Kent Academic Repository

Full text document (pdf)

Citation for published version

Canova, Fabio and Hamidi Sahneh, Mehdi (2017) Are Small-Scale SVARs Useful for Business Cycle Analysis? Revisiting Nonfundamentalness. Journal of the European Economic Association . ISSN 1542-4766.

DOI

https://doi.org/10.1093/jeea/jvx032

Link to record in KAR

http://kar.kent.ac.uk/64222/

Document Version

Author's Accepted Manuscript

Copyright & reuse
Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research
The version in the Kent Academic Repository may differ from the final published version.
Users are advised to check **http://kar.kent.ac.uk** for the status of the paper. **Users should always cite the published version of record.**

Enquiries
For any further enquiries regarding the licence status of this document, please contact:
researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at **http://kar.kent.ac.uk/contact.html**

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

Fabio Canova
BI Norwegian Business School, and CEPR *

Mehdi Hamidi Sahneh
Universidad Carlos III †

April 20, 2017

Abstract

Non-fundamentalness arises when current and past values of the observables do not contain enough information to recover SVAR disturbances. Using Granger causality tests, the literature suggested that several small scale SVAR models are non-fundamental and thus not necessarily useful for business cycle analysis. We show that causality tests are problematic when SVAR variables cross sectionally aggregate the variables of the underlying economy or proxy for non-observables. We provide an alternative testing procedure, illustrate its properties with Monte Carlo simulations, and re-examine a prototypical small scale SVAR model.

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

JEL classification: C5, C32, E5.

*Centre for Applied Macro and 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 the editor (Claudio Michelacci), five anonymous referees, Carlos Velasco, Jesus Gonzalo, Hernan Seoane, Gianni Amisano, Jordi Gali, Davide de Bortoli, Luca Sala, Benjamin Holcblat, Domenico Giannone, Lutz Kilian, Luca Gambetti and Mario Forni and the participants of seminars at BI Norwegian Business School, UPF, Bocconi, Humboldt, Luiss, University of Glasgow, University of Helsinki, and Federal Reserve Bank of New York, Federal Reserve Board, and of the 2016 IAAE and 2016 ESEM conferences for comments and discussions. Canova acknowledges the financial support from the Spanish Ministerio de Economia y Competitividad through the grants ECO2012-33247 and ECO2015-68136-P and FEDER, UE.

1 Introduction

Structural Vector Autoregressive (SVAR) models have been extensively used over the last 30 years to study sources cyclical fluctuations . The methodology hinges on the assumption that structural shocks can be obtained from linear combinations of current and past values of the observables. Non-fundamentalness arises when this is not the case. In a non-fundamental system, structural 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 (2007), 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, this choice 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.

Despite the far-reaching implications it has for applied work, little is known on how to empirically detect non-fundamentalness. Following the lead of Lutkepohl (1991), Giannone and Reichlin (2006) and Forni and Gambetti (2014) (henceforth, FG) suggest 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 in macroeconomics, this result is disturbing. This paper shows that Granger causality diagnostics may lead to spurious results in common and relevant situations.

Why are there problems? Because of small samples, instabilities, identification or interpretation difficulties, one typically uses a small scale SVAR to examine the transmission of relevant disturbances, even if the process generating the data (DGP) features many more variables and shocks. But the shocks recovered by such SVAR systems are linear combinations of a potentially larger set of primitive structural shocks driving the economy. Thus, any variable excluded from the SVAR, but containing information about these primitive disturbances, predicts SVAR shocks (and thus Granger cause the endogenous variables), regardless of whether the model is fundamental or not.

To illustrate the point, suppose we want to measure the effects of technology shocks on economic activity. Small scale SVARs designed for this purpose typically include an aggregate measure of labour productivity, hours, and a few other aggregate variables. Suppose that what drives the economy are sector-specific, serially correlated productivity disturbances. The technology shock recovered from an SVAR will be a linear transformation of current and past sectoral productivity shocks. Since, e.g., sectoral capital or sectoral labour productivity have information about sectoral disturbances, they will predict SVAR technology shocks, both when the model is fundamental and when it is not.

A similar problem occurs when the SVAR features a proxy variable. For example, TFP is latent and typical estimates are obtained from output, capital and hours worked data. If capital and hours worked are excluded from the SVAR, any variable that predicts them will Granger cause estimated TFP, regardless of whether the model is fundamental or not.

In general, whenever a small scale SVAR is used, aggregation rather than non-fundamentalness may be the reason for why Granger causality tests find predictability. Thus, if non-fundamentalness is of interest, it is crucial to have a testing approach which is robust to aggregation and non-observability problems. We propose an alternative procedure, based on ideas of Sims (1972), which is has this property and exploits the fact that, under non-fundamentalness, future SVAR shocks predict a vector of variables excluded from the SVAR.

We perform Monte Carlo simulations using a version of the model of Leeper et al. (2013) as DGP with capital tax, income tax, and productivity disturbances. We assume that the

SVAR includes capital and an aggregate tax variable (or an aggregate tax rate computed from revenues and output data) and show that our approach has good small sample properties. In contrast, spurious non-fundamentalness arises with standard diagnostics. Absent aggregation problems, our approach and a Granger causality test have similar small sample properties.

We re-examine the small scale SVAR employed by Beaudry and Portier (2006) designed to measure the macroeconomic effects of news. We find that the model is fundamental according to our test but non-fundamental according to a Granger causality diagnostic. We show that the rejection of the null with the latter is due to aggregation: once coarsely disaggregated TFP data is used in the SVAR, Granger causality no longer rejects the null of fundamentalness. The dynamics responses to news shocks in the systems with aggregated and disaggregated TFP measures are however similar (see also Beaudry et al. (2015)). Thus, the SVAR disturbances the two systems recover are likely to be similar combinations of the primitive structural shocks and, thus, not necessarily economically interpretable.

Two caveats need to be mentioned. First, our analysis is concerned with Gaussian macroeconomic variables. For non-Gaussian situations, see Hamidi Saneh (2014) or Gourieroux and Monfort (2015). Second, although we focus on SVARs, our procedure also works for SVARMA models, as long as the largest MA root is 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 problem occurs. Section 3 shows why standard tests may fail and propose an alternative approach. Section 4 examines the performance of various procedures using Monte Carlo simulations. Section 5 investigates the properties of a small scale SVAR system. Section 6 concludes.

2 A few example of non-fundamental systems

As Kilian and Lutkepohl (2016) highlighted, the literature has primarily focused on non-fundamentalness driven by a mismatch between agents and econometricians information sets, because of omitted variables (see e.g. Giannone and Reichlin (2006), Kilian and Murphy (2014)), or of the timing of news revelation (see e.g. Leeper et al. (2013) , Forni et al. (2014)). However, there may be other reasons for why it emerges.

First, non-fundamentalness may be intrinsic to the optimization process and to the modelling choices an investigator makes, see e.g. Hansen and Sargent, 1991). Optimizing models producing non-fundamental solutions are numerous; the next example shows one.

Example 1. Suppose the dividend process is $d_t = e_t - a\mathrm{e}_{t-1}$, where $a < 1$, and suppose stock prices are expected discounted future dividends: $p_t = E_t \sum_j \beta^j d_{t+j}$, $0 < \beta < 1$. The equilibrium value of p_t in terms of the dividends innovations is

$$p_t = (1 - \beta a)e_t - ae_{t-1} \tag{2.1}$$

Thus, even though the dividends process is fundamental ($a < 1$), the process for stock prices could be non-fundamental if $|\frac{(1-\beta a)}{a}| < 1$, which occurs when $\frac{1}{1+\beta} < a$. If $a \geq 0.5$, any economically reasonable value of β will make stock prices non-fundamental. On the other hand, if we allow stock prices to have a bubble component e_t^b whose expected value is zero, the vector (e_t, e_t^b) is fundamental for (d_t, p_t), regardless of the value of β. Thus, allowing for bubbles in theory makes a difference as far as recovering dividend shocks from the data. $\square$

Second, non-fundamentalness may be due to non-observability of some of the endogenous variables of a fundamental model. 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:

$$K_t = (1 - \delta)K_{t-1} + ae_t \tag{2.3}$$

If both (K_t, Y_t) are observable this is just a bivariate restricted VAR(1) and e_t is fundamental for both (k_t, y_t). However, if the capital stock is unobservable, (2.2) becomes

$$Y_t - (1-\delta)Y_{t-1} = (1+a)e_t + (1-\delta)e_{t-1} \tag{2.4}$$

Clearly, if $a < 0$ and $|a| < |\delta|$, e_t can not be expressed as a convergent sum of current and past values of Y_t and (2.4) is non-fundamental. In addition, if δ and a are both small, (2.4) has a MA root close to unity and a finite order VAR for Y_t poorly approximates the underlying bivariate process; see also Ravenna (2007), and Giacomini (2013). $\square$

Third, a particular variable selection may induce non-fundamentalness, even if the system is, in theory, fundamental. Hansen and Hodrick (1980) showed that this happens when forecast errors are used in a VAR. The next example shows a less known situation.

Example 3. Consider a standard consumption-saving problem. Let income $Y_t = e_t$ be a white noise. Let $\beta = \frac{1}{R} < 1$ be the discount factor 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 we setup the empirical model in terms of savings, $S_t \equiv Y_t - C_t$, 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. $\square$

In sum, there may be many reasons for why an empirical model may be non-fundamental. Assuming away non-fundamentalness is problematic. Focusing on omitted variable or anticipation problems is, on the other hand, reductive. One ought to have procedures able to detect whether a SVAR is fundamental and, if it is not, whether violations are intrinsic to theory or due to applied investigators choices.

3 The Setup

Because in this section we need to distinguish the structural disturbances driving the fluctuations in the DGP from the shocks a SVAR may recover, we use the convention that " primitive" structural shocks are the disturbances of the DGP and "SVAR" structural shocks those obtained with the empirical model.

We assume that the DGP for the observables can be represented by an n-dimensional vector of stationary variables χ_t driven by $s \geq n$ serial and mutually uncorrelated primitive structural shocks ς_t.

Assumption 1. The vector χ_t satisfies

$$\chi_t = \Gamma(L) C \varsigma_t$$

where C is a $n \times s$ matrix, $\Gamma(L) = \sum_{i=0}^{\infty} \Gamma_i L^i$, $\Gamma_0 = I$, Γ_i's are $(n \times n)$ matrices each i, L is the lag operator, and $\sum_{i=0}^{\infty} \Gamma_i^2 < \infty$.

The DGP in (3) is quite general and covers, for example, stationary dynamics general equilibrium (DSGE) models solved around a deterministic steady state or non-stationary DSGEs solved around a deterministic or a stochastic balanced growth path. Stationarity is assumed for convenience; the arguments we present are independent of whether χ_t stochastically drifts or not. Assumption 1 places mild restrictions on the roots of $\Gamma(L)$. In theory, ς_t could be fundamental for χ_t or not.

Given a typical sample, n the dimension of χ_t is generally large and $\Gamma(L)$ is of infinite dimension. Thus, for estimation and inferential purposes an applied investigator typically confines attention to an m-dimensional vector x_t, where $\mathcal{H}_t^x \subset \mathcal{H}_t^{\chi}$, and $\mathcal{H}_t^j$ is the closed linear span of $\{j_s : s \leq t\}, j_t = (x_t, \chi_t)$ [1].

Assumption 2. The vector x_t is driven by a $m \times 1$ vector of mutually and serially

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

uncorrelated SVAR structural shocks $\varsigma_{x,t}$:

$$\begin{aligned} x_t &= \Gamma_x(L) C_x \varsigma_t \qquad (3.1) \\ &\equiv \Pi(L) u_t = \Pi(L) D \varsigma_{x,t} \qquad (3.2) \end{aligned}$$

where $m < n$, $\Gamma_x(L)$ is an $m \times m$ matrix for every L, C_x is a also a $m \times s$ matrix, $\Pi(L) = \sum_{i=0}^{\infty} \Pi_i L^i$, Π_i are $m \times m$ matrices for each i, $\Pi_0 = I$, $\sum_{i=0}^{\infty} \Pi_i^2 < \infty$, D is an $m \times m$ matrix.

Equation (3.1) covers many cases of interest in macroeconomics. For example, x_t may contain a subset of the variables belonging to χ_t, linear combinations, regression residuals, or forecast errors computed from the elements of χ_t. Thus, the framework includes the case of a variable belonging to the DGP but unobserved and thus omitted from the empirical model (as in example 2); the situation where the DGP has disaggregated variables but the empirical model is set up in terms of aggregated variables; the case where the DGP has an unobservable variable (e.g. total factor productivity) proxied by a linear combination of observables (i.e. output, capital and labor); and the case where all DGP variables are observables (e.g., we have consumption data) but the empirical model contains linear combinations of the observables (i.e. savings as in example 3).

Since the dimension of ς_t is larger than the dimension of x_t, cross-sectional aggregation occurs. That is, the econometrician estimating an SVAR may be able to recover the $m \times 1$ vector $\varsigma_{x,t}$ from the reduced form residuals u_t, but never the $s \times 1$ vector ς_t. For example, the DGP may describe a small open economy subject to external shocks coming from many countries, while the empirical model is specified so that only rest of the world variables are used. If $\Gamma(L)$ has a block exogenous structure, it may be possible to aggregate the vector external shocks into one shock without contamination from other disturbances, see e.g. Faust and Leeper (1988). However, even in this case, it is clearly impossible to recover the full vector of country specific external disturbances.

Next, we provide the definition of fundamentalness for the empirical model (3.2) (see also

Rozanov (1967)) and Alessi et al. (2011)).

Definition 1: An uncorrelated process $\{u_t\}$ is x_t-fundamental if $\mathcal{H}_t^u = \mathcal{H}_t^x$ for all t. It 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.

The empirical 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. Alternatively, the model is fundamental if it is possible to express u_t as a convergent sum of current and past x_t's. 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.

In standard situations, there is a one-to-one mapping between the u_t and ς_t and thus examining the fundamentalness of u_t provides information about the fundamentalness of ς_t. When the mapping is not one-to-one but the relationship between u_t and ς_t has a particular structure, it may be possible to find conditions insuring that when u_t is fundamental for x_t, ς_t is fundamental for χ_t, see e.g. Forni et al. (2009). In all other situations, many of which are of interest, knowing the properties of u_t for x_t may tells us little about the properties of the primitive shocks ς_t for χ_t.

Note that, although $\varsigma_{x,t}$ are linear combination of ς_t, they may still be economically interesting. An aggregate TFP shock may be meaningful, even if the sectoral TFP shocks drive the economy, as long as several sectoral TFP disturbances produce similar dynamics for the variables of the SVAR. On the other hand, it is not generally true that a fundamental shock is necessarily structurally interpretable (this occurs, for example, when the wrong D matrix is used to recover $\varsigma_{x,t}$ from a fundamental u_t).

3.1 Standard approaches to detect non-fundamentalness

Checking whether a Gaussian VAR is fundamental or not is complicated because the likelihood function or the spectral density can not distinguish between a fundamental and a non-fundamental representations. Earlier work by Lippi and Reichlin (1993, 1994) informally compared the dynamics produced by fundamental and selected non-fundamental representations. Giannone and Reichlin (2006) 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 to 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 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 lags of y_t.

While such an approach is useful to examine whether there are variables omitted from the empirical model, it is not clear whether it can reliably detect non-fundamentalness when shock aggregation is present. The reason is that cross-sectional aggregation is not innocuous. For example, Chang and Hong (2006) show that aggregate and sectoral technology shocks behave quite differently and Sbrana and Silvestrini (2010) show that volatility predictions are quite different depending on the degree of cross sectional aggregation of the portfolio one considers. The next example shows that aggregation may lead to spurious conclusions when using Granger causality to test for fundamentalness in small scale SVARs.

Example 4. Suppose the DGP is given by the following trivariate process:

$$\chi_{1t} = \varsigma_{1t} + b_1\varsigma_{1t-1} + a\varsigma_{2t} + a\varsigma_{3t} \tag{3.5}$$

$$\chi_{2t} = a\varsigma_{1t} + \varsigma_{2t} + b_2\varsigma_{2t-1} + a\varsigma_{3t} + \varsigma_{4t} \tag{3.6}$$

$$\chi_{3t} = a\varsigma_{1t} + a\varsigma_{2t} + \varsigma_{3t} + b_3\varsigma_{3t-1} - \varsigma_{4t} \tag{3.7}$$

where $\varsigma_t = [\varsigma_{1t}, \varsigma_{2t}, \varsigma_{3t}, \varsigma_{4t}]' \sim iid(0, \text{diag}(\Sigma_\varsigma))$ and $a \leq 1$.

Suppose an econometrician sets up a bivariate empirical model with $x_{1t} = \chi_{1t}$ and $x_{2t} = 0.5(\chi_{2t} + \chi_{3t})$. Thus, the second variable is an aggregated version of the last two variables of the DGP. The process generating x_t is

$$x_t = \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{bmatrix} 1 + b_1 L & a & a \\ a & 0.5((a+1) + b_2 L) & 0.5((a+1) + b_3 L) \end{bmatrix} \begin{pmatrix} \varsigma_{1t} \\ \varsigma_{2t} \\ \varsigma_{3t} \end{pmatrix} \tag{3.8}$$

Because with two endogenous variables one can recover at most two shocks, the econometrician implicitly estimates:

$$x_t = \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{bmatrix} 1 + b_1 L & a \\ a & 1 + cL \end{bmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \tag{3.9}$$

where $\sigma^2_{u1} = \sigma^2_{\varsigma 1}$. Letting $\rho_0 + \rho_1 L \equiv [0.5(a+1)\ \ 0.5(a+1)] + [0.5b_2\ \ 0.5b_3]L$, and $\hat{\Sigma}_\varsigma = diag\{\sigma^2_{\varsigma 2}, \sigma^2_{\varsigma 3}\}$, c and σ^2_{u2} are obtained from:

$$\text{E}(x_{2t}x'_{2t}) \equiv \gamma(0) = \rho_0 \hat{\Sigma}_\varsigma \rho'_0 + \rho_1 \hat{\Sigma}_\varsigma \rho'_1 = (1 + c^2)\sigma^2_{u2} \tag{3.10}$$

$$\text{E}(x_{2t}x'_{2t-1}) \equiv \gamma(1) = \rho_1 \hat{\Sigma}_\varsigma \rho'_0 = c\sigma^2_{u2} \tag{3.11}$$

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

$$c^2\gamma(1) - c\gamma(0) + \gamma(1) = 0 \tag{3.12}$$

Given $\gamma(0), \gamma(1)$ (3.12) can be used to compute the solution for c and then $\sigma^2_{u2} = c^{-1}\gamma(1)$.

Since u_t in (3.9) is a white noise, it is unpredictable using u_{t-s} (or x_{t-s}), $s > 0$. However, it can be predicted using ς_{t-s}, even when u_t is fundamental. In fact, letting c^* be the fundamental solution of (3.12) and using (3.8) and (3.9) have:

$$\begin{aligned} u_{2t} &= (1+c^*L)^{-1}[\rho_0\hat{\varsigma}_t + \rho_1\hat{\varsigma}_{t-1}] \\ &= \rho_0\hat{\varsigma}_t + c^*\rho_0\hat{\varsigma}_{t-1} + (c^*)^2\rho_0\hat{\varsigma}_{t-2} + (c^*)^3\rho_0\hat{\varsigma}_{t-3} + \cdots \\ &+ \rho_1\hat{\varsigma}_{t-1} + c^*\rho_1\hat{\varsigma}_{t-2} + (c^*)^2\rho_1\hat{\varsigma}_{t-3} + (c^*)^3\rho_1\hat{\varsigma}_{t-4} + \cdots \end{aligned} \tag{3.13}$$

where $\hat{\varsigma} = [\varsigma_{2t}, \varsigma_{3t}]'$. Since χ_{2t-s} and χ_{3t-s} carry information about ς_{t-s}, lags of $y_t = [\chi_{2t}, \chi_{3t}]$ predict u_t, and thus x_t. Notice that in terms of equation (3.3), ς_{4t} plays the role of v_t. $\square$.

To gain intuition for why predictability tests give spurious results notice that (3.13) implies $(1+c^*L)u_{2t} = \rho_0\hat{\varsigma}_t + \rho_1\hat{\varsigma}_{t-1}$. Thus, under aggregation, estimated SVAR shocks are linear functions of current and **past** primitive structural shocks, making them predictable using any variable which has information about the lags of the primitive structural shocks. This occurs even if the VAR is correctly specified (i.e. it is there are sufficient lags to recover u_t as in (3.9)). In standard SVARs without aggregation, the condition corresponding to (3.13) is $u_t = \rho\varsigma_t$. Thus, absent misspecification, lags of y_t will not predict u_t.

Granger causality tests have been used by many as a tool to detect misspecification in small scale VARs. For example, if a serially correlated variable is omitted from the VAR, the u_t the econometrician recovers are serially correlated and thus predictable using any variable correlated with the omitted one, see e.g. Canova et al. (2010). When they are applied to systems like those in example 4, causality tests detect misspecification but for the wrong reason. The VAR system is fundamental, the u_t derived from (3.9) are white noise, but Granger causality tests reject the predictability null because aggregation has created a particular correlation structure in SVAR shocks.

Example 4 also clearly highlights that the concepts of predictable, fundamental, and structural shocks are distinct. The u_t's in (3.9) are predictable, regardless of whether they are fundamental or not. In addition, $u_t = \varsigma_{x,t}$ are structural, in the sense that the responses of x_{1t} to u_t and to ς_{it}, $i = 1, 2, 3$, are similar even u_t are predictable. Finally, u_t may be non-fundamental (if c, the non-fundamental solution of (3.12) is used in (3.13)), even if they are structural.

A similar outcome obtains if the empirical model contains, e.g., an estimated proxy for an observable variable or residuals computed from the elements of χ_t. Suppose $(x_{1t} = \chi_{1t}, x_{2t} = \chi_{1t} - \gamma_1 \chi_{2t} - \gamma_2 \chi_{3t})'$, and γ_1, γ_2 are (estimated) parameters. For example, x_{2t} are Solow residuals and γ_1, γ_2 are the labor and the capital shares. The process generating x_t is:

$$x_t = \begin{bmatrix} 1+b_1L & a & a & 0 \\ (1-\gamma a-(1-\gamma)a)-b_1L & (a-\gamma-a(1-\gamma))-b_2L & (a-\gamma a-(1-\gamma))-b_3L) & -\gamma_1+\gamma_2 \end{bmatrix} \begin{pmatrix} \varsigma_{1t} \\ \varsigma_{2t} \\ \varsigma_{3t} \\ \varsigma_{4t} \end{pmatrix} \tag{3.14}$$

As before, the econometrician estimates (3.9). Also in this situation, u_t is unpredictable using u_{t-s} or x_{t-s}. However, lags of any y_t constructed as noisy linear transformation of $[\chi_{2t}, \chi_{3t}]$ predict u_t, even when it is fundamental for x_t.

In sum, the existence of variables that Granger cause x_t may have nothing to do with fundamentalness. What is crucial to create spurious results is that SVAR shocks linearly aggregate the information contained in current and past primitive structural shocks.

Although to some readers example 4 may look special, it is not. We next formally show that predictability obtains, in general, under linear cross-sectional aggregation. This together with the fact that small scale SVARs are generally used in business cycle analysis, even when the DGP may feature a large number of primitive structural shocks, should convince skeptical readers of the relevance of example 4. Proposition 1 shows that the class of moving average models is closed with respect to linear transformations and Proposition 2 that aggregated moving average models are predictable.

Proposition 1. Let χ_{1t} be a zero-mean MA(q_1) process:

$$\chi_{1t} = \varsigma_{1t} + \Phi_1 \varsigma_{1t-1} + \Phi_2 \varsigma_{1t-2} + \cdots + \Phi_{q_1} \varsigma_{1t-q_1} \equiv \Phi(L)\varsigma_{1t} \tag{3.15}$$

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

$$\chi_{2t} = \varsigma_{2t} + \Psi_1 \varsigma_{2t-1} + \Psi_2 \varsigma_{2t-2} + \cdots + \Psi_{q_2} \varsigma_{2t-q_2} \equiv \Psi(L)\varsigma_{2t} \tag{3.16}$$

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

$$x_t = \chi_{1t} + \gamma \chi_{2t} = u_t + \Pi_1 u_{t-1} + \Pi_2 u_{t-2} + \cdots + \Pi_q u_{t-q} \equiv \Pi(L) u_t \tag{3.17}$$

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

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

Proposition 2. Let x_t be an m-dimensional process obtained as in Proposition 1. Then ς_{1t-s} and ς_{2t-s}, $s \geq 1$ Granger cause x_t.

Proof: It is enough to show that

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

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

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

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

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

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

Thus, although u_t in (3.17) is unpredictable given own lagged values, it can be predicted using lagged values of ς_{1t} and ς_{2t} because the information contained in the histories of ς_{1t} and ς_{2t} is not optimally aggregated into u_t.

While the analysis is so far 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 systems

$$\begin{aligned} x_{1,t} &= u_{1t} \\ x_{2,t} &= u_{1t} + u_{2t} - 3u_{2t-1} \end{aligned} \tag{3.18}$$

$$\begin{aligned} x_{1,t} &= u_{1t} - 2u_{2t-1} \\ x_{2,t} &= u_{1t-1} + u_{2t-1} \end{aligned} \tag{3.19}$$

Both systems are non-fundamental - the determinants of the MA matrix are $1 - 3L$, and $L(1 - 2L)$ respectively, and they both vanish 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_{1t} can be obtained from $x_{1,t}$, in the second system no individual shock can not be obtained from linear combinations of current and past x_t's. □

A necessary condition for a SVAR shock to be an innovation is that it is orthogonal to the past values of the observables. FG suggest that a shock derived as in the first system of example 5 is fundamental if it is unpredictable using (orthogonal to the) lags of the principal components obtained from variables belonging to the econometrician's information 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, orthogonality tests are, in general, insufficient to assess fundamentalness. Second, as it is clear from example 5, when one shock can be recovered, it is not the shock that creates non-fundamentalness in the first place. Finally, an orthogonality test has the same shortcomings as a Granger causality test. It will reject the null of unpredictability of a SVAR shock using disaggregated variables or factors providing noisy information about them, when the SVAR shock is a linear combinations of primitive disturbances, 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 we expect to have better properties in the situations of interest in this paper. 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

$$u_t = x_t - \sum_{j=1}^{r} \omega_j x_{t-j} \tag{3.20}$$

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

$$u_t^* = \theta(L)u_t \tag{3.21}$$

where $\theta(L)$ is a Blaschke matrix [2]. Thus, the relationship between 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 a two-sided polynomial, y_t depends on current, past and *future* values of u_t^*. This proves the following proposition.

Proposition 3. The system (3.2) is fundamental if $u_{t+j}^*,\ j \geq 1$ fails to predict y_t.

Example 6. To illustrate proposition 3, let $x_t = (1 - 2.0L)u_t$, then:

$$x_t = (1 - 2.0L)\frac{(1 - 0.5L)}{(1 - 2.0L)}\frac{(1 - 2.0L)}{(1 - 0.5L)}u_t \equiv (1 - 0.5L)u_t^* \tag{3.22}$$

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

$$\begin{aligned} y_t &= (1 - 0.5L)\frac{(1 - 0.5L)}{(1 - 2.0L)}u_t^* + (1 - 0.6L)v_t \\ &= \sum_{j=0}^{\infty}(1/2)^j((1 - 0.5L)^2 u_{t+j}^*) + (1 - 0.6L)v_{t-j} \end{aligned} \tag{3.23}$$

Two points about our testing procedure need to be stressed. First, Sims (1972) has shown that x_t is exogenous with respect to y_t if future values of x_t do not help to explain y_t. Similarly here, a VAR system is fundamental if future values of x_t (u_t) do not help to predict the variables y_t, excluded from the empirical model. Thus, although the null tested here and with Granger causality is the same, aggregation/non-observability problems may make the testing results different. Second, our approach is likely to have better size properties, when SVAR shocks are linear functions of lags of primitive shocks, because y_t generally contains more information than x_t - under fundamentalness, future values of u_t will not predict y_t. Note also that our test is sufficiently general to detect non-fundamentalness due to structural causes, omitted variables, or the use of proxy indicators.

[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 small sample properties of traditional predictability tests and of our new procedure, we carry out a simulation study using a version of the model of Leeper et al. (2013), with two sources of tax disturbances. The representative household maximizes:

$$\mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \log(C_t) \tag{4.1}$$

subject to

$$C_t + (1-\tau_{t,k})K_t + T_t \leq (1-\tau_{t,y})A_t K_{t-1}^{\alpha} = (1-\tau_{t,y})Y_t \tag{4.2}$$

where $C_t,\ K_t,\ Y_t,\ T_t,\ \tau_{t,k}$ and $\tau_{t,y}$ denote time-t consumption, capital, output, lump-sum transfers, investment tax and income tax rates, respectively; A_t is a technology disturbance and E_t is the conditional expectation operator. To keep the setup tractable, we assume full capital depreciation. The government sets tax rates randomly and adjusts transfers to satisfy $T_t = \tau_{t,y}Y_t + \tau_{t,k}K_t$. The Euler equation and the resource constraints are:

$$\begin{aligned}
\frac{1}{C_t} &= \alpha\beta\mathrm{E}_t\Big[\frac{(1-\tau_{t+1,y})}{(1-\tau_{t,k})}\frac{1}{C_{t+1}}\frac{A_{t+1}K_t^{\alpha}}{K_t}\Big] && (4.3)\\
C_t + K_t &= A_t K_{t-1}^{\alpha} && (4.4)
\end{aligned}$$

Log linearizing, combining (4.3) and (4.4), we have

$$\hat{K}_t = \alpha\hat{K}_{t-1} + \sum_{i=0}^{\infty}\theta^i \mathrm{E}_t\hat{A}_{t+i+1} - \kappa_k\sum_{i=0}^{\infty}\theta^i\mathrm{E}_t\hat{\tau}_{t+i,k} - \kappa_y\sum_{i=0}^{\infty}\theta^i\mathrm{E}_t\hat{\tau}_{t+i+1,y} \tag{4.5}$$

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

We posit that technology and investment tax shocks are iid: $\hat{A}_t = \varsigma_{t,A}, \hat{\tau}_{t,k} = \varsigma_{t,k}$; and that the income tax shock is a MA(1) process: $\hat{\tau}_{t,y} = \varsigma_{t,y} + b\varsigma_{t-1,y}$. Then (4.5) is:

$$\hat{K}_t = \alpha \hat{K}_{t-1} + \varsigma_{t,a} - \kappa_k \varsigma_{t,k} - \kappa_y b \varsigma_{t,y} \tag{4.6}$$

We assume that an econometrician observes $\hat{K}_t$ and an aggregate tax variable:

$$\hat{\tau}_t = \omega \hat{\tau}_{t,y} + \hat{\tau}_{t,k} = \varsigma_{t,k} + \omega(\varsigma_{t,y} + b\varsigma_{t-1,y}) \tag{4.7}$$

where ω controls the relative weight of income taxes in the aggregate. Alternatively, one can assume that investment and income tax revenues are both observables, but an econometrician works with a weighted sum of them. If $(\hat{K}_t, \hat{\tau}_t)$ are the variables the econometrician uses in the VAR, our design covers both the cases of aggregation and of a relevant latent variable. In fact, the DGP for the observables is:

$$\begin{bmatrix} (1-\alpha L)\hat{K}_t \\ \hat{\tau}_t \end{bmatrix} = \begin{bmatrix} 1 & -\kappa_k & -\kappa_y b \\ 0 & 1 & \omega(1+bL) \end{bmatrix} \begin{bmatrix} \varsigma_{t,a} \\ \varsigma_{t,k} \\ \varsigma_{t,y} \end{bmatrix} \equiv \Gamma_x(L) C_x \varsigma_t \tag{4.8}$$

while the process recoverable by the econometrician is

$$\begin{bmatrix} (1-\alpha L)\hat{K}_t \\ \hat{\tau}_t \end{bmatrix} = \begin{bmatrix} 1 & \rho \\ 0 & 1+cL \end{bmatrix} \begin{bmatrix} u_{t,1} \\ u_{t,2} \end{bmatrix} \equiv \Pi(L) u_t \tag{4.9}$$

where $\sigma_1^2 = \sigma_a^2$ while c, σ_2^2, ρ are obtained from:

$$\begin{aligned} c^2 - c((1+b^2)/b + \sigma_k^2/(\omega^2 b \sigma_y^2)) + 1 &= 0 && (4.10) \\ \sigma_2^2 &= b\omega^2 \sigma_y^2 / c && (4.11) \\ \rho &= -\sqrt{(\omega^2 \kappa_y^2 b^2 \sigma_y^2 + \kappa_k^2 \sigma_k^2)/\sigma_2^2} && (4.12) \end{aligned}$$

By comparing (4.9) and (4.8), one can see that the aggregate tax shock $u_{t,2}$ will produce the same qualitative dynamic response in $\hat{K}_t$ as the investment and the income tax shocks but the scale of the effect will be altered. Depending on the size of ω, the aggregate shock will looks more like the income or the investment tax shock. For the exercises we present, we let $\varsigma_{t,a}, \varsigma_{t,k}, \varsigma_{t,y} \sim iid\ N(0,1)$; set $\alpha = 0.36$, $\beta = 0.99$, $\tau_y = 0.25$, $\tau_k = 0.1, \omega = 1$ and vary b so that $c \in (0.1, 0.8)$ (fundamentalness region) or $c \in (2, 9)$ (non-fundamentalness region).

To perform the tests, we need additional data not used in the empirical model (4.9). We assume that an econometrician observes a panel of 30 time series generated by:

$$(1 - 0.9L)y_{i,t} = \varsigma_{t,a} + \gamma_i \varsigma_{t,y} + (1 - \gamma_i)\varsigma_{t,k} + \xi_{i,t},\ i = 1, \cdots, 30 \tag{4.13}$$

where $\xi_{i,t} \sim\ iid\ N(0, \sigma_\xi^2)$, and γ_i is Bernoulli, taking value 1 with probability 0.5.

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

$$f_t = \sum_{j=1}^{p_1} \phi_j f_{t-j} + \sum_{j=0}^{p_2} \psi_{-j} u_{t-j} + \sum_{j=1}^{q} \psi_j u_{t+j} + e_t \tag{4.14}$$

where f_t is a $s \times 1$ vector of principal components of (4.13) and u_t is estimated using

$$x_t = \sum_{j=1}^{r} \rho_j x_{t-j} + u_t \tag{4.15}$$

where $x_t = (\hat{\tau}_t, \hat{K}_t)'$. The null is $\mathbb{H}_0^{CH} : R\Psi = 0$, where $\Psi = \text{Vec}[\psi_1, \psi_2, \cdots, \psi_q]$, R is a matrix of zeros and ones. We report the results for $p_1 = 4, p_2 = 0, q = 2, r = 4$.

To examine the properties of Granger causality tests, denoted by GC, we employ

$$x_t = \sum_{j=0}^{p_1} \phi_j x_{t-j} + \sum_{j=1}^{p_2} \varphi_j f_{t-j} + e_t \tag{4.16}$$

where again $x_t = (\hat{\tau}_t, \hat{K}_t)'$. The null is $\mathbb{H}_0^{GC} : R\Phi = 0$ where $\Phi = \text{Vec}[\varphi_1, \varphi_2, \cdots, \varphi_{p_2}]$ and R is a matrix of zeros and ones. We report results for $p_1 = 4, p_2 = 2$.

To perform an orthogonality test, denoted by OR, we first estimate (4.15) with $r = 4$. The tax shock, $u_{t,\tau}$, is identified as the only one affecting $\hat{\tau}_t$. Then, in the regression

$$u_{t,\tau} = \sum_{j=1}^{p_2} \lambda_j f_{t-j} + e_t \tag{4.17}$$

the ortogonality null is $\mathbb{H}_0^{OR} : R\Lambda = 0$ where $\Lambda = \text{Vec}[\lambda_1, \lambda_2, \cdots, \lambda_q]$ and R is a matrix of zeros and ones. We report results for $p_2 = 2$.

To maintain comparability, all null hypotheses are tested using an F-test, setting $s = 3$ and $\sigma_\xi^2 = 1$ and no correction for generated regressors in (4.14) and (4.17). The appendix present results for the CH test when other values of p_2, σ_ξ^2, s, and q are used. We set $T = 200$, which is the length of the time series used in section 5, and $T = 2000$.

To better understand the properties of the tests, we also run an experiment with no aggregation problems. Here $\tau_{k,t} = 0, \forall t$, so that the DGP for capital and taxes is

$$\begin{bmatrix} (1-\alpha L)\hat{K}_t \\ \hat{\tau}_t \end{bmatrix} = \begin{bmatrix} 1 & -\kappa_y b \\ 0 & (1+bL) \end{bmatrix} \begin{bmatrix} \varsigma_{t,a} \\ \varsigma_{t,y} \end{bmatrix} \tag{4.18}$$

and the process for the additional data is

$$(1-0.9L)y_{i,t} = \varsigma_{t,a} + \gamma_i \varsigma_{t,y} + \xi_{i,t},\ i = 1, \cdots, n \tag{4.19}$$

The percentage of rejections of the null in 1000 replications when the model is fundamental are in tables 1 and 2. Our procedure is undersized (it rejects less than expected from the nominal size) but its performance of independent of the nominal confidence level and the sample size. Granger causality and orthogonality tests are prone to spurious non-fundamentalness. This is clear when T=2000; in the smaller sample predictability due to aggregation is somewhat harder to detect.

Why are traditional predictability tests rejecting the null much more than one would

Table 1: Size of the tests: aggregation, T=200

	c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	
	10%	1.5	1.3	0.5	0.6	1.4	1.7	1.2	1.7	
CH	5%	0.8	0.5	0.1	0.3	0.3	0.9	0.4	1.1	
	1%	0.1	0.1	0.1	0.2	0.1	0.2	0.1	0.1	
	10%	13.1	15.1	16.5	15.8	19.5	27.4	38.9	55.1	
GC	5%	7.5	8.2	9.5	9.2	11.2	15.5	26.7	42.2	
	1%	2.0	2.7	2.2	3.1	4.3	5.8	10.9	19.6	
	10%	5.2	5.7	5.3	6.2	6.5	6.2	8.5	13.2	
OR	5%	2.9	2.3	2.9	2.5	3.5	2.3	4.2	6.5	
	1%	0.1	0.5	0.6	0.2	0.2	0.7	0.7	1.6	

Notes: The table reports the percentage of rejections of the null hypothesis in 1000 replications when there is aggregation; CH is the test proposed in this paper; OR is the orthogonality test; GC is the Granger causality test. The length of the vector of principal components used in the testing equation is s=3.

Table 2: Size of the tests: aggregation, T=2000

	c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	
	10%	0.1	0.4	0.2	0.1	0.2	0.1	1.9	9.5	
CH	5%	0.1	0.2	0.2	0.1	0.1	0.1	1.0	4.2	
	1%	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.8	
	10%	83.3	86.6	88.8	92.4	98.1	99.9	100	100	
GC	5%	75.3	75.1	76.3	83.9	96.1	99.8	100	100	
	1%	49.7	44.7	46.0	58.7	83.9	98.6	100	100	
	10%	34.0	30.1	29.4	30.2	41.7	52.1	81.0	99.1	
OR	5%	21.4	18.5	18.5	19.0	27.9	36.0	66.4	96.5	
	1%	7.4	6.8	7.3	6.4	9.0	13.0	34.9	81.8	

Notes: The table reports the percentage of rejections of the null hypothesis in 1000 replications when there is aggregation; CH is the test proposed in this paper; OR is the orthogonality test; GC is the Granger causality test. The length of the vector of principal components used in the testing equation is s=3.

Table 3: Size of the tests: no aggregation, T=200

	c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
	10%	2.5	1.7	2.5	1.6	1.5	2.3	2.8	2.6
CH	5%	1.1	0.6	0.6	1.2	0.5	0.8	1.2	1.0
	1%	0.1	0.1	0.2	0.1	0.1	0.1	0.2	0.3
	10%	11.4	10.5	13.3	13.5	10.9	14.8	15.5	28.4
GC	5%	5.6	5.0	6.2	8.2	5.3	7.4	9.2	19.8
	1%	1.3	1.0	1.6	1.6	1.1	2.0	2.4	6.1
	10%	4.4	5.1	5.3	4.7	4.0	6.4	6.2	8.9
OR	5%	1.7	1.3	2.8	2.2	1.3	2.3	2.6	4.9
	1%	0.2	0.1	0.5	0.3	0.1	0.6	0.6	1.8

Notes: The table reports the percentage of rejections of the null hypothesis in 1000 replications when there is no aggregation; CH is the test proposed in this paper; OR is the orthogonality test; GC is the Granger causality test. The length of the vector of principal components used in the testing equation is s=3.

Table 4: Size of the tests: no aggregation, T=2000

	c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
	10%	0.9	1.3	1.0	0.9	0.5	1.0	1.3	6.0
CH	5%	0.3	0.5	0.5	0.3	0.2	0.3	0.5	3.8
	1%	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1.7
	10%	13.2	13.3	15.6	14.8	18.2	26.7	53.8	99.8
GC	5%	7.4	8.0	8.7	9.0	11.1	16.3	41.2	99.3
	1%	1.6	1.9	2.5	3.1	3.0	5.2	19.4	95.3
	10%	3.9	5.2	5.6	4.8	4.2	6.8	8.7	20.9
OR	5%	1.3	3.2	1.7	1.8	1.7	3.2	4.8	17.8
	1%	0.3	0.7	0.4	0.4	0.2	0.4	1.2	6.0

Notes: The table reports the percentage of rejections of the null hypothesis in 1000 replications when there is no aggregation; CH is the test proposed in this paper; OR is the orthogonality test; GC is the Granger causality test. The length of the vector of principal components used in the testing equation is s=3.

Table 5: Power of the tests: aggregation, T=200

	c	2	3	4	5	6	7	8	9
	10%	99.9	100	100	100	100	100	100	100
CH	5%	99.5	100	100	100	100	100	100	100
	1%	98.7	100	100	100	100	100	100	100
	10%	100	100	100	100	100	100	100	100
GC	5%	100	100	100	100	100	100	100	100
	1%	100	100	100	100	100	100	100	100
	10%	100	100	100	100	100	100	100	100
OR	5%	100	100	100	100	100	100	100	100
	1%	100	100	100	100	100	100	100	100

Notes: The table reports the percentage of rejections of the null hypothesis in 1000 replications when there is aggregation; CH is the test proposed in this paper; OR is the orthogonality test; GC is the Granger causality test. The length of the vector of principal components used in the testing equation is s=3.

Table 6: Power of the tests: no aggregation, T=200

	c	2	3	4	5	6	7	8	9
	10%	100	100	100	100	100	100	100	100
CH	5%	100	100	100	100	100	100	100	100
	1%	100	100	100	100	100	100	100	100
	10%	100	100	100	100	100	100	100	100
GC	5%	100	100	100	100	100	100	100	100
	1%	100	100	100	100	100	100	100	100
	10%	100	100	100	100	100	100	100	100
OR	5%	100	100	100	100	100	100	100	100
	1%	100	100	100	100	100	100	100	100

Notes: The table reports the percentage of rejections of the null hypothesis in 1000 replications when there is no aggregation; CH is the test proposed in this paper; OR is the orthogonality test; GC is the Granger causality test. The length of the vector of principal components used in the testing equation is s=3.

expect from the nominal size? The answer is obtained recalling equation (3.13). u_t are linear combinations of current and past values of $\hat{A}_t, \hat{\tau}_{t,k}, \hat{\tau}_{t,y}$ while f_t are linear combinations of $\hat{A}_t, \hat{\tau}_{t,k}, \hat{\tau}_{t,y}$ and $\xi_{i,t}, \imath = 1, \ldots, 30$. Since $\hat{\tau}_{t,k}$ is serially correlated, lags of f_t may help to predict x_t even when lags of x_t are included, in particular, when the draws for γ_i are small.

It is known that Granger causality tests have poor size properties when x_t is persistent, see e.g. Ohanian (1988). Tables 3 and 4 disentangle aggregation from persistence problems: since they have been constructed absent aggregation, they report size distortions due to persistent data. It is clear that, when $b > 0.6$, the size of Granger causality tests is distorted. To properly run such tests, the lag length p_1 of the testing equation must be made function of the (unknown) persistence of the DGP. However, when $b > 0.8$, distortions are present even if $p_1 = 10$. The orthogonality test performs better because it preliminary filters x_t with a VAR. Thus, high serial correlation in x_t is less of a problem.

Comparing the size tables constructed with and without aggregation, one can see that the properties of the CH test do not depend on the presence of aggregation or the persistence of the DGP. On the other hand, aggregation make the properties of Granger causality and orthogonality tests significantly worse.

Tables 5 and 6 report the empirical power of the tests when T=200 with and without aggregation. All tests are similarly powerful to detect non-fundamentalness when it is present, regardless of the confidence level and the nature of the DGP. Although not reported for reasons of space, the power of the three tests is unchanged when T=2000.

The additional tables in the appendix indicate that the size properties of the CH test are insensitive to the selection of three nuisance parameters: the variance of the shocks to the additional data σ_ξ^2, the number of principal components used in the testing equation s, and the number of leads of the first stage residuals used in the testing equation q. On the other hand, the choice of p_2, the number of lags of the first stage residuals used in the testing equations, matters. This is true, in particular, when the persistence of the DGP increases and is due to the fact that with high persistence, r=4 is insufficient to whiten the first stage residual, and the presence of serial correlation in u_t makes its future values spuriously significant. To avoid this problem in practice, we recommend users to specify the testing equation with only leads of u_t. Alternatively, if lags of u_t are included, r should be large to insure that serial correlation in the first stage residuals is negligible.

5 Reconsidering a small scale SVAR

Standard business cycle theories assume that economic fluctuations are driven by surprises in current fundamentals, such as aggregate productivity or the monetary policy rule. Motivated by the idea that changes in expectations about future fundamentals may drive business fluctuations, Beaudry and Portier (2006) study the effect of news shocks on the real economy using a SVAR that contains stock prices and TFP.

Since models featuring news shocks have solutions displaying moving average components, empirical models with a finite number of lags may be unable to capture the underlying dynamics, making the SVARs considered in the literature prone to non-fundamentalness. In addition, Forni et al. (2014) provide a stylized Lucas tree model where perfectly predictable news to the dividend process may induce non-fundamentalness in a VAR 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:

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

where d_t are dividends, p_t are stock prices, $0 < \beta < 1$ is the discount factor. 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' information set, which includes current and past values of structural shocks, is not aligned with the econometrician's information set, which includes current and past values of the growth rate of dividends and stock prices. The fundamental and non-

Table 7: Testing fundamentalness: VAR with TFP growth and stock prices growth.

	PC=3	PC=4	PC=5	PC=6	PC=7	PC=8	PC=9	PC=10
sample 1960-2010								
CH	0.05	0.08	0.03	0.15	0.13	0.08	0.14	0.13
GC	0.02	0.00	0.04	0.01	0.00	0.00	0.00	0.00
Fernald data, sample 1960-2005								
GC(agg)	0.02	0.16	0.22	0.00	0.00	0.02	0.01	0.01
GC(dis)	0.17	0.52	0.54	0.11	0.09	0.17	0.25	0.34
Wang data, sample 1960-2009								
GC(agg)	0.05	0.02	0.11	0.03	0.00	0.00	0.00	0.00
GC(dis)	0.37	0.38	0.51	0.40	0.27	0.28	0.27	0.23

Notes: The table reports the p-value of the tests; CH is the test proposed in this paper; GC is the Granger causality test; the row GC(agg) reports the results of the test using aggregate data, the row GC(dis) the results of the test using disaggregated data; PC is the number of principal component in the auxiliary regression. In CH test the number of leads tested is two and the preliminary VAR has 4 lags. In GC test the lag length of the VAR is chosen with BIC and two lags of the principal components are used in the tests.

fundamental dynamics this model generates in response to news shocks are similar because the root generating non-fundamentalness ($L = -\beta$) is near unity, see also Beaudry et al. (2015). In general, the properties of the SVAR the econometrician considers depend on the process describing the information flows, on the variables observed by the econometrician and those included in the SVAR.

To reexamine the evidence we estimate a VAR with the growth rates 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). Table 7 reports the $p-$values of the tests, varying the number of principal components employed in the auxiliary regression, which enter in first difference in all the tests. In the CH test, the testing model has four lags of the PC and we are examining the predictive power of 2 leads of the VAR residuals. In the GC test the lag length of the VAR is chosen by BIC and two lags of the principal components are used in the tests.

The CH test finds the system fundamental and, in general, the number of PC included in the testing equations does not matter. In contrast, a Granger causality test rejects the null of fundamentalness. Since the VAR includes TFP, which is a latent variable, and estimates are obtained from an aggregated production function, differences in the results could be due to aggregation and/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 was constructed by John Fernald at the Federal Reserve Bank of San Francisco, and is obtained using the methodology of Basu et al. (2013), which produces time series for private consumption TFP, private investment TFP, government consumption and investment TFP and ' net trade" TFP. The second panel of table 7 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 data ends in 2005, the first row of the panel reports the 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, convert the data from annual to quarterly using a polynomial regression and use the growth rate of these three variables together with the growth rate of stock prices in the VAR. The third panel of table 7 presents results obtained with this VAR. Because the data ends in 2009, the first row of the panel reports the p-values of a Granger causality test for the original bivariate system restricted to the 1960-2009 sample.

Granger causality tests applied to the original bivariate system estimated over the two new samples still find the VAR non-fundamental. When the test is used in the VARs with sectoral/industry TFP measures, the null of non-fundamentalness is instead not rejected for all choices of vectors of principal components. Since this result holds when we enter the sectoral/industry TFP variables in level rather than growth rates, when we allow for a break in the TFP series, and when we use only two sectoral/industry TFP variables in the VAR, the conclusion is that a Granger causality test rejects the null in the original VAR because of aggregation problems. The diagnostic of this paper, being robust to aggregation problems, correctly identifies the original bivariate VAR as fundamental.

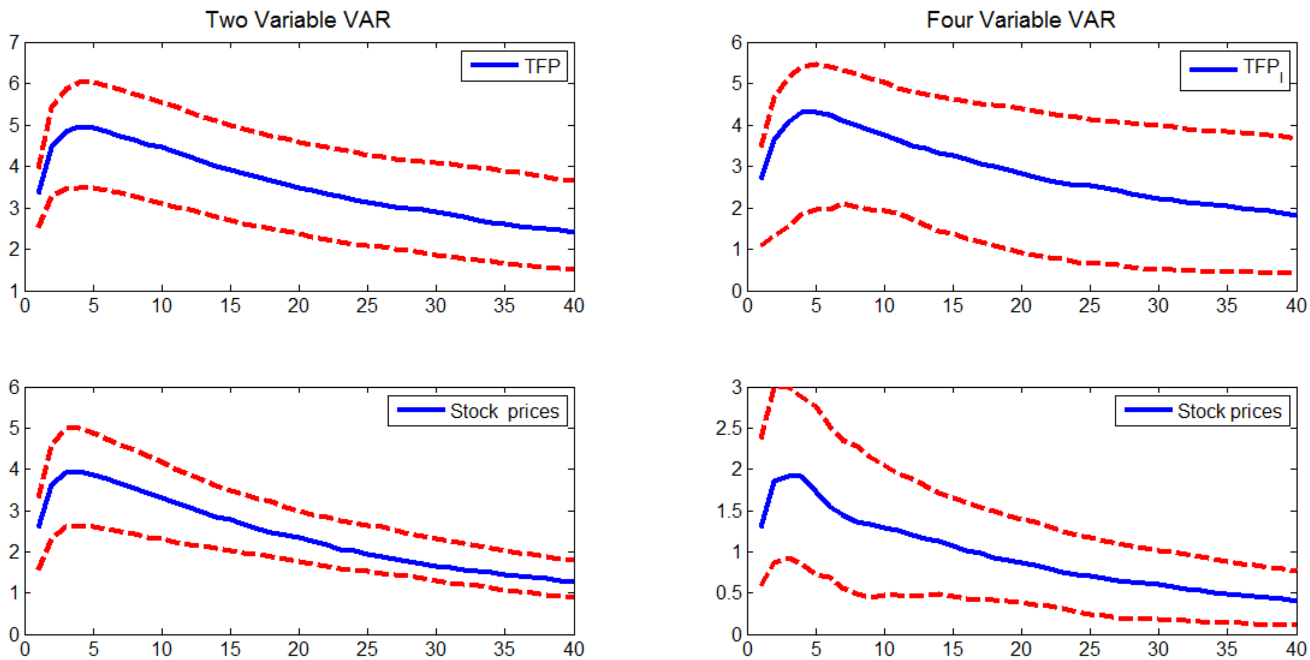
Clearly, if the DGP is a truely sectoral model, the shocks and the dynamics produced by both the bivariate and the four variable VAR systems are likely to be averages of the shocks and dynamics of the primitive economy, which surely includes more than two or four disturbances. The interesting question is whether the news shocks extracted in the two and four variable systems produce different TFP responses.

For illustration, figure 1 reports the responses of stock prices and of TFP to standardized technology news shocks in the original VAR and in the four varibale VAR with Fernald disaggregated TFP measures. For the four variable VAR we only present the responses of investment TFP since the responses of the other two TFP variables are insignificantly different from zero. It is clear that the conditional dynamics in the two systems are qualitatively similar and statistically indistinguishable. Nevertheless, median responses are smaller, uncertainty is more pervasive, and the hump in the TFP response muted in the larger system. Hence, cross sectional aggregation does not change much the dynamics but makes TFP responses artificially large and more precisely estimated. Researchers often construct models to quantitatively match the dynamics induced by shocks in small scale VARs. Figure 1 suggests that the size and the persistence of the structural shocks needed to produce the aggregate evidence are probably smaller than previously agreed upon.

6 Conclusions

Small scale SVAR models are often used in empirical business cycle analyses even though the economic model one thinks has generated the data has a larger number of variables

Figure 1: Responses to technology news shocks



Note: The dotted regions report pointwise 68 % credible intervals; the solid line is the pointwise median response. The x-axis reports quarters, the y-axis the response of the level of the variable in deviation from the predictable path.

and shocks. In this situation, SVAR shocks are linear transformations of current and past primitive structural shocks perturbing the economy. SVAR shocks might be fundamental or non-fundamental, depending on the details of the economy, the information set available to the econometrician, and the variables chosen in the empirical analysis. However, variables providing noisy information about the primitive structural shocks will Granger cause SVAR shocks, even when the SVAR is fundamental. A similar problem arises when SVAR variables proxy for latent variables. We conduct a simulation study illustrating that *spurious non-fundamentalness* may indeed occur when the SVAR used for the empirical analysis is of smaller scale than the DGP of the data.

We propose an alternative testing 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 demonstrate that a Granger causality diagnostic finds that a bivariate SVAR measuring the impact of news is non-fundamental, while our test finds it fundamental. The presence of an aggregated TFP measure in the SVAR explains the discrepancy. When sectoral TFP measures are used, a Granger causality diagnostic also finds the SVAR fundamental.

A few lessons 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. Second, 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, if one is willing to assume that the DGP is a particular structural model, the procedure described Sims and Zha (2006) can be used to check if a particular VAR shock can be recovered from current and past values of the observables, therefore by-passing the need to check for fundamentalness. However, when the DGP is unknown, the structural model one employs misspecified, or the exact mapping from the DGP and the estimated SVAR hard to construct, procedures like ours can help researchers to understand whether small scale SVARs are good starting points to undertake informative business cycle analyses.

References

Lucia Alessi, Matteo Barigozzi, and Marco Capasso. Non-fundamentalness in structural econometric models: A review. *International Statistical Review*, 79:16–47, 2011.

Susanto Basu, John Fernald, Jonas Fisher, and Miles Kimball. Sector-specific technical change. Technical report, National Bureau of Economic Research, 2013.

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

Paul Beaudry and Franck Portier. News-driven business cycles: Insights and challenges. *Journal of Economic Literature*, 52:993–1074, 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, 2007.

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

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

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

Jon Faust and Eric Leeper. When do long run restrictions give reliable results. *Journal of Business and Economic Statistics*, 15:345–353, 1988.

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

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

Mario Forni, Domenico Giannone, Lucrezia Reichlin, and Marco Lippi. Opening the black box: Structural factor models with large cross sections. *Econometric Theory*, 25:1319–1347, 2009.

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

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

Raffaella Giacomini. The relationship between dsge and var models. *in T.B. Fomby, L. Kilian, A. Murphy, eds., Advances in Econometrics, 32, Emerald*, pages 1–25, 2013.

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

Lars Peter Hansen and Thomas J Sargent. Two difficulties in interpreting vector autoregressions. *in L.P. Hansen and T. Sargent, eds, Rational expectations, Westview Press Boulder, CO*, pages 77–120, 1991.

Lutz Kilian and Helmuth Lutkepohl. *Structural Vector Autoregressive Analysis*. Cambridge University Press, 2016.

Lutz Kilian and David Murphy. The role of inventory and speculative trading in global market for crude oil. *Journal of Applied Econometrics*, 29:454–478, 2014.

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

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

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

Helmuth Lutkepohl. *Introduction to Multiple Time Series*. Spinger and Verlag, 1991.

Lee Ohanian. The spurious effect of unit roots on vars. *Journal of Econometrics*, 39:251–266, 1988.

Federico Ravenna. Vector autoregressions and reduced form representations of dsge models. *Journal of Monetary Economics*, 54:2048–2064, 2007.

Y. A. Rozanov. *Stationary Random Processes*. Holden-Day, 1967.

Giacomo Sbrana and Andrea Silvestrini. Aggregation of exponential smoothing processes with an application to portfolio risk evaluation. Technical report, Core discussion paper 2010/39, 2010.

Christopher Sims. Money, income and causality. *American Economic Review*, 62:540–552, 1972.

Christopher Sims and Tao Zha. Does monetary policy generate recessions? *Macroeconomic Dynamics*, 10:231–272, 2006.

Appendix

This appendix reports the size of the CH test when nuisance parameters are varied. We change the number of lags of first stage residuals in the auxiliary regression p_2; the variance of the error in the DGP for the additional variables, σ_ξ^2; the number of principal components used in the auxiliary regressions, s, the number of leads of the first stage residuals in the auxiliary regression q. Tables with power are omitted, since they identical to those reported in the text.

Table A1: Size of the CH test, aggregation, varying p_2

	c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
	10%	11.2	13.5	14.5	13.3	14.8	20.1	29.0	44.1
$p_2 = 4$	5%	2.5	2.3	2.5	2.2	2.9	4.6	6.1	11.9
	1%	1.6	1.9	1.2	1.6	2.2	4.1	6.2	12.3
	10%	10.5	13.2	12.1	12.5	14.1	19.3	27.0	40.8
$p_2 = 2$	5%	5.8	7.1	5.4	6.0	7.6	12.2	15.9	29.7
	1%	1.8	2.0	0.9	1.1	2.1	3.2	5.7	12.5

Notes: The table reports the percentage of rejections of the null hypothesis in 1000 replications when there is aggregation, T=200, and three principal components of the large dataset are considered; p_2 represents the number of lags in the testing equation (4.14).

Table A2: Size of the CH-test, aggregation, varying σ_ξ^2

	c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
	10%	2.20	1.80	1.70	2.10	1.60	2.10	1.80	3.00
$\sigma_\xi^2 = 4$	5%	1.10	0.70	0.40	0.60	0.50	1.00	0.60	0.90
	1%	0.30	0.10	0.10	0.00	0.00	0.20	0.20	0.10
	10%	1.00	0.70	0.20	0.80	0.50	1.50	0.60	1.10
$\sigma_\xi^2 = 0.25$	5%	0.50	0.40	0.10	0.20	0.40	0.50	0.30	0.30
	1%	0.00	0.20	0.00	0.10	0.00	0.00	0.10	0.10

Notes: The table reports the percentage of rejections of the null hypothesis in 1000 replications when there is aggregation, T=200, and three principal components of the large dataset are considered; σ_ξ^2 is the variance of the idiosyncratic error in the DGP for additional data.

Table A3: Size of the CH test, aggregation, varying s

	c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
	10%	1.10	1.10	0.30	0.60	0.80	1.00	1.00	1.70
$s = 2$	5%	0.50	0.50	0.00	0.30	0.40	0.50	0.10	0.60
	1%	0.10	0.10	0.00	0.10	0.00	0.10	0.00	0.10
	10%	1.70	1.80	0.70	1.80	1.40	1.90	1.40	2.50
$s = 4$	5%	0.80	0.70	0.10	0.60	0.50	0.60	0.50	1.10
	1%	0.20	0.10	0.00	0.10	0.00	0.20	0.10	0.10

Notes: The table reports the percentage of rejections of the null hypothesis in 1000 replications when there is aggregation, T=200, and three principal components of the large dataset are considered; s is the length of the vector of factors in the testing equation (4.14).

Table A4: Size of the CH test, aggregation, varying q

	c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
	10%	1.80	3.10	2.40	1.90	2.00	2.60	1.60	3.70
$q = 1$	5%	0.70	1.40	0.80	0.30	0.70	1.50	0.70	2.10
	1%	0.00	0.10	0.00	0.00	0.40	0.10	0.30	0.50
	10%	1.20	0.80	0.50	0.70	0.90	1.20	0.60	1.80
$q = 2$	5%	0.40	0.20	0.20	0.30	0.30	0.50	0.30	0.80
	1%	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.10

Notes: The table reports the percentage of rejections of the null hypothesis in 1000 replications when there is aggregation, T=200, and three principal components of the large dataset are considered; q represents the number of leads in the testing equation (4.14).