# Are Small-Scale SVARs Useful for Business Cycle Analysis? Revisiting Non-Fundamentalness

Fabio Canova

BI Norvegian Business School, CAMP, and CEPR *

Mehdi Hamidi Sahneh

Universidad Carlos III †

January 18, 2016

**Abstract**

Non-fundamentalness arises when observables do not contain enough information to recover the vector of structural shocks. Using Granger causality tests, the literature suggested that many small scale VAR models are non-fundamental and thus not useful for business cycle analysis. We show that causality tests are problematic when VAR variables are cross sectionally aggregated or proxy for non-observables. We provide an alternative testing procedure, illustrate its properties with a Monte Carlo exercise, and reexamine the properties of two prototypical VAR models.

**Keywords:** Aggregation; Non-Fundamentalness; Granger causality, Small scale VARs.

**JEL classification:** C5, C32, E5.

*Center applied and Macro Petroleum economics and Department of Economics, BI Norvegian Business School, Nydalsveien 37, 0484 Oslo, Norway.

†Departamento de Economia, Universidad Carlos III de Madrid, Getafe, 28903, Spain. We thank Carlos Velasco, Jesus Gonzalo, Hernan Seoane, Jordi Gali, Davide de Bortoli, Luca Sala, Benjamin Holcblat, and the participants of seminars at BI Norwegian Business School, UPF, Bocconi, Humboldt for comments and discussions.

# 1 Introduction

Small scale Structural Vector Autoregressive (SVAR) models have been extensively used over the last 30 years to study cyclical fluctuations. The methodology hinges on the assumption that structural shocks can be obtained from linear combinations of the observables. Non-fundamentalness arises when current and past values of the observed variables do not contain enough information to recover the vector of structural shocks. In this case, shocks obtained via standard identification procedures may have little to do with the true disturbances, even when identification is correctly performed, making SVAR evidence unreliable.

Since likelihood or spectral estimation procedures can not distinguish fundamental vs. non-fundamental Gaussian systems (see e.g. Canova (2011), page 114), it is conventional in applied work to rule out all the non-fundamental representations that possess the same second-order structure of the data. However, the focus on fundamental systems is arbitrary. There are rational expectation models (Hansen and Sargent, 1991), optimal prediction models (Hansen and Hodrick, 1980) permanent income models (Fernández-Villaverde et al., 2007), news shocks models (Forni et al., 2014), and fiscal foresight models (Leeper et al., 2013), where optimal decisions may generate non-fundamental solutions. In addition, non observability of certain states or particular choices of observables may make fundamental systems non-fundamental. Hence, it is important to know if a fundamental system is a reasonable starting point to investigate the cyclical properties of the data.

Despite the far-reaching implications the fundamentalness assumption has for applied work, little is known on how to empirically detect non-fundamentalness. Giannone and Reichlin (2006) and Forni and Gambetti (2014) (henceforth, FG) proposed diagnostics based on the idea is that, under fundamentalness, external information should not Granger cause VAR variables. Using such a methodology, FG and Forni et al. (2014) argued that several small scale SVARs are non-fundamental, thus implicitly questioning the economic conclusions that are obtained. Considering the popularity of small scale SVARs, this result is disturbing. In this paper we show that existing diagnostics may be misleading because of aggregation or non-observability problems - we call the phenomenon spurious non-fundamentalness.

Why are there problems? Due to the *curse of dimensionality*, empirical studies consider only VARs with a small number of variables, usually much smaller than the potential set of primitive shocks driving the economy. Because SVARs can recover at most the same number of shocks as observables, estimated structural shocks are linear transformations of primitive shocks. Thus, they will be predictable using variables containing information about the primitive shocks, regardless of whether the VAR is fundamental or not.

For instance, suppose a researcher is interested in the effects of technology shocks on economic activity. Small scale VARs typically include a measure of total factor productivity (TFP) and a few other aggregate variables. Suppose that what drives the economy are sector-specific disturbances. Clearly, the estimated aggregate TFP shock will be a linear transformation of the vector of sectoral technology shocks and will be predictable using sectoral variables, regardless of the fundamentalness properties of the VAR.

A similar problem occurs when VAR variables proxy for unobservables. For example, TFP is latent and estimates are obtained from output, capital and hours worked data. If capital and hours worked are excluded from the VAR, any variable that predicts them will Granger cause estimated TFP, both when the VAR is fundamental and when it is not.

We propose an alternative procedure to detect non-fundamentalness, which is robust to aggregation and non-observability problems, and exploits the fact that if the model is non-fundamental, future VAR shocks predict a vector of variables excluded from the VAR.

We perform a Monte Carlo exercise using a version of fiscal foresight model of Leeper et al. (2013) as DGP with two sources of tax disturbances (capital and income taxes) and a productivity disturbance. We assume that the researcher considers a VAR with an aggregate tax variable (or with an aggregate tax rate computed from revenues and output data) and capital and show that, away from the unit circle, our approach has good size and power properties. In contrast, spurious non-fundamentalness arises with standard diagnostics.

We examine two conventional small scale VARs widely used in the literature. We find that

Beaudry and Portier (2006) model is fundamental according to our test but non-fundamental according to a Granger causality test. We show that the rejection of the null with the latter is due to aggregation problems: once disaggregated TFP data is used in the VAR, the null of fundamentalness is no longer rejected. We also demonstrate that Galí (1999) model is non-fundamental according to both tests and that technology shocks can not be obtained from the growth rate of labour productivity and hours. However, they can be extracted in an augmented VAR which adds to these variables a confidence indicator. The dynamics responses in bivariate and trivariate systems are however similar.

Two important caveats need to be mentioned. First, our analysis is concerned with Gaussian macroeconomic variables. Since most of the small scale VARs considered in the literature are estimated using quarterly data in post great moderation period, Gaussianity is approximately correct. For non-Gaussian situations, see Hamidi Saneh (2014) or Gourieroux and Monfort (2015). Second, although dynamic general equilibrium models deliver VARMA reduced form representations, we focus on VARs since they are much easier to estimate and more often used in the macroeconomic literature. The procedure we suggest also works for VARMA models as long as the MA roots are sufficiently away from unity.

The rest of the paper is organized as follows. Section 2 provides examples of non-fundamental systems and highlights the reasons for why the problem occurs. Section 3 shows why standard tests may be problematic and propose an alternative approach. Section 4 examines the performance of various procedures using Monte Carlo simulations. Section 5 investigates the properties of two small scale VAR systems. Section 6 concludes.

## 2 A few example of non-fundamental systems

While the literature has focused on non-fundamentalness driven by omitted variables, see e.g. Giannone and Reichlin (2006) or Lütkepohl (2012), there are other reasons for why the problem may emerge.

First, non-fundamentalness may be intrinsic to an economic model and to the optimization process, see e.g. Hansen and Sargent (1980, 1991). It may also be due to its information structure. For example, Leeper et al. (2013) show that with anticipatory shocks a standard RBC model could become non-fundamental while Forni et al. (2014) show that a standard asset pricing model may be non-fundamental when news come in a particular way. Finally, it may be due to the use of forecast errors in the VAR, see Hansen and Hodrick (1980). In general, optimizing models producing non-fundamental solutions are numerous. The next example shows one of these situations.

**Example 1.** Suppose the process for dividends is $d_t = e_t - \theta \mathrm{e}_{t-1}$, where $\theta < 1$, and suppose stock prices are expected discounted future dividends: $p_t = E_t \sum_j \beta^j d_{t+j}$. Then, the equilibrium value of $p_t$ in terms of the innovations in the dividend process is

$$p_t = (1 - \beta\theta)e_t - \theta e_{t-1} \tag{2.1}$$

Thus, even though the dividends process is fundamental, the process for stock prices is non-fundamental if $|\frac{(1-\beta\theta)}{\theta}| < 1$, which is the case when $\theta > \frac{1}{1+\beta}$.□

Second, non-fundamentalness may be due to non-observability of some of the endogenous variables. The next example illustrates how this is possible.

**Example 2.** Suppose the production function (in logs) is:

$$y_t = k_t + e_t \tag{2.2}$$

and the law of motion of capital is:

$$k_t = (1 - \delta)k_{t-1} + a_0 e_t \tag{2.3}$$

If both $(k_t, y_t)$ are observable this is just a bivariate restricted VAR(1). However, if the capital stock is not observable, the production function becomes

$$y_t - (1-\delta)y_{t-1} = (1+a_0)e_t + (1-\delta)e_{t-1} \tag{2.4}$$

Clearly, if $a_0 < 0$ and $|a_0| < |\delta|$, $e_t$ can not be recovered from current and past values of $y_t$ and (2.4) is non-fundamental. In addition, if $a_0 > 0$ and $\delta$ and $a_0$ are both small, (2.4) has a MA root close to unity and a finite order VAR for $y_t$ is a very poor approximation of underlying bivariate process, see also Ravenna (2007), and Giacomini (2013). □

Third, particular variable selection may induce non-fundamentalness, even if the system is, in theory, fundamental.

**Example 3.** Consider a standard consumption-saving problem. Let income $y_t = e_t$ be a white noise. Let $\beta = \frac{1}{R} < 1$ and assume quadratic preferences. Then:

$$c_t = c_{t-1} + (1 - R^{-1})e_t \tag{2.5}$$

Thus, growth rate of consumption has a fundamental representation. However, if instead of consumption, we setup the model in terms of savings $s_t = y_t - c_t$, the solution is

$$s_t - s_{t-1} = R^{-1}e_t - e_{t-1} \tag{2.6}$$

and the growth rate of saving is non-fundamental. □

In sum, there may be many reasons for why a VAR system may be non-fundamental. Assuming away non-fundamentalness, as it is done by standard computer packages, is problematic. Focusing on omitted variable problems is, on the other hand, reductive. In general, one ought to have procedures able to detect whether fundamentalness is appropriate and, when it is not, whether violations are due to theory or to applied investigators choices.

# 3 The Methodology

The economy is represented with an $n$-dimensional vector of stationary variables $\chi_t$ driven by $s$ (not necessarily equal to $n$) serial and mutually uncorrelated primitive shocks $\vartheta_t$.

**Assumption 1.** (Economy representation) The vector $\chi_t$ satisfies

$$\chi_t = \Gamma(L)\vartheta_t \tag{3.1}$$

where $\vartheta_t$ is a vector white noise, $\Gamma(L) = \sum_{i=0}^{\infty} \Gamma_i L^i$, $\Gamma_i$'s are $(n \times s)$ matrices each $i$, $L$ is the lag operator, and $\sum_{i=0}^{\infty} \Gamma_i^2 < \infty$.

Assumption 1 can be formally derived using the Wold theorem, assuming a linear representation, and repeatedly applying the projection operator (see Canova (2011), Ch. 4). Given a sample, the number of variables, $n$, is generally large, and to estimate the $\Gamma_i$ matrices an investigator needs to confine attention to an $m$-dimensional sub-vector of $\chi_t$.

**Assumption 2.** (VAR information set) The VAR is set up in terms of $x_t$, a sub-vector of $\chi_t$, driven by a sub-vector $u_t$ of $\vartheta_t$ of dimension $k$ (not necessarily equal to $m$):

$$x_t = \Pi(L)u_t \tag{3.2}$$

where $m \leq n$ and $k \leq s$, $\Pi(L) = \sum_{i=0}^{\infty} \Pi_i L^i$ has no root on the unit circle, and $\sum_{i=0}^{\infty} \Pi_i^2 < \infty$. By construction, $u_t$ are linear combination of current and, possibly, past values of $\vartheta_t$.

Next, we provide the definition of fundamentalness for (3.2) (see Rozanov (1967))

**Definition 1:** An uncorrelated process $\{u_t\}$ is $x_t$-fundamental if $\mathcal{H}_t^u = \mathcal{H}_t^x$ for all $t$, where $\mathcal{H}_t^u$ is the closed linear span of $\{u_s : s \leq t\}$[1]. $\{u_t\}$ is non-fundamental if $\mathcal{H}_t^u \subset \mathcal{H}_t^x$ and $\mathcal{H}_t^u \neq \mathcal{H}_t^x$, for at least one $t$.

Model (3.2) is fundamental if and only if all the roots of the determinant of the $\Pi(L)$ polynomial lie outside the unit circle in the complex plane - in this case $\mathcal{H}_t^u = \mathcal{H}_t^x$, for all $t$.

[1] The linear span is the smallest closed subspace which contains the subspaces.

Fundamentalness is closely related to the concept of invertibility: the latter requires that no root of the determinant of $\Pi(L)$ is on or inside the unit circle. Since we consider stationary variables, the two concepts are equivalent in our framework.

To obtain structural shocks from (3.2) one typically first imposes fundamentalness and estimates a VAR. Then, one identifies and estimates the $u_t$ using theory-based or informational-based restrictions, and constructs the structural MA representation. While it is conventional to assume fundamentalness, it is arbitrary and may lead to incorrect inference.

## 3.1 Standard approaches to detect non-fundamentalness

Checking whether a Gaussian VAR is fundamental is complicated because the likelihood function or the spectral density can not distinguish between a fundamental and a non-fundamental representation. Earlier work by Lippi and Reichlin (1993, 1994) informally compared the dynamics produced by fundamental and selected non-fundamental representations. Giannone and Reichlin (2006) and FG proposed to use Granger causality tests. The procedure works as follows. Suppose we augment $x_t$ with a vector of variables $y_t$

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \Pi(L) & 0 \\ B(L) & C(L) \end{bmatrix} \begin{bmatrix} u_t \\ v_t \end{bmatrix} \tag{3.3}$$

where $v_t$ are specific $y_t$ and orthogonal to $u_t$. Assume that all the roots of the determinant of $B(L)$ are outside the unit circle. If (3.2) is fundamental, $u_t = \Pi(L)^{-1}x_t$, and

$$y_t = B(L)\Pi(L)^{-1}x_t + C(L)v_t \tag{3.4}$$

where $B(L)\Pi(L)^{-1}$ is a one-sided polynomial in the non-negative powers of $L$. Thus, under fundamentalness, $y_t$ is a function of current and past values of $x_t$, but $x_t$ does not depend on $y_t$. Hence, to detect non-fundamentalness one can check whether $x_t$ is predicted by $y_t$.

While such an approach might be useful to check if there are variables omitted from the

VAR - effectively we are testing if the MA representation of the augmented system is lower triangular - it is not clear whether it can reliably detect non-fundamentalness. To know if it is the case, we need to know the mapping between SVAR shocks $u_t$ and primitive shocks $\vartheta_t$.

In small scale VARs the number of variables is smaller than the number of potentially interesting primitive shocks. Hence, it is impossible to obtain primitive shocks from present and past values of observables and applied researchers consider only systems featuring the same number of shocks as observables (m=k). However, these systems aggregate primitive shocks into a smaller number of structural shocks ($m = k < s$). Aggregation is not innocuous. For example, Chang and Hong (2006) show that aggregate and sectoral technology shocks behave quite differently. In general, aggregation may lead to spurious conclusions when testing for fundamentalness, as the next example shows.

**Example 4.** Suppose a scalar $x_t$ is driven by two primitive shocks

$$x_t = \vartheta_{1t} + b\vartheta_{1t-1} + \vartheta_{2t} \tag{3.5}$$

where $\vartheta_{1t} \sim iid(0, \sigma_1^2)$ and $\vartheta_{2t} \sim iid(0, \sigma_2^2)$ are mutually independent at all leads and lags. Clearly, we can't recover $\vartheta_{1t}$ and $\vartheta_{2t}$ from $x_t$. By considering a system with one observable and one structural shock, the econometrician implicitly assumes that $x_t$ has the form:

$$x_t = u_t + cu_{t-1} \tag{3.6}$$

where $u_t$ are iid, and $c$ and $\sigma_u^2$ are obtained from the moment conditions:

$$\mathrm{E}(x_t^2) = (1 + b^2)\sigma_1^2 + \sigma_2^2 = (1 + c^2)\sigma_u^2 \tag{3.7}$$

$$\mathrm{E}(x_t x_{t-1}) = b\sigma_1^2 = c\sigma_u^2 \tag{3.8}$$

These two conditions can be combined to obtain the quadratic equation:

$$bc^2 - \big[(1+b^2) + \frac{\sigma_1^2}{\sigma_2^2}\big]c + b = 0 \tag{3.9}$$

Given $b, \sigma_1^2$ and $\sigma_2^2$, (3.9) has two solutions; a fundamental one $c^*$ , and a non-fundamental one $\tilde{c}^*$. Since $u_t$ is a white noise, it can not be linearly predicted using lagged values of $u_t$ or $x_t$. However, it can predicted using lagged $\vartheta_{1t}$ and $\vartheta_{2t}$, regardless of whether $u_t$ is fundamental for $x_t$ or not. In fact, using (3.5) and (3.6) and the fundamental $c^*$ we have

$$u_t = (1+c^*L)^{-1}\big[\vartheta_{1t} + b\vartheta_{1t-1} + \vartheta_{2t}\big] = (\vartheta_{1t} - c^*\vartheta_{1t-1} + c^{*2}\vartheta_{1t-2} - c^{*3}\vartheta_{1t-3} + \cdots)$$
$$+ b(\vartheta_{1t-1} - c^*\vartheta_{1t-2} + c^{*2}\vartheta_{1t-3} - c^{*3}\vartheta_{1t-4} + \cdots) + (\vartheta_{2t} - c^*\vartheta_{2t-1} + c^{*2}\vartheta_{2t-2} - c^{*3}\vartheta_{2t-3} + \cdots)$$

so that $\mathbb{P}[u_t|\vartheta_{1t-1}, \vartheta_{1t-2}, \cdots, \vartheta_{2t-1}, \vartheta_{2t-2}, \cdots] \neq 0$ where $\mathbb{P}$ is the linear projection operator. Since $u_t$ is predictable, $x_t$ will be predictable. □

This simple setup provides a relevant counterexample to proposition 2 of FG (see page 126) in the sense the existence of variables that Granger cause $x_t$ may have nothing to do with whether the system is fundamental or not. In example 3, the assumptions of FG are satisfied (here there is no measurement error), but lags of $\vartheta_{1t}$ and $\vartheta_{2t}$ predict $x_t$ in both cases, making Granger causality tests irrelevant.

To generalize example 3, we first show that the class of moving average models is closed with respect to linear transformations.

**Proposition 1.** Let $\Theta_{1t}$ be a zero-mean MA($q_1$) process:

$$\Theta_{1t} = \vartheta_{1t} + \Phi_1\vartheta_{1t-1} + \Phi_2\vartheta_{1t-2} + \cdots + \Phi_{q_1}\vartheta_{1t-q_1} \equiv \Phi(L)\vartheta_{1t} \tag{3.10}$$

with $\mathrm{E}(\vartheta_{1t}\vartheta_{1t-j}) = \sigma_1^2$ if $j=0$ and 0 otherwise,and let $\Theta_{2t}$ be a zero-mean MA($q_2$) process:

$$\Theta_{2t} = \vartheta_{2t} + \Psi_1\vartheta_{2t-1} + \Psi_2\vartheta_{2t-2} + \cdots + \Psi_{q_2}\vartheta_{2t-q_2} \equiv \Psi(L)\vartheta_{2t} \tag{3.11}$$

with $\mathrm{E}(\vartheta_{2t}\vartheta_{2t-j}) = \sigma_2^2$ if $j = 0$ and 0 otherwise. Assume that $\Theta_{1t}$ and $\Theta_{2t}$ are independent at all leads and lags. Then

$$x_t = \Theta_{1t} + \Theta_{2t} = u_t + \Pi_1 u_{t-1} + \Pi_2 u_{t-2} + \cdots + \Pi_q u_{t-q} \equiv \Pi(L) u_t \tag{3.12}$$

where $q = \max\{q_1, q_2\}$, and $u_t$ is a white noise process.

*Proof:* The proof follows from Hamilton (1994), page 106.□

Next, we show that although $u_t$ in (3.12) is unpredictable given own lagged values, it can be predicted using lagged values of $\vartheta_{1t}$ and $\vartheta_{2t}$. This happens because the information contained in the histories of $\vartheta_{1t}$ and $\vartheta_{2t}$ is not optimally aggregated in $u_t$.

**Proposition 2.** Let $x_t$ be an $m$-dimensional process obtained as in Proposition 1. Then, lagged $\vartheta_{1t}$ and lagged $\vartheta_{2t}$ Granger causes $x_t$.

*Proof:* It is enough to show that

$$\mathbb{P}\big[x_t | x_{t-1}, x_{t-2}, \cdots, \vartheta_{1t-1}, \vartheta_{1t-2}, \cdots, \vartheta_{2t-1}, \vartheta_{2t-2}, \cdots\big] \neq \mathbb{P}\big[x_t | x_{t-1}, x_{t-2}, \cdots\big]$$

when the model is fundamental. Here $\mathcal{H}_t^x = \mathcal{H}_t^u$, and it suffices to show that $u_t$ is Granger caused by lagged values of $\vartheta_{1t}$ and $\vartheta_{2t}$. That is

$$\mathbb{P}\big[u_t | u_{t-1}, u_{t-2}, \cdots, \vartheta_{1t-1}, \vartheta_{1t-2}, \cdots, \vartheta_{2t-1}, \vartheta_{2t-2}, \cdots\big] \neq \mathbb{P}\big[u_t | u_{t-1}, u_{t-2}, \cdots\big]$$

From Proposition 1, we have that $\Pi(L) u_t = \Phi(L)\vartheta_{1t} + \Psi(L)\vartheta_{2t}$, and therefore

$$u_t = \Pi(L)^{-1}\Phi(L)\vartheta_{1t} + \Pi(L)^{-1}\Psi(L)\vartheta_{2t}$$

where $\Pi(L)^{-1}$ exists since the model is fundamental. Hence, $\Pi(L)^{-1}\Phi(L)$ and $\Pi(L)^{-1}\Psi(L)$

are one-sided polynomial in the non-negative powers of $L$ and

$$\mathbb{P}[u_t|u_{t-1},u_{t-2},\cdots,\vartheta_{1t-1},\vartheta_{1t-2},\cdots,\vartheta_{2t-1},\vartheta_{2t-2},\cdots] = \mathbb{P}[u_t|\vartheta_{1t-1},\vartheta_{1t-2},\cdots,\vartheta_{2t-1},\vartheta_{2t-2},\cdots] \neq 0$$

where the equality follows from $u_t$ being a white noise process.□

Proposition 2 implies that lagged values of disaggregated variables or of factors obtained from a large data set providing signals about primitive shocks can Granger cause aggregated VAR variables. Hence, the null of fundamentalness might be rejected even if it holds true.

While the analysis so far is concerned with the fundamentalness of the vector $u_t$, it is common in the VAR literature to focus attention on just one shock, see e.g. Christiano et al. (1999) or Galí (1999). The next example shows when one can recover a shock from current and past values of the observables, even when the system is non-fundamental.

**Example 5.** Consider the following three MA models

$$\begin{aligned} x_{1,t} &= u_{1t} + 2u_{2t-1} \\ x_{2,t} &= u_{1t} - u_{2t-1} \end{aligned} \tag{3.13}$$

$$\begin{aligned} x_{1,t} &= u_{1t} \\ x_{2,t} &= u_{1t} + u_{2t} - 3u_{2t-1} \end{aligned} \tag{3.14}$$

$$\begin{aligned} x_{1,t} &= u_{1t} - 2u_{2t-1} \\ x_{2,t} &= u_{1t} + u_{2t-1} \end{aligned} \tag{3.15}$$

All three systems are non-fundamental since the determinants of the MA matrix are $-3L$, $1-3L$, and $L(1-2L)$ respectively, and they vanishes for $L<1$. Thus, it is impossible to recover $u_t = (u_{1t}, u_{2t})$ from current and lagged $x_t = (x_{1,t}, x_{2,t})'$. However, while in the first system $u_{2t}$ can be recovered as $u_{2t} = (x_{1,t+1} - x_{2,t+1})/3$ and $u_{1t}$ can be recovered as $u_{1t} = (x_{1,t} + 2x_{2,t})/3$, in the second system only $u_{1t}$ can be recovered from $x_{1,t}$ while in the third system no shock can be recovered from linear combinations of the $x_t$'s. $\square$

A necessary condition for an estimated shock to be an innovation is that it is orthogonal to the past of the observables. FG suggest that a shock derived as in the second system of example 5 is fundamental if it is orthogonal to the lags of the principal components obtained from a large data set. Three important points need to be made about such an approach. First, fundamentalness is a property of a system not of a single shock. Thus, the results of orthogonality tests are insufficient to assess the fundamentalness of a system. Second, as it is clear from the examples, when one shock can be recovered, it is not the one that creates the non-fundamentalness in the first place, unless future information is available, a point stressed also by Leeper et al. (2013). Finally, as shown in the next section, an orthogonality test has the same shortcomings as a Granger causality test: it will reject the correct null when one tests for orthogonality of an aggregate shock with respect to some disaggregated shocks or some factors providing noisy information about them, for exactly the same reasons that Granger causality tests fail.

## 3.2 An alternative approach

In this section we propose an alternative testing approach that does not suffer from aggregation or non-observability problems. To see what the procedure involves suppose we still augment (3.2) with a vector of additional variables $y_t = B(L)u_t + C(L)v_t$. If (3.2) is fundamental, $u_t$ can be obtained as from current and past values of $x_t$

$$u_t = x_t - \sum_{j=1}^{r} \omega_j x_{t-j} \tag{3.16}$$

where $\omega(L) = \Pi(L)^{-1}$ and $r$ is generally finite. Thus, under fundamentalness $y_t$ only depends on current and past values of $u_t$. If instead (3.2) is non-fundamental, $u_t$ can not be recovered from the current and past values of the $x_t$. A VAR econometrician can only recover $u_t^* = x_t - \sum_{j=1}^r \omega_j^* x_{t-j}$ where $\omega(L)^* = \Pi(L)^{-1}\theta(L)^{-1}$ which is related to $u_t$ via

$$u_t^* = \theta(L) u_t \tag{3.17}$$

where $\theta(L)$ is a Blaschke matrix [2]. Then, the relationship between the additional variables $y_t$ and the shocks recovered by the econometrician is $y_t = B(L)\theta(L)^{-1}\theta(L)u_t + C(L)v_t \equiv B(L)^* u_t^* + C(L)v_t$. Since $B(L)^*$ is generally two-sided polynominal, $y_t$ depends on current, past and future values of $u_t^*$. This proves the following.

**Proposition 3.** Model (3.2) is fundamental if $u_{t+j}^*,\ j \geq 1$ do not Granger cause $y_t$.

**Example 6.** To illustrate proposition 3, let $x_t = (1-2L)u_t$, which can be rewritten as

$$x_t \;=\; (1-2L)\frac{(1-0.5L)}{(1-2L)}\frac{(1-2.0L)}{(1-0.5L)}u_t \equiv (1-0.5L)u_t^* \tag{3.18}$$

where $u_t^* = \frac{(1-2.0L)}{(1-0.5L)}u_t$ and $\frac{(1-2.0L)}{(1-0.5L)}$ is a Blaschke matrix. Let $y_t = (1-0.5L)u_t + (1-0.6L)v_t$. Then

$$\begin{aligned} y_t &= (1-0.5L)\frac{(1-0.5L)}{(1-2.0L)}u_t^* + (1-0.6L)v_t \\ &= \sum_{j=0}^{\infty}(1/2)^j((1-0.5L)^2 u_{t+j}^*) + (1-0.6L)v_{t-j} \end{aligned} \tag{3.19}$$

A few points about our procedure need to be stressed. First, the null we consider differs from the ones used in the literature. Second, our approach is likely to be robust to linear transformations of primitive shocks, because when $y_t$ is the principal components of a large data set, it includes more information than the VAR, and thus should not be Granger caused by future VAR shocks under fundamentalness. Third, our setup is sufficiently general to deal with non-fundamentalness due to structural causes, omitted or proxy variables.

[2] Blaschke matrices are complex-valued filters. The main property of Blaschke matrices is that they take orthonormal white noises into orthonormal white noises. See Lippi and Reichlin (1994) for more details.

# 4 Some Monte Carlo evidence

To evaluate the properties of our procedure, we carry out a simulation study using a version of the fiscal foresight model of Leeper et al. (2013), with two sources of tax disturbances. In this economy the representative household maximizes:

$$\mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \log(C_t)$$

subject to

$$C_t + (1-\tau_{t,k})K_t + T_t \leq (1-\tau_{t,y})A_t K_{t-1}^{\alpha} = (1-\tau_{t,y})Y_t$$

where $C_t,\ K_t,\ Y_t,\ T_t,\ \tau_{t,k}$ and $\tau_{t,y}$ denote time-$t$ consumption, capital, output, lump-sum taxes, investment tax (or subsidy) and the income tax rate, respectively; $A_t$ is an exogenous technology shock and $\mathrm{E}_t$ is the expectation operator conditional on time t information. To keep the setup tractable, we assume full capital depreciation and inelastic labor supply. The government sets tax rates parametrically and adjusts lump-sum transfers to satisfy $T_t = \tau_{t,y}Y_t + \tau_{t,k}K_t$. The Euler equation and the resource constraints are:

$$\frac{1}{C_t} = \alpha\beta \mathrm{E}_t\Big[\frac{(1-\tau_{t+1,y})}{(1-\tau_{t,k})}\frac{1}{C_{t+1}}\frac{A_{t+1}K_t^{\alpha}}{K_t}\Big] \tag{4.1}$$

$$C_t + K_t = A_t K_{t-1}^{\alpha} \tag{4.2}$$

Log linearizing the two equations, combining (4.1) and (4.2), and assuming that the technology shock is *iid*, we have that

$$k_t = \alpha k_{t-1} + \epsilon_{t,A} - \kappa_k \sum_{i=0}^{\infty} \theta^i \mathrm{E}_t \hat{\tau}_{t+i,k} - \kappa_y \sum_{i=0}^{\infty} \theta^i \mathrm{E}_t \hat{\tau}_{t+i+1,y} \tag{4.3}$$

where $\kappa_k = \frac{\tau_k(1-\theta)}{(1-\tau_k)}$, $\kappa_y = \frac{\tau_y(1-\theta)}{(1-\tau_y)}$, $\theta = \alpha\beta\frac{1-\tau_y}{1-\tau_k}$, and lower case letters denote percentage deviations from steady states, $k_t \equiv \log(K_t) - \log(K)$, $a_t \equiv \log(A_t) - \log(A)$, $\hat{\tau}_{t,k} \equiv \log(\tau_{t,k}) - \log(\tau_k)$ and $\hat{\tau}_{t,y} \equiv \log(\tau_{t,y}) - \log(\tau_y)$.

We posit that income tax shocks are unanticipated, but agents receive a one period in advance signal of investment tax shocks. Thus, the equilibrium dynamic of capital is:

$$k_t = \alpha k_{t-1} + \vartheta_{t,A} - \kappa_k \Big\{ \vartheta_{t-1,k} + \theta \vartheta_{t,k} \Big\} - \kappa_y \vartheta_{t,y} \tag{4.4}$$

We assume that the econometrician only observes an aggregate tax variable and does not have access to data on investment and income tax rates separately. Alternatively, one can assume that the econometrician observes sub-categorical taxes, but she is only interested in working with a weighted sum of them. Thus, our design covers both the cases of aggregation and of a relevant latent variable. Proposition 1 implies that (4.4) can be written as:

$$k_t = \alpha k_{t-1} + u_{t,A} - u_{t,\tau} - c u_{t-1,\tau} \tag{4.5}$$

where $u_{t,\tau}$ is an aggregate iid tax shock and $c$ is obtained from the quadratic equation:

$$\theta c^2 - \big[ (1+\theta^2) + (\kappa_y^2 \sigma_y^2 / \kappa_k^2 \sigma_k^2) \big] c + \theta = 0. \tag{4.6}$$

The equilibrium conditions for capital and the aggregate tax rate are then:

$$\begin{bmatrix} \hat{\tau}_t \\ (1-\alpha L) k_t \end{bmatrix} = \begin{bmatrix} 1 + cL & 0 \\ -1 - cL & 1 \end{bmatrix} \begin{bmatrix} u_{t,\tau} \\ u_{t,A} \end{bmatrix} = \Pi(L) u_t \tag{4.7}$$

The determinant of $\Pi(L)$ vanishes for $L = -\frac{1}{c}$ and $c$ is a function of $\tau_k$ and $\tau_y$ through $\theta$ and $\kappa_k, \kappa_y$. Thus, by varying $\tau_k$, (4.7) is fundamental or non-fundamental. For the exercises we present, we let $\vartheta_{t,A}, \vartheta_{t,k},$ and $\vartheta_{t,y} \sim iid\ N(0,1)$, set $\alpha = 0.36$, $\beta = 0.95$, $\tau_y = 0.25$ and select $\tau_k$ so that the two roots of (4.6) vary in the range $c^* \in (0.1, 0.8)$ (fundamentalness region) and $\tilde{c}^* \in (1.25, 10)$ (non-fundamentalness region). Note that, in the latter region, capital responses to a tax shock are humped shaped; in the former they peak at zero.

To perform the tests, we need additional data not used in the VAR. We assume that an econometrician observes a panel of $n = 30$ time series, generated by:

$$y_{i,t} = \gamma_i \vartheta_{t,A} + (1-\gamma_i)\vartheta_{t,\tau} + (\epsilon_{t,y} + \vartheta_{t,k})\xi_{i,t},\ i = 1, \cdots, n \tag{4.8}$$

where $\xi_{i,t} \sim$ *iid* $N(0, 0.5)$ is a variable specific measurement error, and $\gamma_i$ is a Bernoulli random variable taking value 1 with probability 0.5. We set $T = 200$, so that our sample length resembles the one typically employed in quarterly macroeconomic exercises.

The properties of our procedure, denoted by $CH$, are examined with the regression:

$$f_t = \phi_0 + \phi_1 f_{t-1} + \cdots + \phi_p f_{t-p} + \psi_0 u_t + \psi_{-1} u_{t-1} + \cdots + \psi_{-p} u_{t-p} + \psi_1 u_{t+1} + \cdots + \psi_q u_{t+q} + e_t \tag{4.9}$$

where $f_t$ is a vector of principal components of (4.8) and $u_t$ is obtained estimating

$$x_t = \omega_0 + \omega_1 x_{t-1} + \cdots + \omega_r x_{t-r} + u_t \tag{4.10}$$

where $x_t = (\hat{\tau}_t, k_t)'$. The null that $u_{t+i},\ i = 1, \cdots, q$ are irrelevant in (4.9) is:

$$\mathbb{H}_0^{CH} : R\Psi = 0 \tag{4.11}$$

where $\Psi = \text{Vec}[\psi_1, \psi_2, \cdots, \psi_q]$ and $R$ is a matrix of zeros and ones, and can be checked using a standard F-test. We report the results for $p = 4$ and $q = 2$, when $r = 4$ and three principal components are included in $f_t$ vector. The appendix present results obtained with other values of the nuisance parameters $p$, $q$, and $r$.

To examine the properties of Granger causality tests, denoted by $FG_{GC}$, we employ

$$x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varphi_1 f_{t-1} + \cdots + \varphi_q f_{t-q} + e_t \tag{4.12}$$

where again $x_t = (\hat{\tau}_t, k_t)'$. The null that $f_{t-i},\ i = 1, \cdots, q$ are irrelevant in (4.12) is

$$\mathbb{H}_0^{GC} : R\Phi = 0 \tag{4.13}$$

where $\Phi = \text{Vec}[\varphi_1, \varphi_2, \cdots, \varphi_q]$ and $R$ is a matrix of zeros and ones. To test (4.13) we use the out-of-sample procedure of Gelper and Croux (2007). In particular, we split the $T$ observations in two parts, and set $T_1 = T/2$. $T_1$ observations are used to estimate the parameters of (4.12). Once this is done, a sequence of $T_2$ one-step-ahead forecasts are constructed from two models: an unrestricted one (4.12), and a restricted one, where the null is imposed. Let the matrix of one-step-ahead forecast errors of the unrestricted model be $e_U$ and of the restricted model be $e_R$. The null is evaluated using the mean squared forecast errors $\text{MSFE} = \log\left(\frac{|e_R' e_R|}{|e_U' e_U|}\right)$. Since the asymptotic distribution of MSFE under the null is unknown, p-values are obtained by a residual based bootstrap method. As in FG, we report results when $q = p$ and $p$ is selected using the BIC criteria.

To perform the orthogonality test, denoted by $FG_{OR}$, we first estimate the VAR:

$$x_t = \omega_0 + \omega_1 x_{t-1} + \cdots + \omega_r x_{t-r} + u_t \tag{4.14}$$

where $x_t = (\hat{\tau}_t, k_t)'$, and where $r$ is determined with BIC. The tax shock, $u_{t,\tau}$, is identified as the only one affecting the tax rate in the long run. Then in the regression

$$u_{t,\tau} = \lambda_1 f_{t-1} + \cdots + \lambda_q f_{t-q} + e_t \tag{4.15}$$

the null hypothesis that $u_{t,\tau}$ and $f_{t-i},\ i = 1, \cdots, q$ are orthogonal corresponds to

$$\mathbb{H}_0^{OR} : R\Lambda = 0 \tag{4.16}$$

where $\Lambda = \text{Vec}[\lambda_1, \lambda_2, \cdots, \lambda_q]$ and $R$ is a matrix of zeros and ones. This null can be tested

Table 1: Testing for fundamental representation: Size

| | $c^*$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 |
|---|---|---|---|---|---|---|---|---|---|
| 1% | $CH$ | 0.8 | 1.2 | 1.0 | 0.8 | 1.0 | 0.9 | 1.1 | 1.1 |
| | $FG_{OR}$ | 0.1 | 1.8 | 22.4 | 84.2 | 98.8 | 99.3 | 99.3 | 99.3 |
| | $FG_{GC}$ | 0.9 | 1.1 | 2.3 | 13.8 | 58.0 | 91.6 | 97.3 | 97.1 |
| 5% | $CH$ | 5.5 | 5.6 | 5.0 | 4.9 | 4.7 | 4.7 | 5.3 | 5.5 |
| | $FG_{OR}$ | 1.4 | 9.3 | 47.2 | 93.9 | 99.7 | 99.5 | 99.8 | 99.5 |
| | $FG_{GC}$ | 4.9 | 5.6 | 9.8 | 32.9 | 81.3 | 98.6 | 99.6 | 99.5 |
| 10% | $CH$ | 10.8 | 10.6 | 9.1 | 9.4 | 9.3 | 9.9 | 10.3 | 10.4 |
| | $FG_{OR}$ | 4.5 | 16.3 | 59.3 | 97.5 | 99.8 | 99.5 | 99.8 | 99.5 |
| | $FG_{GC}$ | 9.2 | 12.2 | 16.9 | 47.1 | 88.8 | 99.4 | 99.8 | 99.8 |

Notes: The table reports the percentage of rejections of the null hypothesis in 1000 replications when three principal components of the large dataset are considered; $CH$ is the test proposed in this paper; $FG_{OR}$ is the orthogonality test of Forni and Gambetti (2014); $FG_{GC}$ is the Granger causality test of Forni and Gambetti (2014).

using a standard F-test. As in FG we set $q = r$.

The percentage of rejections of the null of each procedure in 1000 replications when the model is fundamental are in Table 1. Our procedure has good empirical size properties and the performance of the test is independent of the confidence level employed. By contrast, Granger causality and orthogonality tests are prone to spurious non-fundamentalness. The rejection rate becomes large when the weight on last period information increases and deteriorates dramatically when $c^*$ exceeds 0.4-0.5. The presence of a threshold is intuitive. From (4.4) and (4.5) we have $(1 + cL)u_{\tau,t} = \kappa_k(\theta + L)\vartheta_{t,k} + \kappa_y\vartheta_{t,y}$ or

$$u_{\tau,t} = \sum_j c^j(\kappa_k(\theta + L)\vartheta_{t-j,k} + \kappa_y\vartheta_{t-j,y}) \equiv Z_1(L)\vartheta_{t,k} + Z_2(L)\vartheta_{t,y} \qquad (4.17)$$

If $c^*$ is small, $Z_{1j}, Z_{2j}$ tends to zero fast as $j \to \infty$ making predictability hard to detect.

The performance of Granger causality and orthogonality tests is very poor because BIC, which we use to choose the VAR lag length, underestimates the lag length needed to replicate the properties of the DGP when $c^* \in [0.4, 0.8]$. If one would use AIC, which chooses a longer

Table 2: Testing for fundamental representation: Power

| | $\tilde{c}^*$ | 10 | 5 | 3.33 | 2.5 | 2.0 | 1.66 | 1.43 | 1.25 |
|---|---|---|---|---|---|---|---|---|---|
| | $CH$ | 100 | 100 | 100 | 100 | 100 | 100 | 98.4 | 40.4 |
| 1% | $FG_{OR}$ | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | $FG_{GC}$ | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | $CH$ | 100 | 100 | 100 | 100 | 100 | 100 | 99.8 | 62.6 |
| 5% | $FG_{OR}$ | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | $FG_{GC}$ | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | $CH$ | 100 | 100 | 100 | 100 | 100 | 100 | 99.9 | 75 |
| 10% | $FG_{OR}$ | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | $FG_{GC}$ | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Notes: The table reports the percentage of rejections of the null hypothesis in 1000 replications when three principal components of the large dataset are considered; $CH$ is the test proposed in this paper; $FG_{OR}$ is the orthogonality test proposed by Forni and Gambetti (2014); $FG_{GC}$ is the Granger causality test proposed by Forni and Gambetti (2014).

lags for all $c^*$, the performance would be less dramatic but still poor. For example, if $c^* = 0.4$, the rejection rate is 60.7 rather than 93.9 for the orthogonality test and 20.9 rather than 32.9 for the Granger causality test at the 5 percent nominal level.

Table 2 reports the empirical power. All tests are similarly powerful against local departures from the null, regardless of the confidence level. However, the CH test looses power as $\tilde{c}^*$ gets close to one. Thus, it is important to check the magnitude of the roots of the system before applying the test in empirical applications.

The additional tables in the appendix indicate that the size and the power of the CH test are insensitive to nuisance parameter problems. In particular, the lag length of the preliminary VAR does not matter, the number of lags of the principal components and of VAR residuals in the auxiliary regression is irrelevant, and the properties of the test are independent of the number of leads of VAR residuals we include and test for.

It is also interesting to examine how our procedure performs when there are no aggregation problems. The DGP is the same as before, but we now set $\tau_y = 0$. Table 3, which

Table 3: Testing for fundamental representation: No aggregation

| $\tilde{c}^*$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 |
|---|---|---|---|---|---|---|---|---|
| 1% | 0.6 | 0.7 | 0.8 | 0.8 | 0.9 | 1.2 | 1.1 | 1.0 |
| 5% | 5.4 | 5.4 | 5.5 | 5.6 | 5.5 | 5.2 | 6.6 | 7.3 |
| 10% | 11.4 | 10.7 | 10.4 | 10.7 | 10.7 | 10.5 | 11.6 | 12.0 |
| | | | | | | | | |
| $\tilde{c}^*$ | 10 | 5 | 3.33 | 2.5 | 2.0 | 1.66 | 1.43 | 1.25 |
| 1% | 100 | 100 | 100 | 100 | 100 | 100 | 89.4 | 68.6 |
| 5% | 100 | 100 | 100 | 100 | 100 | 100 | 93.8 | 70.1 |
| 10% | 100 | 100 | 100 | 100 | 100 | 100 | 99.4 | 85.4 |

reports the size and the power in this situation, indicates that the properties of the CH test are independent of whether aggregation is present or not.

# 5 Reconsidering two small scale VARs

Standard business cycle theories assume that economic fluctuations are driven by surprises in current fundamentals, such as aggregate productivity or demand disturbances. Motivated by the idea that changes in expectations about future fundamentals may drive business fluctuations, Beaudry and Portier (2006) find that positive news shocks have a positive impact on output, consumption, investment and hours worked. These responses are at odds with those obtained in a Real Business Cycle (RBC) model: here, positive news about future technology increase consumption, while hours, output and investment fall.

Since models featuring news shocks have solutions displaying moving average components, finite order models may be unable to capture the underlying dynamics, making the VARs considered in the literature prone to non-fundamentalness. In addition, Forni et al. (2014) provide a stylized Lucas tree model where predictable news to the dividend process may induce non-fundamentalness in a system comprising the growth rate of stock prices and the growth rate of dividends. The solution of their model, when news come two periods in advance, is:

$$\begin{bmatrix} \Delta d_t \\ \Delta p_t \end{bmatrix} = \begin{bmatrix} L^2 & 1 \\ \frac{\beta^2}{1-\beta} + \beta L & \frac{\beta}{1-\beta} \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \equiv C(L)u_t \tag{5.1}$$

where $d_t$ are dividends, $p_t$ are stock prices, $0 < \beta < 1$ is the discount factor. Here, since $|C(L)|$ vanishes for $L = 1$ and $L = -\beta$, $u_t$ is non-fundamental for $(\Delta d_t, \Delta p_t)$. Intuitively, this occurs because agents observe $u_{1t}$ at $t$, but the shock affects dividends at $t+2$. Thus, agents' information set, which includes current and past values of structural shocks, does not match the econometrician's information set, which includes current and past values of the growth rate of dividends and stock prices. As noted by Beaudry et al. (2015), the fundamental and non-fundamental dynamics this model generates are similar. This is because the root generating non-fundamentalness is near unity. In general, the properties of the solution depend on the details of the economy, on the process describing the information flows, and the variables observed by the econometrician.

We estimate a VAR with the growth rate of capacity adjusted TFP and of stock prices for the period $1960Q1$ to $2010Q4$, both of which are taken from Beaudry and Portier (2014) and we use the same principal components as in Forni et al. (2014). The VAR system has no root larger than 0.75; thus power problems are not relevant. Table 4 reports the $p-$values of the tests for different numbers of principal components, which enter in first difference in all the tests. In the CH test, the testing model has four lags of the PC and of the VAR residuals and we are examining the predictive power of 2 leads of the VAR residuals.

The $CH$ test finds the system fundamental, regardless of the number of PC included in the testing equations. In contrast, a Granger causality test soundly rejects the null of fundamentalness. Since the VAR includes TFP, which is latent and aggregated, differences in the conclusions could be due to aggregation or non-observability problems.

To verify this possibility we consider a VAR where in place of utilization adjusted aggregated TFP we consider two different utilization adjusted sectoral TFP measures. The first is obtained using the methodology of Basu et al. (2013), which produces time series for

Table 4: Testing fundamentalness: VAR with TFP and stock prices.

| | PC=3 | PC=4 | PC=5 | PC=6 | PC=7 | PC=8 | PC=9 | PC=10 |
|---|---|---|---|---|---|---|---|---|
| CH, sample 1960-2010 | | | | | | | | |
| | 0.12 | 0.13 | 0.28 | 0.48 | 0.23 | 0.11 | 0.18 | 0.04 |
| $FG_{GC}$, sample 1960-2010 | | | | | | | | |
| | 0.05 | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $FG_{GC}$, Fernand 1960-2005 data | | | | | | | | |
| Aggregate | 0.02 | 0.04 | 0.23 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 |
| Sectorial | 0.07 | 0.21 | 0.30 | 0.02 | 0.00 | 0.01 | 0.03 | 0.05 |
| $FG_{GC}$ Wang 1960-2009 data | | | | | | | | |
| Aggregate | 0.05 | 0.03 | 0.15 | 0.04 | 0.00 | 0.00 | 0.00 | 0.01 |
| Sectorial | 0.36 | 0.49 | 0.38 | 0.21 | 0.19 | 0.27 | 0.31 | 0.28 |

Notes: The table reports the p-value of the tests; $CH$ is the test proposed in this paper; $FG_{GC}$ is the Granger causality test proposed by Forni and Gambetti (2014); $PC$ is the number of principal component in the auxiliary regression. In CH test the number of leads tested is two.

private consumption TFP, private investment TFP, goverment consumption and investment TFP and ' net trade" TFP. The second panel of table 4 presents results obtained in a VAR which includes consumption TFP (obtained aggregating private and public consumption), investment TFP (obtained aggregating private and public investments) and net trade TFP, all in log growth rates and the growth rate of stock prices. Because the sample for this data terminates in 2005, the panel also reports p-values of a Granger causality test for the original bivariate system restricted to the 1960-2005 sample. As an alternative, we use the utilization adjusted industry TFP data constructed by Christina Wang at the Federal Reserve Bank of Boston. We reaggregate industry TFPs into manufacturing, services and 'others' sectors, and use the growth rate of these three variables together with the log of stock prices in the VAR [3]. The third panel of table 4 presents results obtained with this VAR. Because the sample ends in 2009, the panel also reports p-values of a Granger causality test for the bivariate system restricted to the 1960-2009 sample.

A Granger causality test applied to the original bivariate system estimated over the two

[3]Because the input of both methodologies are annual series, we interpolate TFP series from annual to quartely using a polynominal regression procedure.

new samples confirms that the system is non-fundamental. However, when it is applied to the VAR with the first set of sectorial TFP measures, the null of non-fundamentalness is not rejected in four of the eight cases and, for all choices of PC, the p-values increase relative to the baseline system with only aggregated TFP. When the test is applied to the VAR with the second set of sectorial TFP measures, the system turns out to be fundamental. Since this holds when we enter the three sectoral TFP variables in level rather than growth rates, when we allow for a break in the TFP series, and when we exclude the 'others' sector TFP variable from the VAR, the conclusion is that a Granger causality diagnostic rejects the null in the original VAR because of aggregation problems. Instead, the diagnostic of this paper, being robust to aggregation problems, correctly identifies the original small scale VAR as fundamental. Nevertheless, one should remember that even if the original system is fundamental, its shocks and dynamics are likely to be complicated averages of the shocks and dynamics of the primitive economy, which surely includes more than two disturbances.

The seminal paper of Galí (1999) investigated the comovements of labour productivity and hours in response to technology shocks. Since a negative conditional correlation is found, and since a textbook RBC model predicts a positive conditional correlation, the author concludes that technology shocks can not be the main sources of macroeconomic fluctuations. The paper has spurred a large literature refuting, see e.g. Christiano et al. (2003), or confirming the conclusions, see e.g. Canova et al. (2010). However, since small scale VARs are typically used, omitted variables, non-observability, or anticipatory information about productivity developments could potentially induce non-fundamentalness.

We focus on the bivariate specification used in the literature, which include the growth rate of the log of labor productivity ($\Delta x_t$) and of the log of hours ($\Delta n_t$):

$$\begin{bmatrix} \Delta x_t \\ \Delta n_t \end{bmatrix} = \begin{bmatrix} C_{11}(L) & C_{12}(L) \\ C_{21}(L) & C_{22}(L) \end{bmatrix} \begin{bmatrix} u_t^z \\ u_t^m \end{bmatrix} = C(L) \begin{bmatrix} u_t^z \\ u_t^m \end{bmatrix} \tag{5.2}$$

where $u_t^z$ and $u_t^m$ are assumed to be orthogonal at all leads and lags and typically interpreted as technology and non-technology shocks, respectively. Fundamentalness is imposed assuming that the determinant of $C(L)$ has no roots inside the unit circle. Our task is to check whether such an assumption is reasonable.

The sample $1960Q1$ to $2010Q4$ is used to match the time span that our large dataset covers. The labor input is the log of US hours of all persons in the nonfarm business sector; labor productivity is constructed from this variable and the log of US real GDP. The largest root of the VAR is 0.685, so that power problems are not relevant also in this case. To conduct the tests, we use the same principal components used in the previous VAR.

Table 5 reports the results. Different columns represent the number of principal components used in the testing equation which enter in first difference in all tests. Both CH and Granger causality tests find the system non-fundamental. Thus, aggregation/non observability problems do not seem relevant.

Recall that even if the system is non-fundamental, we may be able to recover technology shocks, which is all that matter here. The technology disturbance is identified as the only disturbance with a permanent impact on labor productivity, i.e., $C_{12}(1) = 0$. The second panel of table 5 reports the results of applying our test and the orthogonality test to the technology shock recovered this way. As mentioned, we are not testing for fundamentalness here; but only if there is enough information in the data to recover the technology shock. Both tests reject also this possibility.

To verify whether the selection of the variables included in the VAR matters, we repeat the latter exercise using a VAR with the growth rate of labor productivity and the unemployment rate. The third panel of table 5 indicates that in a number of cases technology shocks are now recoverable from current and past values of the observables. Thus, it seems that it is the use total hours worked in the VAR that causes problems.

In a representative agent RBC model what matters, however, is total hours, not the unemployment rate. Thus, if one is interested in using the data to evaluate this model, a system with the unemployment rate is far from ideal. Is there any way to add variables to

Table 5: Testing fundamentalness: VAR with labor productivity and hours

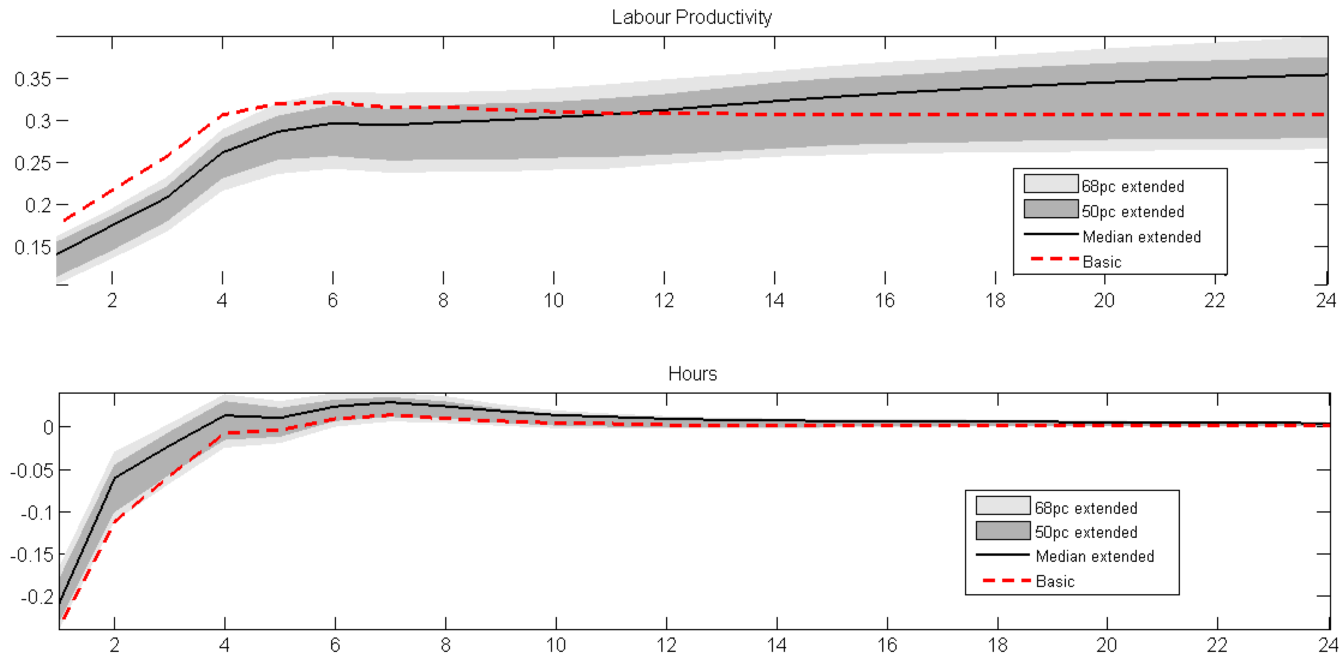
| | PC=3 | PC=4 | PC=5 | PC=6 | PC=7 | PC=8 | PC=9 | PC=10 |
|---|---|---|---|---|---|---|---|---|
| $\Delta(y/n), \Delta n$ System | | | | | | | | |
| CH | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $FG_{GC}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $\Delta(y/n), \Delta n$ Technology shock | | | | | | | | |
| CH | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $FG_{OR}$ | 0.06 | 0.02 | 0.04 | 0.03 | 0.05 | 0.02 | 0.03 | 0.03 |
| $\Delta(y/n), \Delta u$ Technology shock | | | | | | | | |
| CH | 0.15 | 0.02 | 0.02 | 0.05 | 0.14 | 0.07 | 0.07 | 0.07 |
| $FG_{OR}$ | 0.32 | 0.15 | 0.04 | 0.09 | 0.16 | 0.01 | 0.02 | 0.02 |
| $\Delta(y/n), \Delta n, \Delta CI$ Technology shock | | | | | | | | |
| CH | 0.56 | 0.14 | 0.11 | 0.24 | 0.23 | 0.40 | 0.46 | 0.38 |
| $FG_{OR}$ | 0.50 | 0.38 | 0.37 | 0.52 | 0.59 | 0.51 | 0.63 | 0.59 |

Notes: The table reports the p-value of the tests; $CH$ is the test proposed in this paper; $FG_{GC}$ is the Granger causality test proposed by Forni and Gambetti (2014); $PC$ is the number of principal component in the testing equations. In CH test the number of leads tested is 2. The extended system includes the growth rate of the consumer confidence indicator.

the original system so that current and past values of the observables allow us to recover the technology shocks? To study this possibility, we include the growth rate of a confidence indicator in the VAR. The last panel of table 5 shows that a technology shock recovered from this trivariate system is indeed a function of the current and past values of the observables. Since the confidence indicator proxies for the current (and future) state of the economy, crucial information may be missing from the original VAR.

We plot in figure 1 the responses of labor productivity and hours to technology shocks obtained in bivariate and trivariate models. While the responses have similar shape, those obtained in the bivariate system tend to overestimate the initial negative response of hours and the initial positive response of labor productivity.

Figure 1: Responses to technology shocks




*Note:* The shaded regions report pointwise 50 and 68% credible intervals. The $x$-axis reports quarters, the $y$-axis the response of the level of the variable in deviation from the predictable path.

# 6 Conclusions

Small scale VAR models are typically used in empirical business cycle analyses. Thus, the structural shocks which are obtained are linear transformations of primitive structural shocks. Aggregate shocks might be fundamental or non-fundamental, depending on the details of the economy, the information set, and the variables chosen in the empirical analysis. If one observes some variables that provide noisy information about the primitive shocks, Granger causality (orthogonality) tests will reject the null of no causality (no orthogonality), even when the shocks are fundamental. A similar problem arises when VAR variables proxy for latent variables. A simulation study illustrates that *spurious non-fundamentalness* occurs when primitive shocks are aggregated and a small scale VAR is used.

We propose an alternative procedure which has the same power properties as existing diagnostics when non-fundamentalness is present, but does not face aggregation or non-observability problems when the system is fundamental. We also show that the procedure is robust to specification issues and to nuisance features.

We show that a Granger causality diagnostic finds that a bivariate model measuring the impact of TFP news is non-fundamental, while our test finds it fundamental. We show that the use of an aggregated TFP measure in the VAR explains the discrepancy. When sectorial TFP measures are used, a Granger causality diagnostic also finds the VAR fundamental.

We also show that both tests find a bivariate VAR system with labour productivity growth and hours growth non-fundamental and that the technology shocks can not be recovered from current and past values of the observables. This last conclusion changes when the growth rate of a confidence indicator enters the VAR. Hence, lack of information may be the reason for why technology shocks can't be obtained from the original system.

A few lessons that can be learned from our paper. First, Granger causality tests may give misleading conclusions when testing for fundamentalness whenever aggregation or non-observability problems are present. The test proposed in this paper is robust to these issues. Second, tests focusing on omitted variable problems may fail to identify the properties of small scale systems when the source of non-fundamentalness is theoretical or due to investigators' choice. Third, to derive reliable conclusions, one should have fundamentalness tests that are insensitive to specification and nuisance features. The test proposed in this paper satisfies both criteria; those present in the literature do not. Finally, to make business cycle analysis advance on solid ground, it is important to check that the VAR system contains enough information to recover structural shocks from current and past values of the observables. The test proposed in this paper facilitates this ”reality” check.

# References

Susanto Basu, John Fernald, Jonas Fisher, and Miles Kimball. Sector-specific technical change. 2013.

Paul Beaudry and Franck Portier. Stock prices, news, and economic fluctuations. *American Economic Review*, 96(4):1293–1307, 2006.

Paul Beaudry and Franck Portier. News-driven business cycles: Insights and challenges. *Journal of Economic Literature*, 52(4):993–1074, 2014.

Paul Beaudry, Patrick Feve, Alain Guay, and Franck Portier. When is non-fundamentalness in vars a real problem? an application to news shocks. Technical report, University of Toulouse, 2015.

Fabio Canova. *Methods for Applied Macroeconomic Research.* Princeton University Press, 2011.

Fabio Canova, Claudio Michelacci, and David Lopez Salido. The effects of technology shocks on hours and output: A robustness analysis. *Journal of applied Econometrics*, 25(5):775–773, 2010.

Yongsung Chang and Jay H. Hong. Do technological improvements in the manufacturing sector raise or lower employment? *American Economic Review*, 96(1):352–368, 2006.

Larry Christiano, Martin Eichenbaum, and Charles Evans. Monetary policy shocks: What have we learned and to what end? *Handbook of Macroeconomics, volume 1*, pages 65–145, 1999.

Lawrence J Christiano, Martin Eichenbaum, and Robert Vigfusson. What happens after a technology shock? Technical report, National Bureau of Economic Research, 2003.

Jess Fernández-Villaverde, Juan F. Rubio-Ramírez, Thomas J. Sargent, and Mark W. Watson. ABCs (and Ds) of Understanding VARs. *American Economic Review*, 97(3):1021–1026, 2007.

Mario Forni and Luca Gambetti. Sufficient information in structural vars. *Journal of Monetary Economics*, 66:124–136, 2014.

Mario Forni, Luca Gambetti, and Luca Sala. No news in business cycles. *The Economic Journal*, 124(581):1168–1191, 2014.

Jordi Galí. Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations? *American Economic Review*, 89(1):249–271, 1999.

Sarah Gelper and Christophe Croux. Multivariate out-of-sample tests for granger causality. *Computational statistics & data analysis*, 51(7):3319–3329, 2007.

Raffaella Giacomini. The relationship between dsge and var models. *Advances in Econometrics*, 31, 2013.

Domenico Giannone and Lucrezia Reichlin. Does information help recovering structural shocks from past observations? *Journal of the European Economic Association*, 4(2-3): 455–465, 2006.

Christian Gourieroux and Alain Monfort. Revisiting identification and estimation in structural varma models. Technical report, CREST, 2015.

Mehdi Hamidi Saneh. Are shocks obtained from vars fundamental? Technical report, Universidad Carlos III, 2014.

James Douglas Hamilton. *Time series analysis*. Princeton university press Princeton, 1994.

Lars Hansen and Robert Hodrick. Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *Journal of Political Economy*, 88:829–853, 1980.

Lars Peter Hansen and Thomas J Sargent. Formulating and estimating dynamic linear rational expectations models. *Journal of Economic Dynamics and Control*, 2:7–46, 1980.

Lars Peter Hansen and Thomas J Sargent. Two difficulties in interpreting vector autoregressions. *Rational expectations econometrics*, pages 77–120, 1991.

Eric M Leeper, Todd B Walker, and Shu-Chun Susan Yang. Fiscal foresight and information flows. *Econometrica*, 81(3):1115–1145, 2013.

Marco Lippi and Lucrezia Reichlin. The dynamic effects of aggregate demand and supply disturbances: Comment. *The American Economic Review*, 83(3):644–652, 1993.

Marco Lippi and Lucrezia Reichlin. Var analysis, nonfundamental representations, blaschke matrices. *Journal of Econometrics*, 63(1):307–325, 1994.

Helmut Lütkepohl. Fundamental problems with nonfundamental shocks. Technical report, DIW Berlin, German Institute for Economic Research, 2012.

Federico Ravenna. Vector autoregressions and reduced form representations of dsge models. *Journal of monetary economics*, 54(7):2048–2064, 2007.

Y. A. Rozanov. *Stationary random processes.* Holden-Day, 1967.

# Appendix

This appendix reports tables with the results of the Monte Carlo exercise using the CH test. From the baseline values, the three nuisance parameters: $r$, the number of lags in the preliminary VAR; $p$, the number of lags of the PC and the VAR residuals in the auxiliary regression; and $q$, the number of leads of the VAR residuals tested in the null.

Table 6: Testing for fundamental representation: Size

| | $c^*$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 |
|---|---|---|---|---|---|---|---|---|---|
| Preliminary VAR,$r = 2$ lags | | | | | | | | | |
| | 1% | 1.2 | 1.2 | 1.1 | 1.1 | 0.9 | 0.8 | 1.2 | 1.7 |
| $CH$ | 5% | 5.1 | 4.3 | 4.6 | 4.7 | 5.2 | 5.4 | 5.9 | 7.0 |
| | 10% | 10.6 | 9.9 | 9.8 | 9.7 | 9.8 | 10.1 | 10.8 | 11.8 |
| Preliminary VAR,$r = 3$ lags | | | | | | | | | |
| | 1% | 1.1 | 1.2 | 0.9 | 0.8 | 0.8 | 0.9 | 1.2 | 1.4 |
| $CH$ | 5% | 5.3 | 4.6 | 4.1 | 4.4 | 4.6 | 4.5 | 5.2 | 6.2 |
| | 10% | 11.0 | 11.0 | 9.7 | 9.9 | 11.0 | 11.2 | 12.2 | 13.4 |
| Preliminary VAR, $r = 5$ lags | | | | | | | | | |
| | 1% | 1.1 | 0.9 | 0.9 | 0.9 | 1.0 | 1.1 | 1.2 | 1.6 |
| $CH$ | 5% | 5.7 | 5.1 | 4.5 | 5.0 | 4.8 | 5.1 | 5.5 | 5.8 |
| | 10% | 10.9 | 9.8 | 9.4 | 9.5 | 9.7 | 9.6 | 9.9 | 10.5 |
| Preliminary VAR, $r = 6$ lags | | | | | | | | | |
| | 1% | 1.0 | 0.8 | 0.8 | 0.9 | 0.9 | 1.1 | 1.3 | 1.2 |
| $CH$ | 5% | 5.4 | 4.6 | 4.5 | 4.4 | 4.0 | 4.1 | 4.4 | 4.9 |
| | 10% | 10.6 | 11.0 | 9.3 | 9.2 | 9.0 | 9.0 | 9.6 | 10.0 |

Table 7: Testing for fundamental representation: Power

| | $\tilde{c}^*$ | 10 | 5 | 3.33 | 2.5 | 2.0 | 1.66 | 1.43 | 1.25 |
|---|---|---|---|---|---|---|---|---|---|
| Preliminary VAR,$r = 2$ lags | | | | | | | | | |
| $CH$ | 1% | 100 | 100 | 100 | 100 | 100 | 100 | 99.2 | 56.8 |
| | 5% | 100 | 100 | 100 | 100 | 100 | 100 | 99.8 | 78.2 |
| | 10% | 100 | 100 | 100 | 100 | 100 | 100 | 99.9 | 85.8 |
| Preliminary VAR $r = 3$ lags | | | | | | | | | |
| $CH$ | 1% | 100 | 100 | 100 | 100 | 100 | 100 | 97.6 | 40.0 |
| | 5% | 100 | 100 | 100 | 100 | 100 | 100 | 99.6 | 63.0 |
| | 10% | 100 | 100 | 100 | 100 | 100 | 100 | 99.9 | 74.4 |
| Preliminary VAR, $r = 5$ lags | | | | | | | | | |
| $CH$ | 1% | 100 | 100 | 100 | 100 | 100 | 100 | 98.7 | 41.1 |
| | 5% | 100 | 100 | 100 | 100 | 100 | 100 | 99.9 | 64.1 |
| | 10% | 100 | 100 | 100 | 100 | 100 | 100 | 99.9 | 76.2 |
| Preliminary VAR $r = 6$ lags | | | | | | | | | |
| $CH$ | 1% | 100 | 100 | 100 | 100 | 100 | 100 | 98.4 | 42.6 |
| | 5% | 100 | 100 | 100 | 100 | 100 | 100 | 99.7 | 67.4 |
| | 10% | 100 | 100 | 100 | 100 | 100 | 100 | 99.9 | 76.6 |

Table 8: Testing for fundamental representation: Size

| | $c^*$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 |
|---|---|---|---|---|---|---|---|---|---|
| Conditioning on 4 lags of PC, testing $q = 2$ leads | | | | | | | | | |
| $CH$ | 1% | 0.8 | 1.2 | 1.0 | 0.8 | 1.0 | 0.9 | 1.1 | 1.1 |
| | 5% | 5.5 | 5.6 | 5.0 | 4.9 | 4.7 | 4.7 | 5.3 | 5.5 |
| | 10% | 10.8 | 10.6 | 9.1 | 9.4 | 9.3 | 9.9 | 10.3 | 10.4 |
| Conditioning on 4 lags of PC, testing $q = 4$ leads | | | | | | | | | |
| $CH$ | 1% | 1.0 | 1.0 | 1.0 | 1.0 | 0.9 | 1.3 | 1.3 | 1.5 |
| | 5% | 6.5 | 6.1 | 6.1 | 6.4 | 5.8 | 6.1 | 6.2 | 6.6 |
| | 10% | 12.6 | 12.0 | 12.1 | 11.6 | 11.6 | 12.1 | 12.4 | 13.0 |
| Conditioning on 4 lags of PC, testing $q = 6$ leads | | | | | | | | | |
| $CH$ | 1% | 1.1 | 1.2 | 1.1 | 1.0 | 1.2 | 1.2 | 1.5 | 1.5 |
| | 5% | 7.2 | 6.5 | 5.8 | 5.5 | 5.4 | 5.7 | 5.8 | 6.2 |
| | 10% | 12.3 | 11.8 | 11.1 | 11.0 | 11.5 | 12.3 | 11.3 | 11.3 |

Table 9: Testing for fundamental representation: Power

| | $c^*$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 |
|---|---|---|---|---|---|---|---|---|---|
| Conditioning on 4 lags of PC, testing $q = 2$ leads | | | | | | | | | |
| $CH$ | 1% | 100 | 100 | 100 | 100 | 100 | 100 | 98.4 | 39.8 |
| | 5% | 100 | 100 | 100 | 100 | 100 | 100 | 99.8 | 62.9 |
| | 10% | 100 | 100 | 100 | 100 | 100 | 100 | 99.9 | 75.0 |
| Conditioning on 4 lags of PC, testing $q = 4$ leads | | | | | | | | | |
| $CH$ | 1% | 100 | 100 | 100 | 100 | 100 | 100 | 99.8 | 49.5 |
| | 5% | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 73.0 |
| | 10% | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 82.3 |
| Conditioning on 4 lags of PC, testing $q = 6$ leads | | | | | | | | | |
| $CH$ | 1% | 100 | 100 | 100 | 100 | 100 | 100 | 99.6 | 47.1 |
| | 5% | 100 | 100 | 100 | 100 | 100 | 100 | 99.8 | 72.6 |
| | 10% | 100 | 100 | 100 | 100 | 100 | 100 | 99.9 | 82.2 |

Table 10: Testing for fundamental representation: Size

| | $c^*$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 |
|---|---|---|---|---|---|---|---|---|---|
| Conditioning on 3 lags of PC and 4 lags of $u_t$, testing $q = 2$ leads | | | | | | | | | |
| $CH$ | 1% | 0.7 | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 | 1.0 | 1.2 |
| | 5% | 5.3 | 5.1 | 5.0 | 4.8 | 4.4 | 4.7 | 4.8 | 5.5 |
| | 10% | 11.1 | 10.7 | 9.9 | 9.5 | 9.7 | 10.1 | 10.4 | 10.9 |
| Conditioning on 4 lags of PC and 4 lags of $u_t$, testing $q = 2$ leads | | | | | | | | | |
| $CH$ | 1% | 0.8 | 1.2 | 1.0 | 0.8 | 1.0 | 0.9 | 1.1 | 1.1 |
| | 5% | 5.5 | 5.5 | 5.0 | 4.9 | 4.7 | 4.7 | 5.3 | 5.5 |
| | 10% | 10.8 | 10.9 | 9.1 | 9.4 | 9.3 | 9.9 | 10.1 | 10.4 |
| Conditioning on 5 lags of PC and 4 lags of $u_t$, testing $q = 2$ leads. | | | | | | | | | |
| $CH$ | 1% | 1.2 | 1.2 | 1.0 | 1.0 | 1.0 | 0.9 | 1.0 | 1.0 |
| | 5% | 6.0 | 5.5 | 4.5 | 4.8 | 5.4 | 4.1 | 4.9 | 6.0 |
| | 10% | 11.4 | 11.0 | 9.8 | 9.8 | 10.2 | 10.4 | 10.3 | 11.6 |

Table 11: Testing for fundamental representation: Power

| | $c^*$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 |
|---|---|---|---|---|---|---|---|---|---|
| Conditioning on 3 lags of PC and 4 lags of $u_t$, testing $q = 2$ leads | | | | | | | | | |
| | 1% | 100 | 100 | 100 | 100 | 100 | 100 | 98.7 | 40.2 |
| $CH$ | 5% | 100 | 100 | 100 | 100 | 100 | 100 | 99.8 | 64.0 |
| | 10% | 100 | 100 | 100 | 100 | 100 | 100 | 99.9 | 76.0 |
| Conditioning on 4 lags of PC and 4 lags of $u_t$, testing $q = 2$ leads. | | | | | | | | | |
| | 1% | 100 | 100 | 100 | 100 | 100 | 100 | 98.4 | 39.8 |
| $CH$ | 5% | 100 | 100 | 100 | 100 | 100 | 100 | 99.8 | 62.9 |
| | 10% | 100 | 100 | 100 | 100 | 100 | 100 | 99.9 | 75.0 |
| Conditioning on 5 lags of PC and 4 lags of $u_t$ , testing $q = 2$ leads. | | | | | | | | | |
| | 1% | 100 | 100 | 100 | 100 | 100 | 100 | 96.9 | 32.4 |
| $CH$ | 5% | 100 | 100 | 100 | 100 | 100 | 100 | 99.6 | 56.0 |
| | 10% | 100 | 100 | 100 | 100 | 100 | 100 | 99.9 | 67.5 |