



Avoiding pitfalls in using structural VARs to estimate economic models [☆]

Martial Dupaigne ^{a,b}, Patrick Fève ^{b,c,d,*}, Julien Matheron ^d

^a *University of Western Brittany, France*

^b *GREMAQ—Université de Toulouse 1, Manufacture des Tabacs, Aile J.-J. Laffont, 21 allée de Brienne, 31000 Toulouse, France*

^c *IDEL, France*

^d *Banque de France, Research Division, Paris, France*

Received 17 November 2005; revised 6 June 2006

Available online 8 February 2007

Abstract

Structural Vector Autoregressions with a differenced specification of hours (DSVAR) suggest that productivity shocks identified using long-run restrictions lead to a persistent and significant decline in hours worked. Economists have interpreted this evidence as showing that standard business cycle models in which a positive technology shock leads to a rise in hours are inconsistent with the data. In this paper we argue that such a conclusion is unwarranted because model's data and actual data are not treated symmetrically. To illustrate this problem, we estimate and test a flexible-price DSGE model with non-stationary hours using Indirect Inference on impulse responses of hours and output after technology and non-technology shocks. We find that, once augmented with a moderate amount of real frictions, the model can mimic well impulse responses obtained from a DSVAR on actual data. Using this model as a data generating process, we show that our estimation method is less subject to bias than a method that would directly compare theoretical responses with responses from the DSVAR.

© 2006 Elsevier Inc. All rights reserved.

JEL classification: E24; E32

Keywords: DSVARs; Long-run restrictions; DSGE models; Non-stationary hours; Indirect Inference

[☆] The views expressed herein are those of the authors and not necessarily those of the Banque de France.

* Corresponding author.

E-mail address: patrick.feve@univ-tlse1.fr (P. Fève).

0. Introduction

This paper combines dynamic stochastic general equilibrium (DSGE) models and Structural Vector Autoregressions (SVARs) to characterize the joint dynamics of hours worked and labor productivity.

The literature on the subject is currently burgeoning. In a provocative paper, Galí (1999) argues that the response of hours to a technology shock is persistently and significantly negative in the US as well as in other G7 countries (see also Galí, 2004 regarding the euro area). This claim is based on SVARs of the growth rates of labor productivity and hours with long-run restrictions. The negative response of hours is interpreted as invalidating a whole class of business cycle models, namely technology-driven flexible price ones (Galí and Rabanal, 2004 and Francis and Ramey, 2005). These SVARs results and their interpretation have raised major controversies. On the one hand, Christiano et al. (2004) provide evidence in favor of a level-stationary specification of hours worked (LSVAR), as opposed to the difference-stationary (or DSVAR) specification in Galí (1999).¹ On the other hand, Erceg et al. (2005) and Chari et al. (2005) use DSGE models estimated on US data as Data Generating Processes (DGP) to evaluate the identification of technology shocks provided by DSVARs. Chari et al. (2005) show that a DSVAR model leads to a negative response of hours under a business cycle model in which hours respond positively. They conclude that SVARs with long-run restrictions are not useful guides for developing business cycle theories.²

The contribution of the present paper is twofold. First, we develop a DSGE model with non-stationary hours. In contrast to related studies, our model is in close conformity with the DSVAR specification advocated by Galí (1999): hours being non-stationary, the DSVAR is the correct specification to investigate. The model also embeds time non-separability in leisure choices. We select the parameters of the model so that a DSVAR estimated on simulated data from the model would mimic a DSVAR estimated on actual data. More precisely, the method used in this paper picks the parameters that minimize the discrepancy between the simulated Impulse Response Functions (IRFs)—those from the DSVAR estimated on model's data—and the empirical IRFs—those from the DSVAR estimated on actual data. While our model is technology-driven, features flexible prices and displays a positive response of hours to a technology shock, it can generate simulated IRFs that look like the Galí (1999) finding. This calls into question the conclusions reached by Galí (1999) and Francis and Ramey (2005).

Second, we illustrate the merits of our approach through a comparison with a popular alternative methodology used by Rotemberg and Woodford (1997), Christiano et al. (2005a), and Altig et al. (2005), among others. This alternative methodology selects parameter values so that the theoretical IRFs are as close as possible to the empirical IRFs obtained from a DSVAR. We work through an example and show that this methodology delivers severely biased estimates of the model's structural parameters, and underperforms the method we suggest.

The methodology used in this paper is close in spirit to Sims's (1989) general principle, subsequently adopted by Cogley and Nason (1995): treating models and data symmetrically. With such an approach, the IRFs from the DSVAR are given no a priori structural interpretation. They are just moments from the data, and the DSVAR is similar to a filter.³ However, these moments

¹ With their LSVAR specification, Christiano et al. (2004) show that hours increase after a permanent technology shock.

² See also Cooley and Dwyer (1998) criticism of SVARs.

³ The DSVAR plays the role of the *auxiliary model* in Gouriéroux et al. (1993).

can be useful to evaluate a theoretical model, provided the associated model's quantitative implications vary appropriately with the parameters of interest (loosely speaking, it is the analog of the likelihood function not being flat in the relevant dimension of the parameter space). Indeed, our results suggest that using IRFs from a DSVAR as moments to be matched allows us to discriminate between different DSGE models. In particular, real frictions (here, under the form of habit persistence) are essential to mimic the data.

The direct comparison of theoretical responses and IRFs from a DSVAR would only make sense if the DSVAR properly recovers structural shocks and the associated dynamic propagation (see Kehoe, 2006). In what follows, we label Direct Inference (or direct approach) the estimation of structural parameter based on the distance between theoretical responses and IRFs from a SVAR model estimated on actual data. The danger of this approach is obvious: if the SVAR is biased, as shown by Erceg et al. (2005) and Chari et al. (2005), a direct evaluation of a DSGE model will prove highly misleading. We investigate this issue taking our estimated model as the data generating process. We then compare the relative performances of the direct approach with our Indirect Inference methodology. To begin with, we generate data from our model and estimate model's parameters using these competing approaches. Doing this a large number of times allows us to obtain a population of parameter estimates under each estimation methods, which we can compare to the true value and to each other. We show that the parameters we get with Direct Inference can be rather far from the parameters we fixed for simulation. In contrast, the parameters obtained by Indirect Inference are very close to their true values.

To ensure that our DSGE model satisfies the key identifying assumptions of the DSVAR, we follow Chang et al. (2007) and consider two permanent shocks in our theoretical framework. In addition to a permanent technology shock, the model includes a permanent preference shock which yields non-stationary hours. Importantly, this model preserves the property that only technology shocks exert a permanent effect on labor productivity. Non-stationary hours and the long-run impact of technology shocks on labor productivity are perfectly in accordance with the DSVAR adopted by Galí (1999) and Galí and Rabanal (2004).

Modeling hours as a non-stationary process is of course debatable, since hours per capita admit a physical constraint. As most of the literature does, we ignore this constraint in our quantitative investigation. Hence, the resulting dynamics must be viewed as a local approximation. Note that from an empirical point of view, this property is supported by the data.

The paper is organized as follows. In Section 1, we reproduce Galí's (1999) findings obtained from a DSVAR model estimated on quarterly post war US data. The resulting IRFs will be treated as the moments of interest. In Section 2, we expound our DSGE model with non-stationary hours. Section 3 is devoted to an exposition of our quantitative exercise based on Indirect Inference. Section 4 provides a comparison between this indirect approach and the popular direct estimation. The last section briefly concludes.

1. Estimated impulse responses for US data

The first step in our procedure is to reproduce the results obtained by Galí (1999), and define the moments that will be matched in our quantitative exercise. To do so, we first consider the following VAR(p) model

$$\mathbf{z}_t = \mathbf{A}_1 \mathbf{z}_{t-1} + \dots + \mathbf{A}_p \mathbf{z}_{t-p} + \boldsymbol{\varepsilon}_t, \quad E\{\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'\} = \boldsymbol{\Sigma}, \quad (1)$$

with $\mathbf{z}_t = (\Delta x_t, \Delta n_t)'$, where x_t is logged labor productivity, n_t is logged total hours worked per capita, and Δ is the first difference operator.⁴ To simplify the exposition, we abstract from constant terms in the above VAR model. This specification with hours worked in first difference is used by Galí (1999, 2004, 2005), Galí and Rabanal (2004), and Francis and Ramey (2005). We follow Galí and Rabanal (2004) and set p to 4.

This VAR model is used to identify technology and non-technology shocks. Let us define $\mathbf{B}(L) = (\mathbf{I}_2 - \mathbf{A}_1 L - \dots - \mathbf{A}_p L^p)^{-1}$, so that

$$\mathbf{z}_t = \mathbf{B}(L)\boldsymbol{\varepsilon}_t,$$

with \mathbf{I}_2 the identity matrix. Now, we assume that the canonical innovations $\boldsymbol{\varepsilon}_t$ are linear combinations of the structural shocks $\boldsymbol{\eta}_t$ for some non singular matrix \mathbf{S} , i.e. $\boldsymbol{\varepsilon}_t = \mathbf{S}\boldsymbol{\eta}_t$. As usual, we impose an orthogonality assumption on the structural shocks which, together with a scale normalization, implies $\mathbf{E}\boldsymbol{\eta}_t\boldsymbol{\eta}_t' = \mathbf{I}_2$. This gives us three constraints out of the four needed to completely identify \mathbf{S} .

To setup the last identifying constraint, let us define $\mathbf{C}(L) = \mathbf{B}(L)\mathbf{S}$. Given the ordering of \mathbf{z}_t , we simply require that $\mathbf{C}(1)$ be lower triangular, so that only technology shocks can affect the long-run level of labor productivity. This amounts to imposing that $\mathbf{C}(1)$ is the Cholesky factor of $\mathbf{B}(1)\boldsymbol{\Sigma}\mathbf{B}(1)'$. Given consistent estimates of $\mathbf{B}(1)$ and $\boldsymbol{\Sigma}$, we easily obtain an estimate for $\mathbf{C}(1)$. Retrieving \mathbf{S} is then a simple task using the formula $\mathbf{S} = \mathbf{B}(1)^{-1}\mathbf{C}(1)$. The impulse responses are then deduced from the VMA(∞) representation

$$\mathbf{z}_t = \mathbf{B}(L)\mathbf{B}(1)^{-1}\mathbf{C}(1)\boldsymbol{\eta}_t \quad (2)$$

with $\boldsymbol{\eta}_t = (\boldsymbol{\eta}_{1,t}, \boldsymbol{\eta}_{2,t})'$, where $\boldsymbol{\eta}_{1,t}$ is the identified technology shock, whereas $\boldsymbol{\eta}_{2,t}$ is the non-technology one. The standard-errors of the IRFs are computed numerically using the δ -function method (see Hamilton, 1994).⁵

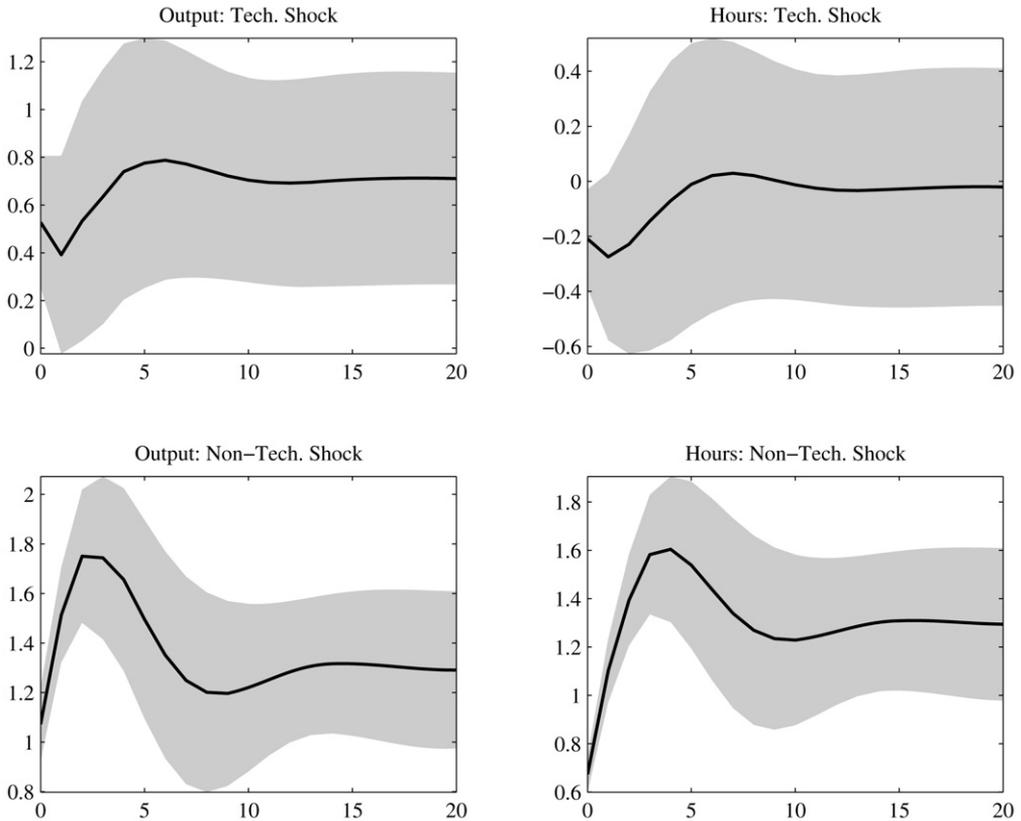
The IRFs are reported on Fig. 1. Several salient features emerge from this figure. Output increases permanently in response to either a technology shock or a non-technology shock, hours do not respond asymptotically to technology shocks, while non-technology shocks generate a long-run effect on hours. The impulse responses of output and hours to the non-technology shock are persistent and hump-shaped. These responses are estimated precisely for each horizon. In addition, as in Galí (1999), hours worked decrease significantly on impact after a technology shock. Moreover, the negative effect is rather persistent.

Confronted with these DSVAR evidence, business cycle students have one of two choices. Either they take them at face value, thus discarding a host of DSGE models previously thought to be useful, as suggested by Galí (1999) and Francis and Ramey (2005); or they can treat these IRFs as what they are, i.e. moments from the data that need not be given a structural interpretation, especially so if the DSVAR is highly biased, as forcefully put forth by Chari et al. (2005) and, to a lesser extent, by Erceg et al. (2005).

The purpose of this paper is to provide a quantitative tool that simply treats these IRFs as arbitrarily-chosen moments from the data with no necessary interpretation, thus preventing business cycle students from reaching incorrect conclusions as to the dynamic effects of potentially ill-identified (technology) shocks. Doing so requires first that we develop a DSGE model with

⁴ Our quarterly data for the period 1948:1–2002:4 are constructed as follows. Labor productivity is defined as non-farm business output divided by non-farm business hours worked; hours worked are non-farm business hours worked divided by civilian population over the age of 16.

⁵ We obtain virtually identical results using a simple bootstrap method.



Note: The IRFs are obtained from a DSVAR with four lags estimated on productivity growth and hours growth. The data are for the Non-Farm Business sector, over the sample period 1948:1–2002:4. The solid line represents the point estimate and the gray area corresponds to the 90% asymptotic confidence interval.

Fig. 1. IRFs of output and hours to technology and non-technology shocks.

non-stationary hours which is completely compatible with the DSVAR specification and the identification scheme proposed by Galí (1999). The next section describes the model.

2. A DSGE model with non-stationary hours

We build a flexible price DSGE model with non-stationary hours, due to a permanent preference shock (B_t), along the lines of Chang et al. (2007). In addition, the model includes a random walk productivity shock (Z_t).

Following Bouakez and Kano (2006), we assume that labor supply choices are time non-separable: the service flows from leisure are a function of current and once-lagged leisure choices. The intertemporal expected utility function of the representative household is given by

$$E_0 \sum_{t=0}^{\infty} \beta^t \left\{ U(C_t) - V\left(\frac{N_t}{B_t}, \frac{N_{t-1}}{B_{t-1}}\right) \right\},$$

where

$$U(C_t) = \log(C_t),$$

$$V\left(\frac{N_t}{B_t}, \frac{N_{t-1}}{B_{t-1}}\right) = \frac{1}{1+\chi} \left[\frac{N_t/B_t}{(N_{t-1}/B_{t-1})^a} \right]^{1+\chi}, \quad \chi > 0.$$

Here $\beta \in (0, 1)$ denotes the discount factor and E_0 is the expectation operator conditional on the information set available as of time $t = 0$. C_t is the consumption at date t and N_t represents the household's labor supply. When $a < 0$, the model implies that labor supply is intertemporally substitutable, while $a > 0$ implies intertemporal complementarities. The labor supply N_t is subject to a permanent preference shock B_t , that follows the stochastic process

$$\log(B_t) = \log(B_{t-1}) + \sigma_B \varepsilon_{B,t},$$

where $\sigma_B > 0$, and $\varepsilon_{B,t}$ is *iid* with zero mean and unit variance. As noticed by Galí (2005), this shock can be an important source of fluctuations, as it accounts for permanent shifts in the marginal rate of substitution between goods and labor (see Hall, 1997).

The representative firm uses capital K_t and labor N_t to produce the homogeneous final good Y_t . The technology is represented by the following constant returns-to-scale Cobb–Douglas production function

$$Y_t = K_t^{1-\alpha} (Z_t N_t)^\alpha,$$

where $\alpha \in (0, 1)$. Z_t is assumed to follow an exogenous process of the form

$$\log(Z_t) = g_Z + \log(Z_{t-1}) + \sigma_Z \varepsilon_{Z,t},$$

where $\sigma_Z > 0$ and $\varepsilon_{Z,t}$ is *iid* with zero mean and unit variance. The constant g_Z is a drift term in the random walk process of Z_t . 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. Finally, the final good can be either consumed or invested

$$Y_t = C_t + I_t.$$

The first order conditions implied by the model are standard, except that relating to labor choices, which can be written as

$$\frac{\partial V(N_t/B_t, N_{t-1}/B_{t-1})}{\partial N_t} + \beta E_t \frac{\partial V(N_{t+1}/B_{t+1}, N_t/B_t)}{\partial N_t} = U'(C_t) \alpha \frac{Y_t}{N_t}.$$

As usual, this condition equates the marginal disutility of work to the marginal utility of consuming the marginal product of labor. In our setup, past hours worked affect current utility. Therefore, the marginal disutility of labor involves past, current and expected hours worked. In the intertemporal complementarity case (respectively substitutability), current increases in labor supply reduce (respectively raise) next period's marginal disutility of labor.

In this model, the labor supply shock B_t has a permanent effect on hours as well as on output, consumption, investment, and capital. In addition, Z_t has a long-run effect on Y_t , C_t , K_t , and I_t , but no long-run effect on N_t . Accordingly, it is important to notice that in our model Z_t alone can have a long-run effect on labor productivity. Hence, this model is perfectly compatible with the first difference specification for hours and the identification assumptions used in the DSVAR previously estimated.⁶

⁶ In the DGPs considered in Erceg et al. (2005) and Chari et al. (2005), hours worked are stationary. Facing data that originate from these models, the natural statistical representations are VARs with hours in level, not in first-differences.

To obtain a stationary equilibrium these variables must be rescaled as follows

$$\tilde{N}_t = \frac{N_t}{B_t}, \quad \tilde{Y}_t = \frac{Y_t}{Z_t B_t}, \quad \tilde{C}_t = \frac{C_t}{Z_t B_t}, \quad \tilde{I}_t = \frac{I_t}{Z_t B_t}, \quad \tilde{K}_{t+1} = \frac{K_{t+1}}{Z_t B_t}.$$

With these transformations, a well-behaved deterministic steady state exists, which is akin to what would stem from a standard neoclassical growth model. The approximate solution of the model is computed from a log-linearization of the stationary equilibrium conditions around this deterministic steady state using the numerical algorithm of Anderson and Moore (1985).

Non-stationary hours receive empirical support. First, standard univariate unit root tests (ADF, KPSS) lead in general to accept the hypothesis of a unit-root in hours worked (Galí and Rabanal, 2004; Francis and Ramey, 2005; Christiano et al., 2004).⁷ Second, Galí (2005) regresses hours on the consumption–output ratio (in logs). When the data are unfiltered, he finds that this regression has little explanatory power. However, with data in first-differences, the fit is greatly improved, consistently with basic neoclassical theory. Christiano et al. (2005b) run a closely related regression and obtain a highly autocorrelated residual, suggesting the presence of permanent labor-supply wedges.

However, modeling hours as a non-stationary process is highly controversial. Obviously, the maximal number of hours that a person can work a day is bounded. Hence, no model that takes into account this physical constraint can yield a unit root in the logarithm of hours. The specification of the utility function ignores any physical bound: it is always possible to find a sequence of B_t such that N_t will eventually exceed any positive limit \bar{N} , even if B_t is a stationary shock. In addition, log-linear solutions ignore this upper bound. This problem is all the more reinforced as we assume that B_t follows a random walk. Thus, our model as well as our solution procedure yield dynamics that should be viewed as a *local approximation*, a view that is widespread in the literature with similar utility functions.

3. Indirect Inference using the DSVAR

In this section, we consider the IRFs of output and hours previously estimated as moments from the data that we ask our DSGE model to reproduce. We follow the methodology advocated by Sims (1989) and Cogley and Nason (1995), which consists in treating models and actual data symmetrically. In our setup, this implies that the DSGE parameters should be picked so that IRFs from the DSVAR on model's simulated data match as well as possible the IRFs from the same DSVAR on US data. As a consequence, it does not matter whether IRFs from actual data are structural or not: they are just moments of the data with possibly little or no economic content.

A legitimate question arises: Why these moments? Either the DSVAR does a good job of identifying the true structural responses, in which case these IRFs are a legitimate object to be matched. Or the DSVAR delivers a distorted image of the true structural responses, in which case these moments can still be useful, provided their DSGE counterparts vary appropriately as one varies the parameters of interest. The potential power of our approach is that even if the DSVAR is completely biased and distorted, the DSGE parameters can still be consistently estimated. In this sense our approach is very much in the spirit of the Indirect Inference approach or the simulated method of moments approach.

⁷ The multivariate tests (ADF with covariates and encompassing tests) in Christiano et al. (2004) seem to be more favorable to a stationary representation of hours.

Let us now describe the way our method is implemented. We partition the model parameters θ into two groups $\theta = \{\theta_1, \theta_2\}$. The first group, $\theta_1 = \{\chi, g_Z, \beta, \alpha, \delta\}$, is calibrated prior to estimation. The parameter χ is set to one, as in Altig et al. (2005). The growth rate of Z_t , g_Z , is equal to 1.005. We set $\beta = 1.03^{-0.25}$, which implies a steady state annualized real interest rate of 3 percent. We set $\alpha = 0.64$, that implies a steady state capital's share in output equal to 36%, as in Cooley and Prescott (1995). Finally, we set $\delta = 0.025$, which implies an annual rate of depreciation of capital equal to 10 percent.

Given these a priori values, we first investigate the response of hours to a positive technology shock. We conduct this exercise for several values of the habit parameter a , having no prior on this parameter. For parameter values consistent with a saddle-path configuration, hours never respond negatively to such a shock. In order to illustrate the mechanism implied by habits in leisure, Fig. 2 reports the response of hours to a 1% technology shock for positive and negative values of a . When $a < 0$, as in Kydland and Prescott (1982), agents are willing to substitute intertemporally leisure after a technology shock, so as to exploit the increase in productivity. This implies a substantial response of hours on impact, though a smaller persistence. To the contrary, when $a > 0$, the labor supply is complementary in adjacent periods, due to habit formation. In this case, agents are less willing to increase their labor supply, conducting to a smaller response of hours on impact and a gradual increase of hours over time.

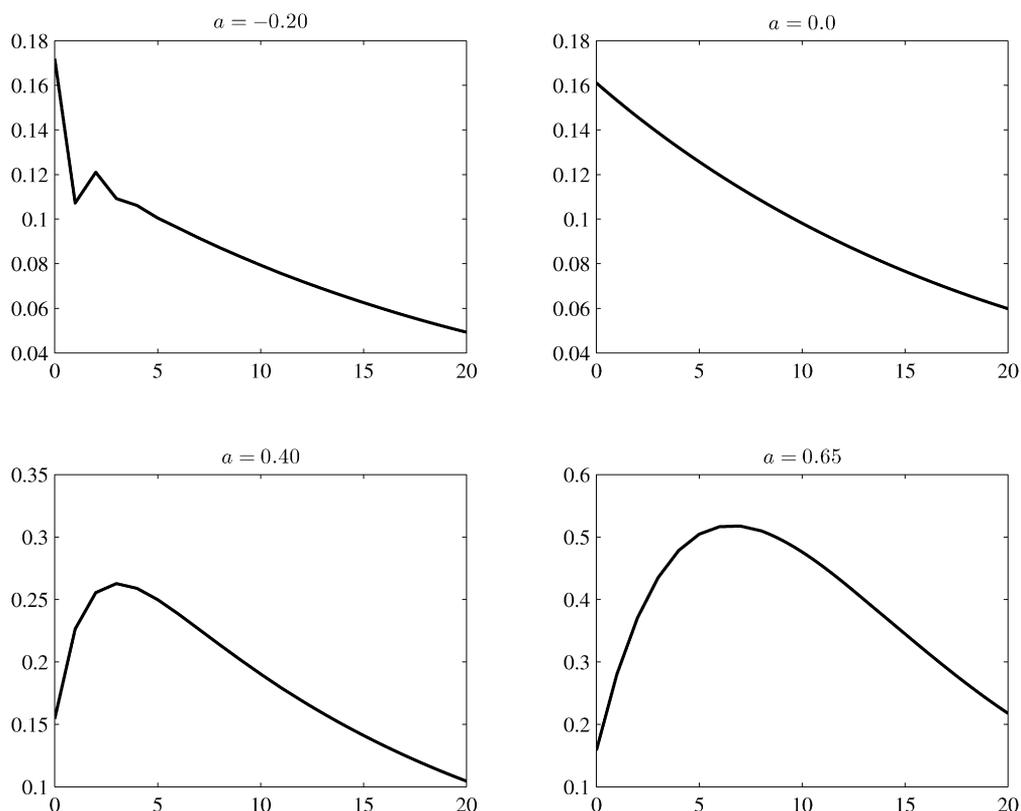


Fig. 2. Model impulse responses of hours to technology shocks.

The second group of model parameters is $\theta_2 = \{a, \sigma_Z, \sigma_B\}$. As suggested by the Sims–Cogley–Nason approach, these three parameters are estimated using an indirect estimation. The empirical IRFs of hours from the DSVAR are used as an auxiliary criterion that allows us to identify and estimate θ_2 through simulations. More precisely, let $\hat{\psi}_T$ denote a vector of DSVAR-based IRFs estimated using equation (2) from actual data on a sample of size T , and $\tilde{\psi}_{T,S}(\theta)$ the average over S simulations of DSVAR-based IRFs estimated from artificial data with the same sample size. More details regarding the implementation of the Indirect Inference approach are provided in Appendix A.

The Indirect Inference estimator of θ_2 minimizes the following quadratic form

$$J(\theta_2) = \mathbf{g}'_{T,S} \mathbf{W}_T \mathbf{g}_{T,S}$$

where $\mathbf{g}_{T,S} = [\hat{\psi}_T - \tilde{\psi}_{T,S}(\theta_2)]$ and \mathbf{W}_T is a symmetric nonnegative matrix that depends on the data.

When the number of IRFs exceeds the number of structural parameters, a global specification test can be conducted in the lines of Hansen (1982). The test statistics $\mathcal{J} = T \frac{S}{1+S} J(\theta_2)$ is asymptotically distributed as a chi-square, with degrees of freedom equal to the number of over-identifying restrictions (see Gouriéroux et al., 1993).

The model is estimated using the IRFs of output and hours to the two shocks. Notice that in this case, for a sufficiently long horizon, we implicitly estimate the DSGE model so as to reproduce the likelihood of the DSVAR on actual data. This estimation method is in fact analog to a full information approach. Thus, if the model is not rejected by the data, this means that the DSGE implies a DSVAR equivalent to that on actual data. That this empirical DSVAR can or cannot be given a structural interpretation is completely irrelevant.

In practice, we estimate θ_2 using the first 21 point estimates of the responses of output and hours to each shock in the DSVAR. $S = 50$ simulations were used for a sample size equal to 219 quarters, as in the data. Simulated innovations are redrawn from the same seeds for each function evaluation. In order to reduce the effect of initial conditions, the simulated samples include 200 initial points which are subsequently discarded in the estimation. The minimization of the simulated criterion function is carried out using the sequential dynamic programming algorithm provided by the MATLAB Optimization Toolbox.

In order to investigate the ability of the chosen moments to discriminate between different versions of our DSGE model, we perform two experiments. In the first one, we shut down habits in leisure ($a = 0$). We will later refer to this case as the *no habit* model. In the second, we estimate all the parameters in θ_2 . We will refer to this setup as the *habit* model.

Estimation results are reported in Table 1. The first column of results reports parameter estimates for the *no habit* model. The second column concerns the *habit* model. In each case, we also report the associated \mathcal{J} -statistics as well as the contribution of technology shocks to the fluctuations of output and hours at business cycle frequencies.

As shown in Table 1, the *no habit* model is rejected by the data (see the P-value of the \mathcal{J} -statistics). In contrast, the *habit* model passes the over-identification test. The labor supply parameter a is significantly positive, indicating that labor supply is subject to intertemporal complementarities, in accordance with previous results from Eichenbaum et al. (1988), Bover (1991), Wen (1998), and Bouakez and Kano (2006). Our estimation suggests that today's leisure significantly reduces leisure services in the subsequent time period. The estimated value of the standard-error of the technology shock, σ_Z , is equal to 1.33%. The point estimate is close to what

Table 1
Results from Indirect Inference

	No Habit Model	Habit Model
a	–	0.6383 (0.0185)
σ_Z	0.0149 (0.0006)	0.0133 (0.0004)
σ_B	0.0101 (0.0002)	0.0076 (0.0003)
\mathcal{J}	223.3 [0.00]	81.9 [45.03]
$V(y/\varepsilon_Z)$ (in %)	88.4	93.6
$V(n/\varepsilon_Z)$ (in %)	20.8	53.9

Note: Standard-errors in parentheses; P -values in brackets. $S = 50$ simulations were used in the estimation procedure. The latter is based on the IRFs of output and hours to both shocks in the DSVAR model. The first 21-point estimates of each IRFs are used as the moments to be matched. $V(y/\varepsilon_Z)$ and $V(n/\varepsilon_Z)$ are the contributions of technology shocks to output and hours fluctuations at business cycle frequencies, respectively. These contributions are computed in the DSGE model at the parameters estimates. The business cycle component of each series are obtained from the HP filter with the smoothing parameter set to 1600.

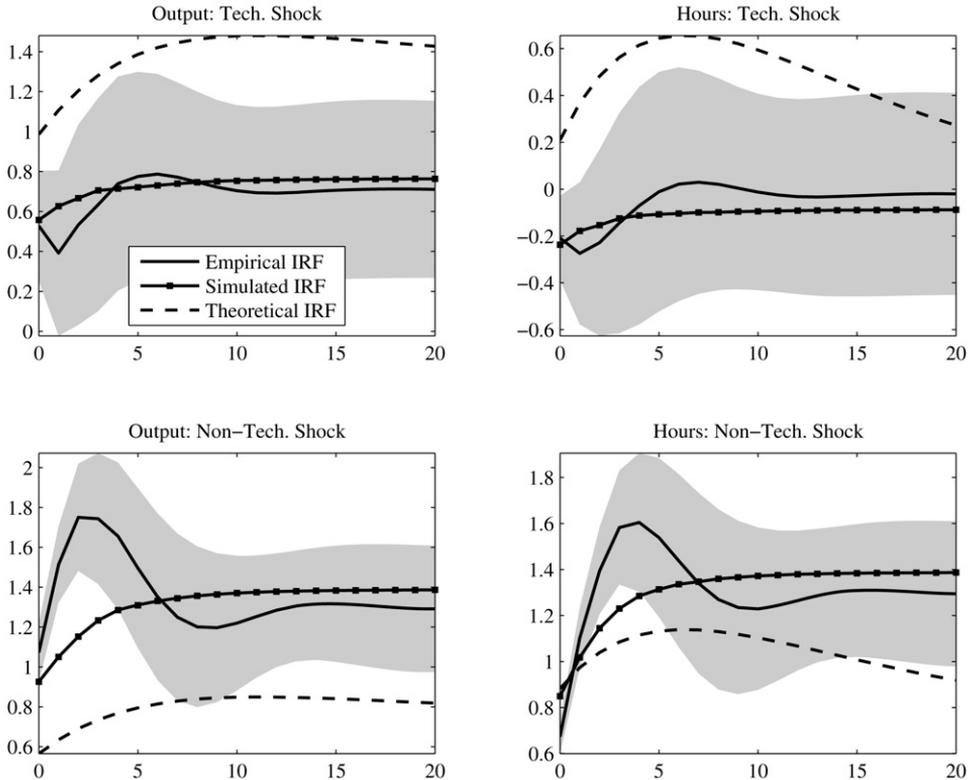
would obtain in simple growth accounting exercises over our data set,⁸ and does not differ from previous studies (Prescott, 1986; Hansen, 1997 and Erceg et al., 2005). The estimated standard-error of the preference shock innovation, σ_B , equals 0.76%. This value is similar to what Chang et al. (2007) obtain (0.6%) when they estimate a DSGE model with non-stationary hours using Bayesian techniques on NFB data. Christiano et al. (2005b) estimate a first order autoregressive representation for this type of variable. They obtain an autoregressive coefficient of 0.986, and a 0.56% standard-error of the innovation. Their results are supportive of our own estimation.

These estimation results make the following point: the IRFs from the DSVAR are useful moments, whether or not they can be given a structural interpretation. Here, these moments unambiguously allow us to discriminate between the two versions of the DSGE model. Next we evaluate the structural content of these IRFs in the *habit* model. From now on, we discard the *no habit* model which has been rejected by the data.

Figure 3 displays the VAR-based IRFs of output and hours to technology and non-technology shocks under actual data (solid line) together with their confidence interval (gray area). It also reports analog IRFs obtained by estimating the same DSVAR on simulated data from the DSGE model (solid line with squares). Finally, it reports the theoretical IRFs implied by the DSGE model (dashed line). This set of IRFs sheds light on the possible biases induced by DSVARs.

Figure 3 makes clear that simulated IRFs match well the empirical IRFs (see the solid lines and solid lines with squares), which is a further illustration of the fact that the model passes the over-identification test. Beyond this, the figure shows the huge bias induced by the DSVAR. Indeed, the DSVAR concludes that hours decline persistently in response to a technology shock even though the data used for estimation are generated from a model that predicts the opposite

⁸ In this type of calculation, we use the calibrated values of δ and a series of aggregate investment (private investment plus durable goods expenditures) to construct a capital stock series assuming an initial capital–output ratio of 6. Using the calibrated value of α , we determine the Solow residual from observed output and hours and the calculated capital stock. We obtain $\sigma_z = 0.0169$.



Note: The plain line corresponds to the empirical IRFs obtained from a DSVAR on actual data. The gray area is the associated 90% asymptotic confidence interval. The line with squares corresponds to the IRFs obtained from a DSVAR on the DSGE model simulated data. The dashed line corresponds to the theoretical IRFs implied by the DSGE model.

Fig. 3. Estimation results with Habit Model.

(see the solid lines with squares and dashed lines). Similar results obtain regarding the response of output to a technology shock. Symmetrically, the DSVAR overestimates the contribution of non-technology shocks to the dynamics of output and hours.

The biases induced by the DSVAR may originate from several sources:

- (i) hours are over-differenced (Erceg et al., 2005);
- (ii) average labor productivity is a poor proxy for total factor productivity at business cycle frequencies (Chang and Hong, 2006);
- (iii) the estimation of DSVARs is subject to small sample biases, especially with long-run restrictions (see Faust and Leeper, 1997);
- (iv) a structural VAR with a finite number of lags may poorly approximate the dynamics of DSGE models (Chari et al., 2005).

In our setup, the first source of bias is irrelevant. In contrast, when we use the *habit* model and conduct counterfactual experiments (not reported), we confirm that the other three sources quantitatively matter, more specifically the latter.

After having considered impulse response functions, we now turn to another set of statistics to describe the properties of the DSVAR, namely the variance decomposition of the forecast

Table 2
Selected US moments

Moments	US Data	Habit Model
σ_y	2.234	1.750
σ_c/σ_y	0.580	0.572
σ_i/σ_y	2.391	2.204
σ_n/σ_y	0.840	0.892
σ_x/σ_y	0.492	0.581
ρ_y	0.828	0.805
ρ_c	0.846	0.748
ρ_i	0.858	0.844
ρ_n	0.893	0.833
ρ_x	0.699	0.675
Corr(x, n)	0.096	-0.122

Note: Moments are obtained using the HP filter with the smoothing parameter equal to 1600. σ_j for $j = \{y, c, i, n, x\}$ is the standard-error of output, consumption, investment, hours and labor productivity, respectively. ρ_j for $j = \{y, c, i, n, x\}$ is the first order autoregressive parameters for the same variables. Corr(x, n) denotes the correlation between labor productivity and hours worked. Consumption is defined as the sum of real consumption expenditures on non-durable goods and services. Investment is the sum of real consumption expenditures on durable goods and real private fixed domestic investment. All variables are per capita.

errors. The last two lines of Table 1 report the contribution of technology shocks to the variance of output and hours in the estimated DSGE model. They show that technology shocks are the dominant impulse at business cycle frequencies (they account for more than 90% of the variance of output and more than half of the variance of hours). These results are in sharp contrast with the predictions drawn from the DSVAR estimated on actual data (see Section 1). In this case, the variance of output explained by technology shocks is 7%, and it is 5% for hours. This large difference confirms that care should be taken when interpreting results from DSVAR models.

Using the Indirect Inference principle, we have shown that it is possible to select a model that matches well a certain set of moments. Table 2 reports key business cycle moments for output, consumption, investment, hours, and labor productivity in US data and under the habit model. This table shows that the model with habits in leisure looks like the one generating US data in many dimensions. Therefore, the habit model provides a good data generating process to conduct counterfactual experiments.

4. An alternative methodology

In the previous section, the Indirect Inference estimation and testing of our *habit* model showed that the DSVAR model delivers biased IRFs. A natural question emerges: What would happen if we were to take the DSVAR evidence at face value? Such an approach would lead a researcher to estimate and test a DSGE model by comparing directly its theoretical IRFs with those from the DSVAR. This empirical strategy has been adopted by many business cycle students (see Rotemberg and Woodford, 1997; Christiano et al., 2005a; Altig et al., 2005). Their approach critically relies on the ability of SVARs to properly identify the aggregate effects of structural shocks. We find that this approach leads to severely biased estimates of parameters. We argue that this approach is dominated by the Sims–Cogley–Nason approach.

To assess this empirical issue, let us consider the following experiment. We use our DSGE model with habit in leisure as our data generating process. The model is fed with the parameter

vectors θ_1 and θ_2 previously calibrated and estimated. As documented in Tables 1 and 2, this model provides an empirically plausible description of US data. We proceed as follows. For one realization of the two structural shocks, we generate artificial data from this model. A DSVAR is then estimated on these simulated data, and the IRFs are deduced according to the procedure outlined in Section 1. These IRFs are taken as “structural moments” and we ask which combination of our DSGE parameters $\theta_2 = (a, \sigma_Z, \sigma_B)'$ can generate theoretical IRFs that match as well as possible the simulated IRFs. In this experiment, the objects to be matched are the responses of output and hours to both identified shocks.⁹ This estimation delivers one particular vector $\hat{\theta}_2^i$. We repeat this exercise one thousand times, thus generating a population of DSGE parameters $\{\theta_2^i\}$, $i = 1, \dots, 1000$, that can be compared to the true value, actually used to generate the data.

The results of this simulation exercise are displayed in the top panel of Table 3. The table reports the true values of the parameters used to generate the data, as well as the mean and median values of parameter estimates. In addition, the table reports the standard deviation of the estimated parameters as well as the empirical probability of finding a parameter lower than the true value.

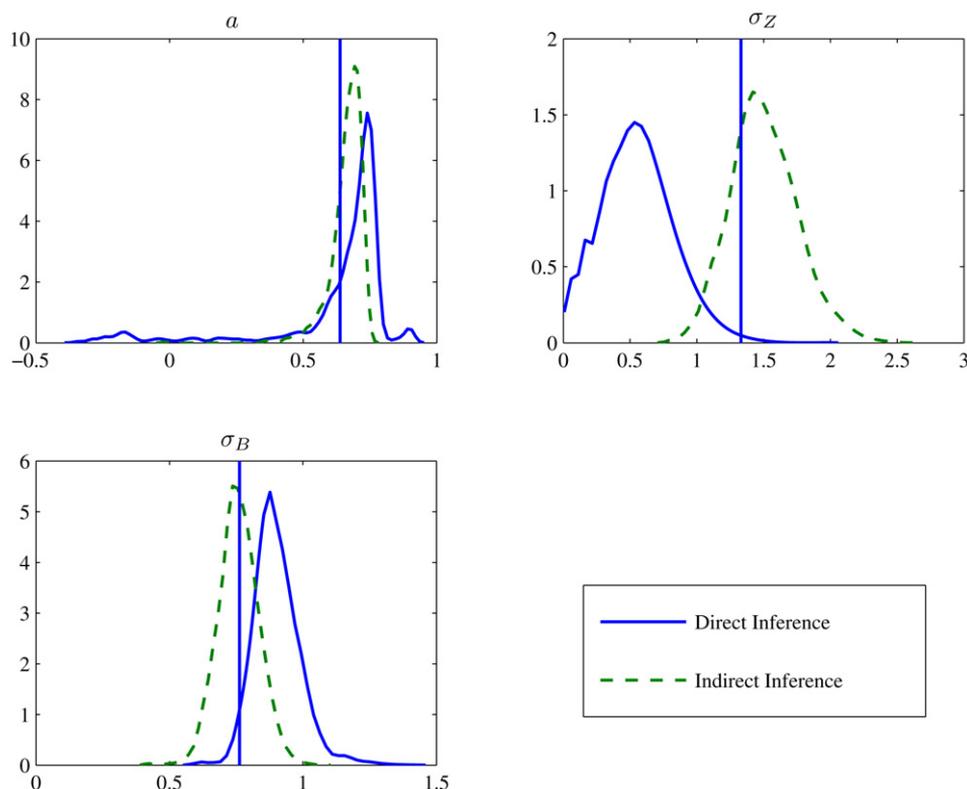
The first line of Table 3 shows that a , the habit parameter, is estimated with a relatively small bias. On the contrary, the Direct Inference method severely underestimates the standard deviation of technology shocks, σ_Z , and overestimates that of preference shocks, σ_B . This is an obvious consequence of the bias in the DSVAR. Regarding the technology shock, the reason why we obtain these results is simple. As shown in Fig. 3, the DSVAR underestimates the true responses of output and hours to a technology shock. This leaves no other choice to the direct estimation approach but to select a small standard error for these shocks. At the same time, the DSVAR

Table 3
Simulation results

Parameters	True value ($\theta_{2,o}$)	Direct Inference			
		Mean	Median	S.E.	$\Pr(\hat{\theta}_2 < \theta_{2,o})$
a	0.6383	0.6320	0.7137	0.2330	0.2530
σ_Z	0.0133	0.0053	0.0053	0.0027	0.9980
σ_B	0.0076	0.0090	0.0089	0.0009	0.0300
Parameters	True value ($\theta_{2,o}$)	Indirect Inference			
		Mean	Median	S.E.	$\Pr(\hat{\theta}_2 < \theta_{2,o})$
a	0.6383	0.6618	0.6744	0.0615	0.2460
σ_Z	0.0133	0.0150	0.0149	0.0025	0.2380
σ_B	0.0076	0.0076	0.0076	0.0007	0.5360

Note: The length of simulated series is 219 after discarding 200 initial observations. These results are obtained from 1000 experiments. For the Direct Inference, the parameters θ_2 are estimated such that the IRFs from the DSGE model are as close as possible to the IRFs from the DSVAR estimated from simulated data. For the Indirect Inference, the parameters θ_2 are estimated such that the average IRFs from the DSVAR under the DSGE model are as close as possible to the IRFs from the DSVAR estimated from simulated data. In this latter case, the Indirect Inference uses 10 simulations with a different seed from that used for generating our data in the first place. In both cases, the DSVAR includes four lags.

⁹ Proponents of this approach usually focus only on a limited number of shocks and specialize their DSGE models accordingly. The attractiveness of this strategy is that it is not necessary to specify the whole stochastic structure of the model. So as to compare the performances of this approach with our Indirect Inference results, we adopt a more demanding view and ask our model to reproduce the responses of key variables to all shocks in the DSVAR.



Note: The length of simulated series is 219 after discarding 200 initial observations. These results are obtained from 1000 experiments. For the Direct Inference, the parameters θ_2 are estimated such that the IRFs from the DSGE model are as close as possible to the IRFs from the DSVAR estimated from simulated data. For the Indirect Inference, the parameters θ_2 are estimated such that the average IRFs from the DSVAR under the DSGE model are as close as possible to the IRFs from the DSVAR estimated from simulated data. In this latter case, the Indirect Inference uses 10 simulations with a different seed from that used for generating our data in the first place. In both cases, the DSVAR includes four lags. Empirical densities are obtained from a Gaussian kernel smoother. The vertical line in each panel corresponds to the true parameter value.

Fig. 4. Empirical density of $\theta_2 = (a, \sigma_Z, \sigma_B)$.

overestimates the responses of output and hours to non-technology shocks. As a consequence, the direct estimation will tend to overestimate the true standard error of the labor supply shock.

Figure 4 displays the empirical distribution of the parameter estimates, obtained using standard kernel methods.¹⁰ For comparison purpose, the vertical lines report the parameters true values. Notice that an important mass at zero is found for σ_Z . This happens because a large number of draws are such that the responses of hours in the DSVAR are very negative, and those of output weakly positive. Since the DSVAR-based responses are considered as structural, the only way to match these moments is to set the standard deviation of technology shocks to zero (hence a zero response of hours).

¹⁰ A Gaussian kernel smoother was used to compute the parameters distribution, enforcing a positive support in the cases of σ_B and σ_Z and treating all data points symmetrically.

Very different results emerge from the Indirect Inference approach, the population of DSGE parameters being computed as above. Instead of trying to find a parameter vector that allows the model's IRFs to match directly their DSVAR counterparts, we resort to the Indirect Inference approach, as expounded in the previous section. The simulation seed used in the estimation stage is of course different from that used for generating our data in the first place. The results are reported in the bottom panel of Table 3. As has already been demonstrated in simpler dynamic models (see Gouriéroux et al., 1993), the Indirect Inference approach performs relatively well. This is especially true for the estimated standard errors of the structural shocks, which are in close conformity with the true values. This is further illustrated in Fig. 4.

The Indirect Inference results include the small sample bias an applied researcher would be confronted to, as well as simulation biases. Hence Indirect Inference may serve as a benchmark estimation method. According to our experiments, the Direct Inference method underperforms this benchmark.

5. Conclusion

The identification of the response of hours worked after a technology shock using SVARs has renewed the debate on the relative contributions of various shocks to the business cycle. More precisely, the DSVAR approach documents a striking evidence against the standard RBC model: after a positive technology shock, hours worked decrease. For researchers that use the DSVAR approach to discriminate between competing theories of the business cycle, this evidence suggests to abandon the RBC model in favor of models with important (real) frictions and (nominal) rigidities, and to scale down the role of technology shocks. However, Chari et al. (2005) and Erceg et al. (2005) forcefully demonstrate that DSVARs are severely biased and, if taken at face value, can lead to incorrect inference. Accordingly, their results cast serious doubts on DSVARs as useful guides for developing business cycle models.

In this paper, we argue that, though distorted and sometimes fallacious, DSVARs can still prove useful for business cycle students, provided that two conditions are met:

- (i) the model and actual data are treated symmetrically;
- (ii) the IRFs on data simulated from the DSGE model vary appropriately with the parameters of interest.

The paper also illustrates that a failure to treat model and data symmetrically can lead to misleading conclusions. In particular, in our experiment, we find that the contribution of technology shocks is far too small under an estimation approach seeking to select the DSGE parameters by directly matching theoretical IRFs with IRFs drawn from a DSVAR estimated on actual data.

Acknowledgments

We would like to thank F. Canova, L.J. Christiano, F. Collard, M. Eichenbaum, A. Guay, L.P. Hansen, H. Le Bihan, E. McGrattan, F. Mihoubi, F. Portier, F. Tripier, and H. Uhlig as well as two anonymous referees for helpful remarks and comments. This paper has also benefited from discussions during presentations at IGIER (Milan), Banque de France (Paris), GREQAM (Marseille), Schumpeter (Berlin), Paris 1 Panthéon-Sorbonne/Paris-Jourdan, as well as Banque de France Workshop "Monetary Models and Aggregate Dynamics", 2005 SED Annual Meeting

(Budapest), 2005 EEA Congress (Amsterdam) and 2006 North American Summer Meeting of the Econometric Society (Minneapolis). The traditional disclaimer applies.

Appendix A. Implementation of Indirect Inference

The estimation method is implemented as follows.

Step 1: We estimate a q -dimensional vector of IRFs, denoted $\hat{\psi}_T$, from actual data, where q denotes the number of selected impulse responses (horizon \times number of selected IRFs).

Step 2: From the model’s solution, and given the vector of structural parameters, $\theta = (\theta_1, \theta_2)$, and initial conditions on capital, labor and the shocks, S simulated paths for productivity and employment, denoted $\tilde{x}_T^i(\theta)$, $\tilde{n}_T^i(\theta)$, $i = 1, \dots, S$, are recursively computed.

Step 3: From these simulations, we estimate a VAR model from the simulated data $\tilde{z}_t^i = (\Delta \tilde{x}_t^i, \Delta \tilde{n}_t^i)'$,

$$\tilde{z}_t^i = \tilde{\mathbf{A}}_1^i \tilde{z}_{t-1}^i + \dots + \tilde{\mathbf{A}}_p^i \tilde{z}_{t-p}^i + \tilde{\epsilon}_t^i, \quad E\{\tilde{\epsilon}_t^i \tilde{\epsilon}_t^{i'}\} = \tilde{\Sigma}^i, \quad i = 1, \dots, S,$$

with the same number of lags ($p = 4$) as in the DSVAR on actual data from step 1. We then compute the associated vector of IRFs, denoted $\tilde{\psi}_T^i(\theta)$ ($i = 1, \dots, S$) using the exact same long-run restrictions as in step 1

$$\tilde{z}_t^i = \tilde{\mathbf{B}}^i(L) \tilde{\mathbf{B}}^i(1)^{-1} \tilde{\mathbf{C}}^i(1) \tilde{\eta}_t^i,$$

and we construct their average over simulations

$$\tilde{\psi}_{T,S}(\theta) = \frac{1}{S} \sum_{i=1}^S \tilde{\psi}_T^i(\theta).$$

Step 4: An Indirect Inference estimate $\tilde{\theta}_{2,T,S}$ for θ_2 minimizes the quadratic form

$$J(\theta_2) = \mathbf{g}'_{T,S} \mathbf{W}_T \mathbf{g}_{T,S},$$

where $\mathbf{g}_{T,S} = (\hat{\psi}_T - \tilde{\psi}_{T,S}(\theta_2))$ and \mathbf{W}_T is a symmetric nonnegative matrix defining the metric.

Steps 2 to 4 are conducted repeatedly until convergence—i.e. until a value of θ_2 that minimizes the objective function is obtained. Let ψ_0 denote the pseudo-true value of ψ and $\theta_{2,0}$ the pseudo-true value of θ_2 . Under standard regularity conditions (see Gouriéroux et al., 1993), for S held fixed and as T goes to infinity, $\sqrt{T}(\tilde{\theta}_{2,T,S} - \theta_{2,0})$ is asymptotically normally distributed, with a covariance matrix equal to $(1 + S^{-1})(\mathbf{D}'_{\theta} \mathbf{W}_T \mathbf{D}_{\theta})^{-1}$ where $\mathbf{D}_{\theta} = \partial \mathbf{g}_{T,S} / \partial \theta_2$.

A preliminary consistent estimates of the weighting matrix \mathbf{W}_T is required for the computation of $\tilde{\theta}_{2,T,S}$. It may be directly based on actual data, and corresponds to the inverse of the covariance matrix of $\sqrt{T}(\hat{\psi}_T - \psi_0)$, which is obtained from step 1. Here, \mathbf{W}_T^{-1} is a diagonal matrix with the sample variances of $\sqrt{T}(\hat{\psi}_T - \psi_0)$ along the diagonal. This choice of weighting matrix ensures that θ_2 is effectively chosen so that $\tilde{\psi}_{T,S}(\theta_2)$ lies as much as possible inside the confidence intervals of $\hat{\psi}_T$.

For the sake of identification, we impose that the number of IRFs exceeds the number of structural parameters. This enables us to conduct a global specification test in the lines of Hansen (1982), denoted $\mathcal{J} = TSJ(\theta_2)/(1 + S)$, which is asymptotically distributed as a chi-square, with degrees of freedom equal to the number of over-identifying restrictions ($q - \dim \theta_2$).

References

- Altig, D., Christiano, L., Eichenbaum, M., Lindé, J., 2005. Firm-specific capital, nominal rigidities and the business cycle. Working paper No. 11034. NBER.
- Anderson, G., Moore, G., 1985. A linear algebraic procedure for solving linear perfect foresight models. *Economics Letters* 17, 247–252.
- Bouakez, H., Kano, T., 2006. Learning-by-doing or habit formation? *Review of Economic Dynamics* 9 (3), 508–524.
- Bover, O., 1991. Relaxing intertemporal separability: A rational habit model of labor supply estimated from panel data. *Journal of Labor Economics* 9, 85–100.
- Chang, Y., Hong, J.H., 2006. Do technological improvements in the manufacturing sector raise or lower employment? *American Economic Review* 96 (1), 352–368.
- Chang, Y., Doh, T., Schorfheide, F., 2007. Non-stationary hours in a DSGE model. *Journal of Money, Credit, and Banking*. In press.
- Chari, V., Kehoe, P., McGrattan, E., 2005. A critique of structural VARs using real business cycle theory. Working paper 631. Federal Reserve Bank of Minneapolis, Research Department.
- Christiano, L., Eichenbaum, M., Vigfusson, R., 2004. What happens after a technology shock. Working paper No. 9819. NBER.
- Christiano, L., Eichenbaum, M., Evans, C., 2005a. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* 113 (1), 1–45.
- Christiano, L., Eichenbaum, M., Vigfusson, R., 2005b. Assessing structural VARs. Mimeo. Northwestern University.
- Cogley, T., Nason, J.M., 1995. Output dynamics in real-business-cycle models. *American Economic Review* 85 (3), 492–511.
- Cooley, T., Dwyer, M., 1998. Business cycle analysis without much theory: A look at structural VARs. *Journal of Econometrics* 83 (1–2), 57–88.
- Cooley, T., Prescott, E., 1995. Economic growth and business cycles. In: Cooley, T. (Ed.), *Frontiers of Business Cycle Research*. Princeton Univ. Press, Princeton.
- Eichenbaum, M., Hansen, L., Singleton, K., 1988. A time series analysis of representative agent models of consumption and leisure choice under uncertainty. *Quarterly Journal of Economics* 103 (1), 51–78.
- Erceg, C., Guerrieri, L., Gust, C., 2005. Can long-run restrictions identify technology shocks. *Journal of the European Economic Association* 3 (4), 1237–1278.
- Faust, J., Leeper, E., 1997. When do long-run identifying restrictions give reliable results? *Journal of Business and Economic Statistics* 15 (3), 345–353.
- Francis, N., Ramey, V., 2005. Is the technology-driven real business cycle hypothesis dead? Shocks and aggregate fluctuations revisited. *Journal of Monetary Economics* 52 (8), 1379–1399. November.
- Galí, J., 1999. Technology, employment and the business cycle: Do technology shocks explain aggregate fluctuations? *American Economic Review* 89 (1), 249–271.
- Galí, J., 2004. On the role of technology shocks as a source of business cycles: Some new evidence. *Journal of the European Economic Association* 2 (2–3), 372–380.
- Galí, J., 2005. Trends in hours, balanced growth and the role of technology in the business cycle. *Federal Reserve Bank of St. Louis Review* 87 (4), 459–486.
- Galí, J., Rabanal, P., 2004. Technology shocks and aggregate fluctuations: How well does the RBC Model fit postwar US data? NBER Macroeconomics Annual. The MIT Press.
- Gouriéroux, C., Monfort, A., Renault, E., 1993. Indirect Inference. *Journal of Applied Econometrics (Supplement)* 8, S85–S118.
- Hall, R., 1997. Macroeconomic fluctuations and the allocation of time. *Journal of Labor Economics* 15 (1), 223–250.
- Hamilton, J., 1994. *Time Series Analysis*. Princeton Univ. Press.
- Hansen, L., 1982. Large sample properties of generalized method of moments. *Econometrica* 50, 1029–1054.
- Hansen, G., 1997. Technical progress and aggregate fluctuations. *Journal of Economic Dynamics and Control* 21, 1005–1023.
- Kehoe, P., 2006. How to advance theory with structural VARs: Use the Sims–Cogley–Nason approach. Working paper No. 12575. NBER.
- Kydland, F., Prescott, E., 1982. Time-to-build and aggregate fluctuations. *Econometrica* 50 (6), 1345–1370.
- Prescott, E., 1986. Theory ahead of business cycle measurement. *Federal Reserve Bank of Minneapolis. Quarterly Review* 10, 9–22.

- Rotemberg, J., Woodford, M., 1997. An optimization-based econometric framework for the evaluation of monetary policy. In: Bernanke, S., Rotemberg, J. (Eds.), *National Bureau of Economic Research Macroeconomics Annual 1997*. MIT Press, Cambridge, MA.
- Sims, C., 1989. Models and their use. *American Journal of Agricultural Economics* 67, 489–494.
- Wen, Y., 1998. Can a real business cycle model pass the Watson test? *Journal of Monetary Economics* 42, 185–203.