Sensitivity of Impulse Responses to Small Low Frequency Co-movements: Reconciling the Evidence on the Effects of Technology Shocks

(Not-for-publication Appendix)

Nikolay Gospodinov	Alex Maynard	Elena Pesavento
Concordia University and CIREQ	University of Guelph	Emory University

October 2010

1 Additional Monte Carlo Simulation Results

This online appendix provides some additional simulation results that are not reported in the paper. In particular, we present results with (i) data simulated from models with structural breaks and (ii) data simulated from a real business cycle (RBC) model.

1.1 Structural Breaks

The simulation results reported in the paper are obtained under the maintained hypothesis of an underlying low frequency co-movement in the data for $\delta \neq 0$, which is implicitly assumed to be structural in nature. We find that this leads to a bias in the IRFs of the differenced VAR. Fernald (2007) provides evidence that the co-movement is due to a similar pair of breaks occurring in both series, giving rise to a common high-low-high pattern. Treating the similarity of the breaks in the two series as coincidental or driven by exogenous factors that should not be used in identifying technology shocks, he shows that this type of low frequency co-movement can result in misleading IRFs from the levels VAR. This contrasts with the conclusions of the previous section.

The simulations discussed below are intended to provide some insight into the reasons underlying the differences between our results and the results in Fernald (2007). They illustrate that the critical difference does not hinge on whether the low frequency co-movement is due to common structural breaks. Rather, it is instead Fernald's (2007) assumption that the observed low frequency correlation is coincidental that leads to his conclusion. This differs from the models in our paper, in which this correlation is treated as a true underlying feature of the data process that is subject to the LR identification restriction.

Below we consider two observationally equivalent structural break models that give rise to the common high-low-high pattern observed by Fernald (2007). We first present the results from a co-break model in which the similar magnitude and timing of the breaks is driven by a common underlying component. We then consider a model in which the commonality of the breaks is coincidental (Fernald, 2007) or exogenously imposed (Francis and Ramey, 2009). Both models are equally consistent with the observed data.

The co-break model has the following form

$$\begin{pmatrix} \triangle l_t \\ h_t \end{pmatrix} = \begin{bmatrix} 0 \\ (1-\rho)\alpha(D_{1t}+D_{2t}) \end{bmatrix} + \begin{bmatrix} 0 & -\gamma(1-\rho) \\ 0 & \rho \end{bmatrix} \begin{pmatrix} l_{t-1} \\ h_{t-1} \end{pmatrix} + \Psi(L)^{-1} \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}, \qquad (1)$$

where $D_{1t} = 1$ if $t \in [1, T/3]$ and 0 otherwise, $D_{2t} = 1$ if $t \in [2T/3, T]$ and 0 otherwise, $(u_{1,t}, u_{2,t})' \sim iidN(0, \Sigma)$ and Σ and $\Psi(L)$ are the same as above. In the reported simulations, $\rho = 0.95$, $\gamma = -0.8$, $\alpha = 5$ and $T = 250$. The pattern of the mean break in hours worked is calibrated to match the U-shape of the actual series. Note that this model generates structural breaks in the mean of both series since the mean break in hours worked is transmitted to labor productivity growth through the low frequency correlation parameter $\delta = -\gamma(1-\rho)$.

As in the main simulation experiment, we assess the finite-sample behavior of IRF estimates obtained from the level specification, differenced specification, and levels specification with HP detrended productivity growth. The Monte Carlo results are presented in Figure 1. As before, removing the low frequency component (either by differencing of hours or HP-detrending of productivity growth) leads to substantial deviations of the IRF estimates from their true values. This is not surprising because these prior transformations of the data remove potentially important (now co-breaking) information regarding the long-run properties of the data.

These results illustrate two important points. First, they indicate that the central insights developed analytically in the model without breaks carry over to the case with breaks. Secondly, they demonstrate that even if one accepts the evidence favoring a common high-low-high break sequence in both series, this does not lead to a forgone conclusion in favor of Gali's (1999) original results. We can, however, obtain results supporting Fernald's (2007) conclusions, if we impose his additional assumption that the similarity of the breaks is coincidental in nature. Thus, we now turn to an alternative break model that also generates structural breaks in both series but shuts off completely the feedback of h_{t-1} to $\triangle l_t$ through the low-frequency correlation coefficient γ (by setting $\gamma = 0$). In order to stay as close as possible to the observed features of the U.S. data, we generate structural breaks in hours and labor productivity growth that happen at the same time. Since these breaks are not generated by some underlying mechanism inside the model, we refer to them as coincidental and possibly spurious. More specifically, the model has the following structure

$$\begin{pmatrix} \triangle l_t \\ h_t \end{pmatrix} = \begin{bmatrix} (1-\rho)\alpha D_{1t} + (1-\rho)\alpha D_{2t} \\ (1-\rho)\alpha D_{1t} + (1-\rho)\alpha D_{2t} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \rho \end{bmatrix} \begin{pmatrix} l_{t-1} \\ h_{t-1} \end{pmatrix} + \Psi(L)^{-1} \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}, \qquad (2)$$

where the values of the parameters and break dummies are the same as specified above.

The results for the IRFs with data generated from model (2) are plotted in Figure 2. As expected, removing the low frequency co-movement due to coincidental or exogenous structural breaks that are not an inherent feature of the structural model produces IRF estimates that are very close to the true IRFs. Similarly, the levels specification, which mistakenly treats the coincidental breaks as a true low frequency component to which the LR identification restriction should apply, produces biased IRF estimates. Nonetheless, the true IRF is still within the 95% confidence bands.

Estimating and removing the breaks will not solve the dichotomy. If the breaks are coincidental as in model (2), then removing the breaks prior to identification will remove any low frequency co-movement and the IRFs from both the VAR in levels and first differences will be approximately unbiased, regardless of the size of the largest root. If instead, the breaks are a true underlying feature of the data process as in (1), then removing the breaks will remove potentially important information regarding the long-run properties of the data and the resulting IRFs can be highly inaccurate. In results available upon request, we confirm this intuition, even when allowing a priori knowledge of the break points. The results are similar to those shown in the third panel of Figures 1 and 2 employing the HP filter.

As these simulations make clear, the mere existence of a common high-low-high specification is not sufficient to conclude in favor of either specification. In order to provide results consistent with Fernald (2007), we must impose his assumption that the timing of the breaks is coincidental. While from a historical perspective, Fernald (2007) provides some convincing arguments for why this may indeed be the case ex-post, we note that the ex-ante probability of such similar, yet completely unrelated, break sequences is rather low. In the co-breaking model (1), to which our earlier insights carried over, the similarity of the two breaks is to be expected, since there is a natural mechanism relating the two breaks. Alternatively, the bivariate system might be regarded as misspecified and an improved inference procedure could result from estimating a larger model as in Erceg, Guerrieri and Gust (2005).

1.2 RBC model

It is interesting to see if our main conclusions in the paper continue to hold if the data are simulated from a dynamic general equilibrium model, in which the persistence and the low frequency co-movements between the variables are implicitly determined. To investigate this, we follow Chari, Kehoe and McGrattan (2008) and Christiano, Eichenbaum and Vigfusson (2006) by generating data from a real business cycle model. The true impulse response functions implied from this structural model are then compared to the estimated impulse responses from a finite-order VAR model. In particular, we use the two-shock CKM specification described in Christiano, Eichenbaum and Vigfusson (2006) as a data generating mechanism (see Christiano, Eichenbaum and Vigfusson, 2006, for details).

Several features of the RBC should be emphasized. First, the RBC model used for simulating the data imposes a unit root on technology while hours worked implied by the model are stationary but highly persistent. As a result, any low frequency co-movements in the model should arise from the persistence of the variables and not from structural breaks. Second, the RBC model implies a VARMA (infinite-order VAR) structure for $(\triangle l_t, h_t)'$ and fitting a finite-order VAR model to $(\triangle l_t, h_t)'$ results in biased estimates of the impulse response functions (Chari, Kehoe and McGrattan, 2008; Christiano, Eichenbaum and Vigfusson, 2006; Ravenna, 2007). Although there exist methods for correcting this misspecification bias (for instance, Christiano, Eichenbaum and Vigfusson, 2006), we do not pursue this avenue, since our primary focus in this paper is on the bias that arises in the differenced specification from omitting a possible low frequency co-movement, regardless of whether or not there is an additional source of bias due to lag truncation.

We generate 1,000 samples of 180 observations each and consider both the levels and differenced VAR specifications with four lags. Figure 3 displays time plots of a typical pair of synthetic sequences of demeaned hours and detrended labour productivity generated from the simulated RBC model.[1] The co-movement of the series is similar to that in the actual data which is plotted, for visual comparison, in Figure 4. Likewise, Figure 5 displays HP trends of the simulated labor productivity growth and hours worked from the RBC model. The figure again shows a similar low frequency co-movement to that of the real data (Figure 2 in the paper). Thus, the calibrated RBC model appears to produce a low frequency co-movement, similar to the one found in the empirical data. In conjunction with the lag truncation bias, this may help to explain why it produces the large discrepancies in the IRFs of the differenced and level specifications discussed below.

We now turn to the simulated IRFs. Again we consider structural VARs with hours in both differences and levels. Since hours worked is a highly persistent variable, it is tempting to subject this variable to a unit root pre-test and depending on the outcome to model hours either in levels or first differences. Thus, we also report the results from this pre-testing procedure in which the decision of modeling h in levels or first differences is based on an ADF test with 4 lags at 5% significance level.

The results from the three specifications are presented in Figure 6. As reported elsewhere (Christiano, Eichenbaum and Vigfusson, 2006, for example), the IRF estimates from the levels VAR suffer from an upward bias that is caused by approximating the true VARMA process by a short-order VAR. Using an estimate of the long-run variance matrix as suggested by Christiano, Eichenbaum and Vigfusson (2006) can substantially reduce this bias, although the sampling uncertainty associated with the IRF estimates remains large.[2] As in the previous simulation design, the differenced specification reduces the sampling uncertainty but completely misses the true impulse response due to the omission of important low frequency information. The true impulse response

[1] To avoid cherry-picking, we used the last of the 1,000 synthetic series from our simulation. Comparison to other draws indicated that it was not atypical.

[2] Christiano, Eichenbaum and Vigfusson (2006) also find that this bias is substantially smaller after relaxing the assumptions on the variance of the measurement error in the CKM specification. In addition, they consider a specification with wage and price frictions for which the sampling uncertainty is much reduced.

falls entirely outside the 90% Monte Carlo confidence bands obtained from the differenced specification. Due to the relatively high persistence of hours worked, the pre-testing procedure has difficulties rejecting the unit root null and leads to only small improvements over the differenced specification. The estimates are slightly less biased and the confidence bands are wider reflecting the uncertainty regarding the presence of a unit root in hours worked.

In summary, the simulated data from the RBC model show low frequency co-movements similar to those found in the empirical data and produce IRFs in which the levels and differenced specifications give widely divergent conclusions. Therefore, although the lag-truncation bias may also play an important role when the data is generated from a dynamic general equilibrium model, we nonetheless re-confirm the central role of the low frequency co-movement in explaining the discrepancy between the IRFs from level and difference specifications.[3]

References

[1] Chari, V., Kehoe, P., and McGrattan, E. (2008), "Are Structural VARs with Long-Run Restrictions Useful in Developing Business Cycle Theory?" *Journal of Monetary Economics*, 55, 1337-1352.

[2] Christiano, L., Eichenbaum, M., and Vigfusson, R. (2006), "Assessing Structural VARs," in *NBER Macroeconomics Annual*, eds. D. Acemoglu, K. Rogoff and M. Woodford, Cambridge: MIT Press.

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

[4] Fernald, J. (2007), "Trend Breaks, Long-Run Restrictions, and Contractionary Technology Improvements," *Journal of Monetary Economics*, 54, 2467-2485.

[5] Gali, J. (1999), "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" *American Economic Review*, 89, 249-271.

[6] Ravenna, F. (2007), "Vector Autoregressions and Reduced Form Representations of DSGE Models," *Journal of Monetary Economics*, 54, 2048-2064.

[3] Further supporting this conclusion, some results with other specifications of the RBC model (available from the authors upon request) do not seem to produce a negative low frequency co-movement between productivity and hours and also result in smaller discrepancies between the estimated impulse response functions from the levels and differenced specifications.

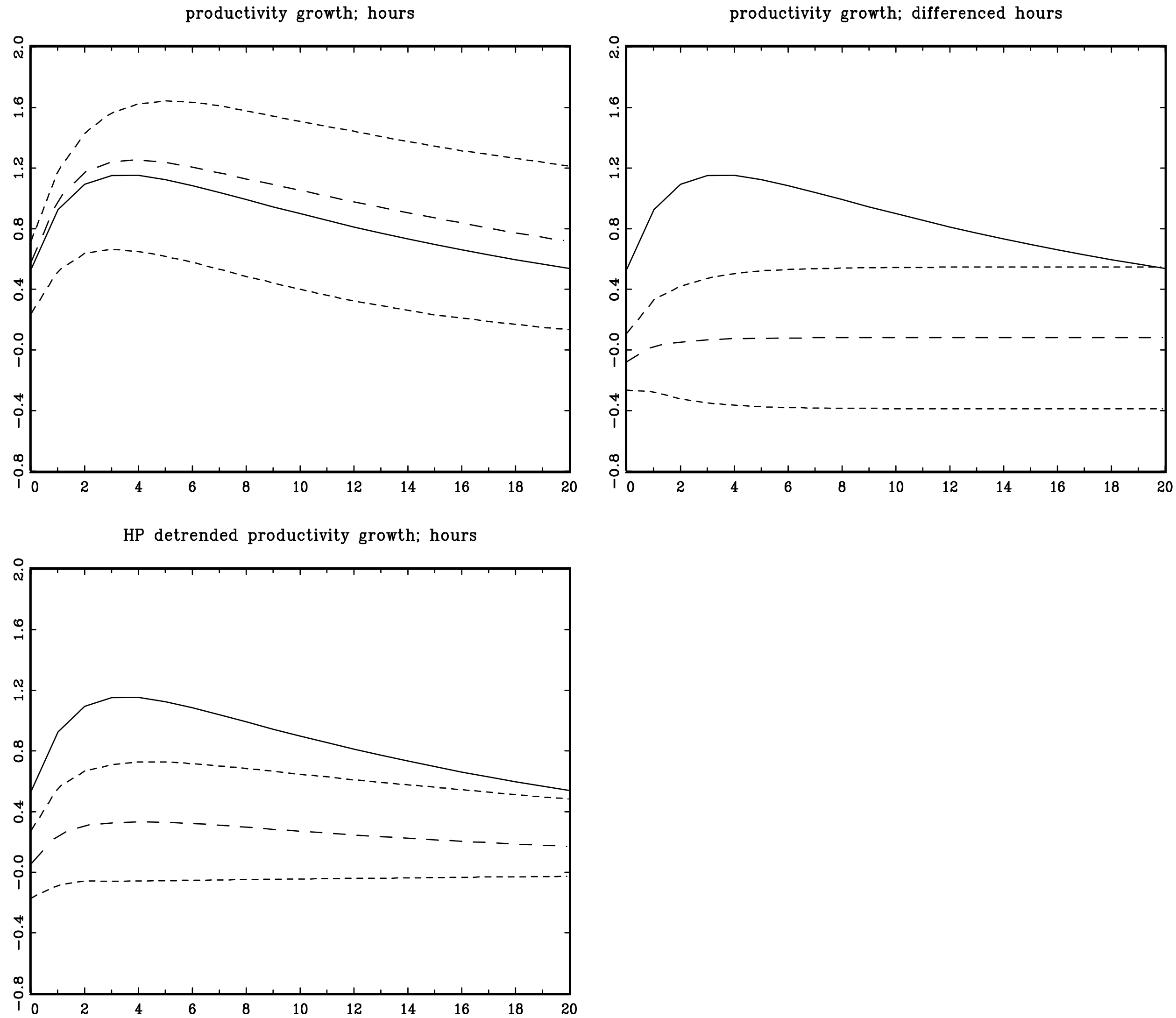

Figure 1. Response of hours to a positive technology shock (long-run identification) with data simulated from the structural break model (1), where $\rho = 0.95$, $\gamma = -0.8$, $\alpha = 5$ and $T = 250$. Solid line: true IRF; long dashes: median Monte Carlo IRF estimate; short dashes: 95% Monte Carlo confidence bands.

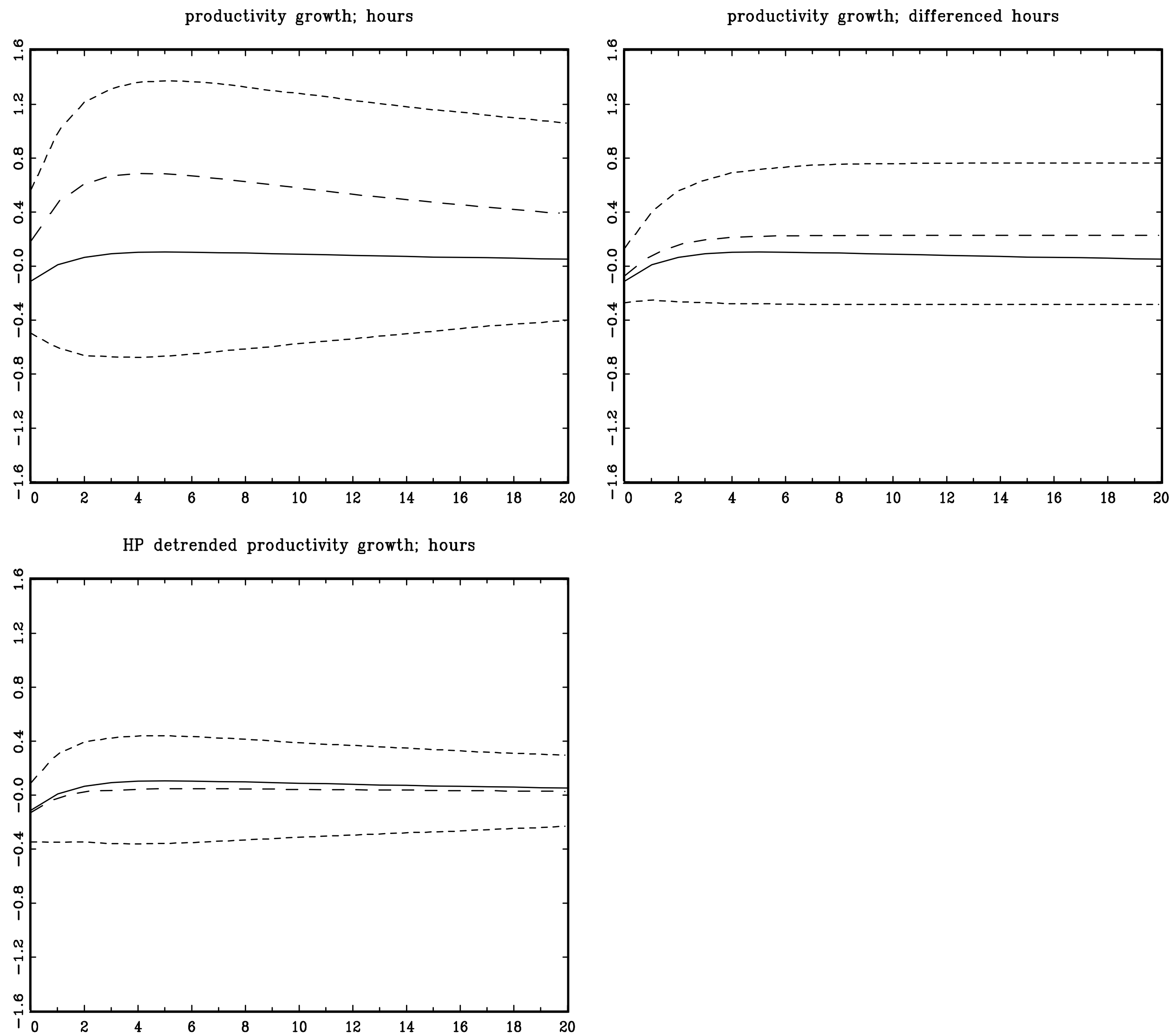

Figure 2. Response of hours to a positive technology shock (long-run identification) with data simulated from the structural break model (2), where $\rho = 0.95$, $\alpha = 5$ and $T = 250$. Solid line: true IRF; long dashes: median Monte Carlo IRF estimate; short dashes: 95% Monte Carlo confidence bands.

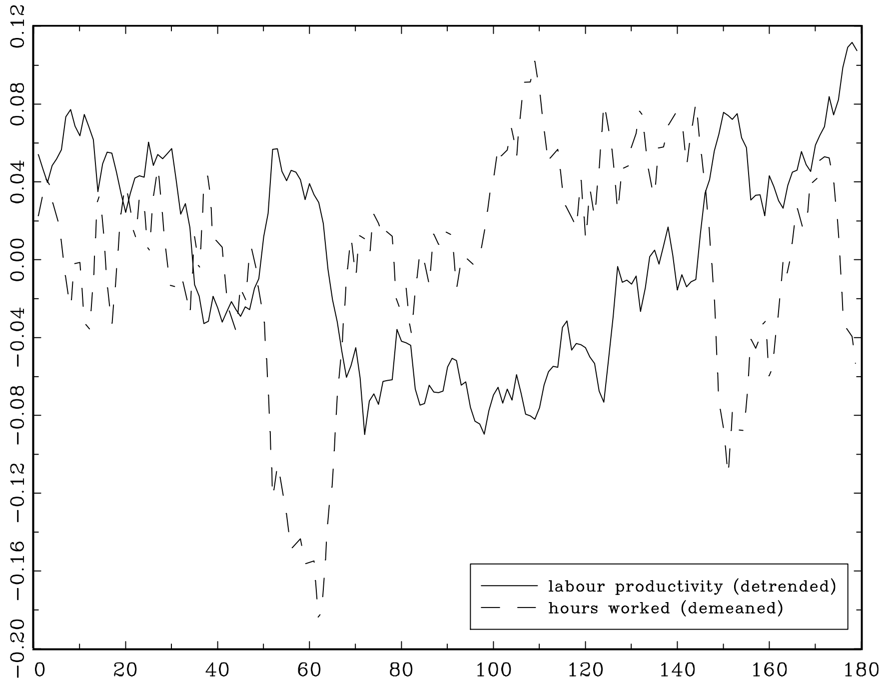

Figure 3. Detrended labour productivity and demeaned hours worked; simulated data from the CKM specification of the RBC model (Christiano, Eichenbaum and Vigfusson, 2006).

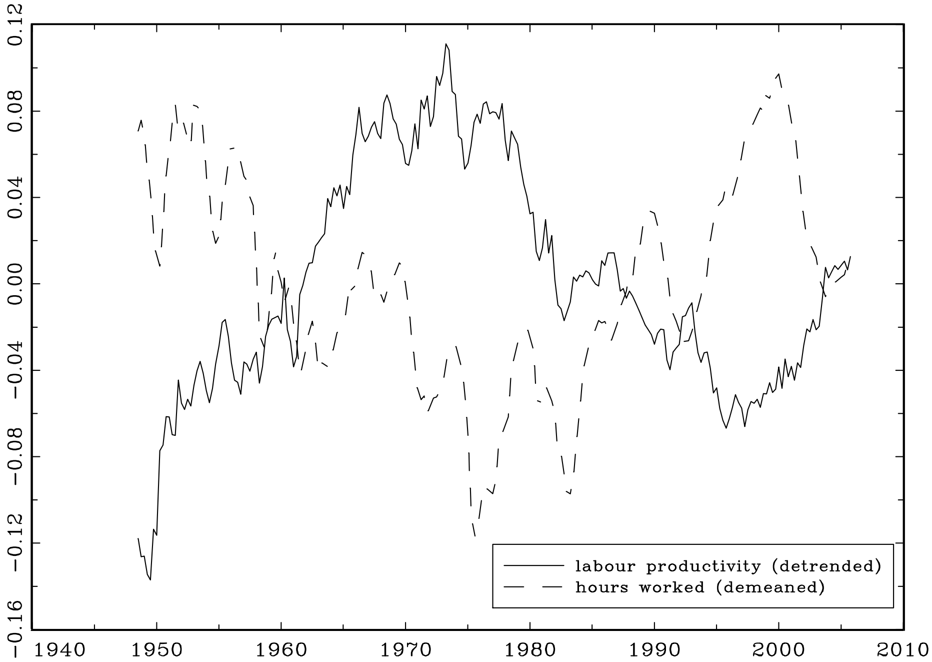

Figure 4. Detrended labour productivity and demeaned hours worked, U.S. data 1948Q2 - 2005Q3.

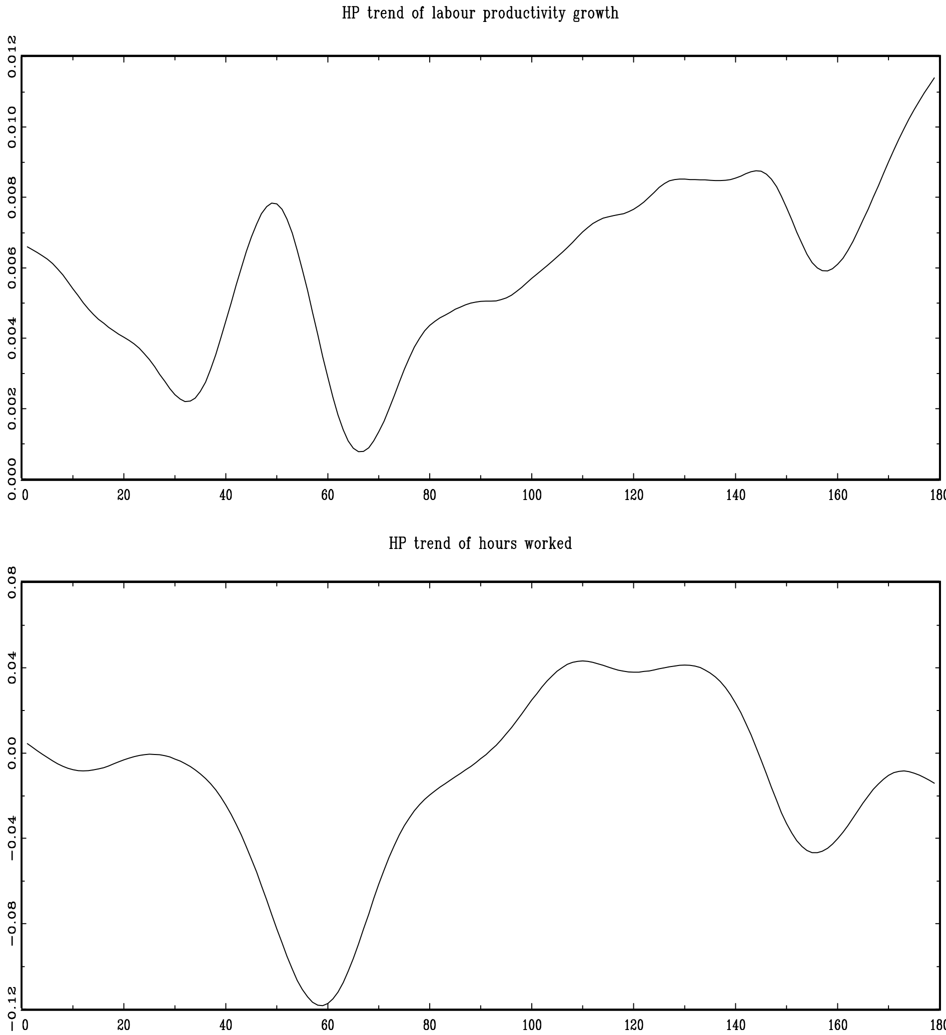

Figure 5. HP trends of labour productivity growth (top graph) and hours worked (bottom graph), simulated data from the CKM specification of the RBC model (Christiano, Eichenbaum and Vigfusson, 2006).

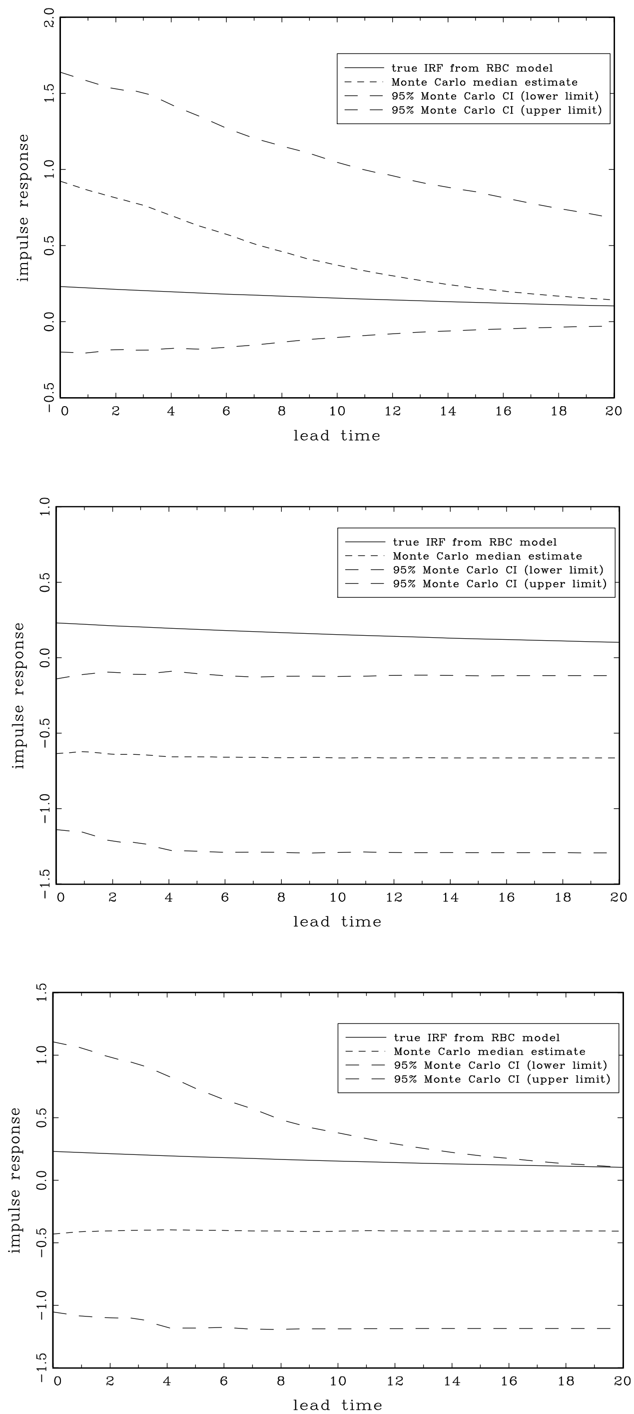

Figure 6. Monte Carlo IRF estimates and 95% confidence bands from the levels (top graph), differenced (middle graph) and pre-test (bottom graph) VAR specifications on simulated data (1,000 samples of length 180) from the CKM specification of the RBC model (Christiano, Eichenbaum and Vigfusson, 2006).