions are misleading.

5.2 Validating the theory

Let us come now to specifications S3 and S10. Figure 6 plots the results for the four-variable specification. As before, the solid black line is the average of the posterior distribution, the dark and light gray areas represent the 68% and 90% probability intervals, respectively and the blue dashed line is the impulse response function of the economic model. For both specifications, the theoretical impulse response functions have the correct signs and lie within the 90% bands for all variables. However, the sudden reaction of TFP at lag 4 is clearly at odds with the VAR estimates, where the empirical impulse response function increases gradually, according to a typical S-shape. Moreover, its long-run effect is much larger than the theoretical one. The reaction of hours worked is anticipated with respect to the empirical one, whereas the converse is true for the interest rate.

Figure 7 plots the results for the seven-variable specification. The above results are confirmed. The long-run reaction of TFP is now very close to the lower bound of the 68% posterior probability interval. The reaction of GDP is almost identical to the empirical one. Notice however that, were the long-run effect on TFP be larger, the effect on GDP and consumption would likely be overstated by the model.

Let us now focus on investment and inflation, two variables not present in specification S3. The signs of the impulse response functions are correct. However, the reaction of investment is understated by the model in the short run. Moreover, the effect on inflation predicted by the model at horizon 30 is zero, whereas the empirical one is positive.

Our overall evaluation is that the model performs reasonably well but clearly understates the long-run effects of technology news on TFP.

6 Conclusions

Structural VARs are well and alive. Existing validity conditions for structural VAR analysis are unnecessarily restrictive. A single shock of interest can be recovered, either exactly or with good approximation, even if the true model is non-fundamental and the Poor Man's Condition does not hold, and even if the number of variables is smaller than the number of structural shocks driving the macroeconomy.

For the validity of a VAR, the relevant question is not whether we have fundamentalness or not, but whether the VAR conveys enough information to recover the shock of interest and the related impulse response functions. VAR deficiency for a given shock can be measured by δ_i, i.e. the fraction of unexplained variance of the linear projection of the i-th shock onto the VAR information set.

If a VAR is used to validate a macroeconomic model, δ_i can be used as a guidance for the choice of the VAR specification. If δ_i is large, the VAR specification in inappropriate; if δ_i is small, the VAR can provide a reliable assessment of the models.

For DSGE models including news or foresight shocks, non-fundamentalness is endemic. Such models are often regarded as incompatible with VARs, in that a VAR representation in the structural shocks does not exist. Hence we illustrate our ideas by conducting a validation exercise with a news shock DSGE model. We show that a few VAR specifications exhibit a deficiency very close to zero, despite non-fundamentalness. We find that the DSGE model performs reasonably well in fitting the impulse response functions derived from US data.

References

[1] Barsky, R. and E. Sims, 2012. Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence, American Economic Review 102, 1343-77.

[2] Beaudry, P., Fève, P., Guay, A. and Portier, F., 2015. When is non-fundamentalness in VAR a real problem? An application to news shocks. NBER Working Paper No. 21466.

[3] Beaudry, P. and Portier, F., 2014. News-driven business cycles: insights and challenges. Journal of Economic Literature, 52(4), pp. 993-1074.

[4] Blanchard O.J., G. Lorenzoni and J.P. L'Huillier, 2013. News, Noise, and Fluctuations: An Empirical Exploration, American Economic Review, vol. 103(7), pp. 3045-70.

[5] Canova, F., 2002. Validating monetary DSGE models through VARs, CEPR Discussion Papers 34-42.

[6] Canova, F., 2007. Methods for applied macroeconomic research. Princeton: Princeton University Press.

[7] Canova, F. and Paustian, M., 2011. Business cycle measurement with some theory, Journal of Monetary Economics 58, pp. 345361.

[8] Canova, F. and Sahneh, M. H., 2017. Are small-scale SVARs useful for business cycle analysis? Revisiting non-fundamentalness, Journal of the European Economic Association, jvx032.

[9] 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, pp. 1337-1352.

[10] Chahrour, R. and K. Jurado, 2016. News or Noise? The Missing Link, Boston College Working Papers in Economics 917, Boston College Department of Economics, revised 02 Nov 2017.

[11] Chahrour, R. and K. Jurado, 2017. Recoverability, Boston College Working Papers in Economics 935, Boston College Department of Economics.

[12] Chen, B., Choi, J. and Escanciano, J. C., 2017. Testing for fundamental vector moving average representations, Quantitative Economics, 8(1), pp. 149-180.

[13] Christiano, L. J., M. Eichenbaum, Vigfusson, R., 2007. Assessing structural VARs. NBER Macroeconomics Annual 2006, vol. 21, pp. 1-106.

[14] Christiano, L. J., Motto, R., and Rostagno, M. 2014. Risk shocks, American Economic Review 104, pp. 27-65.

[15] Fernandez-Villaverde, J. and Rubio-Ramirez, J. F., 2006. Economic and VAR shocks: what can go wrong?, Journal of the European Economic Association, vol. 4(2-3), pp. 466-474.

[16] Fernandez-Villaverde J., Rubio-Ramirez, J. F., Sargent, T. J. and Watson, M. W., 2007. ABCs (and Ds) of understanding VARs, American Economic Review 97, pp. 1021-1026.

[17] Fève, P. and Jidoud, A., 2012. Identifying news shocks from SVARs. Journal of Macroeconomics, 34(4), pp. 919932.

[18] Forni, M. Gambetti, L., 2014. Sufficient information in structural VARs, Journal of Monetary Economics, vol. 66(C), pp. 124-136.

[19] Forni, M. Gambetti, L., 2016. Government spending shocks in open economy VARs. Journal of International Economics, vol. 99, pp. 68-84.

[20] Forni, M., Gambetti, L., Lippi, M. and Sala, L., 2017a. Noisy news in business cycles. American Economic Journal: Macroeconomics, vol. 9(4), pp. 122-152.

[21] Forni, M., Gambetti, L., Lippi, M. and Sala, L., 2017b. Noise bubbles. The Economic Journal, 127, pp. 1940-1976.

[22] Forni, M., Giannone, D., Lippi, M. and Reichlin, L., 2009. Opening the black box: structural factor models with large cross-sections. Econometric Theory 25, pp. 1319-1347.

[23] Forni, Gambetti, L., Sala, L., 2014. No news in business cycles. The Economic Journal 124, pp. 1168-1191.

[24] Franchi, M. and Paruolo, P., 2015. Minimality of state space solutions of DSGE models and existence conditions for their VAR representation, Computational Economics 46, pp. 613-626.

[25] Giacomini, R., 2013. The relationship between DSGE and VAR models. Advances in Econometrics, vol. 32: VAR models in macroeconomics: new developments and applications: essays in honor of Christopher A. Sims.

[26] Giannone, D., Reichlin, L., 2006. Does information help recovering structural shocks from past observations? Journal of the European Economic Association 4, 455-465.

[27] Giannone, D., Reichlin, L., and Sala, L., 2006. VARs, common factors and the empirical validation of equilibrium business cycle models. Journal of Econometrics 132, 257-279.

[28] Haario, H., Saksman, E., and Tamminen, J., 2001. An adaptive Metropolis algorithm, Bernoulli 7, 223-242.

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

[30] Leeper, E.M., Walker, T.B., Yang, S.S., 2013. Fiscal foresight and information flows. Econometrica 81, pp. 1115-1145.

[31] Lippi, M., Reichlin, L., 1993. The dynamic effects of aggregate demand and supply disturbances: comment. American Economic Review 83, pp.644-652.

[32] Lippi, M., Reichlin, L., 1994a. Diffusion of technical change and the decomposition of output into trend and cycle, Review of Economic Studies, vol. 61(1), pp. 19-30.

[33] Lippi, M., Reichlin, L., 1994b. VAR analysis, non-fundamental representation, Blaschke matrices. Journal of Econometrics 63, pp. 307-325.

[34] Lorenzoni, G., 2009. A Theory of Demand Shocks, American Economic Review, 99, 2050-84.

[35] Mertens, K., Ravn, M.O., 2010. Measuring the impact of fiscal policy in the face of anticipation: a structural VAR approach. Economic Journal 120, pp. 393-413.

[36] Ravenna, F., 2007. Vector autoregressions and reduced form representations of DSGE models. Journal of Monetary Economics 54, pp. 2048-2064.

[37] Rozanov, Yu., 1967. Stationary random processes. San Francisco: Holden Day.

[38] Sala, L., 2015. DSGE models in the frequency domain. Journal of Applied Econometrics 30, pp. 219-240.

[39] Schmitt-Grohe, S., Uribe, M., 2012. What's news in business cycles. Econometrica, 80(6), pp. 2733-2764.

[40] Sims, C. A., 1989. Models and their uses. American Journal of Agricultural Economics, 71, pp. 489-494.

[41] Sims, C. A., Zha, T., 2006. Does monetary policy generate recessions? Macroeconomic Dynamics, 10(02), pp. 231-272.

[42] Sims, E.R., 2012. News, non-invertibility, and structural VARs. Advances in Econometrics 28, pp. 81-136

[43] Soccorsi, S., 2016. Measuring non-fundamentalness for structural VARs, Journal of Economic Dynamics and Control 71, pp. 86-101.

[44] Yang, S.S., 2008. Quantifying tax effects under policy foresight. Journal of Monetary Economics 52, pp. 1557-1568.

Appendix A: The DSGE model

The model follows closely Blanchard, L'Huillier and Lorenzoni (2013). The preferences of the representative household are given by the utility function:

$$E_t \left[\sum_{t=0}^{\infty} \beta^t \left(\log\left(C_t - hC_{t-1}\right) - \frac{1}{1+\varsigma} \int_0^1 N_{jt}^{1+\varsigma} dj \right) \right],$$

C_t is consumption, the term hC_{t-1} captures internal habit formation, and N_{jt} is the supply of specialized labor of type j. The household budget constraint is

$$P_t C_t + P_t I_t + T_t + B_t + P_t C(U_t)\bar{K}_{t-1} = R_{t-1}B_{t-1} + Y_t + \int_0^1 W_{jt}N_{jt}dj + R_t^k K_t,$$

where P_t is the price level, T_t is a lump sum tax, B_t are holdings of one period bonds, R_t is the one period nominal interest rate, Y_t are aggregate profits, W_{jt} is the wage of specialized labor of type j, N_{jt}. R_t^k is the capital rental rate.

Households choose consumption, bond holdings, capital utilization, and investment each period so as to maximize their expected utility subject to the budget constraint and a standard no-Ponzi condition. Nominal bonds are in zero net supply, so market clearing in the bonds market requires $B_t = 0$.

The capital stock $\bar{K}_t$ is owned and rented by the representative household and the capital accumulation equation is

$$\bar{K}_t = (1-\delta)\bar{K}_{t-1} + D_t \left[1 - G(I_t/I_{t-1})\right] I_t,$$

where δ is the depreciation rate, D_t is a stochastic investment-specific technology parameter, and G is a quadratic adjustment cost in investment

$$G(I_t/I_{t-1}) = \chi(I_t/I_{t-1} - \Gamma)^2/2,$$

where Γ is the long-run gross growth rate of TFP. The model features variable capacity utilization: the capital services supplied by the capital stock $\bar{K}_{t-1}$ are $K_t = U_t\bar{K}_{t-1}$, where U_t is the degree of capital utilization and the cost of capacity utilization, in terms of current production, is $C(U_t)\bar{K}_{t-1}$, where $C(U_t) = U_t^{1+\zeta}/(1+\zeta)$.

The investment-specific shock $d_t = \log D_t$ follows the stochastic process:

$$d_t = \rho_d d_{t-1} + \varepsilon_{dt}.$$

ε_{dt} and all the variables denoted with ε from now on are i.i.d. shocks.

Consumption and investment are in terms of a final good which is produced by competitive final good producers using the CES production function

$$Y_t = \left(\int_0^1 Y_{jt}^{\frac{1}{1+\mu_{pt}}} dj \right)^{1+\mu_{pt}}$$

which employs a continuum of intermediate inputs. Y_{jt} is the quantity of input j employed and μ_{pt} captures a time-varying elasticity of substitution across goods, where $\log(1+\mu_{pt}) = \log(1+\mu_p) + m_{pt}$ and m_{pt} follows the process $m_{pt} = \rho_p m_{pt-1} + \varepsilon_{pt} - \psi_p \varepsilon_{pt-1}$.

The production function for intermediate good j is

$$Y_{jt} = (K_{jt})^{\alpha} (A_t L_{jt})^{1-\alpha},$$

where K_{jt} and L_{jt} are, respectively, capital and labor services employed. The technology parameter $a_t = \log(A_t)$ follows the process

$$\begin{aligned} a_t &= a_{t-1} + \varepsilon_{t-4} + (1-L)T_t \\ T_t &= \rho T_{t-1} + v_t, \end{aligned}$$

where ε_t is a *news* shock that is known to agents at time t, but will be reflected in a_t at time $t+4$ and the part T_t is a persistent, but temporary, surprise technology shock.

BLL (2013) treat explicitly the constant term in TFP growth by letting $A_t = \Gamma^t e^{a_t}$, but calibrate $\Gamma = 1$.

Intermediate good prices are sticky with price adjustment as in Calvo, 1983. Each period intermediate good firm j can freely set the nominal price P_{jt} with probability $1-\theta_p$ and with probability θ_p is forced to keep it equal to P_{jt-1}. These events are purely idiosyncratic, so θ_p is also the fraction of firms adjusting prices each period.

Labor services are supplied to intermediate good producers by competitive labor agencies that combine specialized labor of types in $[0,1]$ using the technology

$$N_t = \left[\int_0^1 N_{jt}^{\frac{1}{1+\mu_{wt}}} dj \right]^{1+\mu_{wt}},$$

where $\log(1+\mu_{wt}) = \log(1+\mu_w) + m_{wt}$ and m_{wt} follows the process $m_{wt} = \rho_w m_{wt-1} + \varepsilon_{wt} - \psi_w \varepsilon_{wt-1}$.

The presence of differentiated labor introduces monopolistic competition in wage setting as in Erceg, Henderson and Levin, 2000. Specialized labor wages are also sticky and set by the household. For each type of labor j, the household can freely set the price W_{jt} with probability $1-\theta_w$ and has to keep it equal to W_{jt-1} with probability θ_w.

Market clearing in the final good market requires

$$C_t + I_t + C(U_t)\bar{K}_{t-1} + G_t = Y_t.$$

Market clearing in the market for labor services requires $\int L_{jt} dj = N_t$.

Government spending is set as a fraction of output and the ratio of government spending to output is $G_t/Y_t = \psi + g_t$, where g_t follows the stochastic process

$$g_t = \rho_g g_{t-1} + \varepsilon_{gt}.$$

Monetary policy follows the interest rate rule

$$r_t = \rho_r r_{t-1} + (1 - \rho_r)\left(\gamma_\pi \pi_t + \gamma_y \hat{y}_t\right) + q_t,$$

where $r_t = \log R_t - \log R$ and $\pi_t = \log P_t - \log P_{t-1} - \pi$, π is the inflation target, $\hat{y}_t$ is defined below and q_t follows the process

$$q_t = \rho q_{t-1} + \varepsilon_{qt}.$$

The model is solved and a log-linear approximation around a deterministic steady-state is computed.

Given that TFP is non-stationary, some variables need to be normalized to ensure stationarity. We define $\hat{c}_t$ as

$$\hat{c}_t = \log(C_t/A_t) - \log(C/A),$$

where C/A denotes the value of C_t/A_t in the deterministic version of the model in which A_t grows at the constant growth rate Γ. Analogous definitions apply to the quantities $\hat{y}_t, \hat{k}_t, \hat{k}_t, \hat{\imath}_t$. The quantities N_t and U_t are already stationary, so $n_t = \log N_t - \log N$, and similarly for u_t. For nominal variables, it is necessary to take care of non-stationarity in the price level, so: $\hat{w}_t = \log(W_t/(A_t P_t)) - \log(W/(AP)), r_t^k = \log(R_t^k/P_t) - \log(R^k/P), m_t = \log(M_t/P_t) - \log(M/P), r_t = \log R_t - \log R, \pi_t = log(P_t/P_{t-1}) - \pi$.

Finally, for the Lagrange multipliers: $\hat{\lambda}_t = \log(\Lambda_t A_t) - \log(\Lambda A), \hat{\phi}_t = \log(\Phi_t A_t/P_t) - \log(\Phi A/P)$. Φ_t is the Lagrange multiplier on the capital accumulation constraint. The hat is only used for variables normalized by A_t.

The first order conditions can be log-linearized to yield

$$\begin{aligned}\hat{\lambda}_t \quad = \quad & \frac{h\beta\Gamma}{(\Gamma - h\beta)(\Gamma - h)} E_t \hat{c}_{t+1} - \frac{\Gamma^2 + h^2\beta}{(\Gamma - h\beta)(\Gamma - h)} \hat{c}_t + \frac{h\Gamma}{(\Gamma - h\beta)(\Gamma - h)} \hat{c}_{t-1} + \\ & + \frac{h\beta\Gamma}{(\Gamma - h\beta)(\Gamma - h)} E_t[\Delta a_{t+1}] - \frac{h\Gamma}{(\Gamma - h\beta)(\Gamma - h)} \Delta a_t\end{aligned}$$

$$\hat{\lambda}_t = r_t + E_t[\hat{\lambda}_{t+1} - \Delta a_{t+1} - \pi_{t+1}]$$

$$\hat{\phi}_t = (1 - \delta)\beta\Gamma^{-1} E_t[\hat{\phi}_{t+1} - \Delta a_{t+1}] + (1 - (1 - \delta)\beta\Gamma^{-1}) E_t[\hat{\lambda}_{t+1} - \Delta a_{t+1} + r_{t+1}^k]$$

$$\hat{\lambda}_t = \hat{\phi}_t + d_t - \chi\Gamma^2 \left(\hat{\imath}_t - \hat{\imath}_{t-1} + \Delta a_t\right) + \beta\chi\Gamma^2 E_t \left(\hat{\imath}_{t+1} - \hat{\imath}_t + \Delta a_{t+1}\right)$$

$$r_t^k = \zeta u_t$$

$$m_t = \alpha r_t^k + (1-\alpha)\hat{w}_t$$

$$r_t^k = \hat{w}_t - \hat{k}_t + n_t$$

Log-linearizing the accumulation equation for capital and the equation for capacity utilization, yields

$$\begin{aligned} \hat{k}_t &= u_t + \hat{\bar{k}}_{t-1} - \Delta a_t \\ \hat{\bar{k}}_t &= (1-\delta)\Gamma^{-1}\left(\hat{\bar{k}}_t - \Delta a_t\right) + \left(1 - (1-\delta)\Gamma^{-1}\right) d_t + \hat{\imath}_t. \end{aligned}$$

Approximating and aggregating the intermediate goods production function over producers and using the final good production function yields

$$\hat{y}_t = \alpha \hat{k}_t + (1-\alpha) n_t$$

Market clearing in the final good market yields

$$(1-\psi)\hat{y}_t = \frac{C}{Y}\hat{c}_t + \frac{I}{Y}\hat{\imath}_t + \frac{R^k K}{PY} u_t + g_t$$

C/Y, I/Y and $R^k K/(PY)$ are all equilibrium ratios in the deterministic version of the model in which A_t grows at the constant rate Γ.

Aggregating individual optimality conditions for price setters yields the Phillips curve

$$\pi_t = \beta E_t \pi_{t+1} + \kappa m_t + \kappa m_{pt}$$

where $\kappa = (1-\theta_p\beta)(1-\theta_p)/\theta_p$.

Finally, aggregating individual optimality conditions for wage setters yields

$$\begin{aligned} \hat{w}_t \;\; = \;\; & \frac{1}{1+\beta}\hat{w}_{t-1} + \frac{\beta}{1+\beta}E_t\hat{w}_{t+1} - \frac{1}{1+\beta}(\pi_t + \Delta a_t) + \frac{\beta}{1+\beta}E_t(\pi_{t+1} + \Delta a_{t+1}) - \\ & -\kappa_w \left(\hat{w}_t - \zeta n_t + \hat{\lambda}_t + \kappa_w m_{wt}\right) \end{aligned}$$

where $\kappa_w = \frac{(1-\theta_w\beta)(1-\theta_w)}{\theta_w(1+\beta)\left(1+\zeta\left(1+\frac{1}{\mu_w}\right)\right)}$.

The log-linear model is estimated using Bayesian methods. Some parameters were calibrated using the mean values estimated in BLL. Table 5 reports the calibrated parameters.

Variables used in the estimation are the growth rates of output, consumption, investment and real wages, hours, the inflation rate and the federal funds rate (for details, see Appendix B). The choice of priors is very similar to the one used by BLL. Exception is made for the AR coefficients of the shocks, assumed here to be Normal with mean equal to 0 and standard deviation equal to 0.5 (0.4 for the coefficient ρ related to the transitory technology component) and for σ_d assumed here to be distributed as an Inverse Gamma with mean equal to 5 and standard deviation equal to 1.5.

We use an adaptive MCMC random walk Metropolis-Hastings algorithm (Haario et al., 2001) to obtain the posterior distribution. Table 6 summarizes the priors and the posterior estimates of the parameters.

Appendix B: Data and data treatment

The data set includes US quarterly data on Total Factor Productivity, real per-capita GDP, real per-capita consumption of non-durables and services, real per-capita investment, real wages, per-capita hours worked, the federal funds rate and the inflation rate. The time span is 1954Q3-2015Q2, so that we have 243 time observations. TFP data are taken from the website of the Federal reserve Bank of San Francisco. The series is adjusted for capital utilization. Since data are provided in quarter-on-quarter growth rates, we took the cumulated sum to get level data. All other data are taken from the FRED data base. The original GDP series is real GDP in billions of chained 2009 dollars. Consumption is obtained as the sum of nominal personal consumption expenditures for services, divided by its implicit price deflator, and nominal personal consumption expenditures for nondurable goods, divided by its implicit price deflator. Investment is the sum of nominal private fixed investment, divided by its implicit price deflator, and nominal personal consumption expenditures for durable goods, divided by its implicit price deflator. The real wage is obtained from the BLS series "nonfarm business sector: compensation per hour", divided by the GDP deflator. Hours worked are the BLS series named "nonfarm business sector: hours of all persons". We divided GDP, consumption, investment and hours by civilian noninstitutional population (aged 16 years or more) to get per-capita figures, took the logs and multiplied by 100 so that the numbers appearing on the vertical axis are quarter-on-quarter variations expressed in percentage points. The federal funds rate is the monthly effective federal funds rate; we averaged monthly figures to get quarterly frequency and transformed the data to get quarterly rates in percentage points ($25 \log(1 + r_t/100)$). Inflation is the first difference of the log of the GDP implicit price deflator multiplied by 100 to get figures expressed in percentage points.

Tables and Figures

Horizon	0	1	4	16
y_t, median estimate	0.00	0.10	0.12	0.12
y_t, true	0.00	0.11	0.12	0.12
r_t, median estimate	0.41	0.45	0.45	0.45
r_t, true	0.86	0.48	0.45	0.45

Table 1: Fraction of forecast error variance accounted for by the monetary policy shock in the empirical simulation of Example 1.

Period	< 2 yrs	$2 <$yrs< 8	> 8 yrs
y_t, median estimate	0.13	0.06	0.05
y_t, true	0.13	0.06	0.06
r_t, median estimate	0.48	0.29	0.28
r_t, true	0.48	0.29	0.28

Table 2: Spectral variance decomposition at different periodicities accounted for by the monetary policy shock in the empirical simulation of Example 1.

Shocks of interest	$\delta(1)$	$\delta(4)$	$\delta(1000)$
Demand shock, d_t	0.8904	0.8889	0.8889
Monetary shock v_t	0.0000	0.0000	0.0000

Table 3: The measure of informational deficiency $\delta(K)$ for Example 1.

Shocks of interest	$\delta(1)$	$\delta(4)$	$\delta(1000)$
Shock ε_t	0.0347	0.0344	0.0342
Shock d_t	0.9732	0.9687	0.9653
Shock e_t	0.4891	0.2558	0.0899

Table 4: The measure of informational deficiency $\delta(K)$ for Example 2.

Calibrated parameters	
ζ (elasticity of k utilization)	2.07
χ (I adj. cost)	5.5
h (habit persistence)	0.53
ς (inverse Frish elast.)	3.98
θ_w (W stickiness)	0.87
θ_p (P stickiness)	0.88
γ_π (π in Taylor rule)	1.003
γ_y (Y gap in Taylor rule)	0.0044
μ_p (SS P markup)	0.3
μ_w (SS W markup)	0.05
α (coeff. in prod. function)	0.19
Γ (TFP growth)	1
ψ (G/Y)	0.22
δ (K depreciation)	0.025
β (discount factor)	0.99

Table 5: Calibrated parameters. We use the posterior mean values estimated by BLL.

Estimated parameters		
Parameter	Prior	Mean
ρ_r (i smoothing)	$Beta(0.5, 0.2)$	0.57 [0.51 0.62]
ρ (temp. technology)	$\mathcal{N}(0.0 , 0.4)$	0.96 [0.95 0.97]
ρ_q (monetary)	$\mathcal{N}(0.0 , 0.5)$	0.19 [0.09 0.29]
ρ_d (I specific)	$\mathcal{N}(0.0, 0.5)$	0.68 [0.60 0.76]
ρ_p (P markup)	$\mathcal{N}(0.0, 0.5)$	0.81 [0.74 0.87]
ρ_w (W markup)	$\mathcal{N}(0.0, 0.5)$	0.95 [0.91 0.98]
ρ_g (G)	$\mathcal{N}(0.0, 0.5)$	0.98 [0.97 0.99]
ψ_p (MA in P mkup)	$Beta(0.5, 0.2)$	0.49 [0.30 0.65]
ψ_w (MA in W mkup)	$Beta(0.5, 0.2)$	0.96 [0.94 0.98]
σ_ε (permanent tech.)	$I\Gamma(0.5, 1.0)$	0.98 [0.90 1.06]
σ_v (temporary tech.)	$I\Gamma(1.0, 1.0)$	1.28 [1.18 1.39]
σ_q (monetary)	$I\Gamma(0.15, 1.0)$	0.26 [0.24 0.28]
σ_d (I specific)	$I\Gamma(5.0, 1.5)$	4.84 [4.06 5.67]
σ_p (p markup)	$I\Gamma(0.15, 1.0)$	0.14 [0.12 0.17]
σ_w (w markup)	$I\Gamma(0.15, 1.0)$	0.40 [0.36 0.43]
σ_g (gov exp.)	$I\Gamma(0.5, 1.0)$	0.52 [0.48 0.56]
Posterior value at mean		-1424

Table 6: Parameter estimates - mean. In brackets, the 5% and the 95% percentile of the posterior distribution.

specification	shocks						
	news	temp. tech.	price mkup	wage mkup	gov't exp.	inv. spec.	mon. pol.
S1	0.298	0.143	0.968	1.000	0.981	1.000	0.813
S2	0.957	0.357	0.984	1.000	0.996	0.127	0.983
S3	0.007	0.004	0.869	1.000	0.224	0.847	0.209
S4	0.006	0.003	0.812	0.431	0.191	0.840	0.199
S5	0.265	0.124	0.144	0.405	0.191	0.839	0.779
S6	0.703	0.271	0.788	0.205	0.362	0.029	0.206
S7	0.705	0.263	0.190	0.418	0.479	0.877	0.000
S8	0.690	0.258	0.175	0.161	0.349	0.023	0.000
S9	0.000	0.000	0.781	0.175	0.000	0.007	0.194
S10	0.000	0.000	0.101	0.132	0.000	0.004	0.000

Table 7: The measure of informational deficiency δ_i, for each shock $i = 1, \ldots, 7$, for the VAR specifications S1-S10 listed in Section 5.

specification	$\delta(1)$	$\delta(2)$	$\delta(4)$	$\delta(12)$	$\delta(1000)$
S3	0.3349	0.0213	0.0204	0.0135	0.0072
S4	0.3332	0.0211	0.0198	0.0127	0.0062
S9	0.3220	0.0148	0.0139	0.0069	0.0003
S10	0.3178	0.0140	0.0132	0.0067	0.0001

Table 8: The measure of informational deficiency $\delta(K)$, $K = 1, 2, 4, 12, 1000$, for the news shock, specifications S3,S4, S9, S10 in Section 5.

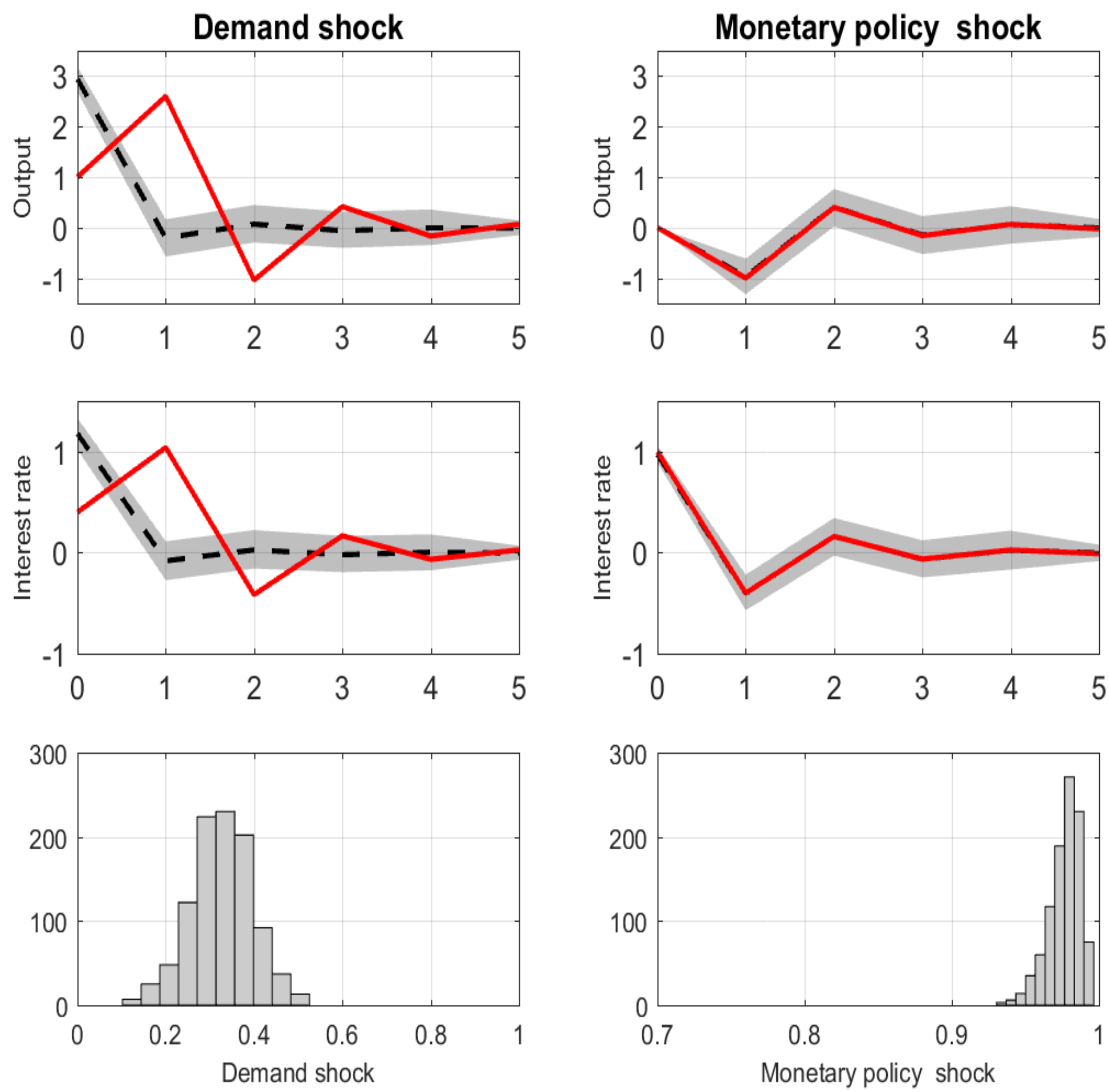

Figure 1: Impulse response functions of model (19) with $\alpha = 3$, $\gamma = 0.4$, $\beta = 1$ and standard normal shocks. Red solid lines: true impulse response functions. Black dashed lines: median of estimated impulse response functions. Grey area contains the 90% confidence interval computed from the Monte Carlo simulations. Lower panels report the distributions of the correlation coefficients between the estimated shocks and the true shocks (demand shock, d_t: left column; monetary policy shock, v_t: right column).

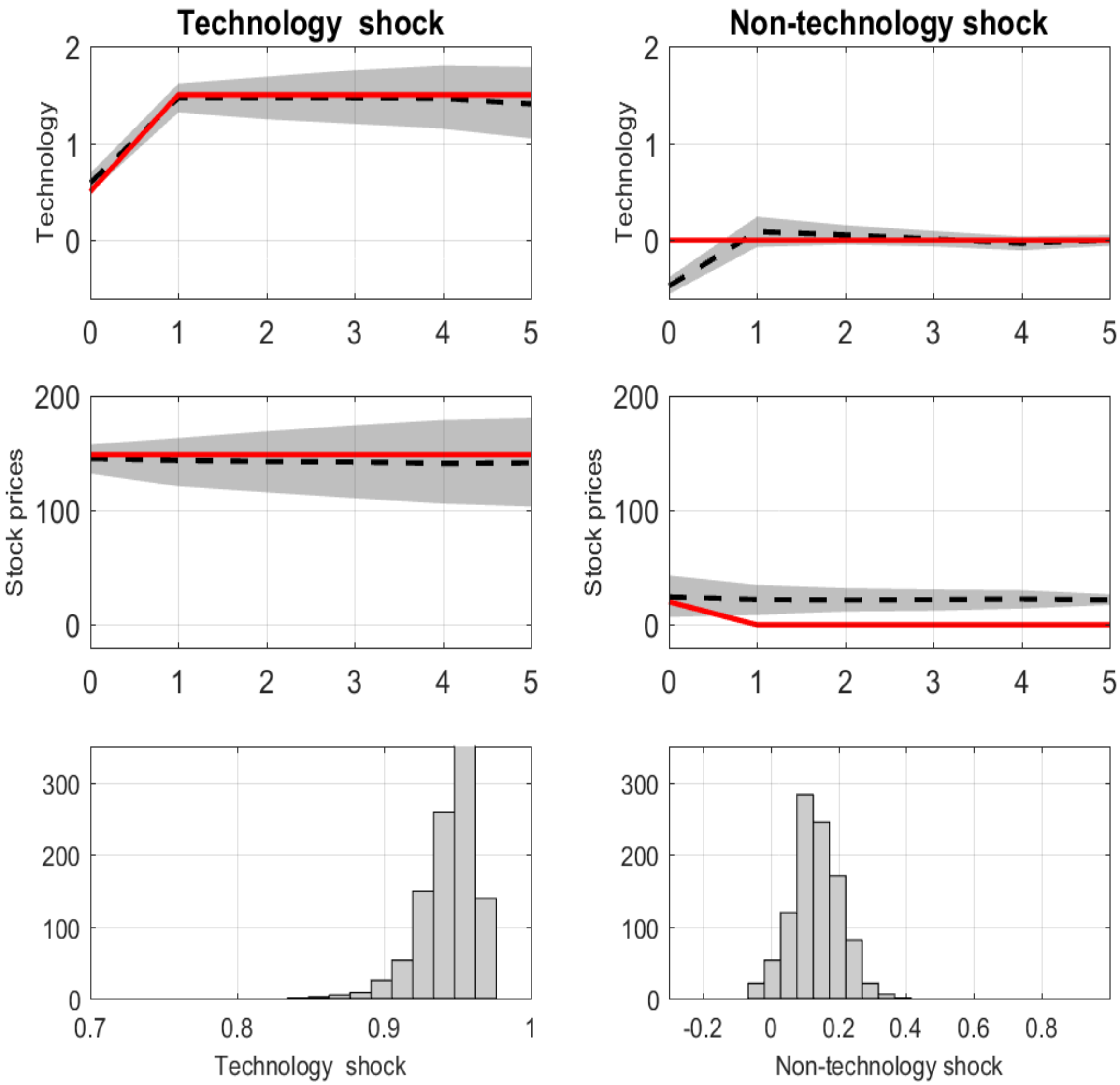

Figure 2: Impulse response functions of model (21) with $\beta = 0.99$, $\alpha = 0.5$, $\gamma = 20$, $\theta = 0.5$ and standard normal shocks. Red solid lines: true impulse response functions. Black dashed lines: median of estimated impulse response functions. Grey area contains the 90% confidence interval computed from the Monte Carlo simulations. Lower panels report the distributions of the correlation coefficients between the estimated shocks and the true shocks (ε_t, technology shock: left column; d_t, temporary shock to prices: right column).

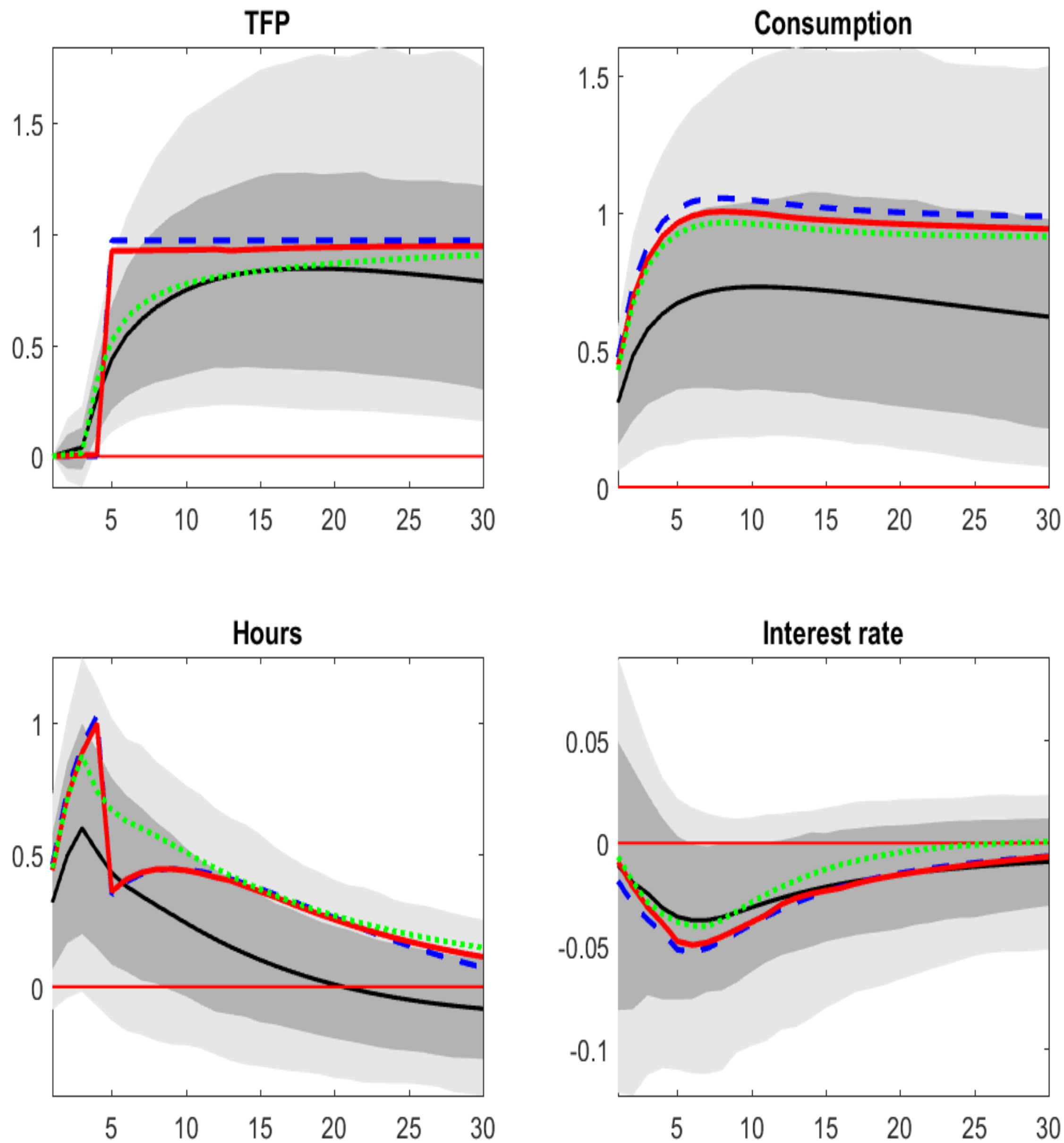

Figure 3: Impulse responses of the Bayesian VAR, specification S3, estimated with artificial data generated from the economic model. The blue dashed lines are the theoretical impulse response functions. The red thick solid lines are obtained with 5000 time observations and 12 lags. The green dotted lines are obtained with 5000 time observations and 4 lags. The black thin solid lines are obtained with 243 time observations and 4 lags. The dark gray and light gray areas are the 68% and 90% confidence bands obtained with 4 lags and 243 observations.

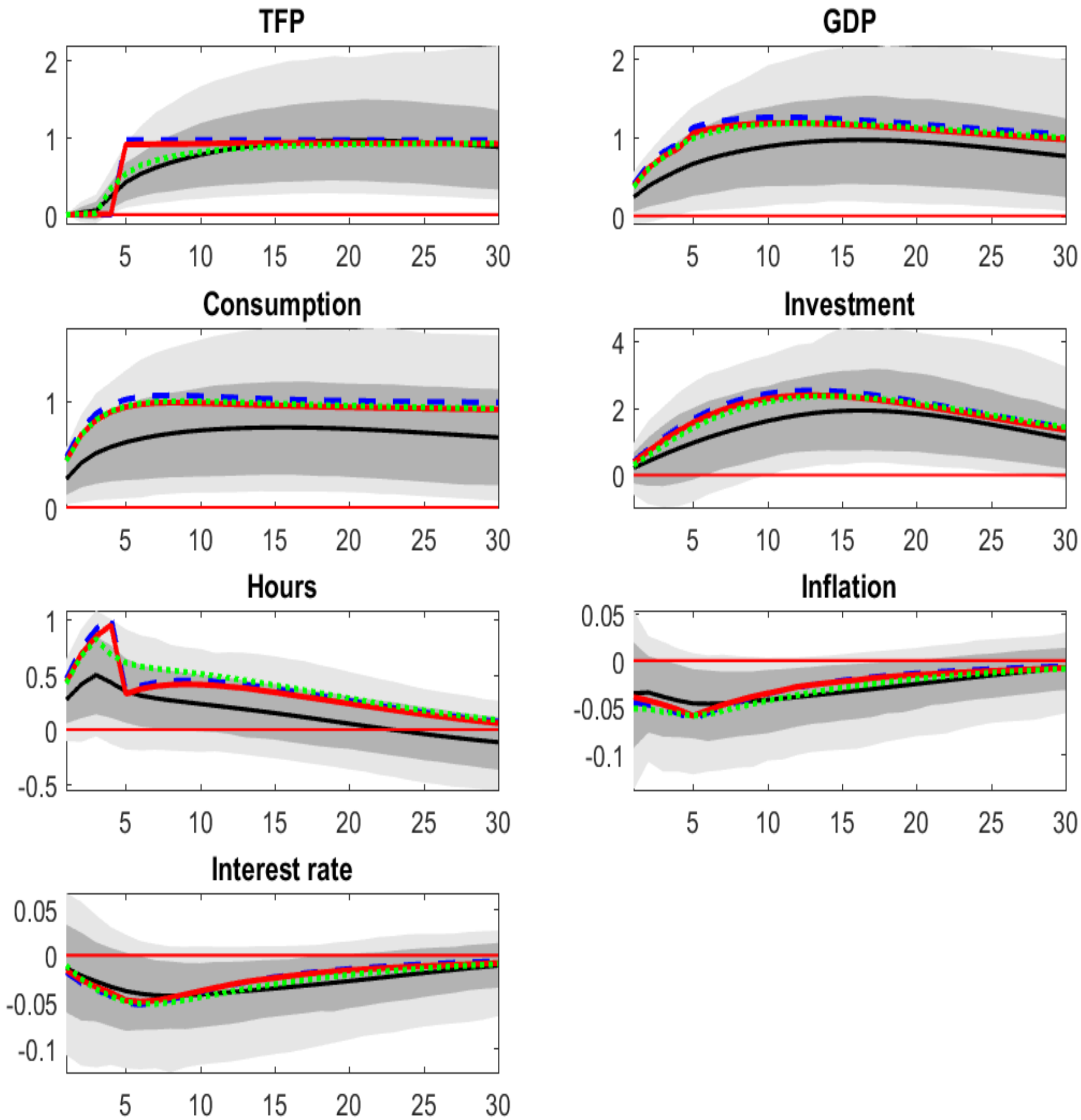

Figure 4: Impulse responses of the Bayesian VAR, specification S10, estimated with artificial data generated from the economic model. The blue dashed lines are the theoretical impulse response functions. The red thick solid lines are obtained with 5000 time observations and 12 lags. The green dotted lines are obtained with 5000 time observations and 4 lags. The black thin solid lines are obtained with 243 time observations and 4 lags. The dark gray and light gray areas are the 68% and 90% confidence bands obtained with 4 lags and 243 observations.

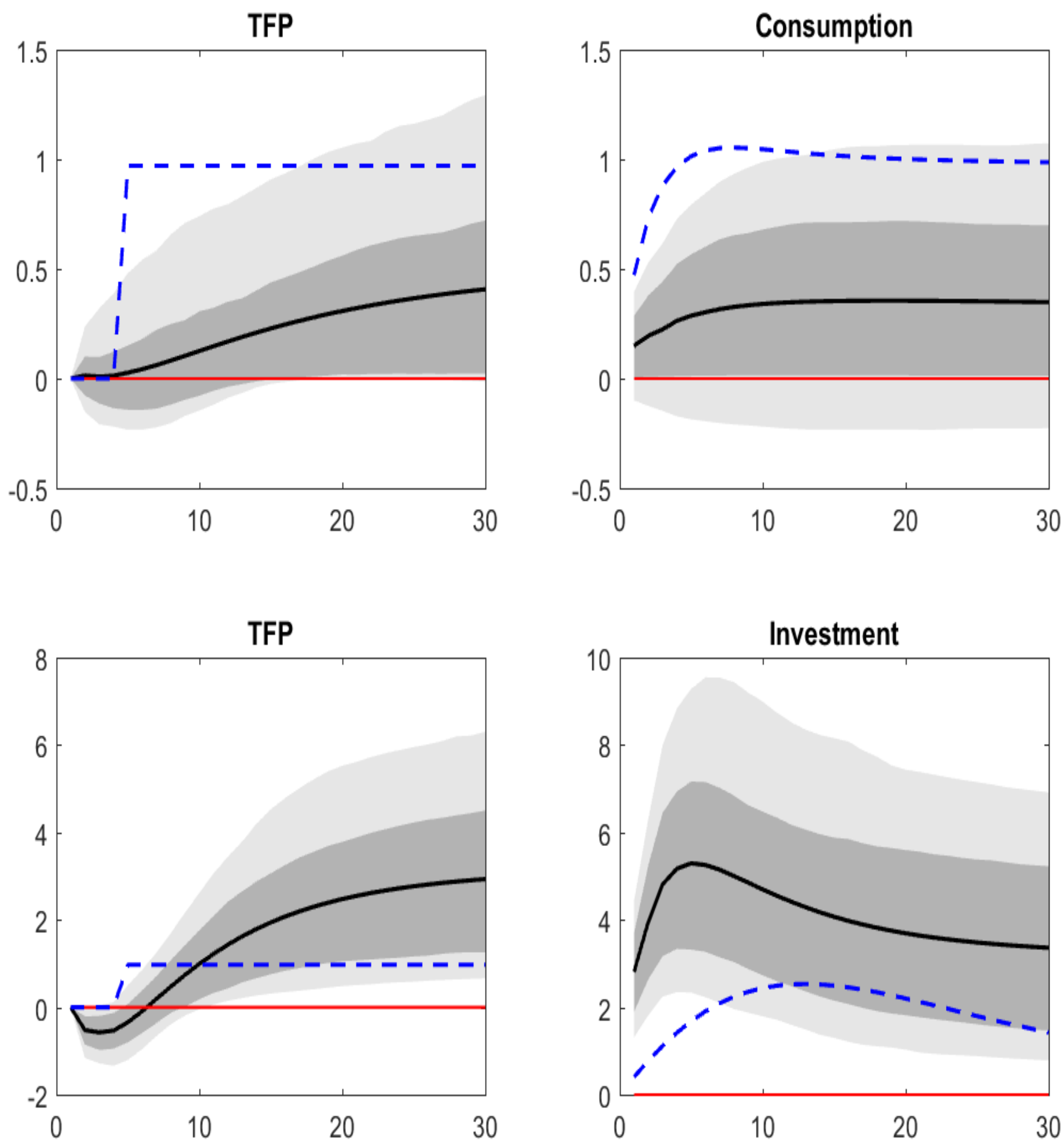

Figure 5: Impulse responses of the Bayesian VAR with US data, for specification S1 (TFP and investment, top panels) and specification S2 (TFP and consuption, bottom panels). The dashed lines are the theoretical impulse response functions. The black solid lines are averages of the posterior distribution (500 draws). The dark gray and light gray areas are the 68% and 90% confidence bands.

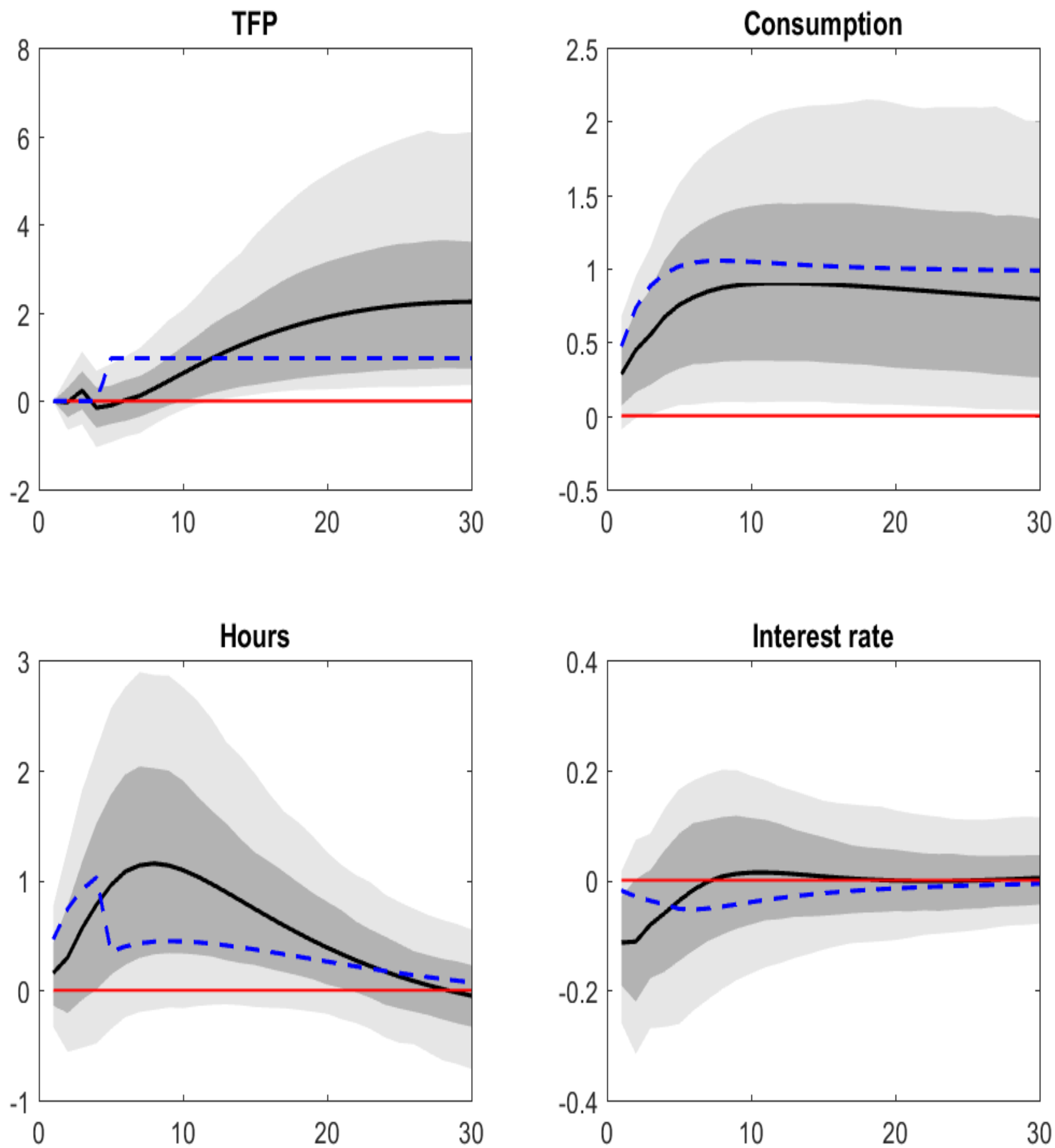

Figure 6: Impulse responses of the Bayesian VAR, specification S3, estimated with real US data. The dashed lines are the theoretical impulse response functions. The black solid lines are the empirical impulse response functions (averages of the posterior distribution, 500 draws). The dark gray and light gray areas are the 68% and 90% confidence bands.

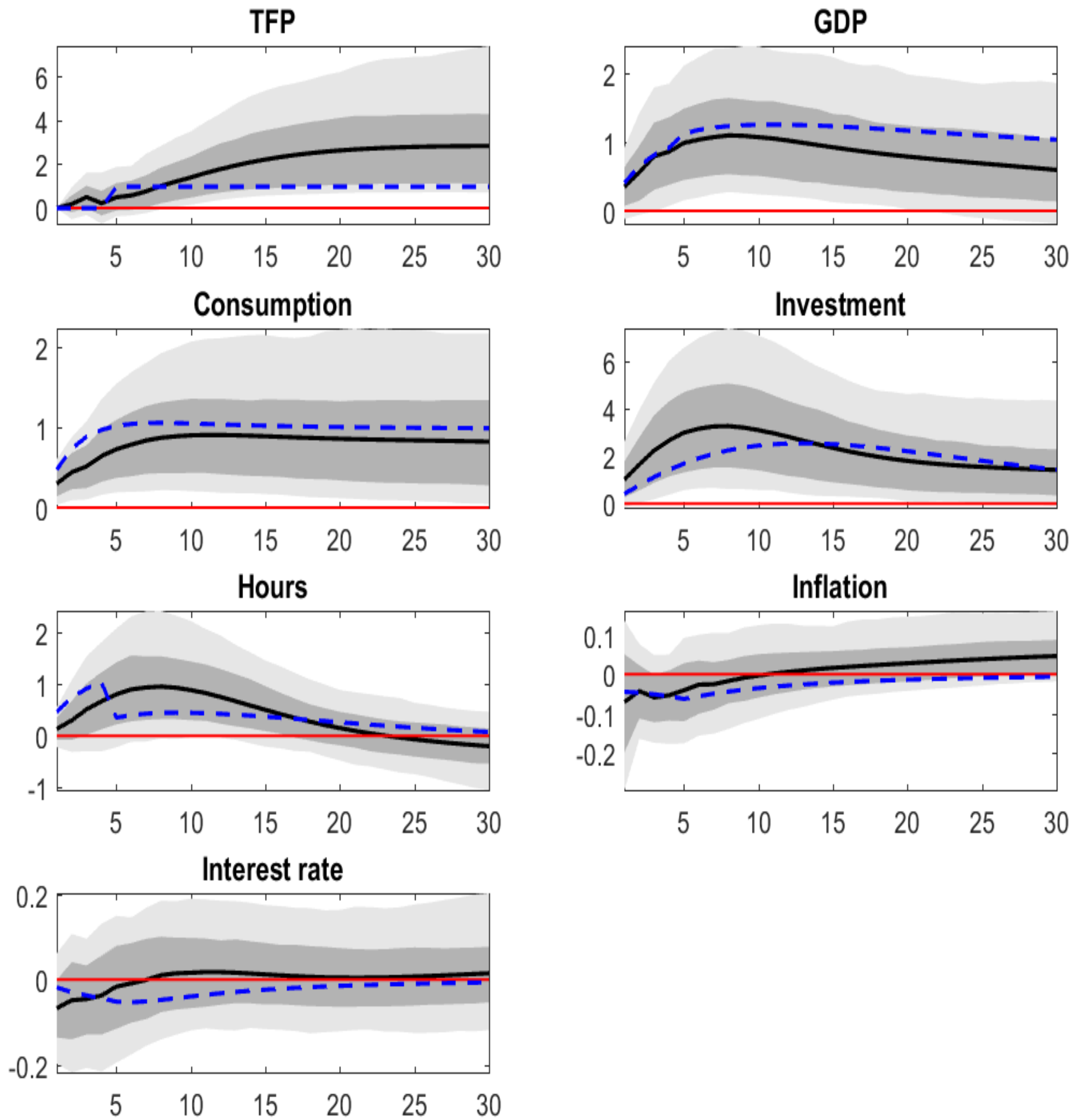

Figure 7: Impulse responses of the Bayesian VAR, specification S10, estimated with real US data. The dashed lines are the theoretical impulse response functions. The black solid lines are the empirical impulse response functions (averages of the posterior distribution, 500 draws). The dark gray and light gray areas are the 68% and 90% confidence bands.