Understanding the Effect of Technology Shocks in SVARs with Long-Run Restrictions

Jeremy Chaudourne
Université du Québec à Montréal and CIRPÉE

Patrick Fève*
TSE (University of Toulouse I-Capitole, GREMAQ, IDEI and IUF)

Alain Guay
Université du Québec à Montréal, CIRPÉE and CIREQ

January, 2014

Abstract

This paper studies the statistical properties of impulse response functions in structural vector autoregressions (SVARs) with a highly persistent variable as hours worked and long–run identifying restrictions. The highly persistent variable is specified as a nearly stationary persistent process. Such process appears particularly well suited to characterize the dynamics of hours worked because it implies a unit root in finite sample but is asymptotically stationary and persistent. This is typically the case for per capita hours worked which are included in SVARs. Theoretical results derived from this specification allow to explain most of the empirical findings from SVARs which include U.S. hours worked.

Keywords: SVARs, long–run restrictions, locally nonstationary process, technology shocks, hours worked

JEL Class.: C32, E32

*Address: GREMAQ–Université de Toulouse I, manufacture des Tabacs, bât. F, 21 allée de Brienne, 31000 Toulouse. emails: chaudourne.jeremy@courrier.uqam.ca, patrick.feve@tse-fr.eu and guay.alain@uqam.ca. We would like to thank C. Otrok (editor) as well as an anonymous referee for constructive remarks. We also thank J. Fernald, J.P. Florens, A. Marcet and V. Ramey for valuable suggestions. This paper has benefited from helpful discussions during presentations at various seminars and conferences. The traditional disclaimer applies.

Introduction

The dynamic response of hours worked to a permanent technology shock identified from SVARs is a controversial issue. The specification of hours in SVARs and the measurement of the productivity appear crucial in the explanation of apparent conflictual results (see e.g. Galí, 1999, Francis and Ramey, 2005, Basu, Fernald and Kimball, 2006, Christiano, Eichenbaum and Vigfusson, 2004a, 2004b, Chari, Kehoe and McGrattan, 2008 and Fève and Guay, 2009). The aim of this paper is to explain these conflicting results in SVARs using a simple statistical framework with hours worked modelized as a locally nonstationary process. Our key result is that low–frequency movements in hours can contaminate the identification of technology shocks when labor productivity is included in SVARs.

In the first part of the paper, we characterize empirical evidences obtained from SVARs with US data. These findings will then be used as basic facts to be explained by our statistical framework. These results also echo those obtained in Galí (1999), Francis and Ramey (2005), Basu, Fernald and Kimball (2006), Christiano, Eichenbaum and Vigfusson (2004a) and (2004b). Using different productivity measures, we compute the dynamic responses of hours worked after a technology improvement. The main results are the following. First, when a "purified" measure of total factor productivity (TFP) constructed by Basu, Fernald and Kimball (2006) is included in SVARs, the specification of hours does not matter as we obtain almost the same dynamic responses: hours decrease in short–run and thus display a positive hump–shaped pattern. Second, when labor productivity is used, the response of hours with the LSVAR model (hours are specified in level) is positive, but it becomes persistently negative with the DSVAR model (hours are specified in first difference) and both differ from the ones obtained with the "purified" measure of total factor productivity. Third, confidence bands with the LSVAR model are fairly large and thus not informative.

Some existing works have tried to explain conflicting results obtained with LSVAR and DSVAR models. Christiano, Eichenbaum and Vigfusson (2004a) and Gospodinov (2010) establish that the large sampling uncertainty with the LSVAR model can arise from a weak instrument problem when the largest root is unity or near the nonstationary boundary (stationary in small sample but asymptotically nonstationary). A LSVAR model for such process thus leads to an inconsistent estimator of the technology shocks. Using also a nearly non-stationary SVAR, Gospodinov, Maynard and Pesavento (2011) explain that the difference in the empirical impulse responses by a small low–frequency comovements between hours worked and productivity growth which appears in the LSVAR specification but disappears in the DSVAR specification.

In this paper, we investigate an alternative finite sample approximation for a highly persistent variable as hours worked. We consider a characterization of the highly persistent process as a nearly stationary persistent process. This characterization implies that the variable is nonstationary in finite sample but asymptotically stationary persistent. This has the interesting property to be compatible with empirical evidence and theoretical implications. In finite sample, the presence of a unit root is usually not rejected for per capita hours worked series. This reflects the inability in finite sample to discrimate between an I(0) and an I(1) process for a highly persistent variable. However, the maximal number of hours that a person can work in a day is bounded (see Francis and Ramey, 2005) and therefore its is asymptotically stationary. No model that takes into account this physical constraint can yield a unit root process for the logarithm of hours (see Chari, Kehoe and McGrattan, 2008, for a discussion).[1] The nearly stationary persistent process is also in connection with equilibrium business cycle models, in which hours worked are stationary but display high degree of serial correlation (see e.g. Erceg, Guerrieri and Gust, 2005, Christiano, Eichenbaum and Vigfusson, 2006, Chari, Kehoe and McGrattan, 2008, Fève and Guay, 2009).[2] SVAR models with a nearly stationary persistent variable seem to offer an interesting alternative representation for studying the properties of the SVAR estimators and the associated dynamic responses in a representation including a highly persistent variable, such as hours worked. This is the objective of this paper.

The second part of the paper studies the properties of estimators and impulse response function (IRF) to a permanent shock (*i.e.* a technology shock), when the nonstationary variable (a TFP measure or the labor productivity) has asymptotically an exact unit root and the other variable (hours worked) follows a nearly stationary persistent process. We show that the estimated responses from the LSVAR model are biased in finite sample if the measure of productivity is contaminated by low frequency movements in hours. However, if the econometrician uses a "purified" measure of TFP, the bias disappears. We also derive the asymptotic distribution for the structural parameters of the LSVAR model. The estimators are asymptotically consistent, but display a nonstandard limiting distribution. This explains the large confidence interval for the dynamic responses in the LSVAR setup. As for the LSVAR model, the estimated responses obtained from the DSVAR model are biased in finite sample if the productivity measure (i.e. the labor productivity) includes the low–frequency movements in hours. This bias occurs because two shocks have a permanent effect on labor productivity in finite sample, violating the long–run identification scheme. This finding explains the huge difference in the estimated IRF between the LSVAR and DSVAR models. When the econometrician uses a "purified" measure of TFP instead of the labor productivity, the DSVAR can yield close to consistent estimates. In this case, the long–run restriction is valid, because the low–frequency

[1] A stationary stochastic representation does not prevent that hours will eventually exceed any positive limit, but the problem is all the more reinforced if they contain a unit root for which the variance is unbounded.

[2] Chang, Doh and Schorfheide (2007) consider a DSGE model with nonstationary hours and show that a level specification of hours must be preferred to a difference specification when the model includes real frictions in the form of labor adjustment costs.

movements in hours does not contaminate the measure of productivity. This result explains the empirical findings that show that the LSVAR and DSVAR models yield close dynamic responses of hours when a "purified" measure of TFP is included in the VAR model. While our statistical approximation is different than the one proposed by Gospodinov, Maynard and Pesavento (2011), both are valuable explanations of the difference between results from LSVAR and DSVAR. The explanation advocated by Gospodinov, Maynard and Pesavento (2011) is based on low–frequency comovements present in a level representation but disappearing in the DSVAR while our explanation is based on contamination by low-frequency comovements affecting differently the LSVAR and DSVAR models.

As an additional contribution, the alternative framework adopted here can explain why the dynamic responses from LSVAR and DSVAR with a "purified" measure of TFP differ from the ones identified with the LSVAR and DSVAR using labor productivity. Moreover, this can also explain why the estimated dynamic responses from LSVAR and DSVAR with labor productivity are biased in small sample as shown using simulations (Chari, Kehoe and McGrattan, 2008 and Fève and Guay, 2010) and differ from the ones with a "purified" measure of TFP.[3] We are then in position to propose a consistent explanation of the three empirical findings discussed earlier from SVARs with U.S. data and existing simulation results.

The paper is organized as follows. In a first section, we reports estimated dynamic responses with US data. Section 2 presents the statistical framework and an illustrative economic model to motivate the adopted specification of hours worked. Section 3 analyzes finite sample and asymptotic behaviors of SVARs with a nearly stationary persistent process. The last section concludes. Proofs are reported in Appendix.[4]

1 Figures with US Data

In this section, we report some figures on the short–run responses of hours worked in various SVARs estimated with actual US data. SVARs include three different measures of productivity that are then used for long–run identification. In each case, this variable is assumed to have an exact unit root.[5] The productivity variables considered successively are the Solow residual, a "purified" measure of TFP and the labor productivity. All the productivity measures are specified in logs and in first difference. Data are borrowed from

[3] Christiano, Eichenbaum and Vigfusson (2006) obtained dynamic responses with a small bias for LSVAR using labor productivity. At the first sight, this could appear at odds with our explanation. However, this result can be explained by the small portion of the labor productivity's variance explained by the non-technology shock thanks to their parameters configuration used to perform their simulations. In fact, for the parameter values used by Christiano, Eichenbaum and Vigfusson (2006), the resulting labor productivity is closed to be an adequate measure of the TFP. This also holds for some results in Fève and Guay (2010) (see Table 2, case with $\rho_\chi = .99$ and $\sigma_\chi/\sigma_\chi = .5$) where parameter values are close to the ones used by Christiano, Eichenbaum and Vigfusson (2006).

[4] See the Online Appendix for additional materials.

[5] The unit root on the level of each measure is not rejected, whereas the unit root on the first difference is rejected at conventional level.

Basu, Fernald and Kimball (2006). The data frequency is annual and covers the sample period 1949–1996.[6] Our measure of hours worked is the log of non–farm business hours, divided by the population 16 and over. The data used in this empirical analysis are reported in Figure 1. The three measures of productivity display similar business cycle patterns, but the "purified" measure is less volatile than the other two. In addition, the Solow residual and the labor productivity growth are imperfectly correlated with the "purified" measure (the correlation coefficient between the growth rates is 0.31 and 0.57, respectively). As shown in the right hand side of Figure 1, the log of hours displays persistent fluctuations over the sample period. The estimated autocorrelation function suggests that hours worked display a high serial correlation. We also perform an Augmented Dickey Fuller (ADF) test of unit root. We regress the growth rate of hours on a constant, lagged level and two lags of the first difference. The ADF test statistic is equal to -1.26 and the null hypothesis of unit root cannot be rejected at 5%. This finding suggests that hours are nonstationary and thus must be specified in first difference in SVARs. However, it is well known that unit root tests have lower power in small sample against stationary alternatives, so hours are specified either in level or in first difference in SVARs.

The four first SVARs (with hours either in level or in first difference) use direct measures of TFP and correspond to the ones adopted by Basu, Fernald and Kimball (2006) and Christiano, Eichenbaum and Vigfusson (2004b). The SVAR model with labor productivity growth and the log of hours in level is the one adopted by Christiano, Eichenbaum and Vigfusson (2004a), whereas the SVAR with the log of hours in first difference is the one adopted by Galí (1999), Galí and Rabanal (2004) and Francis and Ramey (2005). In all cases, we consider a bivariate VAR model and we impose the long restriction *à la* Blanchard and Quah (1989) that technology shock is the only variable that can have a permanent effect on the Solow residual, the "purified" measure of TFP and the labor productivity. Following Basu, Fernald and Kimball (2006) and Christiano, Eichenbaum and Vigfusson (2004b), the lag length for each VAR model is two. Results are found to be robust to other choices of the lag length. The confidence interval is then obtained from standard bootstrap techniques with 1000 replications.[7] The results are reported in Figures 2 and 3. For comparison purpose, we set the same scale for each estimated responses. We will consider two sets of dynamic responses. First, we focus on the dynamic responses of hours after a technology improvement. Second, we inspect the response of the three productivity measures to a technology shock.

Let us first consider the dynamic responses of hours, as they constitute a central and debatable empirical issue (see Figure 2). When the Solow residual and the "purified" measure of TFP are included in the SVAR model, the specification of hours (in level or in first difference) does not matter a lot. On impact, hours

[6] Very recently, Fernald (2012) proposes a quarterly, utilization–adjusted series on TFP.

[7] See Inoue and Kilian (2002) for the justification of bootstrap method in our context.

decrease, but after two years the response becomes persistently positive and displays an hump–shaped pattern.[8] Interestingly, the discrepancy of IRFs (between the level and the first difference specification) is less pronounced when a "purified" measure of TFP is considered instead of the Solow residual. This finding means that when a proper measure of productivity is available, consistent estimates of the dynamic responses of hours worked can be obtained without wondering what is the proper specification of hours. When the growth rate of labor productivity is included in the SVAR model, things however change dramatically. Indeed, the LSVAR and the DSVAR models predict opposite conclusions about the dynamic responses of hours and for both specifications the dynamic responses differ from the ones obtained by the "purified" measure of TFP. Except on impact, the LSVAR model displays a positive hump–shaped response whereas the DSVAR model implies a persistent decrease in hours (see Galí, 1999, Christiano, Eichenbaum and Vigfusson, 2004a and Francis and Ramey, 2005). Finally, in the case of a level specification of hours, the confidence interval is so wide that the estimated IRFs of hours are not significantly different from zero at any horizon. This is especially true when the labor productivity growth is included in the LSVAR model. We also obtain fairly large confidence bands with direct productivity measures. When hours worked are specified in first difference, the confidence interval remains wide but the dynamic responses appear more precisely estimated, especially in the very short run.[9]

Let us now consider the estimated dynamic responses of each productivity measure to a technology improvement. The responses are reported in Figure 3. The responses are comparable for the three productivity measures, because each of them will permanently adjusts in the long–run after a permanent technology shock. The "purified" measure of TFP shifts up almost immediately and the Solow residual after one period. Notice that the specification of hours does not matter, as the dynamic responses with these two measures are very similar. This finding is again in contrast with those obtained from labor productivity. The LSVAR and DSVAR models display different responses in the short–run. The DSVAR model implies a quick adjustment of labor productivity to its new long–run value. With the LSVAR model, the labor productivity adjusts very gradually and persistently. The long–run effect of the technology measure can be directly obtained from the Cholesky decomposition of the long–run covariance matrix of the variables. Results are reported in Table 1. For each measure, the long effect is almost similar whatever the specification of hours.[10] Notice however, that the discrepancy is more pronounced in the case of the labor productivity. Again, the "puri-

[8] See also Vigfusson (2004) for similar empirical findings.

[9] We can also complete these figures by computing the correlation of the technology innovations identified from LSVAR and DSVAR models, for each measure of productivity. The higher correlation is obtained in the case of the "purified" measure of TFP, 0.9860, and the lower is obtained when labor productivity is included in SVARs, 0.9082. The Solow residual provides intermediate results with a correlation of 0.9831.

[10] The DSVAR model produces larger long–run responses than the LSVAR model with every productivity measure.

fied" implies the smaller difference and the Solow residual intermediate results. Another important aspect is the sizeable difference between the "purified" measure and the labor productivity. For example, with the DSVAR specification, the long–run effect is 1.39% with the "purified" measure and 2.32% with the labor productivity. Using this latter variable yields larger long–run effects of a technology improvement.

2 The Statistical Framework and an Illustrative Economic Model

We first present the specification of a nearly stationary persistent process. Second, we connect this statistical representation to the dynamics of hours worked obtained from a standard RBC model.

2.1 SVARs with a Nearly Stationary Persistent Process

In this section we present and study our proposed specification of a highly persistent process. This is obtained by parameterizing it as a nearly stationary persistent process. Phillips (1987) and Chan and Wei (1987), among others, considered nearly unit-root process to investigate the asymptotic power of the unit-root tests under a sequence of local alternatives. Since we are interested by a highly persistent process which is asymptotically stationary, we consider a sequence of local alternatives such that the process is locally nonstationary but asymptotically stationary and persistent.

Let us first introduce this nearly stationary persistent process parametrization with a simple example. For the aim of exposition, suppose the following univariate process:

$$
\begin{aligned}
(1-\rho L)\Delta x_t &= u_t - \delta_T u_{t-1} \\
\delta_T &= \left(1 - c/\sqrt{T}\right),
\end{aligned}
$$

with $0 < \rho < 1$, $c > 0$ and u_t is a white noise. As T increasing to infinity and for a high value of ρ, this process becomes a stationary persistent process whereas, in finite sample, the process is characterized by an unit root. This process is locally nonstationary but asymptotically stationary and persistent. This characterization of the highly persistent process has two advantages. First, it suitably represents the time series behavior of variables for which usual unit root tests cannot reject the nonstationarity in small sample. Second, although highly persistent, the variable is necessarily characterized by an asymptotic stationary process. Pantula (1991), Perron and Ng (1996) and Ng and Perron (2001) consider a simplified version of this process with $\rho = 0$ but to investigate the performance of unit-root tests.[11] It is important to understand that this specification adopted here must not be interpreted as a literal description of the data but as a device

[11] Obviously, the parametrization adopted by these authors is not suitable in our case since such process is asymptotically a white noise.

to approximate the behavior of a highly persistent variable in small sample which is necessarily stationary asymptotically.

More generally, we are interested in a bivariate representation $X_t = (\Delta X_{1t}, X_{2t})$ for $t = 1, ..., T$, where the variable X_{1t} contains an exact unit root and X_{2t} is a highly persistent variable. Both variables in the vector X_t are asymptotically second order stationary and they admit asymptotically the following Wold decomposition

$$X_t = C(L)\varepsilon_t, \tag{1}$$

where $C(L) = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j}$ and $C_0 = I_2$, $\sum_{j=0}^{\infty} C_j^2 < \infty$. ε_t is a vector of white noise processes with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t \varepsilon_t') = \Sigma$. The deterministic part is omitted to simplify the presentation without altering the results below. Now consider a Structural Moving Average (SMA) representation for X_t:

$$X_t = A(L)\eta_t, \tag{2}$$

where $A(L) = \sum_{j=0}^{\infty} A_j$ and $\eta_t = (\eta_{1t}, \eta_{2t})'$ is the vector of orthogonal structural shocks with $E(\eta_t \eta_t') = \Omega$ a diagonal matrix. A common identification assumption is $\Omega = I_2$, the variance of the structural shocks is then normalized to one.

Given representations (1) and (2), the error terms ε_t from the reduced form are related to the structural error terms η_t as follows: $\varepsilon_t = A_0 \eta_t$ which implies that $\Sigma = A_0 A_0'$. The SMA representation is identified through the identification of the relationship matrix A_0. We focus here on the identification strategy based on long-run restrictions such as proposed by Blanchard and Quah (1989). The identification scheme uses the long-run variance-covariance matrix of the reduced from (1) and the structural form (2) which are related by $C(1)\Sigma C(1)' = A(1)A(1)'$ and $A_0 = C(1)^{-1}A(1)$. Typically a lower triangular structure is imposed to the long-run impact matrix $A(1)$ which can be easily obtained using a Choleski decomposition of the long-run variance-covariance matrix $C(1)\Sigma C(1)'$. In the case where two variables are included in X_t, the first structural shock is the only one shock that can have a permanent effect on the first variable.

Now consider for a finite sample of T observations a structural characterization of the highly persistent variable X_{2t} as a nearly stationary persistent process:

$$\begin{aligned} \Delta X_{2t} &= a_{21}(L)\Delta\eta_{1t} + a_{22}(L)\left(1 - \left(1 - \frac{c}{\sqrt{T}}\right)L\right)\eta_{2t} \\ &= a_{21}(L)\Delta\eta_{1t} + \tilde{a}_{22,T}(L)\eta_{2t}, \end{aligned}$$

where $\tilde{a}_{22,T}(L) = a_{22}(L)\left(1 - \left(1 - \frac{c}{\sqrt{T}}\right)L\right)$. By the Beveridge-Nelson decomposition, this can be rewritten as

$$\Delta X_{2t} = a_{21}(L)\Delta\eta_{1t} + \tilde{a}_{22,T}(1)\eta_{2t} + \tilde{a}^*_{22,T}(L)\Delta\eta_{2t}$$

with $\tilde{a}_{22,T}(1) = a_{22}(1)\frac{c}{\sqrt{T}}$ and $\tilde{a}^*_{22,T}(L)(1-L) = \tilde{a}_{22,T}(L) - \tilde{a}_{22,T}(1)$. Let us examine in more details this specification. The SMA bivariate representation contains a difference stationary process ΔX_{1t} and a highly persistent process such that:

$$\begin{bmatrix} \Delta X_{1t} \\ \Delta X_{2t} \end{bmatrix} = \begin{bmatrix} a_{11}(L) & a_{12}(L)(1-L) \\ a_{21}(L)(1-L) & \tilde{a}_{22,T}(L) \end{bmatrix} \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \end{bmatrix}. \tag{3}$$

and only η_{1t} has a permanent effect on the first variable. The second variable X_{2t} is asymptotically stationary but locally nonstationary for values of c greater than zero.[12] The value of c controls the amplitude of the local nonstationarity. Moreover, for appropriate values of $a_{21}(L)$ and $a_{22}(L)$, the process is asymptotically persistent but stationary. We also consider the case that the second shock can have locally a permanent effect on the first variable X_{1t}. We show below that this can occur when the first variable is a linear function of the persistent variable X_{2t}. For a fixed T, the corresponding characterization of the first variable is

$$\begin{aligned} \Delta X_{1t} &= a_{11}(L)\eta_{1t} + a_{12}(L)\left(1 - \left(1 - \frac{c}{\sqrt{T}}\right)L\right)\eta_{2t} \\ &= a_{11}(L)\eta_{1t} + \tilde{a}_{12,T}(L)\eta_{2t}, \end{aligned}$$

where $\tilde{a}_{12,T}(L) = a_{12}(L)\left(1 - \left(1 - \frac{c}{\sqrt{T}}\right)L\right)$ and c is the same as in specification of the variable X_{2t}. Here, the identification scheme based on the long run restriction that only the first shock has a permanent effect is still valid asymptotically but does not hold for a finite T. The resulting SMA bivariate representation is:

$$\begin{bmatrix} \Delta X_{1t} \\ \Delta X_{2t} \end{bmatrix} = \begin{bmatrix} a_{11}(L) & \tilde{a}_{12,T}(L) \\ a_{21}(L)(1-L) & \tilde{a}_{22,T}(L) \end{bmatrix} \begin{bmatrix} \eta_{1t} \\ \eta_{2t}. \end{bmatrix}. \tag{4}$$

An illustrative RBC model presented below motivates the proposed SMA representations (3) and (4).

2.2 An illustrative RBC model

The SMA representations (3) and (4) nest several business cycle models wherein TFP contains an exact unit root. To see this, let us consider the following simple dynamic model. The intertemporal expected utility function of the representative household is given by

$$E_t \sum_{i=0}^{\infty} \beta^i \left\{ \log(C_{t+i}) + B\left(1 - H_{t+i}\right) \right\},$$

where B is a positive scale parameter, $\beta \in (0,1)$ denotes the discount factor and E_t is the expectation operator conditional on the information set available as of time t. C_t is the consumption at t and H_t represents the household's labor supply. Time endowment is normalized to unity. To ease the computation

[12]The case of c smaller than zero are not allowed to exclude nonfundamentalness representations of ΔX_{2t} in finite sample.

of the solution, we assume that utility is linear in leisure (see Hansen, 1985). The representative firm uses capital K_t and labor H_t to produce the homogeneous final good Y_t. The technology is represented by the following constant returns–to–scale Cobb–Douglas production function

$$Y_t = K_t^{\theta} \left(Z_t H_t \right)^{1-\theta},$$

where $\theta \in (0,1)$. The TFP, represented by Z_t, is assumed to follow an exogenous process of the form

$$\log(Z_t) = \log(Z_{t-1}) + (\gamma_z - 1) + \sigma_z \eta_{zt},$$

where $\gamma_z > 1$ is the gross growth rate of TFP, $\sigma_z > 0$ and η_{zt} is *iid* with zero mean and unit variance. The capital stock evolves according to the law of motion

$$K_{t+1} = (1-\delta) K_t + I_t,$$

where $\delta \in (0,1)$ is the constant depreciation rate and I_t denotes investment. Finally, the final good can be either consumed or invested

$$Y_t = C_t + I_t.$$

The approximate solution of the model is obtained from a log–linearization of the stationary equilibrium conditions around their deterministic steady state (see Appendix A for more details). This gives the following equation for hours:

$$(1 - \varphi L)\widehat{h}_t = \nu \sigma_z \eta_{zt} \tag{5}$$

where a hat represents the relative deviation from steady–state. The parameter ν is positive and $\varphi \in (0,1)$ is the stable root of the log–linear version of the model. Hours worked increase after a technology improvement and go back steadily toward their steady–state. Notice that despite its simplicity, the model implies that hours worked can display highly persistent fluctuations when φ is close to one,[13] but they are stationary since $\varphi < 1$ (i.e. φ is the stable root of the model).

Getting a near–stationary persistent process as a solution of the business cycle model is a complicated task because the time series properties of this process depends on the size of the sample T. This is not the case of any typical business cycle model, wherein the size of the sample does not enter as a parameter into the model'solution. Here we consider that the persistence in hours is a combination of the model's property (see equation (5) that is free of T) and a measurement error (an equation that heavily depends on T).[14]

[13] The parameter φ exceeds 0.98 for a standard calibration of the model.

[14] We closely follow the approach developed in Ireland (2004). The method consists in combining the power of DSGE model with the flexibility of unconstrained time-series models. Ireland (2004) notably shows that the measurement equation associated to a prototypical business cycle model displays a high level of serial correlation. Measurement issues have been already deeply considered for the estimation of DSGE models (see the references in Ireland, 2004).

Let $\{h_t\}_{t=1}^{t=T}$, the hours worked observed by the econometrician for a given finite sample T. Hours evolves according to the following measurement equation

$$h_t = \widehat{h}_t + h_t^c \tag{6}$$

where h_t is the observed realizations of hours worked in log and in deviation from its mean, $\widehat{h}_t$ is given by equation (5) (the solution of the business cycle model) and h_t^c represents a low-frequency measurement error. This low-frequency measurement error is assumed to follow a near–stationary persistent process of the form

$$\begin{aligned}
\phi(L)\Delta h_t^c &= \left(1 - \left(1 - \frac{c}{\sqrt{T}}\right) L\right) \sigma_c \eta_{ct} \\
&= \frac{c}{\sqrt{T}} \sigma_c \eta_{ct} + \left(1 - \frac{c}{\sqrt{T}}\right) \sigma_c \Delta \eta_{ct} \qquad (7)
\end{aligned}$$

where $c > 0$, $\sigma_c > 0$ and and η_{ct} is *iid* with zero mean and unit variance. In addition, η_{ct} is assumed to be orthogonal to η_{zt} (contemporaneously and for all leads and lags). The polynomial $\phi(L)$ has all its roots outside the unit circle. For T finite h_t^c contains a unit root. When T goes to infinity for a given positive c, the process (7) reduces to

$$\phi(L) h_t^c = \sigma_c \eta_{ct}.$$

So, in a finite sample, observed hours are nonstationary for $c > 0$. This characterization of low-frequency measurement error tries to capture the difficulty with an insufficient span of available data to consistently estimate the mean reverting property of the underlying process. The series seems nonstationary for the sample on hands while the underlying process is stationary as hours worked per capita. This specification supposes that the highly persistent component which is hardly approximated in finite sample depends on a shock orthogonal to the technology shock (ex: a labor wedge shock). This can also capture low-frequency movements in the standard measure of hours worked corresponding to sectoral shifts in hours and the changing age composition of the working-age population (Francis and Ramey, 2009). According to these authors, these low-frequency comovements must be removed from the labor productivity growth and hours worked series.

Using the model's solution (5), the measurement equation (6) and the measurement error (7), hours in first difference admit the following moving–average representation

$$\begin{aligned}
\Delta h_t &= \Delta \widehat{h}_t + \Delta h_t^c \\
&= \left(\frac{\nu}{1 - \varphi L}\right) \sigma_z \Delta \eta_{zt} + \frac{c \phi(L)^{-1}}{\sqrt{T}} \sigma_c \eta_{ct} + \phi(L)^{-1} \left(1 - \frac{c}{\sqrt{T}}\right) \sigma_c \Delta \eta_{ct} \qquad (8)
\end{aligned}$$

From (8), we obtain that the technology shock has no long-run effect on hours, whereas the low frequency shock on the measurement errors has a permanent effect on hours in finite sample.

Using this equation and other model's variables, we can now present some illustrative examples for SMA representations (3) and (4). Denoting $\eta_{1t} = \eta_{zt}$ and $\eta_{2t} = \eta_{ct}$, equation (8) can be rewritten in the form of the process for X_{2t}, where

$$\begin{aligned} a_{21}(L) &= \left(\frac{\nu}{1-\varphi L}\right)\sigma_z, \\ \tilde{a}_{22,T}(L) &= \frac{c\phi(L)^{-1}}{\sqrt{T}}\sigma_c + \phi(L)^{-1}\left(1-\frac{c}{\sqrt{T}}\right)\sigma_c(1-L) \end{aligned}$$

This type of representation for X_{2t} will be always maintained, *i.e.* hours worked represent the locally nonstationary variable under study. The variable X_{1t} can be interpreted in several ways, but examples that we consider maintain the central assumption that X_{1t} contains an exact unit root.

Example 1 : ΔX_{1t} is the growth rate of TFP. In this case, the process for X_{1t} reduces to $a_{11}(L) = \sigma_z$ and $a_{12}(L) = 0$. However, ΔX_{1t} can be a contaminated measure of TFP, but the contamination has no long–run effect in finite sample. This case can be handle by imposing $a_{11}(L) = \sigma_z$ and $a_{12}(L) = a_{12}^{(0)}$, where $a_{12}^{(0)}$ is a non–zero scalar. A more general formulation for $a_{12}(L)$ can be also considered to account for long–lasting contamination effects.

Example 2 : ΔX_{1t} is the output growth. Using the solution of the business cycle model (see Appendix A), the output growth is given by

$$\Delta y_t = \sigma_z \eta_{zt} - \mu \frac{\sigma_z \Delta\eta_{zt}}{1-\varphi L} \tag{9}$$

where μ is a constant parameter. This implies that $a_{11}(L) = \sigma_z - \mu\frac{\sigma_z(1-L)}{1-\varphi L}$ and $a_{12}(L) = 0$. Notice that a very similar formulation can be obtained if we replace output growth by consumption or investment growth. Measurement errors can be easily included by assuming that $a_{12}(L)$ is non–zero. The central assumption that we maintain is that these measurement errors have no long–run effects on output in a finite sample. In a more complex model, Δy_t can be affected by an other permanent shock than the technology shock (see Galí, 1999).

Example 3 An interesting case is when the nonstationary variable considered in the SVAR model is contaminated by the highly persistent variable. For example, the productivity variable used in most of SVARs with long–run restrictions is growth rate of the labor productivity (see Galí, 1999, Christiano, Eichenbaum and Vigfusson, 2004a, Francis and Ramey, 2005). To illustrate this, consider the case where the econometrician constructs a measure of productivity growth using the output growth given by (9) and the observed change in hours given by (8):

$$\Delta y_t - \Delta h_t = \sigma_z \eta_{zt} - \left(\frac{\mu+\nu}{1-\varphi L}\right)\sigma_z \Delta\eta_{zt} - \frac{c\phi(L)^{-1}}{\sqrt{T}}\sigma_c \eta_{ct} - \phi(L)^{-1}\left(1-\frac{c}{\sqrt{T}}\right)\sigma_c \Delta\eta_{ct} \qquad (10)$$

In this case, $\Delta X_{1t} = \Delta y_t - \Delta h_t$ with $\Delta X_{2t} = \Delta h_t$. Consequently, labor productivity is contaminated by the low–frequency component of hours. In other words, the variable X_{1t} is now function of the nearly stationary persistent process X_{2t}. In finite sample, the two shocks η_{zt} and η_{ct} affect permanently the labor productivity. Francis and Ramey (2009) have already investigated this issue by showing that long–run identifying assumption used in empirical works are valid only when labor productivity are defined in efficiency unit. When they are not, labor productivity depends both on TFP and the ratio of efficiency hours to aggregate hours. In our statistical setup, the effect of this ratio in finite sample is captured by the near stationarity in the measurement errors. It is important to note here that output growth does not share the low-frequency movements of hours precluding a cointegration relationship.

3 Estimation and Inference

We now consider estimation and inference with the two specifications of the SVAR model. In practice, the reduced-form moving average representation is retrieved by performing a finite order VAR on the data. Suppose now that the structural moving average representation can be characterized or approximated in small sample by a finite VAR of order p.[15] Consider, the following reduced form VAR(p):

$$D(L)X_t = \varepsilon_t \quad \text{where} \quad D(L) = I - \sum_{i=1}^{p} D_i L^i = \begin{bmatrix} 1-\sum_{i=1}^{p} d_{11}^{(i)} L^i & -\sum_{i=1}^{p} d_{12}^{(i)} L^i \\ -\sum_{i=1}^{p} d_{21}^{(i)} L^i & 1-\sum_{i=1}^{p} d_{22}^{(i)} L^i \end{bmatrix}.$$

By multiplying both sides by a matrix $B_0 = \begin{bmatrix} 1 & -b_{12}^{(0)} \\ -b_{21}^{(0)} & 1 \end{bmatrix} = A_0^{-1}$, we obtain the VAR in function of the structural shocks as follows: $B(L)X_t = \eta_t$ with $B(L) = B_0 D(L)$. More explicitly, the first variable is given by:

$$\Delta X_{1t} = \left(\sum_{i=1}^{p} d_{11}^{(i)} L^i - b_{12}^{(0)} \sum_{i=1}^{p} d_{21}^{(i)} L^i\right)\Delta X_{1t} + \left(\sum_{i=1}^{p} d_{12}^{(i)} L^i + b_{12}^{(0)}\left[1-\sum_{i=1}^{p} d_{22}^{(i)} L^i\right]\right) X_{2t} + \eta_{1t}.$$

Imposing the structural long-run impact matrix $A(1)$ to be lower triangular implies that $B_0 D(1)$ is also lower triangular by $A(1) = D(1)^{-1}A_0$. The long-run multiplier of the variable X_{2t} on ΔX_{1t} is then zero.

[15]By considering the order $p \to \infty$ (at some rate) and $\frac{p}{T} \to 0$, Lewis and Reinsel (1985) show that a multivariate infinite autoregression can be arbitrarily approximated by a finite VAR of order p. Furthermore, under Assumption 1 in Inoue and Kilian (2002), the least-squares estimators of the VAR parameters are asymptotically normal.

Imposing this constraint yields

$$\begin{aligned}\Delta X_{1t} &= \left(\sum_{i=1}^{p} d_{11}^{(i)} L^i - b_{12}^{(0)} \sum_{i=1}^{p} d_{21}^{(i)} L^i\right) \Delta X_{1t} + b_{12}^{(0)} \Delta X_{2t} + \sum_{i=1}^{p-1} \widetilde{b}_{12}^{(i)} L^i \Delta X_{2t} + \eta_{1t} \\ &= b_{11}(L) \Delta X_{1t-1} + b_{12}^{(0)} \Delta X_{2t} + \widetilde{b}_{12}(L) \Delta X_{2t-1} + \eta_{1t}\end{aligned}$$

with $b_{12}^{(0)} = -d_{12}(1)/(1 - d_{22}(1))$ and $\widetilde{b}_{12}^{(i)} = -\sum_{j=i+1}^{p} \left(d_{12}^{(j)} - b_{12}^{(0)} d_{22}^{(j)}\right)$. The second equation is:

$$X_{2t} = \left(b_{21}^{(0)} \left[1 - \sum_{i=1}^{p} d_{11}^{(i)} L^i\right] + \sum_{i=1}^{p} d_{12}^{(i)} L^i\right) \Delta X_{1t} + \left(\sum_{i=1}^{p} d_{22}^{(i)} L^i - b_{21}^{(0)} \sum_{i=1}^{p} d_{12}^{(i)} L^i\right) X_{2t} + \eta_{2t}$$

which can be rewritten as

$$\begin{aligned}X_{2t} &= \left(b_{21}^{(0)} \left[1 - \sum_{i=1}^{p} d_{11}^{(i)} L^i\right] + \sum_{i=1}^{p} d_{12}^{(i)} L^i\right) \Delta X_{1t} + b_{22}(1) X_{2t-1} + \sum_{i=1}^{p-1} \widetilde{b}_{22}^{(i)} L^i \Delta X_{2t} + \eta_{2t} \\ &= b_{21}^{0} \Delta X_{1t} + b_{21}(L) \Delta X_{1t-1} + b_{22}(1) X_{2t-1} + \widetilde{b}_{22}(L) \Delta X_{2t-1} + \eta_{2t}\end{aligned}$$

where $b_{22}(1) = d_{22}(1) - b_{21}^{(0)} d_{12}(1)$ and $\widetilde{b}_{22}^{(i)} = -\sum_{j=i+1}^{p} b_{22}^{(j)} = -\sum_{j=i+1}^{p} \left(d_{22}^{(j)} - b_{12}^{(0)} d_{12}^{(j)}\right)$.

3.1 The LSVAR Model

Let us first consider the LSVAR specification. With this specification, the second variable X_{2t} is included in the VAR in level. To simplify the exposition, suppose that the initial condition is fixed to zero (i.e. $X_{20} = 0$). The structural highly persistent process is rewritten in level as:

$$X_{2t} = a_{21}(L)\eta_{1t} + a_{22}(1) \frac{c}{\sqrt{T}} \sum_{i=1}^{t} \eta_{2i} + \tilde{a}_{22,T}^{*}(L)\eta_{2t}. \tag{11}$$

For a finite T and $c > 0$, the second structural shock has a permanent effect on the variable X_{2t}. The finite measure of the element $(2,2)$ of the matrix $A(1)$ is then very sensitive to the value of the parameter c. For T going to infinity and $c > 0$, the second term at the RHS of eq. (11) does not disappear asymptotically but converges toward a Brownian motion. Thus,

$$a_{22}(1) \frac{c}{\sqrt{T}} \sum_{i=1}^{[Tr]} \eta_{2i} \xrightarrow{L} a_{22}(1) c W(r)$$

where $W(r)$ is a standard Brownian motion for $r \in [0,1]$, $t = [Tr]$ and $[\cdot]$ means the integer part. Consequently, even though $\frac{1}{\sqrt{T}} \to 0$ as T is going to infinity, the variance of X_{2t} is still function on the second term and in particular depends on the value c. We will show below that this introduces an asymptotic non standard distribution for impulse responses resulting from the LSVAR. For the first variable ΔX_{1t}, when the specification (3) is true, the long-run restriction that only the first structural shock has a permanent effect on the first variable is valid for all finite T. When the first variable is contaminated by the highly persistent process X_{2t} as labor productivity, the restrictions is violated for a finite T as in specification (4) which introduces a potential finite sample bias of the long-run matrix $A(1)$ estimator. This finite sample bias is transmitted to the estimator of A_0 and the resulting impulse responses by the following relationships: $A(0) = C(1)^{-1}A(1)$ and $A(L) = C(L)A(0)$. This explains why dynamic responses from LSVAR using a "purified" measure of TFP differ from LSVAR using labor productivity and why the latter yields biased dynamic responses as shown in simulations by Chari, Kehoe and Mcgrattan (2008) and Fève and Guay (2010).

Let us now examine more precisely the asymptotic properties of the LSVAR. We are interested by the estimation of the two following equations.

$$\Delta X_{1t} = b_{11}(L)\Delta X_{1t-1} + b_{12}^{(0)}\Delta X_{2t} + \widetilde{b}_{12}(L)\Delta X_{2t-1} + \eta_{1t} \tag{12}$$

$$X_{2t} = b_{21}^{0}\Delta X_{1t} + b_{21}(L)\Delta X_{1t-1} + b_{22}(1)X_{2t-1} + \widetilde{b}_{22}(L)\Delta X_{2t-1} + \eta_{2t}. \tag{13}$$

We rewrite the model as function of $b_{12}^{(0)}$ and $b_{21}^{(0)}$ whose asymptotic properties determines the limiting behavior of the impulse responses at the impact and b_{22}^{*} defined as $b_{22}(1) - 1$ above which undermines the asymptotic properties of the dynamic of impulse responses. Thus, the model can be rewritten up to $o_p(1)$ terms as

$$\Delta\widetilde{X}_{1t} = b_{12}^{(0)}\Delta\widetilde{X}_{2t} + \eta_{1t} \tag{14}$$

$$\Delta\widetilde{X}_{2t} = b_{21}^{(0)}\Delta\widetilde{X}_{1t} + b_{22}^{*}\widetilde{X}_{2t-1} + \eta_{2t} \tag{15}$$

where $\Delta\widetilde{X}_{1t}$, $\Delta\widetilde{X}_{2t}$ and $\widetilde{X}_{2t-1}$ are defined as the residuals of the projection of these variables on the predetermined variables $W_{t-1} = (\Delta X_{1t-1}, \ldots, \Delta X_{1t-p}, \Delta X_{2t-1}, \ldots, \Delta X_{2t-p-1})'$ and $b_{22}^{*} = b_{22}(1) - 1$. To study the properties of the impulse responses, we consider the IV estimator of the SVAR model proposed by Shapiro and Watson (1988). Christiano, Eichenbaum and Vigfusson (2004a) and Gospodinov (2010) employed this estimator to analyze the cases where the second variable is difference stationary or a nearly nonstationary process, respectively. The IV estimator of $b_{12}^{(0)}$ with X_{2t-1} as instrument is given by the following expression:

$$\widehat{b}_{12}^{(0)} = \frac{\frac{1}{T}\sum_{t=2}^{T} X_{2t-1}\Delta\widetilde{X}_{1t}}{\frac{1}{T}\sum_{t=2}^{T} X_{2t-1}\Delta\widetilde{X}_{2t}} = \frac{\frac{1}{T}\sum_{t=2}^{T} X_{2t-1}\left[b_{12}^{(0)}\Delta\widetilde{X}_{2t} + \eta_{1t}\right]}{\frac{1}{T}\sum_{t=2}^{T} X_{2t-1}\Delta\widetilde{X}_{2t}}$$

or equivalently

$$\widehat{b}_{12}^{(0)} - b_{12}^{(0)} = \frac{\frac{1}{T}\sum_{t=1}^{T} X_{2t-1}\eta_{1t}}{\frac{1}{T}\sum_{t=1}^{T} X_{2t-1}\Delta\widetilde{X}_{2t}} + o_p(1). \tag{16}$$

Consider now the estimation of the structural parameter $b_{21}^{(0)}$ in equation (15). Since η_{1t} and η_{2t} are orthogonals, the residuals $\widehat{\eta}_{1t} = \Delta\widetilde{X}_{1t} - \widehat{b}_{12}^{(0)}\Delta\widetilde{X}_{2t}$ can be used as instrument for the endogenous variable $\Delta\widetilde{X}_{1t}$. Thus, $\widehat{\eta}_{1t} = \eta_{1t} - \left(\widehat{b}_{12}^{(0)} - b_{12}^{(0)}\right)\Delta\widetilde{X}_{2t}$. Let us now define $z_t = (\widehat{\eta}_{1t}, X_{2t-1})'$ and $x_t = (\Delta\widetilde{X}_{1t}, \widetilde{X}_{2t-1})'$. The IV estimator of $\beta = \left(b_{21}^{(0)}, b_{22}^{*}\right)'$ is then given by $\widehat{\beta} = \left[\frac{1}{T}\sum_{t=2}^{T} z_t x_t'\right]^{-1}\left[\frac{1}{T}\sum_{t=2}^{T} z_t \Delta\widetilde{X}_{2t}\right]$. The following theorem provides the asymptotic properties for the IV estimator of $b_{12}^{(0)}$, $b_{21}^{(0)}$ and b_{22}^{*}.

Theorem 1 *Under the structural model (3) or (4) and assumptions in Lemma 1, the IV estimator* $\widehat{b}_{12}^{(0)} \xrightarrow{p} b_{12}^{(0)}$ *and* $\widehat{\beta} \xrightarrow{p} \beta$ *converges to the true value. Thus,*

$$\sqrt{T}\left(\widehat{b}_{12}^{(0)} - b_{12}^{(0)}\right) \xrightarrow{L} \frac{\left(\psi_{2,0}^{1/2}\xi_1 + \vartheta_1\right)}{a_{22}^{(0)} b_{22}^{*}\widetilde{\psi}_2},$$

$$\sqrt{T}\left(\widehat{b}_{21}^{(0)} - b_{21}^{(0)}\right) \xrightarrow{L} \frac{\xi_2}{a_{11}^{(0)}} - \frac{\left(\psi_{2,0}^{1/2}\xi_1 + \vartheta_1\right)}{a_{22}^{(0)} b_{22}^{*}\widetilde{\psi}_2}.$$

$$\sqrt{T}\left(\widehat{b}_{22}^{*} - b_{22}^{*}\right) \xrightarrow{L} \frac{a_{12}^{(0)}\left(\psi_{2,0}^{1/2}\xi_1 + \vartheta_1\right)}{a_{11}^{(0)}\widetilde{\psi}_2} - \left[\frac{a_{12}^{(0)} b_{22}^{*}}{a_{11}^{(0)}} - \frac{\psi_{2,0}^{1/2}}{\widetilde{\psi}_2}\right]\xi_2 + \frac{\vartheta_2}{\widetilde{\psi}_2}$$

where $\vartheta_1 = a_{22}(1)c\int_0^1 W_2(r)dW_1(r)$, $\vartheta_2 = a_{22}(1)c\int_0^1 W_2(r)dW_2(r)$ *with* $\xi = (\xi_1, \xi_2)' \sim \mathcal{N}(0, I_2)$, $W_1(r)$ *and* $W_2(r)$ *are two independent Brownian motions and* $\widetilde{\psi}_2$ *and* $\psi_{2,0}^{1/2}$ *are defined in the Appendix.*

Theorem 1 establishes that the IV estimator of $b_{12}^{(0)}$, $b_{21}^{(0)}$ and b_{22}^{*} are consistent. This result is not surprising since as T goes to infinity, X_{2t} is a second order stationary variable. The second set of result shows that the asymptotic distribution of $\sqrt{T}(\widehat{b}_{12}^{(0)} - b_{12}^{(0)})$ is a mixture of Gaussian distributions for c greater than zero. The first term of the asymptotic distribution is the usual asymptotic distribution for a stationary variable X_{2t}. The term ϑ_1 has a Gaussian distribution conditional on $W_2(\cdot)$ since $W_1(\cdot)$ is independent of $W_2(\cdot)$. This component produces wider confidence interval compared to the standard case with stationary variables by increasing the parameter c controlling the local nonstationarity. Theorem 1 also shows that the asymptotic distribution of the IV estimator of $b_{21}^{(0)}$ is a function of the asymptotic distribution of $b_{12}^{(0)}$ and then shares the same asymptotic properties. This is due to the use of the instrument which depends on the estimator $\widehat{b}_{12}^{(0)}$. Finally, the asymptotic distribution of the IV estimator of persistence parameter b_{22}^{*} is non standard and depends on ϑ_1 and ϑ_2. The term ϑ_2 is of an unit–root type distribution providing fat tails in the asymptotic distribution. Larger is the value of c, wider is the non standard confidence interval.

In the light of the business cycle model, Theorem 1 states that whatever the variable used for X_{1t} (TFP, output or a proper measure of labor productivity), the LSVAR provides consistent estimators. The definition of X_{1t} does not matter a lot for the asymptotic distribution. To see this, let us consider the case where ΔX_{1t} is the growth rate of TFP, *i.e.* $\Delta X_{1,t} = \Delta z_t$. This implies that $a_{12}^{(0)} = 0$ and thus $b_{12}^{(0)} = 0$. In this case, the limiting distribution of $\widehat{b}_{12}^{(0)}$ and $\widehat{b}_{21}^{(0)}$ are left unaffected. The sole difference concerns $\widehat{b}_{22}^{*}$ for which the limiting distribution does not include the random variables ζ_1 and υ_1. However, the asymptotic distribution is still non standard as it depends on υ_2. Again, larger values of c tends to increase the confidence interval for $\hat{b}_{22}^{*}$.

We now examine the asymptotic behavior of the impulse response function to the permanent shock (e.g, η_{1t}). According to the SMA representation (2), let $a_{kj}^{(l)} = \frac{\partial X_{k,t+l}}{\partial \eta_{jt}}$ be the impulse response function at l periods ahead for a normalized structural shock j for the variable X_k. Since the estimators of $b_{12}^{(0)}$, $b_{21}^{(0)}$ and b_{22}^{*} are consistent according to Theorem 1 and other parameters in the VAR are also consistently estimated, we can show that estimates of impulse responses $a_{kj}^{(l)}$ resulting from the VAR are also consistent. However, their asymptotic distributions are non standard due to the local nonstationarity of X_{2t}. A simple example can give insight about the effect of the local nonstationary on the asymptotic distribution of the impulse responses. Consider the following VAR(1) model

$$\begin{aligned} \Delta X_{1t} &= b_{12}^{(0)} \Delta X_{2t} + \eta_{1t} \\ X_{2t} &= b_{21}^{(0)} \Delta X_{1t} + b_{22} X_{2t-1} + \eta_{2t}. \end{aligned}$$

For instance, in this case the estimator of the impulse response function of the first shock on the second variable X_{2t}, this is given by

$$\widehat{a}_{21}^{(l)} = \frac{\widehat{b}_{21}^{(0)} \left(\widehat{b}_{22} - \widehat{b}_{12}^{(0)} \widehat{b}_{21}^{(0)} \right)^{l}}{(1 - \widehat{b}_{12}^{(0)} \widehat{b}_{21}^{(0)})^{l+1}}.$$

and $\widehat{b}_{22}^{*} = \widehat{b}_{22} - 1$. At the impact (*i.e.* $l = 0$), the asymptotic distribution depends only on the asymptotic distribution of the parameters $b_{12}^{(0)}$ and $b_{21}^{(0)}$ which is a scaled mixed of Gaussian distributions depending on c. For the propagation of the shocks (*i.e.* $l > 0$), the asymptotic distribution depend also on the asymptotic distribution of b_{22}^{*} which is of the unit-root distribution inducing fat-tails distribution for the impulse responses. In the more specific case where ΔX_{1t} is the growth rate of the TFP, $b_{12}^{(0)} = 0$, the estimator of the impulse response function of the first shock on the second variable X_{2t} is given by the following simple expression

$$\widehat{a}_{21}^{(l)} = \widehat{b}_{21}^{(0)} \left(\widehat{b}_{22}^{*} - 1 \right)^{l}.$$

3.2 The DSVAR Model

Let us now consider the DSVAR specification. From (4), the finite sample measure of the long-run impact is given by:

$$A_T(1) = \begin{bmatrix} a_{11}(1) & a_{12}(1)c/\sqrt{T} \\ 0 & a_{22}(1)c/\sqrt{T} \end{bmatrix}.$$

When the first variable ΔX_{1t} is not function of the second variable as in (3), the term $a_{12}(1)c/\sqrt{T}$ is equal to zero and the long-run restriction that only the first shock has a permanent effect on the first variable is valid. When, the first variable corresponds to labor productivity growth, since $\Delta X_{1t} = \Delta y_t - \Delta h_t$ and $\Delta X_{2t} = \Delta h_t$ implies $a_{12}(1)c/\sqrt{T} = -a_{22}(1)c/\sqrt{T}$. The linear dependence between the labor productivity growth and Δh_t induces negative low frequency co-movements between these variables. In contrast to small low-frequency co-movements discussed in Gospodinov, Maynard and Pesavento (2011) which are present in the LSVAR specification but disappear in the DSVAR specification, the negative co-movements here are present in the DSVAR specification.[16] In particular, the negative low-frequency co-movements between labor productivity and hours in the DSVAR which are absent in a DSVAR including TFP and hours explain the difference in the estimated impulse responses between these two specifications.

As $T \to \infty$, the long-run restriction that only the first shock has a permanent effect on the first variable is valid. However, this long-run restriction measured by the matrix $A(1)$ is violated for a finite T and this matrix is upper triangular instead of lower triangular. Provided $c > 0$ and T fixed, two shocks have a long–run effect on X_{1t}, the permanent shock η_{1t} and the non–permanent η_{2t}. Suppose now that we use, as usual, a long-run identification scheme to uncover η_{1t}. It follows that the long–run effect of η_{2t} on X_{1t} will be attributed to η_{1t}, leading to over–estimate the contribution of η_{1t}. Notice that when c increases, *i.e.* when the variable X_{2t} becomes more and more persistent, the effect of η_{2t} on X_{1t} increases, because the variable X_{1t} is contaminated by X_{2t}. Indeed, the second shock η_{2t} will have a permanent effect on X_{2t} in finite sample. When the size of the local nonstationary alternative c increases, the shock η_{2t} will have larger permanent effect on X_{2t}. Again, because the long–run identification will wrongly attribute to η_{1t} the permanent effect on X_{1t} of η_{2t}, this identification scheme will conclude that the shock η_{1t} will have a permanent effect on X_{2t} whose sign will depend on the sign of $a_{12}(1)$. Observe also as $T \to \infty$, only the element (1,1) is different from zero and the matrix $A_\infty(1)$ is singular resulting from the over-differentiation of X_{2t}.

The next proposition characterizes more precisely the finite sample measure at zero frequency when a

[16] Using estimated DSVAR parameters, we can compute low-frequency co-movements between labor productivity growth and hours worked in difference using the non-farm business sector data for a given interval of frequencies and, in particular, for the frequency zero and its neighborhood. In fact, it exists a non negligible negative low-frequency correlation between labor productivity growth and hours worked in difference. This is consistent with our specification.

lower triangular structure is imposed to the matrix $A_T(1)$ for the DSVAR model.

Proposition 1 *Consider a DSVAR model with variables ΔX_{1t} and ΔX_{2t} defined in eq. (4). The finite sample measure of the long-run impact of the structural shocks by using the Choleski decomposition is given by the following lower triangular matrix :*

$$chol\left(A_T(1)A_T(1)'\right) = \begin{bmatrix} \left(a_{11}(1)^2 + a_{12}(1)^2c^2/T\right)^{1/2} & 0 \\ \frac{a_{12}(1)a_{22}(1)c^2/T}{(a_{11}(1)^2+a_{12}(1)^2c^2/T)^{1/2}} & \left(a_{22}(1)^2c^2/T - \frac{a_{12}(1)^2a_{22}(1)^2c^4/T^2}{a_{11}(1)^2+a_{12}(1)^2c^2/T}\right)^{1/2} \end{bmatrix}.$$

where $chol(\cdot)$ is the Choleski decomposition.

Imposing that the second shock has no long-run effect on the variable X_{1t} yields a theoretical (population) impulse response function of the permanent shock η_{1t} that is biased. The finite sample bias introduced by wrongly imposing a lower triangular matrix depends on the parameter $a_{12}(1)$. First, the finite sample measure of the long-run impact of the permanent shock on the first variable is over–stated since $\left(a_{11}(1)^2 + a_{12}(1)^2c^2/T\right)^{1/2} > a_{11}(1)$ for $a_{12}(1) \neq 0$. Second, the finite sample measure of the long–run impact of the first structural shock η_{1t} on the second variable is of the same sign as $a_{12}(1)$. For a negative value of $a_{12}(1) = -a_{22}(1)$, which corresponds to the case of labor productivity, the finite sample measure of the long run impact is negative. The contamination of the finite sample measure of the long–run impact is then transmitted to the estimator of the relationship matrix A_0 by the expression : $A_0 = C(1)^{-1}A(1)$. This result is of importance since it establishes that the local nonstationarity of per capita hours worked may lead to a downward bias in the estimated dynamic responses from a DSVAR model (when $a_{12}(1)$ is negative), despite the fact that a first difference specification seems to be adequate in finite sample.

We can easily translate the results of the above Proposition in terms of the business cycle model. First, consider that ΔX_{1t} is the growth of labor productivity given by (10) and the econometrician estimates a VAR model with a difference specification for observed hours Δh_t. From equations (10) and (8), the finite sample measure of the long–run impact matrix is given by:

$$A_T(1) = \begin{bmatrix} \sigma_z & -\frac{c\phi(1)^{-1}}{\sqrt{T}}\sigma_c \\ 0 & \frac{c\phi(1)^{-1}}{\sqrt{T}}\sigma_c \end{bmatrix}$$

Now, applying a Choleski decomposition to the long–run covariance matrix $A_T(1)A_T(1)'$, we obtain the long–run effect of each structural shocks as identified by the DSVAR model:

$$chol(A_T(1)A_T(1)') = \begin{bmatrix} \left(\sigma_z^2 + \frac{c^2\phi(1)^{-2}\sigma_c^2}{T}\right)^{1/2} & 0 \\ -\frac{c^2\phi(1)^{-2}\sigma_c^2}{T\left(\sigma_z^2+\frac{c^2\phi(1)^{-2}\sigma_c^2}{T}\right)^{1/2}} & \left(\frac{c^2\phi(1)^{-2}\sigma_c^2}{T} - \frac{c^4\phi(1)^{-4}\sigma_c^4}{T^2\left(\sigma_z^2+\frac{c^2\phi(1)^{-2}\sigma_c^2}{T}\right)}\right)^{1/2} \end{bmatrix}$$

First, looking at the (1,1) entry of the decomposition, it appears that the long–run effect of identified technology shock on labor productivity is over–estimated in finite sample as long as $c > 0$. Second, inspecting the (2,1) entry, the DSVAR model will predict a negative long–run effect of technology shock on hours under a business cycle model in which this long–run effect is zero. When T is large enough regarding c, the long–run effect tends to zero.

Another interesting results from Proposition 1 is the case where $a_{12}(L) = 0$. It follows immediately that the long–run response of X_{2t} to η_{1t} is zero, whatever or not this variable follows a local non–stationary process. This can be illustrated again, using the business cycle model. Consider now that the econometrician uses a perfect measure of the TFP.[17] In this case, $\Delta X_{1t} = \Delta z_t$. From (8), the finite sample measure of the long–run impact matrix is now given by:

$$A_T(1) = \begin{bmatrix} \sigma_z & 0 \\ 0 & \frac{c\phi(1)^{-1}}{\sqrt{T}}\sigma_c \end{bmatrix}.$$

The Choleski decomposition of the long–run covariance matrix yields

$$chol(A_T(1)A_T(1)') = \begin{bmatrix} \sigma_z & 0 \\ 0 & \frac{c\phi(1)^{-1}}{\sqrt{T}}\sigma_c \end{bmatrix}.$$

The long–run responses of the TFP and hours obtained from the DSVAR model are consistent with the business cycle model, even in finite sample. The parameter c does not affect the estimated long–run response of TFP and hours to a technology shock. This results is of importance because it corresponds to the case where $b_{12}^{(0)} = 0$. As shown below, in this case, the short–run response obtained from a DSVAR are unbiased and the dynamic responses can display small bias if a sufficient number of lags are included in the VAR model. Coupled with estimated long–run response, our results indicate that DSVAR can almost properly uncover the dynamic responses of hours worked when the econometrician use a perfect measure of TFP.

To study in more details the properties of the estimators resulting for the DSVAR, the corresponding estimated reduced form VAR(p) for both variables in difference is given by:

$$\mathbf{D}(\mathbf{L})\mathbf{X}_t = \varepsilon_t$$

where the vector $\mathbf{X}_t$ is now defined as $\mathbf{X}_t = (\Delta X_{1t}, \Delta X_{2t})'$. By multiplying both sides by $B_0 = \begin{bmatrix} 1 & -b_{12}^{(0)} \\ -b_{21}^{(0)} & 1 \end{bmatrix} = A_0^{-1}$, we obtain the VAR in function of the structural shocks: $\mathbf{B}(L)\mathbf{X}_t = \eta_t^*$ with $\mathbf{B}(L) = B_0\mathbf{D}(L)$. Imposing the structural long-run impact matrix to be lower triangular implies that $B_0\mathbf{D}(1)$ is also lower triangular. The long-run multiplier of the variable ΔX_{2t} on ΔX_{1t} is then zero.

[17] Similar results for long–run responses hold when measurement errors has no long–run effect on this measure in finite sample.

Imposing this constraint yields for the first equation,

$$\Delta X_{1t} \quad = \quad \mathbf{b}_{11}(L)\Delta X_{1t-1} + b_{12}^{(0)}\Delta^2 X_{2t} + \widetilde{\mathbf{b}}_{12}(L)\Delta^2 X_{2t-1} + \eta_{1t}^* \tag{17}$$

and for the second equation:

$$\Delta X_{2t} \quad = \quad b_{21}^0 \Delta X_{1t} + \mathbf{b}_{21}(L)\Delta X_{1t-1} + \mathbf{b}_{22}(1)\Delta X_{2t-1} + \widetilde{\mathbf{b}}_{22}(L)\Delta^2 X_{2t-1} + \eta_{2t}^*. \tag{18}$$

Asymptotically, the LSVAR is correctly specified while the DSVAR is misspecified. To study the asymptotic properties of the estimators obtained with the DSVAR specification, we rewrite the correctly specified LSVAR's eq. (12) such that the variables X_{2t} and its lags appear in second difference as function of the parameter $b_{12}^{(0)}$. Thus

$$\begin{aligned} \Delta X_{1t} &= b_{11}(L)\Delta X_{1t-1} + b_{12}^{(0)}\Delta X_{2t} + \widetilde{b}_{12}(L)\Delta X_{2t-1} + \eta_{1t} \\ &= b_{11}(L)\Delta X_{1t-1} + b_{12}^{(0)}\Delta^2 X_{2t} + \widetilde{b}_{12}(L)\Delta^2 X_{2t-1} + \eta_{1t}^* \end{aligned} \tag{19}$$

with $\eta_{1t}^* = -b_{12}^{(0)}\Delta X_{2t-1} - \widetilde{b}_{12}(L)\Delta X_{2t-2} + \eta_{1t}$. By comparing eq. (17) with eq. (19), we see that the error η_{1t}^* is function of the lagged values of ΔX_{2t}. By rewriting also the second equation of the LSVAR, we can compare with the second equation of the DSVAR. Thus,

$$X_{2t} \quad = \quad b_{21}^0 \Delta X_{1t} + b_{21}(L)\Delta X_{1t-1} + b_{22}(1)X_{2t-1} + \widetilde{b}_{22}(L)\Delta X_{2t-1} + \eta_{2t}$$

and rewriting in difference

$$\Delta X_{2t} \quad = \quad b_{21}^0 \Delta X_{1t} + b_{21}(L)\Delta X_{1t-1} + b_{22}(1)\Delta X_{2t-1} + \widetilde{b}_{22}(L)\Delta^2 X_{2t-1} + \eta_{2t}^* \tag{20}$$

with $\eta_{2t}^* = -b_{21}^0 \Delta X_{1t-1} - b_{21}(L)b_{22}(L)\Delta X_{1t-2} + \eta_{2t} - \eta_{2t-1}$.

As we proceed for the LSVAR, the first equation of the DSVAR is rewritten as:

$$\Delta \widetilde{X}_{1t} = b_{12}^{(0)}\Delta^2 \widetilde{X}_{2t} + \eta_{1t}^* \tag{21}$$

where $\Delta\tilde{X}_{1t}$ and $\Delta^2\widetilde{X}_{2t}$ are defined as the residuals of the projection of these variables on the predetermined variables $(\Delta X_{1,t-1}, \ldots, \Delta X_{1,t-p}, \Delta^2 X_{2t-1}, \ldots, \Delta^2 X_{2t-p-1})$. Using ΔX_{2t-1} as instruments, the IV estimator of $b_{12}^{(0)}$ in equation (21) is then given by the following expression:

$$\widehat{b}_{12}^{(0)} = \frac{\frac{1}{T}\sum_{t=2}^{T}\Delta X_{2t-1}\Delta\widetilde{X}_{1t}}{\frac{1}{T}\sum_{t=2}^{T}\Delta X_{2t-1}\Delta^2\widetilde{X}_{2t}}.$$

Since ΔX_{2t-1} and η_{1t}^* are correlated by eq. (19), the DSVAR estimator of $b_{12}^{(0)}$ is asymptotically biased. Consider now the estimation of the parameters of equation (18). As in the LSVAR case, one uses the residuals $\widehat{\eta}_{1t}^* = \Delta\widetilde{X}_{1t} - \widehat{b}_{12}^{(0)}\Delta^2\widetilde{X}_{2t}$ as instrument for the endogenous variable ΔX_{1t}. Thus, $\widehat{\eta}_{1t}^* = \eta_{1t}^* - \left(\widehat{b}_{12}^{(0)} - b_{12}^{(0)}\right)\Delta^2\widetilde{X}_{2t}$. The estimator is obviously asymptotically biased by the correlation between η_{1t}^* and η_{2t}^* as we can see by expressions derived above. The estimator of $b_{22}(1)$ is also not consistent by the correlation between ΔX_{2t-1} and η_{2t}^*.

Now, suppose that the DSVAR is estimated with a "purified" of TFP. In this case, the first equation of the DSVAR is given by:

$$\Delta X_{1t} \quad = \quad b_{11}(L)\Delta X_{1t-1} + \eta_{1t}^*.$$

The first equation of the DSVAR is now the same as the LSVAR. The IV estimator $\widehat{b}_{12}^{(0)}$ defined above is then consistent and converge to zero since $\eta_{1t}^* = \eta_{1t}$ in this case. The resulting residuals $\widehat{\eta}_{1t}^*$ can be used as instruments in the second equation for the estimation of the parameter $b_{21}^{(0)}$. This also yields a consistent estimator of $b_{21}^{(0)}$ thanks to the absence of correlation between $\widehat{\eta}_{1t}^*$ and η_{2t}^*. However the estimator of $b_{22}(1)$ is still inconsistent. The asymptotic bias of the estimator of $b_{22}(1)$ is function of the following term: $\lim_{T\to\infty}\frac{1}{T}\Delta X_{2t-1}\eta_{2t}^*$. An important part of this bias comes from the unit root in the error term $\Delta\eta_{2t}$ resulting from the overdifferentiation of the second equation of the VAR. This bias can be reduced by increasing the number of lags in the DSVAR model (see Marcet, 2005).

4 Conclusion

This paper studies the statistical properties of impulse response functions in SVARs with a highly persistent variable as hours worked and long–run identifying restrictions. The highly persistent variable is specified as a nearly stationary persistent process. We show that the estimated responses from LSVAR and DSVAR models are biased in finite sample if the measure of productivity is contaminated by low frequency movements in hours. However, if the econometrician uses a "purified" measure of TFP, the bias disappears for the LSVAR and the DSVAR specifications. We also show that the estimators from LSVAR are asymptotically consistent, but display a nonstandard limiting distribution. This explains the large confidence interval for the dynamic responses in the LSVAR setup (see the estimations from US data). This also helps to understand existing simulation results obtained in Chari, Kehoe and McGrattan (2008) and Fève and Guay (2010).

Our findings can serve as useful guideline to improve the reliability of SVAR models with long–run restrictions. First, our theoretical and empirical results suggest that more efforts must be made to obtain proper measures of TFP as done by Basu, Fernald and Kimball (2006) at annual frequency, because including highly persistent variables in SVAR models is less problematic in this case. Second, our findings show that part of the poor performances of SVARs is due to the high persistence of hours. Some previous papers have tried to deal with this problem. Francis and Ramey (2009) construct a corrected measure of per capita hours worked that adjusts for low frequency movements in government employment, schooling, and the aging of population. Fève and Guay (2009) propose a simple two step method. In a first step, a consistent estimator of the technology shock is obtained with a SVAR excluding hours worked. The response of hours that follows a technology improvement is estimated in a second step using different linear projections on the estimated technology innovations. Interestingly, these three aforementioned papers find that both level and first-difference specifications yield very similar dynamic responses in SVARs, *i.e.* a short–run decline followed by a positive hump-shaped response of hours.

References

Basu, S., Fernald, J. and M. Kimball (2006) "Are Technology Improvements Contractionary?", *American Economic Review*, 96(5), 1418–1448.

Blanchard, O.J. and D. Quah (1989) "The Dynamic Effects of Aggregate Demand and Supply Disturbances", *American Economic Review*, 79(4), 655–673.

Chan, N. and C. Wei (1987) "Asymptotic Inference for Nearly Nonstationary AR(1) Process", *Annals of Statistics*, 15, 1050–1063.

Chang, Y., Doh, T. and Schorfheide, F. (2007) "Non-stationary Hours in a DSGE Model", *Journal of Money, Credit and Banking*, 39(6), 1357–1373.

Chari, V., Kehoe, P. and E. Mc Grattan (2008) "A Critique of Structural VARs Using Real Business Cycle Theory", *Journal of Monetary Economics*, 55, 1337–1352.

Christiano, L., Eichenbaum, M. and R. Vigfusson (2004a) "What Happens after a Technology Shock?", NBER Working Paper Number 9819, *revised version 2004*.

Christiano, L., Eichenbaum, M. and R. Vigfusson (2004b) "The Response of Hours to a Technology Shock: Evidence Based on Direct Measures of Technology", *Journal of European Economic Association*, vol. 2, 2–3, 381–395.

Christiano, L., Eichenbaum, M. and R. Vigfusson (2006) "Assessing Structural VARs", NBER Macroeconomics Annual 2006, Volume 21. D. Acemoglu, K. Rogoff and M. Woodford (Eds)

Erceg, C., Guerrieri, L. and C. Gust (2005) "Can Long–Run Restrictions Identify Technology Shocks", *Journal of European Economic Association*, 3, 1237–1278.

Fernald, J. (2012) "A Quarterly, Utilization–Adjusted Series on Total Factor Productivity", Working Paper 2012-19, Federal Reserve Bank of San Francisco.

Fève, P. and A. Guay (2009) "The Response of Hours to A Technology Shock: A Two–Step Structural VAR Approach", *Journal of Money, Credit and Banking*, 41(5), 987–1013.

Fève, P. and A. Guay (2010) "Identification of Technology Shocks in Structural VARs", *Economic Journal*, 120(549), 1284–1318

Francis, N. and V. Ramey (2005) "Is the Technology–Driven Real Business Cycle Hypothesis Dead? Shocks and Aggregate Fluctuations Revisited", *Journal of Monetary Economics*, 52, 1379–1399.

Francis, N., and V. Ramey (2009) "Measures of per Capita Hours and their Implications for the Technology–Hours Debate", *Journal of Money, Credit and Banking*, 41(6), 1071–1097.

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

Galí, J., and P. Rabanal (2004) "Technology Shocks and Aggregate Fluctuations; How Well does the RBC Model Fit Postwar U.S. Data?", *NBER Macroeconomics Annual*, 225–288.

Gospodinov, N. (2010) "Inference in Nearly Nonstationary SVAR Models with Long–Run Identifying Restriction", *Journal of Business & Economic Statistics*, 28(1), 1–12.

Gospodinov, N., A. Maynard and E. Pesavento (2011) “Sensitivity of Impulse Response to Small Low Frequency Co–Movements: Reconciling the Evidence on the Effects of Technology Shocks”, *Journal of Business & Economic Statistics*, 29(4), 455–467.

Hamilton, J. (1994) *Time Series Analysis*, Princeton University Press.

Inoue, A. and L. Kilian (2002) “Bootstrapping Smooth Functions of Slope Parameters and Innovation Variances in VAR(∞) Models”, *International Economic Review*, 43(2), 309–332.

Ireland, P. (2004) “A Method for Taking Models to the Data”, *Journal of Economic Dynamics and Control*, 28(6), 1205–1226.

Lewis, R. and G.C. Reinsel (1985) “Prediction of Multivariate Time Series by Autoregressive Model Fitting”, *Journal of Multivariate Analysis*, 15(1), 393–411.

Marcet, A. (2005) “Overdifferencing VAR’s is Ok”, manuscript.

Ng, S. and P. Perron (2001) “Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power”, *Econometrica*, 69(6), 1519–1554.

Pantula, S. (1991) “Asymptotic Distributions of Unit–Root Tests when the Process is Nearly Stationary”, *Journal of Business & Economic Statistics*, 9(1), 63–71.

Perron, P. and S. Ng (1996) “Useful Modifications to some Unit Root Tests with Dependant Errors and their Local Asymptotic Properties”, *Review of Economic Studies*, 63, 435–463.

Phillips, P.C.B. (1987) “Towards a Unified Asymptotic Theory for Autoregression”, *Biometrika*, 74(3), 535–547.

Shapiro, M. and M. Watson (1988) “Sources of Business Cycles Fluctuations”, *NBER Macroeconomics Annual*, Vol. 3, 111–156, S. Fischer Ed. , MIT Press.

Vigfusson R. (2004) “The Delayed Response to a Technology Shock. A Flexible Price Explanation”, International Finance Discussion Paper Series 2004–810. Washington: Board of Governors of the Federal Reserve System, July.

Appendix

A Log–Linear Solution of the Model

The optimality and equilibrium conditions are given by:

$$
\begin{aligned}
\frac{1}{C_t} &= \beta\left[1-\delta+\theta\frac{Y_{t+1}}{K_{t+1}}\right]\frac{1}{C_{t+1}}\\
BH_t &= (1-\theta)\frac{Y_t}{C_t}\\
K_{t+1} &= (1-\delta)K_t + K_t^{\theta}(Z_tH_t)^{1-\theta} - C_t\\
\Delta\log(Z_t) &= (\gamma_z-1)+\sigma_z\varepsilon_{z,t}.
\end{aligned}
$$

In this model, the technology shock Z_t induces a stochastic trend into output, consumption, investment and capital. Accordingly, to obtain a stationary equilibrium, these variables must be de–trended as follows

$$
\breve{y}_t = \frac{Y_t}{Z_t},\quad \breve{c}_t = \frac{C_t}{Z_t},\quad \breve{i}_t = \frac{I_t}{Z_t},\quad \breve{k}_{t+1} = \frac{K_{t+1}}{Z_t}.
$$

The log–linearization of equilibrium conditions around the deterministic steady state yields

$$
\begin{aligned}
\widehat{\breve{k}}_{t+1} &= \frac{(1-\delta)}{\gamma_z}(\widehat{\breve{k}}_t-\sigma_z\eta_{zt}) + \frac{y}{k}\widehat{\breve{y}}_t - \frac{c}{k}\widehat{\breve{c}}_t && (22)\\
\widehat{h}_t &= \widehat{\breve{y}}_t - \widehat{\breve{c}}_t && (23)\\
\widehat{\breve{y}}_t &= \theta(\widehat{\breve{k}}_t-\sigma_z\eta_{zt}) + (1-\theta)\widehat{h}_t && (24)\\
E_t\widehat{\breve{c}}_{t+1} &= \widehat{\breve{c}}_t + \alpha\beta\frac{y}{k}E_t(\widehat{\breve{y}}_{t+1}-\widehat{\breve{k}}_{t+1}-\sigma_z\eta_{zt+1}), && (25)
\end{aligned}
$$

where $y/k = (\gamma_z-\beta(1-\delta))/(\theta\beta\gamma_z)$ and $c/k = y/k-(\gamma_z+\delta-1)/\gamma_z$. After substitution of (23) into (24), one gets

$$
\widehat{\breve{y}}_t - \widehat{\breve{k}}_t = -\sigma_z\eta_{zt} - \frac{1-\theta}{\theta}\widehat{\breve{c}}_t
$$

Now, using the above expression, equations (22) and (25) rewrite

$$
E_t\widehat{\breve{c}}_{t+1} = \varphi\widehat{\breve{c}}_t \quad \text{with} \quad \varphi = \frac{\gamma_z\theta}{\gamma_z-\beta(1-\theta)(1-\delta)} \in (0,1) \qquad (26)
$$

$$
\begin{aligned}
\widehat{\breve{k}}_{t+1} &= \nu_1\widehat{\breve{k}}_t - \nu_1\sigma_z\eta_{zt} - \nu_2\widehat{\breve{c}}_t\\
&\quad \text{with } \nu_1 = \frac{1}{\beta\varphi} > 1 \quad \text{and} \quad \nu_2 = \frac{\gamma_z(1-\beta\theta^2)-\beta(1-\delta)(1-\theta^2)}{\theta^2\beta\gamma_z} \qquad (27)
\end{aligned}
$$

As $\nu_1 > 1$, (27) must be solved forward

$$
\widehat{\breve{k}}_t = \sigma_z\eta_{zt} + \left(\frac{\nu_2}{\nu_1}\right)\lim_{T\to\infty}E_t\sum_{i=0}^{T}\left(\frac{1}{\nu_1}\right)^i\widehat{\breve{c}}_{t+i} + \lim_{T\to\infty}E_t\left(\frac{1}{\nu_1}\right)^T\widehat{\breve{k}}_{t+T}
$$

Excluding explosive pathes, *i.e.* $\lim_{T\to\infty}E_t(1/\nu_1)^T\widehat{\breve{k}}_{t+T} = 0$, and using (26), one gets the decision rule on consumption:

$$
\widehat{\breve{c}}_t = \left(\frac{\nu_1-\varphi}{\nu_2}\right)\left(\widehat{\breve{k}}_t-\sigma_z\eta_{zt}\right) \qquad (28)
$$

After substituting (28) into (27), the dynamics of capital is given by:

$$\widehat{\breve{k}}_{t+1} = \varphi\left(\widehat{\breve{k}}_t - \sigma_z\eta_{zt}\right) \tag{29}$$

The persistence properties of the model is thus governed by the parameter $\varphi \in (0,1)$, which corresponds to the stable root of the log–linear version of the model. The decision rules of the other (deflated) variables are similar to equation (28). The hours worked are given by

$$\begin{aligned}
\widehat{h}_t &= \widehat{\breve{y}}_t - \widehat{\breve{c}}_t \\
&= -\nu\left(\widehat{\breve{k}}_t - \sigma_z\eta_{zt}\right) \\
&= -\nu\left(-\frac{\varphi}{1-\varphi L}\sigma_z\varepsilon_{z,t-1} - \sigma_z\eta_{zt}\right) \\
&= \nu\left(\frac{\sigma_z\eta_{zt}}{1-\varphi L}\right) \qquad (30)
\end{aligned}$$

where $\nu = \alpha(\nu_1 - \varphi - \theta\nu_2)/(\theta\nu_2)$ is a positive parameter. The latter expression shows that hours worked (or the consumption to output ratio) follows exactly the same stochastic process (an autoregressive process of order one) as the deflated capital $\log(K_t/Z_{t-1})$ in equation (29).

Using the expression for the growth rate of output

$$\Delta y_t = \widehat{\breve{y}}_t - \widehat{\breve{y}}_{t-1} + \sigma_z\eta_{zt},$$

we deduce

$$\Delta y_t = \sigma_z\eta_{zt} - \mu\frac{\sigma_z\Delta\eta_{zt}}{1-\varphi L}, \tag{31}$$

where $\mu = 1-(1-\theta)(\nu_1-\varphi)/(\alpha\nu_2)$.

B Proofs

Lemma 1 *Under the assumptions that $\sum_{i=0}^{\infty} i|a_{kj}^{(i)}| < \infty$ for $k,j = 1,2$ and that $\{\eta_t\}$ is a two dimensional i.i.d. sequence of structural shocks with zero mean, finite fourth moments and $E(\eta_t\eta_t') = I_2$, we get by Proposition 18.1 in Hamilton (1994, p.548)*

$$\begin{aligned}
\frac{1}{T}\sum_{t=2}^{T}\left(c\sum_{i=2}^{t}\eta_{2i-1}\right)\eta_{1t} &\xrightarrow{L} c\int_0^1 W_2(r)dW_1(r) \\
\frac{1}{T}\sum_{t=2}^{T}\left(c\sum_{i=2}^{t}\eta_{2i-1}\right)\eta_{2t} &\xrightarrow{L} c\int_0^1 W_2(r)dW_2(r) \\
\frac{1}{\sqrt{T}}\sum_{t=2}^{T}a_{kj}(L)\eta_{jt-1}\eta_{lt} &\xrightarrow{L} \mathcal{N}\left(0,\sum_{i=0}^{\infty}\left(a_{kj}^{(i)}\right)^2\right)
\end{aligned}$$

where $W_1(r)$ and $W_2(r)$ are two standardized independent Brownian motions $l = 1,2$.

B.1 Proof of Theorem 1

Let us first give the asymptotic variance of X_{2t} with $c > 0$ and $c = 0$. For the case where $c > 0$, we can show that

$$\psi_2 = \lim_{T\to\infty} \frac{1}{T}\sum_{t=2}^{T} X_{2t-1}X_{2t-1} = \sum_{i=0}^{\infty} a_{21}^2 + \sum_{i=0}^{\infty} a_{22}^2 + a_{22}(1)^2 c^2.$$

which depends on the parameter c. We also define $\psi_{2,0}$ as the asymptotic variance of X_{2t} but with $c = 0$.

Under the structural model (3) or (4), the numerator of the IV estimator of $b_{12}^{(0)}$ is given by[18]

$$\frac{1}{T}\sum_{t=2}^{T} X_{2t-1}\eta_{1t} = \frac{1}{T}\sum_{t=2}^{T}\left(a_{21}(L)\eta_{1t-1} + a_{22}(1)\frac{c}{\sqrt{T}}\sum_{j=2}^{t}\eta_{2j-1} + \tilde{a}_{22,T}^{*}(L)\eta_{2t-1}\right)\eta_{1t},$$

where the partial sum verify

$$a_{22}(1)\frac{c}{T^{1+1/2}}\underbrace{\sum_{t=2}^{T}\left(\sum_{j=2}^{t}\eta_{2j-1}\right)\eta_{1t}}_{O_p(T)} \xrightarrow{p} 0.$$

Asymptotically $\tilde{a}_{22,T}^{*}(L) = a_{22}(L)$ and by Lemma 1 this yields

$$\frac{1}{T}\sum_{t=2}^{T} a_{21}(L)\eta_{1t-1}\eta_{1t} \xrightarrow{p} 0 \quad \text{and} \quad \frac{1}{T}\sum_{t=2}^{T} \tilde{a}_{22t}^{*}(L)\eta_{2t-1}\eta_{1t} \xrightarrow{p} 0$$

which implies

$$\frac{1}{T}\sum_{t=2}^{T} X_{2t-1}\eta_{1t} \xrightarrow{p} 0. \tag{32}$$

Let us now examine the denominator. By inverting eq. (14) and (15) and using $B_0^{-1} = A_0$, we get

$$\Delta\widetilde{X}_{1t} = a_{12}^{(0)} b_{22}^{*}\widetilde{X}_{2t-1} + a_{11}^{(0)}\eta_{1t} + a_{12}^{(0)}\eta_{2t} \tag{33}$$

$$\Delta\widetilde{X}_{2t} = a_{22}^{(0)} b_{22}^{*}\widetilde{X}_{2t-1} + a_{21}^{(0)}\eta_{1t} + a_{22}^{(0)}\eta_{2t}. \tag{34}$$

This yields

$$\begin{aligned}\frac{1}{T}\sum_{t=2}^{T} X_{2t-1}\widetilde{\Delta}X_{2t} &= \frac{1}{T}\sum_{t=2}^{T} X_{2t-1}\Delta\widetilde{X}_{2t} \\ &= \frac{1}{T}\sum_{t=2}^{T} X_{2t-1}\left(a_{22}^{(0)} b_{22}^{*}\widetilde{X}_{2t-1} + a_{21}^{(0)}\eta_{1t} + a_{22}^{(0)}\eta_{2t}\right) \xrightarrow{p} a_{22}^{(0)} b_{22}^{*}\widetilde{\psi}_2 \end{aligned} \tag{35}$$

where $\widetilde{\psi}_2 = \lim_{T\to\infty}\frac{1}{T}\sum_{t=2}^{T}\widetilde{X}_{2t-1}\widetilde{X}_{2t-1}$. Since X_{2t} is asymptotically second order stationary and $\widetilde{\psi}_2 \leq \psi_2$ by the projection properties, $\widetilde{\psi}_2$ is then bounded. By combining (32) and (35), we get the result that $\widehat{b}_{12}^{(0)} - b_{12}^{(0)} \xrightarrow{p} 0$.

[18] To simplify, we suppose here that the initial values are zero.

Now we establish the convergence in distribution of $b_{12}^{(0)}$. Thus

$$\sqrt{T}\left(\widehat{b}_{12}^{(0)}-b_{12}^{(0)}\right)=\sqrt{T}\left[\frac{\frac{1}{T}\sum_{t=2}^{T}X_{2t-1}\eta_{1t}}{\frac{1}{T}\sum_{t=2}^{T}X_{2t-1}\Delta\widetilde{X}_{2t}}\right].$$

For the numerator

$$\frac{1}{\sqrt{T}}\sum_{t=2}^{T}X_{2t-1}\eta_{1t} \quad = \quad \frac{1}{\sqrt{T}}\sum_{t=2}^{T}\left[a_{21}(L)\eta_{1t-1}+a_{22}(1)\frac{c}{\sqrt{T}}\sum_{j=2}^{t}\eta_{2j-1}+\tilde{a}_{22,T}^{*}(L)\eta_{2t-1}\right]\eta_{1t}.$$

By using Lemma 1, eq. (35) and noting that $\tilde{a}_{22,T}^{*}(L)\to a_{22}(L)$, we deduce

$$\sqrt{T}\left(\widehat{b}_{12}^{(0)}-b_{12}^{(0)}\right)\xrightarrow{L}\frac{1}{a_{22}^{(0)}b_{22}^{*}\widetilde{\psi}_{2}}\left[a_{22}(1)c\int_{0}^{1}W_{2}(r)dW_{1}(r)+\psi_{2,0}^{1/2}\xi_{1}\right],$$

where ξ_1 is the normal distribution $\mathcal{N}(0,1)$.

Consider now the estimator $\widehat{\beta}$

$$\widehat{\beta}=\left[\frac{1}{T}\sum_{t=2}^{T}z_{t}x_{t}'\right]^{-1}\left[\frac{1}{T}\sum_{t=2}^{T}z_{t}\Delta\widetilde{X}_{2t}\right]=\left[\frac{1}{T}\sum_{t=2}^{T}z_{t}x_{t}'\right]^{-1}\left[\frac{1}{T}\sum_{t=2}^{T}z_{t}(x_{t}'\beta+\eta_{2t})\right].$$

This yields

$$\widehat{\beta}-\beta=\left[\frac{1}{T}\sum_{t=2}^{T}z_{t}x_{t}'\right]^{-1}\left[\frac{1}{T}\sum_{t=2}^{T}z_{t}\eta_{2t}\right]$$

where more explicitly

$$\begin{aligned}\frac{1}{T}\sum_{t=2}^{T}z_{t}x_{t}' &= \begin{bmatrix}\frac{1}{T}\sum_{t=2}^{T}\Delta\widetilde{X}_{1t}\left[\eta_{1t}-(\widehat{b}_{12}^{(0)}-b_{12}^{(0)})\Delta\widetilde{X}_{2t}\right] & \frac{1}{T}\sum_{t=2}^{T}\widetilde{X}_{2t-1}\left[\eta_{1t}-(\widehat{b}_{12}^{(0)}-b_{12}^{(0)})\Delta\widetilde{X}_{2t}\right]\\ \frac{1}{T}\sum_{t=2}^{T}X_{2t-1}\Delta\widetilde{X}_{1t} & \frac{1}{T}\sum_{t=2}^{T}X_{2t-1}\widetilde{X}_{2t-1}\end{bmatrix}\\ &= \begin{bmatrix}G_{11,T} & G_{12,T}\\ G_{21,T} & G_{22,T}\end{bmatrix},\end{aligned}$$

and

$$\frac{1}{T}\sum_{t=2}^{T}z_{t}\eta_{2t}=\begin{bmatrix}\frac{1}{T}\sum_{t=2}^{T}\left[\eta_{1t}-(\widehat{b}_{12}^{(0)}-b_{12}^{(0)})\Delta\widetilde{X}_{2t}\right]\eta_{2t}\\ \frac{1}{T}\sum_{t=2}^{T}X_{2t-1}\eta_{2t}\end{bmatrix}. \tag{36}$$

Let us examine the first term $G_{11,T}$,

$$G_{11,T} \quad = \quad \frac{1}{T}\sum_{t=2}^{T}\Delta\widetilde{X}_{1t}\eta_{1t}-\left(\widehat{b}_{12}^{(0)}-b_{12}^{(0)}\right)\frac{1}{T}\sum_{t=2}^{T}\Delta\widetilde{X}_{1t}\Delta\widetilde{X}_{2t}=\frac{1}{T}\sum_{t=2}^{T}\Delta\widetilde{X}_{1t}\eta_{1t}+o_{p}(1)$$

by $(\widehat{b}_{12}^{(0)}-b_{12}^{(0)})\xrightarrow{p}0$ and $\frac{1}{T}\sum_{t=2}^{T}\Delta\widetilde{X}_{1t}\Delta\widetilde{X}_{2t}=O_p(1)$. Using eq. (33), this gives $\frac{1}{T}\sum_{t=2}^{T}\Delta\widetilde{X}_{1t}\eta_{1t}\xrightarrow{p}a_{11}^{(0)}$ which implies $G_{11,T}\xrightarrow{p}a_{11}^{(0)}$. The term $G_{21,T}=\frac{1}{T}\sum_{t=2}^{T}X_{2t-1}\Delta\widetilde{X}_{1t}\xrightarrow{p}a_{12}^{(0)}b_{22}^{*}\widetilde{\psi}_{2}$ using eq. (33). For the upper right term, we obtain

$$G_{12,T}=\frac{1}{T}\sum_{t=2}^{T}\widetilde{X}_{2t-1}\left[\eta_{1t}-(\widehat{b}_{12}^{(0)}-b_{12}^{(0)})\Delta\widetilde{X}_{2t}\right]=\frac{1}{T}\sum_{t=2}^{T}\widetilde{X}_{2t-1}\eta_{1t}+o_{p}(1)$$

and $\frac{1}{T}\sum_{t=2}^{T}\widetilde{X}_{2t-1}\eta_{1t}\xrightarrow{p}0$ by (32). Finally, $G_{22,T}=\frac{1}{T}\sum_{t=2}^{T}X_{2t-1}\widetilde{X}_{2t-1}\xrightarrow{p}\widetilde{\psi}_2$.

Let us examine the expression (36). We have the following results $\frac{1}{T}\sum_{t=2}^{T}\eta_{1t}\eta_{2t}\xrightarrow{p}0$, $(\widehat{b}_{12}^{(0)}-b_{12}^{(0)})\frac{1}{T}\sum_{t=2}^{T}\Delta\widetilde{X}_{2t}\eta_{2t}\xrightarrow{p}0$ and $\frac{1}{T}\sum_{t=2}^{T}X_{2t-1}\eta_{2t}\xrightarrow{p}0$. We can now conclude that $\widehat{\beta}-\beta\xrightarrow{p}0$.

To establish the convergence in distribution of $b_{21}^{(0)}$ and b_{22}^{*}, we use the following expression

$$\sqrt{T}\begin{bmatrix}\widehat{b}_{21}^{(0)}-b_{21}^{(0)}\\ \widehat{b}_{22}^{*}-b_{22}^{*}\end{bmatrix}=\frac{1}{G_{11,T}G_{22,T}-G_{21,T}G_{12,T}}\begin{bmatrix}G_{22,T} & -G_{12,T}\\ -G_{21,T} & G_{11,T}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{T}}\sum_{t=2}^{T}\left[\eta_{1t}-(\widehat{b}_{12}^{(0)}-b_{12}^{(0)})\Delta\widetilde{X}_{2t}\right]\eta_{2t}\\ \frac{1}{\sqrt{T}}\sum_{t=2}^{T}X_{2t-1}\eta_{2t}\end{bmatrix}. \tag{37}$$

We now examine the first term at the RHS of (37). From the results derived above, we obtain $G_{11,T}G_{22,T}\xrightarrow{p}a_{11}^{(0)}\widetilde{\psi}_2$ and $G_{21,T}G_{12,T}\xrightarrow{p}0$. For the last term of (37), we have $\frac{1}{\sqrt{T}}\sum_{t=2}^{T}\eta_{1t}\eta_{2t}\xrightarrow{L}\xi_2$ where $\xi_2\sim\mathcal{N}(0,1)$ and

$$\sqrt{T}\left(\widehat{b}_{12}^{(0)}-b_{12}^{(0)}\right)\frac{1}{T}\sum_{t=2}^{T}\Delta\widetilde{X}_{2t}\eta_{2t}\xrightarrow{L}\frac{\psi_{2,0}^{1/2}}{b_{22}^{*}\widetilde{\psi}_2}\xi_1+\frac{\vartheta_1}{b_{22}^{*}\widetilde{\psi}_2},$$

since $\frac{1}{T}\sum_{t=2}^{T}\Delta\widetilde{X}_{2t}\eta_{2t}\xrightarrow{p}a_{22}^{(0)}$ with $\vartheta_1=a_{22}(1)c\int_0^1 W_2(r)dW_1(r)$ derived above. Now for the expression $\frac{1}{\sqrt{T}}\sum_{t=2}^{T}X_{2t-1}\eta_{2t}$,

$$\frac{\sqrt{T}}{T}\sum_{t=2}^{T}X_{2t-1}\eta_{2t}=\frac{1}{T}\sum_{t=2}^{T}\sqrt{T}\left[a_{21}(L)\eta_{1t-1}+a_{22}(1)\frac{c}{\sqrt{T}}\sum_{j=2}^{t}\eta_{2j-1}+\tilde{a}_{22,T}^{*}(L)\eta_{2t-1}\right]\eta_{2t}.$$

By using Lemma 1,

$$\frac{\sqrt{T}}{T}\sum_{t=2}^{T}X_{2t-1}\eta_{2t}\xrightarrow{L}\left[a_{22}(1)c\int_0^1 W_2(r)dW_2(r)+\psi_{2,0}^{1/2}\xi_2\right].$$

By collecting these results, we obtain that

$$\sqrt{T}\left(\widehat{b}_{21}^{(0)}-b_{21}^{(0)}\right)\xrightarrow{L}\frac{\xi_2}{a_{11}^{(0)}}-\frac{\psi_{2,0}^{1/2}}{a_{11}^{(0)}b_{22}^{*}\widetilde{\psi}_2}\xi_1-\frac{\vartheta_1}{a_{11}^{(0)}b_{22}^{*}\widetilde{\psi}_2}.$$

Now for $\widehat{b}_{22}^{*}$, we get

$$\sqrt{T}\left(\widehat{b}_{22}^{*}-b_{22}^{*}\right)\xrightarrow{L}\frac{a_{12}^{(0)}\psi_{2,0}^{1/2}}{a_{11}^{(0)}\widetilde{\psi}_2}\xi_1+\frac{a_{12}^{(0)}\vartheta_1}{a_{11}^{(0)}\widetilde{\psi}_2}-\left[\frac{a_{12}^{(0)}b_{22}^{*}}{a_{11}^{(0)}}-\frac{\psi_{2,0}^{1/2}}{\widetilde{\psi}_2}\right]\xi_2+\frac{\vartheta_2}{\widetilde{\psi}_2}$$

where $\vartheta_2=a_{22}(1)c\int_0^1 W_2(r)dW_2$.

B.2 Proof of Proposition 1

According to the structural representation (4), the finite sample measure of the long-run impact is given by:

$$A_T(1)=\begin{bmatrix}a_{11}(1) & a_{12}(1)c/\sqrt{T}\\ 0 & a_{22}(1)c/\sqrt{T}\end{bmatrix}.$$

This matrix is not lower triangular as imposed in the identification procedure of the SVAR. The corresponding long–run variance-covariance matrix is then:

$$A_T(1)A_T(1)' = \begin{bmatrix} a_{11}(1)^2 + a_{12}(1)^2c^2/T & a_{12}(1)a_{22}(1)c^2/T \\ a_{12}(1)a_{22}(1)c^2/T & a_{22}(1)^2c^2/T \end{bmatrix}. \tag{38}$$

Now, we wrongly impose a lower triangular form using a Choleski decomposing on (38). In that respect, we can rewrite the expression above using equation (4.4.12) in Hamilton (1994 p.90) as

$$\begin{aligned} A_T(1)A_T(1)' &= \begin{bmatrix} 1 & 0 \\ \frac{a_{12}(1)a_{22}(1)c^2/T}{a_{11}(1)^2+a_{12}(1)^2c^2/T} & 1 \end{bmatrix} \begin{bmatrix} a_{11}(1)^2 + a_{12}(1)^2c^2/T & 0 \\ 0 & a_{22}(1)^2c^2/T - \frac{a_{12}(1)^2a_{22}(1)^2c^4/T^2}{a_{11}(1)^2+a_{12}(1)^2c^2/T} \end{bmatrix} \\ &\times \begin{bmatrix} 1 & \frac{a_{12}(1)a_{22}(1)c^2/T}{a_{11}(1)^2+a_{12}(1)^2c^2/T} \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

By a Choleski decomposition for $A_T(1)A_T(1)'$ we obtain the lower triangular matrix:

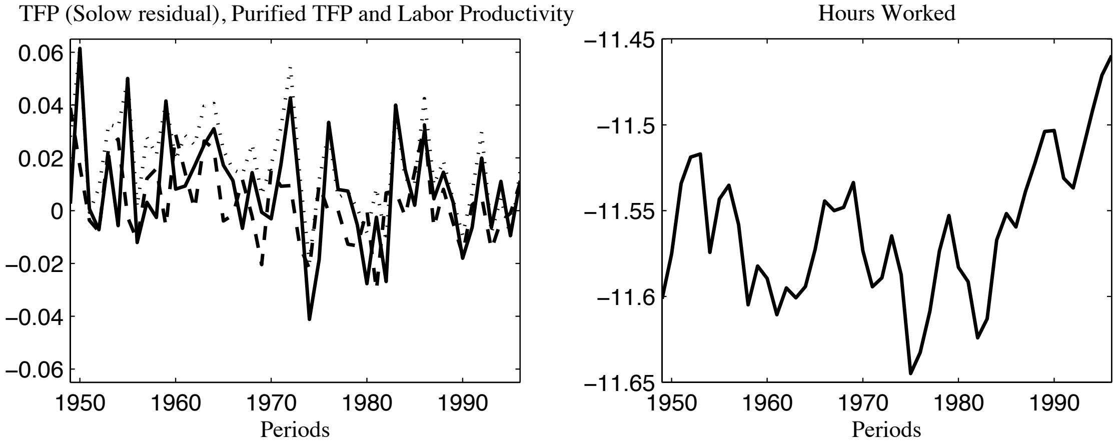
$$\mathit{chol}\left(A_T(1)A_T(1)'\right) = \begin{bmatrix} \left(a_{11}(1)^2 + a_{12}(1)^2c^2/T\right)^{1/2} & 0 \\ \frac{a_{12}(1)a_{22}(1)c^2/T}{\left(a_{11}(1)^2+a_{12}(1)^2c^2/T\right)^{1/2}} & \left(a_{22}(1)^2c^2/T - \frac{a_{12}(1)^2a_{22}(1)^2c^4/T^2}{a_{11}(1)^2+a_{12}(1)^2c^2/T}\right)^{1/2} \end{bmatrix}.$$

Table 1: Long–Run effect of a Technology Improvement on Productivity Measures (in %)

	LSVAR model	DSVAR Model
Solow Residual	1.51	1.63
	[0.90;2.93]	[0.86;2.48]
"Purified" Measure of TFP	1.33	1.39
	[0.80;2.18]	[0.78;2.09]
Labor Productivity	2.05	2.32
	[1.06;3.92]	[1.10;3.77]

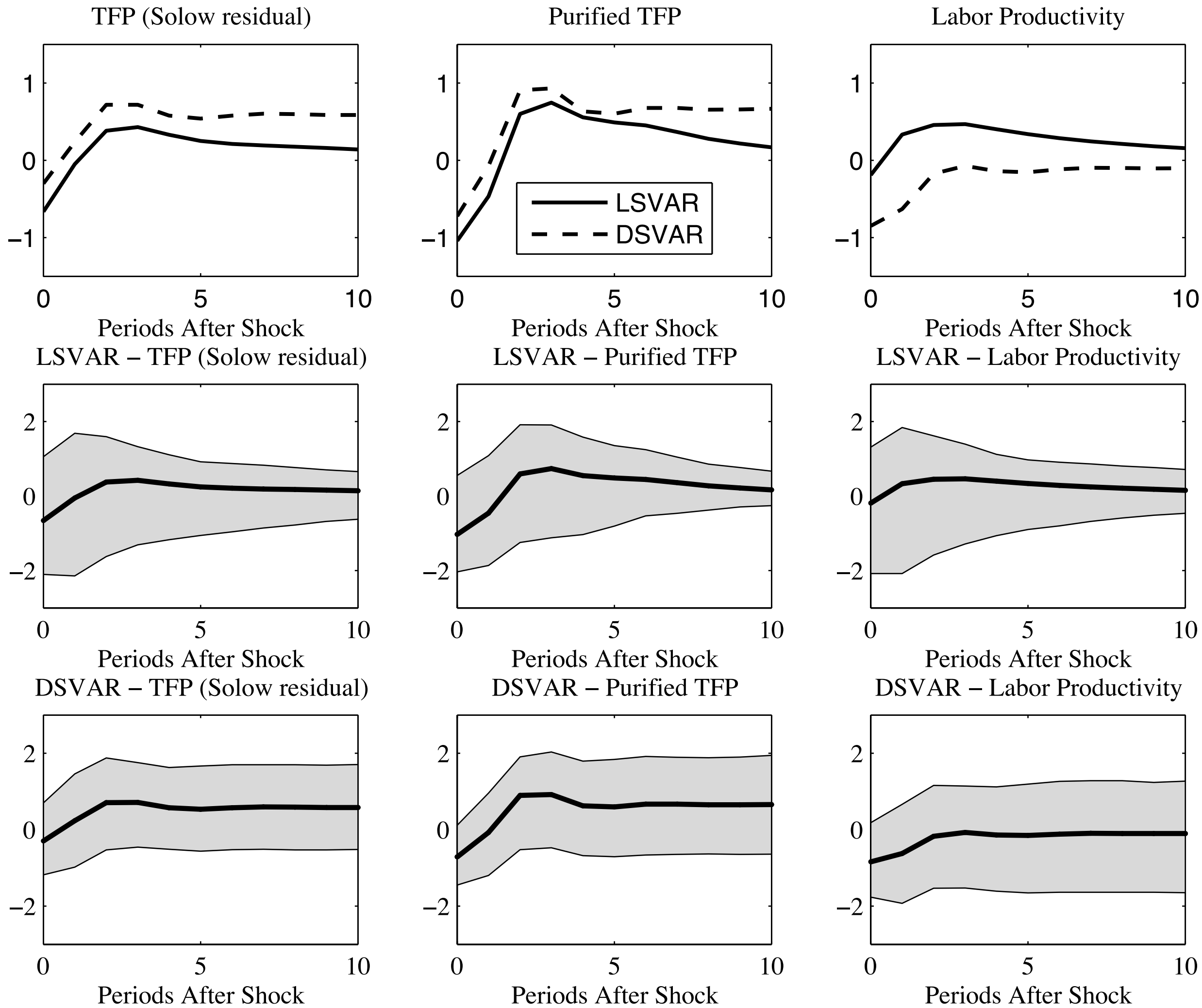
Notes: 95% percent confidence interval in brackets obtained from a standard bootstrap technique with 1000 replications.

Figure 1: US Data



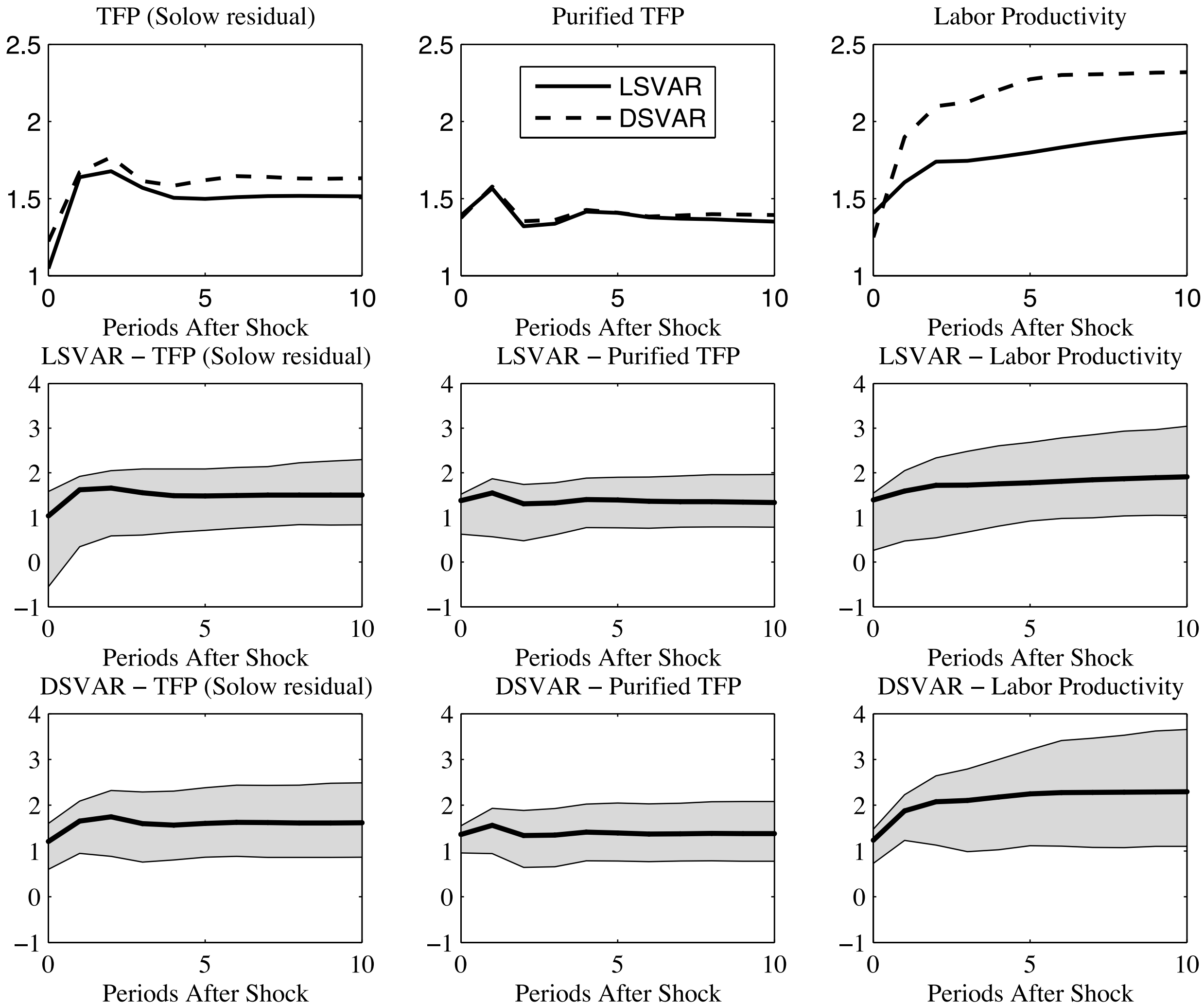
Note: The left hand side of the figure reports different measures of productivity. The solid line refers to the Solow residual, the dashed line to the purified measure of TFP and the dotted line to the labor productivity. All are specified in logs and in first difference. The right hand side reports the log of per capita hours worked. The data are at annual frequency and cover the sample period 1949–1996.

Figure 2: IRFs of Hours Worked to a Technological Improvement



Note: The DSVAR model includes alternatively the growth rate of the Solow residual, the "purified" measure of TFP and the labor productivity, and the log of hours in first difference. The LSVAR model includes alternatively the growth rate of the Solow residual, the "purified" measure of TFP and the labor productivity, and the log of hours in level. The sample period is 1949–1996. Two lags are included in each VAR model. The selected horizon for IRFs is 11. 95% percent confidence interval obtained from a standard bootstrap technique with 1000 replications.

Figure 3: IRFs of Technology Measures to a Technological Improvement



Note: The DSVAR model includes alternatively the growth rate of the Solow residual, the "purified" measure of TFP and the labor productivity, and the log of hours in first difference. The LSVAR model includes alternatively the growth rate of the Solow residual, the "purified" measure of TFP and the labor productivity, and the log of hours in level. The sample period is 1949–1996. Two lags are included in each VAR model. The selected horizon for IRFs is 11. 95% percent confidence interval obtained from a standard bootstrap technique with 1000 replications.