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

Nikolay Gospodinov*
Concordia University and CIREQ

Alex Maynard
University of Guelph

Elena Pesavento
Emory University

February 2010

Abstract

This paper clarifies the empirical source of the debate on the effect of technology shocks on hours worked. Providing theoretical support for previous contentions made in the literature, we find that the contrasting conclusions from levels and differenced VAR specifications can be explained by a small low frequency co-movement between hours worked and labour productivity growth, which is allowed for in the levels specification but is implicitly set to zero in the differenced VAR. This is due to a discontinuity in the solution for the structural coefficients identified by long-run restrictions that exists only when this correlation is present. Consequently, even when the root of hours is very close to one and the low frequency co-movement is quite small, removing it can give rise to biases large enough to account for the empirical difference between the two specifications. While this low frequency correlation has recently been interpreted as evidence against the levels specification, we find that if it is a true property of the correctly identified model, then it can actually lead to substantial biases in the differenced rather than the levels specification. Similar biases result from HP pre-filtering of the data.

Keywords: Technology shocks, impulse response functions, structural VAR, long-run identification, low frequency co-movement.

JEL Classification: C32, C51, E32, E37.

---

*Corresponding author. Department of Economics. Concordia University, 1455 de. Maisonneuve Blvd. West, Montreal, Quebec, H3G 1M8. nikolay.gospodinov@concordia.ca. The authors thank Michelle Alexopolous, Efrem Castelnuovo, Yongsung Chang, Francesco Furlanetto, Leo Michelis, Barbara Rossi, Shinichi Sakata, seminar participants at University of Padova, University of Cyprus, University of British Columbia, Carleton University, Queens University and Louisiana State University and conference participants at the 2008 Far Eastern Meetings of the Econometric Society, the 2008 Meeting of the Canadian Econometric Study Group, ICEEE 2009, CIREQ Time Series Conference and the 2009 Joint Statistical Meetings. Part of this research was done while Elena Pesavento was a Marco Fanno visitor at the Dipartiment di Scienze Economiche at Universita' di Padova. Nikolay Gospodinov gratefully acknowledges financial support from FQRSC and SSHRC.

1 Introduction

An ongoing debate exists regarding the empirical effect of technology shocks on production inputs, such as hours worked. Most standard real business cycle models start with the premise that business cycles result from unexpected changes in production technologies. This has the implication that hours worked and other inputs to production should rise following a positive technology shock. On the other hand, models with frictions, such as sticky prices, often predict an initial fall in hours worked following a productivity shock.[1]

As technology shocks are difficult to measure,[2] they are commonly specified as structural shocks in vector autoregressive (VAR) models which are identified via the long-run (LR) restriction that only technology shocks have a permanent effect on labour productivity (Gali, 1999, for example). This identification scheme, an implication of many modern macroeconomic models, has been widely employed in recent years. However, despite its common acceptance, the qualitative results have proven quite sensitive to other aspects of the VAR specification, particularly whether hours worked are modeled in levels or first differences.

Specifying the VAR in the difference of both hours worked and labour productivity, Gali (1999) and Shea (1999) find that hours worked initially fall following a positive technology shock, a finding which gives support to models with embedded frictions. Other papers have reached similar conclusions (see, for example, Francis and Ramey, 2005; Basu, Fernald and Kimball, 2006; among others) and this has spurred a line of research aimed at developing general equilibrium models (Gali and Rabanal, 2004) or alternative finite-horizon identification schemes (Uhlig, 2004; Francis, Owyang and Roush; 2005) that can account for this empirical finding.

However, maintaining the long-run identification restriction but allowing hours worked to en-

[1] However, Chang, Hornstein, and Sarte (2008) show that once inventories are allowed for, the dynamics in both models become more complicated and it is no longer so simple to distinguish between flexible and sticky price models based on the response of hours worked.

[2] Alexopoulos (2006) and Shea (1999) provide measurements of technological progress based on technology publications and patent data respectively.

ter the model in levels, Christiano, Eichenbaum and Vigfusson (2003, 2006) provide support for the prediction of standard RBC models, with hours worked rising immediately after a positive productivity shock. Christiano, Eichenbaum and Vigfusson (2003) point out that the differenced specification is misspecified unless there is an exact unit root in hours and argue strongly in favor of the levels specification. They find by simulations that the levels specification encompasses the estimated impulse response function of the differenced specification, but not vice-versa.

More recently, Fernald (2007) argues against this interpretation. He reports evidence of two level-shift breaks in both productivity growth and hours worked, occurring in the early 1970s and mid-1990s. This results in a common high-low-high pattern, which Fernald (2007) refers to as a low frequency correlation. He provides both intuition and simulation results in support of the claim that the results from the levels specification are mechanically driven by this common high-low-high pattern. After removing sub-sample means, Fernald (2007) finds that both the levels and differenced specifications produce similar impulse response functions, both qualitatively matching the original results of Gali (1999). As we discuss further below, an important assumption made in his framework is that the similar timing and direction of the breaks in labor productivity and hours is treated as coincidental.

Francis and Ramey (2009) propose an alternative explanation for the low frequency correlation which is based on demographics and employment shifts between private, government, and non-profit sectors. In line with Fernald's (2007) findings, they demonstrate that when this demographic factors are removed, either by a simple Hodrick-Prescott (HP) filter or by more sophisticated demographic adjustments, then both the levels and differenced specifications agree qualitatively with those of Gali (1999). In their framework the correlation is real, as low frequency shifts in both hours worked and productivity are caused by the same demographic and public employment trends. However, they argue that these demographic trends should be removed since they have little do with technology shocks. To lend support to this argument, they analyze a simple macroeconomic model with demographics in which the long-run identifying assumption are violated unless the demographic influences are removed from the data.

Our paper takes up the challenge of reconciling the conflicting empirical findings reported in the literature and contributes to the understanding of this debate in several respects. We demonstrate analytically that the documented extreme sensitivity to different model specifications appears to be due to a discontinuity in the solution for the structural coefficients implied by the long-run restriction. Even when the estimated reduced-form parameters are quite similar, the implied structural parameter and therefore the implied impulse response functions, can be very different when the largest autoregressive root of hours is one and when it is equal to, say, 0.99.

Interestingly, we find, both analytically and by simulations, that this discontinuity appears to arise only in the presence of a low frequency correlation between hours worked and productivity growth. This draws a tight link between the apparently conflicting results of Christiano, Eichenbaum and Vigfusson (2003), who argue that the differenced SVAR is misspecified, and Fernald (2007) who argue that the levels specification is misleading without accounting for structural breaks. Although they point in opposite directions, both sets of results at least implicitly rely on similar low frequency correlations.

The discontinuity that we uncover cautions against basing specification choices in models identified by long-run restrictions on univariate pre-tests of unit roots. It is well known that most unit root tests tend to favor the differenced specification. This is well described by Francis and Ramey (2009, p. 1072) who write that "if one were to rely on econometrics, which fail to reject the presence of a unit root in per capita labor, one would be led to enter labor input in first differences ... However, common sense tells us that per capita labor being a bounded series cannot have a unit root." In fact, due to the sharp discontinuity which occurs at unity, we argue that the unit root pre-testing is unlikely to provide any indication of which specification is preferred and that over-differencing can lead to equally biased results in a local-to-unity specification, against which unit root tests have no consistent power.

Our findings also help to clarify why the literature reports a preponderance of evidence in favor of the original Gali (1999) result, when applying other methods of filtering, such as the Hodrick-Prescott filter or break removal to either series. Just as in the case of differencing, these filters can be seen to remove the low frequency correlation. Thus, these alternative filtering strategies confirm the conclusions of the differenced specification because they share the same primary function as differencing. Our results indicate that in the absence of a low frequency correlation, the long-run restriction is unchanged by differencing. This further explains why the level and differenced specification provide similar results after the low frequency components have been removed by either trend break removal or HP filtering.

More generally, we argue that the difference in conclusions cannot be determined solely on the basis of empirical methods, such as unit-root pre-tests or HP pre-filtering. Instead, the appropriate conclusions that one draws from any of these approaches rests critically on the economic assumptions made about the source of the low frequency correlation. If the long-run identifying assumption holds true and thus these low frequency correlations are treated as a true feature of the data generating process, as is implicitly the case in Christiano, Eichenbaum and Vigfusson (2003), then over-differencing improperly removes this low frequency correlation, thereby corrupting the long-run identification of the difference specification. Such true low frequency co-movement may be plausible if technological changes have long-lasting effects on the underlying structure of the labour market. For example, technological improvements give rise to greater efficiency in household production, leading to increased female labour market participation. Likewise, technological innovations affecting regional transportation or labour search costs, may also have lasting impacts on labour markets. On the other hand, Fernald (2007) provides some convincing arguments for why the similar timing of the structural breaks in productivity and hours may be coincidental, arising from disparate causes. Although Francis and Ramey (2009) instead demonstrate that this low frequency behavior may be driven by common demographic and sectoral employment changes, they argue that it violates the long-run identifying assumption and should thus be treated as low frequency noise. In either case, it is the presence of this low frequency correlation that corrupts the long-run identification and renders the unmodified levels specification misleading.

The popularity of the long-run identification scheme derives in large part from its robustness to model specification, in the sense that it often remains valid under a wide variety of macroeconomic models. However, the implementation of the long-run restriction also relies on the low frequency properties of the data. Our results, which illustrate the possibility of discontinuity in this dependence, reinforce the conclusions from the empirical literature suggesting that the long-run identifying scheme can be far less robust to assumptions on the low frequency properties of the variables. Of course, there may still be many cases in which robust empirical results can be obtained; for example, if the variables in question are obviously stationary or if low frequency correlations are not present. Nevertheless, we echo the recommendation made by Fernald (2007) that empirical researchers should check carefully the robustness of their results to alternate assumptions on the low frequency properties of the data.[3]

The rest of the paper is organized as follows. Section 2 briefly reviews some empirical evidence and provides the intuition behind our findings. In Section 3, we formalize this intuition and present a theoretical model that helps us to identify the possible source of low frequency correlations and derive the implications for the impulse responses identified with long run restrictions. Section 4 presents the results from a Monte Carlo simulation experiment. Section 5 discusses the main implications of our analysis for empirical work and Section 6 concludes.

[3] This issue is distinct from the critique of Erceg, Guerrieri and Gust (2005) and Chari, Kehoe and McGrattan (2008), who argue that finite structural VARs poorly approximate the infinite order models that are implied by economic theory. Christiano, Eichenbaum and Vigfusson (2006) show that the resulting lag-truncation biases can be serious in theory but they tend to be less serious under realistic parameterizations.

2 Illustrative Example and Intuitive Arguments

To put the subsequent discussion in the proper empirical context, we present in Figure 1 the estimated impulse response functions (IRFs) based on the levels and differenced specifications with quarterly U.S. data for the period 1948Q2 - 2005Q3.[4] The difference in the impulse response functions is quite striking. Despite the voluminous recent literature on the effects of technology shock on hours worked, there is still little understanding of how such large quantitative and qualitative differences in the impulse responses can be generated. While the literature attributed these discrepancies to potential biases in both VAR specifications, it is not clear that such biases are large enough in practice to explain such highly divergent results especially in the short run. In fact, we find that it is nearly impossible to justify these differences solely by the behavior of hours worked itself and, in particular, by small deviations of the largest root of hours worked from unity.

It is well known, for example, that over-differencing, and misspecification in general, can lead to biased results. However, what is indeed surprising is that a seemingly very minor, even undetectable, misspecification in the difference specification, may lead to such a substantial bias in the resulting impulse response function. Standard unit root and stationarity tests on hours worked, neither of which reject their respective null hypothesis, provide little guidance regarding this specification choice (Christiano, Eichenbaum and Vigfusson, 2006).[5] Pesavento and Rossi (2005) provide confidence intervals on the largest autoregressive root in hours worked using inversions of four different unit root tests. All four confidence intervals include unity and in two cases, the lower bound on largest on largest root exceeds 0.980 (in the other two cases, it exceeds 0.925). On the face of it, this hardly appears to be a case in which over-differencing would lead to large misspecification errors. In fact, in a reduced-form near unit root model, the specification error committed by over-differencing is second order. Nevertheless, Christiano, Eichenbaum and Vigfusson (2003) find

[4] U.S. data on labour productivity, hours worked in the non-farm business sector and population over the age of 16 from DRI Basic Economics (the mnemonics are LBOUT, LBMN and P16, respectively).

[5] Using a multivariate Bayesian posterior odds procedure, Christiano, Eichenbaum and Vigfusson (2003) find evidence in favor of the levels specification.

quite a large specification error in their calibrated simulations. This provocative result has yet to be satisfactorily explained in an econometric sense.

Another way to look at the problem is to note that the differenced specification ignores possible low frequency co-movements between labour productivity growth and hours worked. Figure 2 reveals that the Hodrick-Prescott trend[6] of labour productivity growth and hours worked exhibit some similarities and suggest that labour productivity growth may inherit its small low frequency trend component from hours worked. On a more intuitive level, if hours worked are a highly persistent, but stationary, process, it is possible that labour productivity growth inherits some small low frequency component from hours without inducing any observable changes in its time series properties.

In fact, as we show later, the seemingly conflicting evidence from the levels and differenced specifications identified with LR restrictions can only be reconciled when these deviations from the exact unit root are accompanied by small low frequency co-movements between labour productivity growth and hours worked. We show that this low frequency co-movement drives a wedge between the levels and differenced specifications with a profound impact on their impulse response functions.

This situation arises when restrictions on the matrix of LR multipliers, which includes low frequency information, are used to identify technology shocks. While the levels specification explicitly estimates and incorporates this low frequency co-movement in the computation of the impulse response functions, the differenced specification restricts this element to be zero. It is important to emphasize that this component could be arbitrarily small and could accompany an autoregressive (AR) root arbitrarily close to one, yet still produce substantial differences in the impulse responses from the two specifications. Therefore, our results also suggest that a pre-testing procedure for a unit root will be ineffective in selecting a model that approximates well the true IRF when hours worked are close to a unit root process. In this case, the pre-testing procedure would favor the differenced specification, which rules out the above mentioned low frequency correlation, with high

[6] Throughout the paper, the value of the smoothing parameter for the HP filter is set to 1,600.

probability. This could in turn result in highly misleading IRF estimates. In the next section, we provide more formal arguments for explaining and reconciling the conflicting empirical evidence from the levels and differenced specifications.

3 Analytical Framework for Understanding the Debate

Our analytical framework and econometric specification is designed to mimic some of the salient features of the data and the implications of the theoretical macroeconomic (in particular, RBC) models. First, we specify labour productivity as an exact unit root process. The RBC model imposes a unit root on technology and the data provide strong empirical support for this assumption. Hours worked exhibit a highly persistent, near-unit root behavior, although the standard RBC model implies that they are a stationary process. Since an exact unit root cannot be ruled out as an empirical possibility, we do not take a stand on this issue and consider both the stationary and unit root cases. However, these different specifications (stationary or nonstationary) either allow for or restrict the low frequency co-movement between hours worked and labour productivity growth. It turns out that this has crucial implications for the impulse response functions.

If hours worked are assumed stationary, the matrix of largest roots of the labour productivity growth and hours worked can contain a non-zero upper off-diagonal element, whose magnitude depends on the closeness of the root of hours worked to one. This, typically fairly small, off-diagonal element can produce substantial differences in the shapes and the impact values of the impulse response functions from models that incorporate (levels specification) and ignore (differenced specification) this component.

Alternatively, in the case of an exact unit root for hours worked, the matrix of largest roots specializes to the identity matrix. In this case, there can be no low frequency co-movement between hours work and labour productivity growth, ruling this out as an explanation for the difference between the two sets of impulse response functions. It is important to note, however, that this explanation is ruled out only in the case of an *exact* unit root. Our results suggest that this small low frequency co-movement can continue to induce large discrepancies between the IRFs of the differenced and levels VARs, even when the largest root is arbitrarily close to and indistinguishable from unity.[7]

In order to complete the model, we need to adopt an identification scheme that allows us to recover the structural parameters and shocks. We follow the literature and impose the long-run identifying restriction that only shocks to technology can have a permanent effect on labour productivity. In addition, we assume that the structural shocks are orthogonal. In the next subsections, we formalize this analytical framework and work out its implications for the impulse response functions based on levels and differenced specifications.

3.1 Reduced-form model

Consider the reduced form of a bivariate vector autoregressive process $\widetilde{y}_t = (l_t, h_t)'$ of order $p+1$

$$\Psi(L)(I - \Phi L)\widetilde{y}_t = u_t, \tag{1}$$

where $\Psi(L) = I - \sum_{i=1}^{p} \Psi_i L^i = \begin{bmatrix} \psi_{11}(L) & \psi_{12}(L) \\ \psi_{21}(L) & \psi_{22}(L) \end{bmatrix}$, $E(u_t | u_{t-1}, u_{t-2}, ...) = 0$, $E(u_t u_t' | u_{t-1}, u_{t-2}, ...) = \Sigma$ and the matrix Φ is expressed in terms of its eigenvalue decomposition as $\Phi = V^{-1}\Lambda V$, where $\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & \rho \end{bmatrix}$ contains the largest roots of the system and $V = \begin{bmatrix} 1 & -\gamma \\ 0 & 1 \end{bmatrix}$ is a matrix of corresponding eigenvectors (see, for example, Pesavento and Rossi, 2006). Simple algebra yields $\Phi = \begin{bmatrix} 1 & \delta \\ 0 & \rho \end{bmatrix}$, where $\delta = -\gamma(1-\rho)$, is the parameter that determines the low frequency co-movement between the variables and ρ denotes the largest root of hours worked. This parameterization, which arises directly from the eigenvalue decomposition of Φ, allows for a small (δ) impact of h_t on l_t, provided that ρ is not *exactly* equal to one. Note that in the exact unit root case, Φ collapses to the identity matrix. The other off diagonal element of V, and therefore of Φ, is set to

[7] This argument can also be formalized in the local-unity setting that we consider in Section 3.4. In this setting, the off-diagonal element must itself be vanishing (i.e. local-to-zero), but nonetheless has a critical impact on the impulse response functions.

zero as it would otherwise imply that hours is $I\left(2\right)$ when $\rho=1$ and $I\left(1\right)$ when $\rho<1$.[8]

It is convenient to rewrite model (1) in Blanchard and Quah's (1989) framework by imposing the exact unit root on productivity so that $\triangle l_t$ is a stationary process. In this case, let $y_t=(\triangle l_t, h_t)'$ and $A(L)=\Psi(L)\begin{bmatrix} 1 & \gamma\left(1-\rho\right)L \\ 0 & 1-\rho L \end{bmatrix}$.[9] Then, the reduced form VAR model is given by

$$A(L)y_t=u_t \tag{2}$$

$$y_t=A_1y_{t-1}+...+A_{p+1}y_{t-p-1}+u_t.$$

The non-zero off-diagonal element $\gamma\left(1-\rho\right)L$ allows for the possibility that a small low frequency component of hours worked affects labour productivity growth. When the low frequency component is removed from either hours worked (Francis and Ramey, 2009, and Gali and Rabanal, 2004) or labour productivity growth (Fernald, 2007), this coefficient is driven to zero and the estimated IRF resembles the IRF computed from the differenced specification. The above parameterization of Φ can be used to explain this result.

3.2 Structural VAR

We denote the structural shocks (technology and non-technology shocks, respectively), by $\varepsilon_t=(\varepsilon_t^z,\varepsilon_t^d)'$, which are assumed to be orthogonal with variances σ_1^2 and σ_2^2, respectively, and relate them to the reduced form shocks by $\varepsilon_t=B_0u_t$, where $B_0=\begin{bmatrix} 1 & -b_{12}^{(0)} \\ -b_{21}^{(0)} & 1 \end{bmatrix}$. Pre-multiplying both sides of (2) by the matrix B_0 yields the structural VAR model

$$B(L)y_t=\varepsilon_t,$$

[8] In principle, the model can also be generalized to include a non-zero (but asymptotically vanishing) feedback from the level of productivity to hours worked. Simple algebra (available from the authors upon request) shows that this parameterization does not affect the subsequent analysis of the impulse response of hours worked to technology shocks under the long-run restriction that non-technology shocks have no permanent effect on productivity. For this reason, we set the lower off diagonal element of Φ to zero without any loss of generality.

[9] It is important to note, however, that the zero lower off diagonal restriction on matrix Φ does not rule out a feedback from productivity *growth* to hours worked in higher-order ($p>0$) VAR models. Thus, it has no implication for the direction of causality implied by the low frequency correlation between hours worked and productivity *growth*. For example, in a VAR(2) model, the lagged productivity growth is allowed to affect hours worked through the possibly non-zero coefficient ψ_{21}.

where $B(L) = B_0 A(L)$.

The matrix of long-run multipliers in the SVAR for y_t is

$$B(I) = \left[\begin{array}{cc} \psi_{11}(1) - b_{12}^{(0)}\psi_{21}(1) & (1-\rho)\left([\gamma\psi_{11}(1) + \psi_{12}(1)] - b_{12}^{(0)}[\gamma\psi_{21}(1) + \psi_{22}(1)]\right) \\ \psi_{21}(1) - b_{21}^{(0)}\psi_{11}(1) & (1-\rho)\left([\gamma\psi_{21}(1) + \psi_{22}(1)] - b_{21}^{(0)}[\gamma\psi_{11}(1) + \psi_{12}(1)]\right) \end{array} \right].$$

Imposing the restriction that non-technology shocks have no permanent effect on labour productivity renders the matrix $B(I)$ lower triangular.[10] For $\rho < 1$, this LR restriction translates into the restriction $b_{12}^{(0)} = [\gamma\psi_{11}(1) + \psi_{12}(1)]/[\gamma\psi_{21}(1) + \psi_{22}(1)]$.

Suppose now that one assumes $\rho = 1$ and let $\triangle\widetilde{y}_t = (\triangle l_t, \triangle h_t)'$. Then, the reduced form specializes to

$$\Psi(L) \triangle \widetilde{y}_t = u_t \tag{3}$$

and the structural form is given by

$$B_0\Psi(L) \triangle \widetilde{y}_t = \varepsilon_t$$

with a long-run multiplier matrix

$$B(I) = \left[\begin{array}{cc} \psi_{11}(1) - b_{12}^{(0)}\psi_{21}(1) & \psi_{12}(1) - b_{12}^{(0)}\psi_{22}(1) \\ \psi_{21}(1) - b_{21}^{(0)}\psi_{11}(1) & \psi_{22}(1) - b_{21}^{(0)}\psi_{12}(1) \end{array} \right].$$

Note that the LR restriction implies that $b_{12}^{(0)} = \psi_{12}(1)/\psi_{22}(1)$ and even if the upper right element of Φ is non-zero, the differenced VAR would ignore any information contained in the levels and implicitly set this element equal to zero.

Once the structural parameter $b_{12}^{(0)}$ is obtained (by plugging consistent estimates of the elements of $\Psi(I)$ from the reduced form estimation), the remaining parameters can be recovered from $B_0 E(u_t u_t')B_0' = E(\varepsilon_t\varepsilon_t')$ or

$$b_{21}^{(0)} = \frac{b_{12}^{(0)}\Sigma_{22} - \Sigma_{12}}{b_{12}^{(0)}\Sigma_{12} - \Sigma_{11}},$$

$$\sigma_1^2 = \Sigma_{11} - 2b_{12}^{(0)}\Sigma_{12} + \left[b_{12}^{(0)}\right]^2 \Sigma_{22}$$

[10] An alternative formulation of the restriction is that $C(I)$ is lower triangular, where $C(L) = B(L)^{-1}$ describes the moving average representation $y_t = C(L)\varepsilon_t$. Simple matrix algebra shows that the two restrictions are equivalent.

and

$$\sigma_2^2 = \Sigma_{22} - 2b_{21}^{(0)}\Sigma_{12} + \left[b_{21}^{(0)}\right]^2 \Sigma_{11},$$

where Σ_{ij} is the $[ij]$th element of Σ. These parameters can be used consequently for impulse response analysis and variance decomposition.

3.3 Implications for impulse response analysis

The impulse response functions of hours worked to a shock in technology can be computed either from the levels specification (Blanchard and Quah, 1989; Christiano, Eichenbaum and Vigfusson, 2006; among others) or the differenced specification (Gali, 1999; Francis and Ramey, 2005). The levels approach will explicitly take into account and estimate a possible non-zero upper off-diagonal element in Φ but it suffers from some statistical problems when hours worked are highly persistent. Christiano, Eichenbaum and Vigfusson (2003) note that the levels specification tends to produce IRFs with large sampling variability that are nearly uninformative for distinguishing between competing economic theories. Gospodinov (2010) shows that this large sampling uncertainty arises from a weak instrument problem when the largest root of hours worked is near the nonstationary boundary. On the other hand, the differenced approach will produce valid and asymptotically well-behaved IRF estimates in the exact unit root case but it ignores any possible low frequency co-movement between hours and labour productivity growth when hours worked is stationary. It can therefore give rise to highly misleading IRFs even for very small deviations from the unit root assumption on hours.

Since $b_{12}^{(0)} = [\gamma\psi_{11}(1) + \psi_{12}(1)]/[\gamma\psi_{21}(1) + \psi_{22}(1)]$ and $b_{12}^{(0)} = \psi_{12}(1)/\psi_{22}(1)$ can produce very different values of $b_{12}^{(0)}$, the IRFs from these two approaches can be vastly different. In fact, because the value of γ does not depend on ρ, these differences can remain large even for $(\rho - 1)$ arbitrarily close, but not equal, to zero. For simplicity, take the first-order model where $\Psi(L) = I$. In this case, the two restrictions set the value of $b_{12}^{(0)}$ to γ and 0, respectively, implying two very different values for $b_{21}^{(0)}$, which, in turn, directly determines the impulse response function, since in the first-order model

$$\theta_{hz}^{(l)} = \frac{\partial h_{t+l}}{\partial \varepsilon_t^z} = \left[\Phi^l B_0^{-1}\right]_{21} = \frac{b_{21}^{(0)}\rho^l}{1 - b_{12}^{(0)} b_{21}^{(0)}}. \tag{4}$$

As it is clear from (4) , the impact effect at $l = 0$ does not depend on the value of ρ as $\rho^0 = 1$, but only on the values of $b_{21}^{(0)}$and $b_{12}^{(0)}$, which themselves depend on $\Psi(1)$ and γ. Focusing the debate on the distance of ρ from one is therefore misleading, provided that ρ is not precisely equal to one.

To visualize the differences in the IRFs from the levels and differenced specifications when Φ is not diagonal, it is instructive to consider the following simplified example. Suppose that the true data generating process is a first-order VAR with $\rho = 0.98$, $\gamma = -1$ (which implies an off-diagonal element $\delta = -\gamma(1-\rho) = 0.02$) and $\Sigma = \begin{pmatrix} 1 & -0.2 \\ -0.2 & 0.8 \end{pmatrix}$. From the above formulas, it can be easily inferred that the true values of the parameters that enter the IRF are $b_{12}^{(0)} = 1$, $b_{21}^{(0)} = 0.75$ and $\sigma_1^2 = 1.4$, whereas the differenced approach uses values of $b_{12}^{(0)} = 0$, $b_{21}^{(0)} = -0.2$ and $\sigma_1^2 = 1$. The IRFs based on the levels (true) and differenced specifications are plotted in Figure 3.

Figure 3 clearly illustrates the large differences in the IRFs from the two specifications that are generated by the presence of a small off-diagonal element δ. Interestingly, the differences between the IRFs do not necessarily disappear as ρ gets closer to one and δ approaches zero. As our analytical framework suggests, they can remain substantial even for values of $\rho - 1$ and $\delta = -\gamma(1-\rho)$ arbitrarily close, but not equal, to zero. This is because, provided that $\rho < 1$, the size of this discrepancy depends on the co-movement through the parameter γ, rather than through either δ or ρ. At a more intuitive level, the reason that the short-horizon IRFs can be highly sensitive to even small low frequency co-movements accompanying small deviations of ρ from one, is that they are identified off of long-run identification restrictions, which depend entirely on the zero frequency properties of the data. As reported below, a similar sensitivity does not arise when short-run identification restrictions are employed.

3.4 An alternative parameterization

The fact that our framework suggests potentially large IRF discrepancies even for values of ρ quite close to one is practically relevant, precisely because this is the case in which unit root tests have the greatest difficulty detecting stationarity. The low power of the unit root test in this case arises because, in small samples, the resulting process for hours may behave more like a unit root process than like a stationary series. This concept has been formalized in the econometrics literature by the near unit root or local-to-unity model, in which $\rho = 1 - c/T$ for $c \geq 0$ is modelled as a function of the sample size T and shrinks towards unity as T increases (Phillips, 1987; Chan, 1988). Naturally, this dependence on the sample size should not be interpreted as a literal description of the data, but rather as a device to approximate the behavior of highly persistent processes in small samples. What makes this modelling device particularly relevant, is that, for small values of the local-to-unity parameter c, it describes a class of alternatives to $\rho = 1$ against which unit root tests have no consistent power. Intuitively, $c = T(1 - \rho)$ can be viewed as measuring the distance of the root from one relative to the sample size. Small values of c correspond to cases in which T is relatively small and ρ is relatively close to one, so that unit root tests have low power and the difference specification is likely to be employed when computing IRFs.

An alternative parameterization of the model in (1) is therefore obtained by modeling the largest root in hours as a local-to-unity process with $\rho = 1 - c/T$ with $c \geq 0$. Then, it follows that $\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 1 - c/T \end{bmatrix}$, $V = \begin{bmatrix} 1 & -\gamma \\ 0 & 1 \end{bmatrix}$ and $\Phi_T = \begin{bmatrix} 1 & -\gamma c/T \\ 0 & 1 - c/T \end{bmatrix}$. In finite samples, as long as $c > 0$, no matter how small, the co-movement between hours and productivity is different than zero, although arbitrary small. The reduced form for $y_t = (\triangle l_t, h_t)'$ is now

$$A(L)y_t = u_t$$

with $A(L) = \Psi(L) \begin{bmatrix} 1 & \gamma\,(c/T)\,L \\ 0 & (1 - L) + (c/T)\,L \end{bmatrix}$. In the unit root case ($c = 0$), Φ_T collapses to the identity matrix, the variables are not cointegrated and there is no feedback from hours to productivity growth. Thus, the impact of h_{t-1} on $\triangle l_t$ is local-to-zero and vanishing at rate $T^{-1/2}$,[11] capturing the notion that the low frequency co-movement between hours and productivity growth must be small if the root of hours is close to unity. Writing the model in the local-to-unity form is also intuitively appealing since the low frequency correlation between h_{t-1} and $\triangle l_t$ is bound to disappear asymptotically, so that hours do not affect productivity growth in the long run.

Under the local-to-unity parameterization, the matrix of long-run multipliers becomes

$$B(I) = \begin{bmatrix} \psi_{11}(1) - b_{12}^{(0)}\psi_{21}(1) & c/T\left([\gamma\psi_{11}(1) + \psi_{12}(1)] - b_{12}^{(0)}[\gamma\psi_{21}(1) + \psi_{22}(1)]\right) \\ \psi_{21}(1) - b_{21}^{(0)}\psi_{11}(1) & c/T\left([\gamma\psi_{21}(1) + \psi_{22}(1)] - b_{21}^{(0)}[\gamma\psi_{11}(1) + \psi_{12}(1)]\right) \end{bmatrix}$$

and the restriction that non-technology shocks have no permanent effect on labour productivity yields $b_{12}^{(0)} = [\gamma\psi_{11}(1) + \psi_{12}(1)]/[\gamma\psi_{21}(1) + \psi_{22}(1)]$ for $c > 0$. Note, that when $c = 0$, the model again specializes to the differenced VAR specification in (3), for which the LR specification implies $b_{12}^{(0)} = \psi_{12}(1)/\psi_{22}(1)$. As a result, the analysis of the shapes of the impulse response functions under the different specifications in Section 3.3 remains unchanged. This confirms the finding that substantial differences in IRFs can arise even within this class of models for which unit root tests are not powerful enough to detect that hours worked is stationary. Thus, the stylized fact that hours worked is indistinguishable from a unit root process does not guarantee that the true IRF will be close to the IRF from the differenced specification.

4 Monte Carlo Experiment

To demonstrate the differences in the IRF estimators with a non-diagonal Φ, we conduct a Monte Carlo simulation experiment. 10,000 samples for $y_t = (l_t, h_t)'$ are generated from the VAR(2) model

$$\left[I - \begin{pmatrix} -0.05 & -0.08 \\ 0.2 & 0.55 \end{pmatrix} L\right]\left[I - \begin{pmatrix} 1 & \delta \\ 0 & \rho \end{pmatrix} L\right]\begin{pmatrix} l_t \\ h_t \end{pmatrix} = \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix},$$

where $\delta = -\gamma(1-\rho)$, $T = 250$, $(u_{1,t}, u_{2,t})' \sim iidN(0, \Sigma)$, $\Sigma = \begin{pmatrix} 0.78 & 0 \\ 0 & 0.55 \end{pmatrix}$, and the parameter values are calibrated to match the empirical shape of the IRF of hours worked to a technology

[11] The level of h affects $\triangle l_t$ through the term $(\gamma c/T)h_{t-1}$ which is $O_p(T^{-1/2})$ since $T^{-1/2}h_{t-1} = O_p(1)$ in the local-to-unity setup.

shock.[12] The lag order of the VAR is assumed known. In addition to the IRF estimates from the levels and differenced specifications, we consider the IRF estimates from a levels specification with HP detrended productivity growth, as in Fernald (2007).

Figures 4 to 7 show simulation results for the IRFs under four different parameter configurations for ρ and γ, all of which lie in a range of values that is potentially consistent with the actual data. The three panels of each figure correspond to the different model specifications: a VAR in productivity growth and hours, a VAR in productivity growth and differenced hours and a VAR in HP detrended productivity growth and hours. For each model we show the true IRF (solid line), the median Monte Carlo IRF estimate (long dashes), and the 95% Monte Carlo confidence bands (short dashes).

In Figure 4 we consider a stationary but persistent process for hours ($\rho = 0.95$), while allowing a small low frequency component of hours worked to enter labour productivity growth ($\delta = 0.04$). As shown in the figure, the VAR in levels (left graph) estimates an IRF that is close, on average, to the true IRF, except for a small bias (see Gospodinov, 2010, for an explanation). On the other hand, the VAR with hours in first differences (middle graph) incorrectly estimates a negative initial impact of the technology shock even though the true impact is positive. These results are in agreement with our discussion in the analytical section.

In Figure 5, we increase the largest root of hours worked from $\rho = 0.95$ to $\rho = 0.97$ and also substantially decrease value of the off-diagonal element from $\delta = 0.04$ to $\delta = 0.015$. This brings the matrix of largest roots closer to an identity matrix for which the differenced specification is correct. Despite these changes, the IRFs shown in Figures 4 and 5 are strikingly similar. This underlines the ability of even a very small low frequency co-movement to drive a qualitatively important wedge between the IRFs based on the levels and differenced models. Likewise, it illustrates that the largest root need not be far from one for this effect to be important.

[12] Note that while the numbers for the short-run dynamics are chosen to match the empirical values estimated from a VAR in levels, in our simulations we also impose $\rho = 1$ and therefore allow both specifications (levels and first differences) to be the true DGP.

The lower panels of Figures 4 and 5 are also interesting. When the HP filter is used to remove the low frequency component from labour productivity growth (Fernald, 2007), the estimated IRF resembles the IRF computed from the differenced specification. The graphs clearly demonstrate that the removal of the low frequency component, by either differencing or HP filtering, eliminates the possibility of any low frequency co-movements between the transformed series and this has a profound influence on the IRFs.[13]

Figure 6 presents the results for the exact unit root case. In this case the matrix of largest roots becomes diagonal, eliminating the low frequency co-movement between hours and productivity growth ($\delta = 0$). Despite some small biases, all median IRF estimates now correctly sign the impact of the technology shock and come close to tracing out the true IRFs. Not surprisingly, the differenced specification is particularly accurate and produces an unbiased estimator with tight confidence intervals. The estimator from the levels specification exhibits both a modest bias that arises from the biased estimation of the largest root of hours and a very large sample uncertainty (Gospodinov, 2010). The estimator from the specification with HP filtered labour productivity growth performs similarly to the differenced estimator, although it is slightly biased and more dispersed.

In Figure 7, we maintain the assumption of a zero off-diagonal element ($\delta = 0$) and return to a persistent but stationary specification for hours worked ($\rho = 0.95$). The median IRFs from all models are again quite similar, both to each other and to the true IRF. In this sense, beside having smaller bias and variance, the basic message from Figures 6 and 7 is similar, despite the fact that hours are nonstationary in Figure 6 but stationary in Figure 7.

[13] After removing the low frequency component, the nature of the IRF changes and it is not completely justifiable to compare the IRFs from the transformed and the original processes. Nevertheless, we still report the IRFs on the same graph to illustrate the economically large differences created by a fairly small off-diagonal element. We also considered the specification when hours worked are HP-filtered as in Francis and Ramey (2009). The behavior of the IRF estimates in this model is similar to the case of HP-filtered productivity growth. Here, we make no argument as to whether the low frequency components should or should not be removed prior to the IRF analysis. Instead, we provide an analytical framework for explaining and reconciling the conflicting results documented in the empirical literature. We further discuss the implications of low frequency filtering in the next section.

In summarizing the results from these four figures, we note that large qualitative differences in median IRFs for the differenced and levels VARs were observed only in Figures 4 and 5, in which there is a small low frequency relationship between hours and labour productivity ($\delta \neq 0$). Neither Figure 6 nor Figure 7 show qualitative differences in the median IRFs from the levels and differenced specifications. Yet, in Figure 6, hours have a unit root, whereas they are stationary in Figure 7. While small sample bias is present and affects the precision of the estimation, our simulations show that, unlike Ergec, Guerrieri and Gust (2005), the small sample bias and persistence of the non-technology shocks alone are not enough to generate the substantial differences in the impulse responses that we find in practice. What Figures 6 and 7 share in common is the absence of the low frequency co-movement of Figures 4 and 5 (i.e. $\delta = 0$). Although the size of the unit root in hours worked has important implications for the sampling distributions of the IRFs, these results suggest that it is the low frequency co-movement that plays the critical role in driving the qualitative differences between the level and differenced specifications.

To better assess the sensitivity of the levels and differenced specifications to different values of ρ and δ, we plot in Figures 8 and 9 the true and estimated responses for various degrees of persistence and low frequency co-movement. Each line represents values for $\gamma = \{-0.5, -0.2, 0, 0.2, 0.5\}$, which correspond to different off-diagonal elements δ depending on the value of ρ (recall that $\delta = -\gamma(1 - \rho)$). Once again, it is clear that, while the level specification explicitly estimates and incorporates the different values for δ in the computation of the impulse response functions, the differenced specification implicitly imposes this element to be zero. This leads to substantial deviations from the true impulse response functions.

We also want to stress that the confidence intervals reported in Figures 4-7 are Monte Carlo confidence intervals, which are infeasible since they utilize knowledge of the true data generating process. The bias in the levels VAR and the misspecification in the first difference regressions result in poor coverage of confidence intervals constructed with standard procedures at medium and long horizons (Pesavento and Rossi, 2006). This is not reflected in our infeasible confidence intervals. At the same time, Figures 4-6 show clearly how a wide range of different estimates for the IRF are possible, and that the sampling uncertainty in the levels VAR is indeed larger. At the same time, except for the cases in which either ρ is exactly one or δ is exactly zero, the true impulse response is never contained in the Monte Carlo confidence bands for the VAR in first differences.

Finally, the differences in the IRFs for the various model specifications are expected to arise only in the case of long-run identification restrictions that are directly affected by the inclusion or the omission of the low frequency component. In order to verify this conjecture, we estimate the IRFs based on a short-run identification (Cholesky decomposition) scheme, with productivity growth ordered first and hours second. While we recognize that imposing short-run restrictions may be rather ad hoc and may lack a solid theoretical justification, Christiano, Eichenbaum and Vigfusson (2006) demonstrate that the short-run identification scheme produces estimates with appealing statistical properties.[14] The results from the three models for $\rho = 0.97$ and $\delta = 0.015$ are presented in Figure 10. Unlike the long-run identification scheme (Figure 5), the IRF estimates for all specifications are very close to the true IRF and fall inside the 95% Monte Carlo confidence bands. This suggests that the short-run identification scheme is robust to the presence or absence of low frequency co-movements, which is not the case with identifying restrictions that are based on long-run information.[15]

[14] Our short-run identifying scheme is used only to illustrate the relative insensitivity of the IRFs to the low frequency co-movement when they are identified by short-run restrictions. We do not advocate its use in practice since it has no clear theoretical justification. See Christiano, Eichenbaum and Vigfusson (2006) for a more sophisticated, model-based, short-run identification scheme.

[15] The main conclusions from this simulation experiment continue to hold if the data are generated from a dynamic general equilibrium model in which the persistence and the low frequency co-movements between the variables are implicitly determined. Simulation results with data from the two-shock CKM specification in Christiano, Eichenbaum and Vigfusson (2006) are available from the authors upon request. These results confirm the poor properties of the impulse responses from the differenced specification and reveal that a pre-testing procedure (ADF test) has difficulties rejecting the unit root null of hours and leads to only small improvements over the differenced specification.

5 Discussion of Results

The analytical and numerical results presented above clearly suggest that some seemingly innocuous transformations of the data can lead to vastly (qualitatively and quantitatively) different policy recommendations. The main objective of this paper is to illustrate and identify the source of these differences. At the same time, several interesting observations and remarks emerge from our analysis that highlight some potential pitfalls in empirical work with structural dynamic models that use highly persistent variables in conjunction with long-run identifying restrictions.

First, it is common practice in macroeconomics to remove low frequency components by applying the HP filter when focusing on business cycle frequencies. For example, Fernald (2007) argues that the low frequency component is not important for business cycle analysis. The effect of technology shocks on hours worked is typically evaluated at business cycle frequency and it is reasonable to assume that the removal of low frequency components will not affect the conclusions. We agree with this position, provided that the structural shocks are identified using short- or medium-run restrictions. We argue that the low frequency component contains important long-run information that, while not directly relevant at business cycle frequencies, affects in a fundamental way the long-run restrictions. Therefore, omitting or explicitly removing the low frequency correlation can result in misspecification of the long-run restriction and hence the business cycle component that is of primary interest to the analysis. In contrast, the low frequency component does not seem to matter for the short-run restrictions and the transformations applied to the data do not affect the impulse responses that they identify, as illustrated in our simulation section.

Although the analogy is not exact, the removal of low frequency components bears some similarities to ignoring the long-run information contained in the error-correction term in cointegrated models. The cointegration information does not directly affect the business cycle analysis but is essential to the long-run equilibrium. If we use short-run restrictions, the cointegration information can be left out without serious consequences. If the data are subjected to differencing (filtering) prior to the analysis, the long-run information contained in the cointegrating relationship will be lost and the long-run restriction will be misspecified, which in turn will give rise to misleading results.

Second, it is well known that a highly persistent linear process often exhibits dynamics that are observationally equivalent to dynamics generated by a long memory, structural break or regime-switching process. Therefore, it is difficult to statistically distinguish between these processes in finite samples and commit to a particular specification. In our context, it is hard to determine if the low frequency component (for example, the U shape in hours worked) and co-movement are spurious or not. Fernald (2007) convincingly illustrates the cost of falsely keeping the low frequency component if this co-movement is spurious. Our results indicate that there is an equally large cost of falsely removing it when the co-movement is a true feature of the correctly identified model. Ultimately, the researcher has to take a stand on whether the long-restriction applies to the original or filtered data. Our analysis in the previous sections provides important information on the sensitivity (robustness) of the different statistical transformations of the data to misspecification of the long-run restriction.

Finally, pre-testing procedures that are used to determine which specification is more appropriate perform poorly, especially when the data are highly persistent. Our analysis suggest that large differences in the IRFs arise even when the largest root is arbitrarily close to one, in which case the pre-testing procedure selects the differenced specification with probability approaching one. Put another way, we find that, when identified by LR restrictions, the IRFs from the differenced specification are not robust to small deviations of the largest root from unity, even when those deviations are too small to be empirically detected.

6 Conclusion

This paper analyzes the source of the conflicting evidence from structural VARs identified by long-run restrictions on the effect of technology shocks on hours worked reported in several recent empirical studies. In this paper, we show analytically that the extreme sensitivity of the results to different model specifications can be explained by a discontinuity in the solution for the structural coefficients implied by the long run restrictions, which arises only in the presence of a low frequency correlation between hours worked and productivity growth. The critical mechanism underlying the difference between the levels and differenced specifications, is that the differenced specification restricts this correlation to zero when solving for structural model, whereas the levels specification allows it to enter in unrestricted manner. Consequently, it may not be surprising that alternative filtering approaches reported in the literature, such as HP filtering and trend-break removal, which remove this low frequency correlation, provide evidence supportive of the differenced VAR. We also demonstrate that low frequency correlations capable of causing strong discrepancies between the two specifications are compatible with autoregressive roots in hours worked that are indistinguishable from one. This sharp discontinuity implies that one cannot rely on univariate unit root tests to resolve this debate, since they are not designed to discriminate between exact and near-unit root models.

Fernald (2007) also highlights the role of an observed low frequency correlation in the data, modeled as a common U-shaped pattern driven by structural breaks, and provides a number of convincing empirical exercises to illustrate its importance. While this insightful analysis clearly demonstrates the empirical link between the low frequency correlation and the conflicting results, to date there has not been a full theoretical understanding of why this low frequency correlation plays such an important role. We fill this gap by showing, in a more general analytic framework, that the key role played by this low frequency correlation is to create a discontinuity between the structural solutions of the differenced and levels specifications.

The existing literature has unambiguously concluded that the low frequency correlation biases the unfiltered levels VAR, but not the differenced or filtered specification (Fernald 2007, Francis and Ramey, 2009). This has provided renewed support for Gali's (1999) influential finding in the face of earlier criticism from Christiano, Eichenbaum and Vigfusson (2003). By contrast, our results indicate that the presence of low frequency correlation does not necessarily indicate a problem with the levels specification. Instead, it can result in very substantial biases in the differenced specification. Indeed, we demonstrate that the biases in this specification observed in the simulations of Christiano, Eichenbaum and Vigfusson (2003) are due to essentially the same low frequency correlation that Fernald (2007) and Francis and Ramey (2009) interpret as contaminating the unfiltered levels specification.

Therefore, we argue that the importance of the low frequency correlation cannot by itself resolve the debate, because, depending on the way it is modeled, it may lead to biases in either the levels or difference specification. However, in conjunction with Fernald (2007) and Francis and Ramey (2009), our results help to clarify the terms of the debate. We demonstrate that if there is a true low frequency correlation in the population model that is correctly identified by the long-run identification restriction, then any procedure which removes this low frequency correlation, whether by differencing, HP filtering, or trend-break removal would result in a very serious bias. In fact, we find that one cannot reproduce the discrepancy in the results with any reasonable probability in a correctly identified model, without introducing such a true population correlation.[16]

The reason that this finding might seem to be at odds with those of Fernald (2007) and Francis and Ramey (2009), who both argue that the levels specification is biased, is that neither model the observed correlation as a true population correlation in a correctly identified model. Fernald (2007) argues that the observed low frequency correlation is purely coincidental, in which a similar

[16] While the levels VAR appears to provide a more reliable framework for analysis in this setup, it may also produce biased and highly variable IRF estimates, especially when the root in hours worked is close or equal to one. Imposing additional restrictions on the model (see, for example, Gospodinov, 2010) can lead to improved inference for the structural parameters and impulse responses.

pair of breaks occur in both series for unrelated reasons, due to historical happenstance. Francis and Ramey (2009) treat the observed correlation as a true population correlation explained by common low frequency trends in demographic and public employment, but argue that the long-run restriction does not hold until these low frequency trends are purged from the data. In our view the debate therefore hinges on the interpretation given to this low frequency correlation. If one has reasons to believe that there is a genuine low frequency co-movement in a correctly identified model, this would support the findings of Christiano, Eichenbaum and Vigfusson (2003). On the other hand, if one is convinced either that the observed correlation is coincidental (Fernald, 2007) or that it is due to factors that violate the identification restriction (Francis and Ramey, 2009) then the results may be interpreted as supporting the earlier findings of Gali (1999). More generally, our results also underline and help to explain the potential sensitivity of long-run identifying schemes to uncertainty regarding low frequency dynamics, even when identifying characteristics at business cycle frequency.

References

[1] Alexopoulos, M. (2006), "Read all about it!! What happens following a technology shock?" *Mimeo*, University of Toronto.

[2] Basu, S., Fernald, J.G., and M.S. Kimball (2006), "Are technology improvements contractionary?" *American Economic Review* 95, 1418-1448.

[3] Blanchard, O., and D. Quah (1989), "The dynamic effects of aggregate demand and supply disturbances," *American Economic Review* 79, 655-673.

[4] Chan, H.C. (1988), "The parameter inference for nearly nonstationary time series," *Journal of the American Statistical Association* 83, 857-862.

[5] Chang, Y., Hornstein, A., and P.-D. Sarte (2008), "On the employment effects of productivity shocks: The role of inventories, demand elasticity, and sticky prices," *Mimeo*, University of Rochester.

[6] Chari, V., Kehoe, P., and E. McGrattan (2008), "Are structural VARs with long-run restrictions useful in developing business cycle theory?" *Journal of Monetary Economics* 55, 1337-1352.

[7] Christiano, L., Eichenbaum, M., and R. Vigfusson (2003), "What happens after a technology shock?" *Board of Governors of the Federal Reserve System, International Finance Discussion Paper* 768.

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

[9] Erceg, C., L. Guerrieri, and C. Gust (2005), "Can long-run restrictions identify technology shocks?", *Journal of the European Economic Association* 3, 1237-1278.

[10] Fernald, J. (2007), "Trend breaks, long-run restrictions, and contractionary technology improvements," *Journal of Monetary Economics* 54, 2467-2485.

[11] Francis, N., Owyang, M., and J. Roush (2005), "A flexible finite-horizon identification of technology shocks," *Federal Reserve Bank of St. Louis*, *Working Paper* 2005-024E.

[12] 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.

[13] Francis, N., and V. Ramey (2009), "Measures of hours per capita with implications for the technology-hours debate," *Journal of Money, Credit, and Banking* 41, 1071-1097.

[14] Gali, J. (1999), "Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations?" *American Economic Review* 89, 249-271.

[15] Gali, J., and P. Rabanal (2004), "Technology shocks and aggregate fluctuations: How well does the RBC model fit postwar U.S. data?" In: B. Bernanke and K. Rogoff, eds., *NBER Macroeconomics Annual*, Cambridge, MA: MIT Press.

[16] Gospodinov, N. (2010), "Inference in nearly nonstationary SVAR models with long-run identifying restrictions," *Journal of Business and Economic Statistics* 28, 1-12.

[17] Pesavento, E., and B. Rossi (2005), "Do technology shocks drive hours up or down? A little evidence from an agnostic procedure," *Macroeconomic Dynamics* 9, 478-488.

[18] Pesavento, E., and B. Rossi (2006), "Small sample confidence intervals for multivariate impulse response functions at long horizons," *Journal of Applied Econometrics* 21, 1135-1155.

[19] Phillips, P.C.B. (1987), "Towards a unified asymptotic theory for autoregression," *Biometrika* 74, 535-547.

[20] Ravenna, F. (2007), "Vector autoregressions and reduced form representations of DSGE models," *Journal of Monetary Economics* 54, 2048-2064.

[21] Shea, J. (1999), "What do technology shocks do?" In: B. Bernanke and J. Rotemberg, eds., *NBER Macroeconomics Annual,* Cambridge, MA: MIT Press.

[22] Uhlig, H. (2004), "Do technology shocks lead to a fall in total hours worked?" *Journal of the European Economic Association* 2, 361–371.

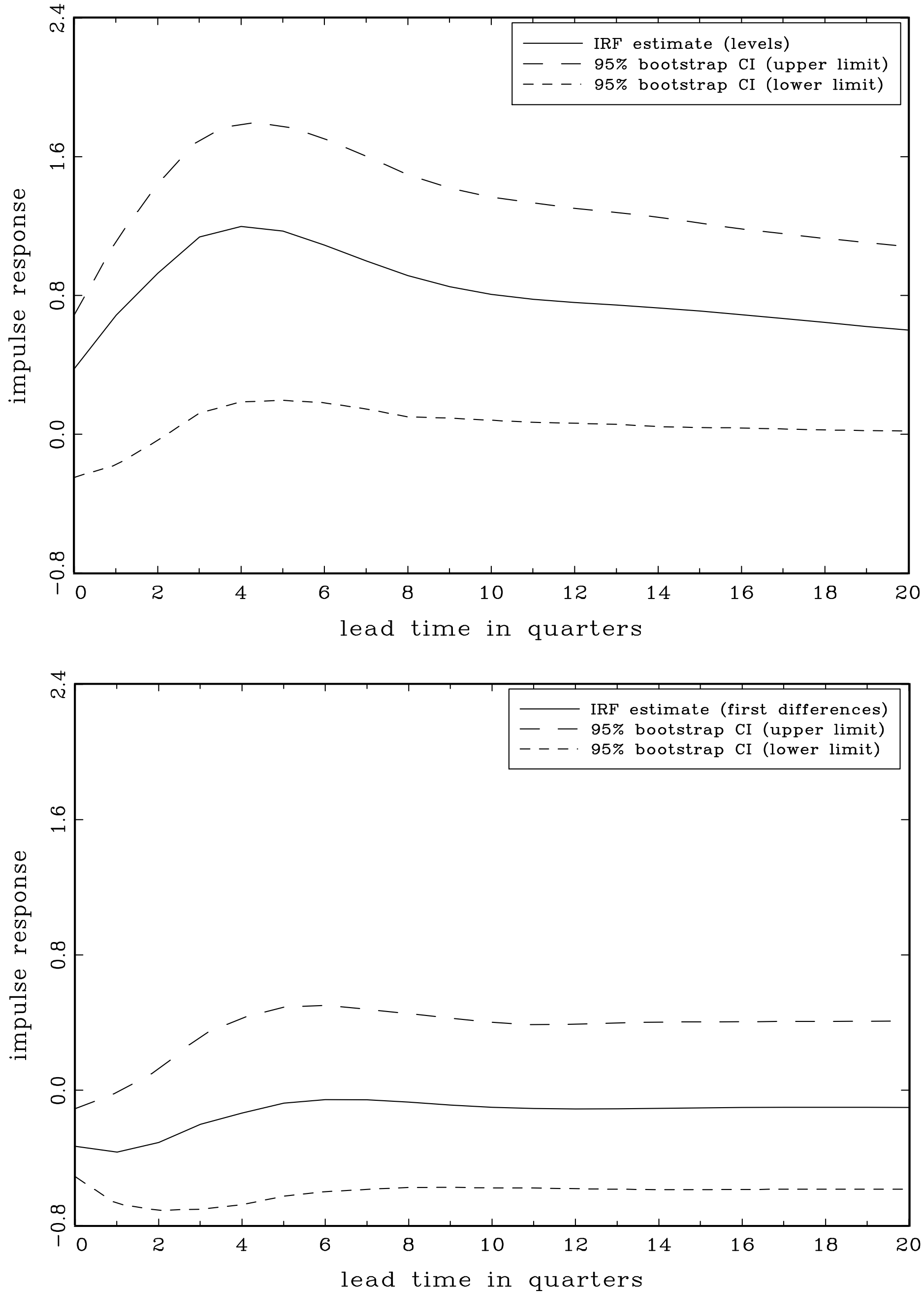

Figure 1. Response of hours worked to a 1% positive technology shock, U.S. data 1948Q2 - 2005Q3. Top graph: hours worked in levels; Bottom graph: hours worked in first differences.

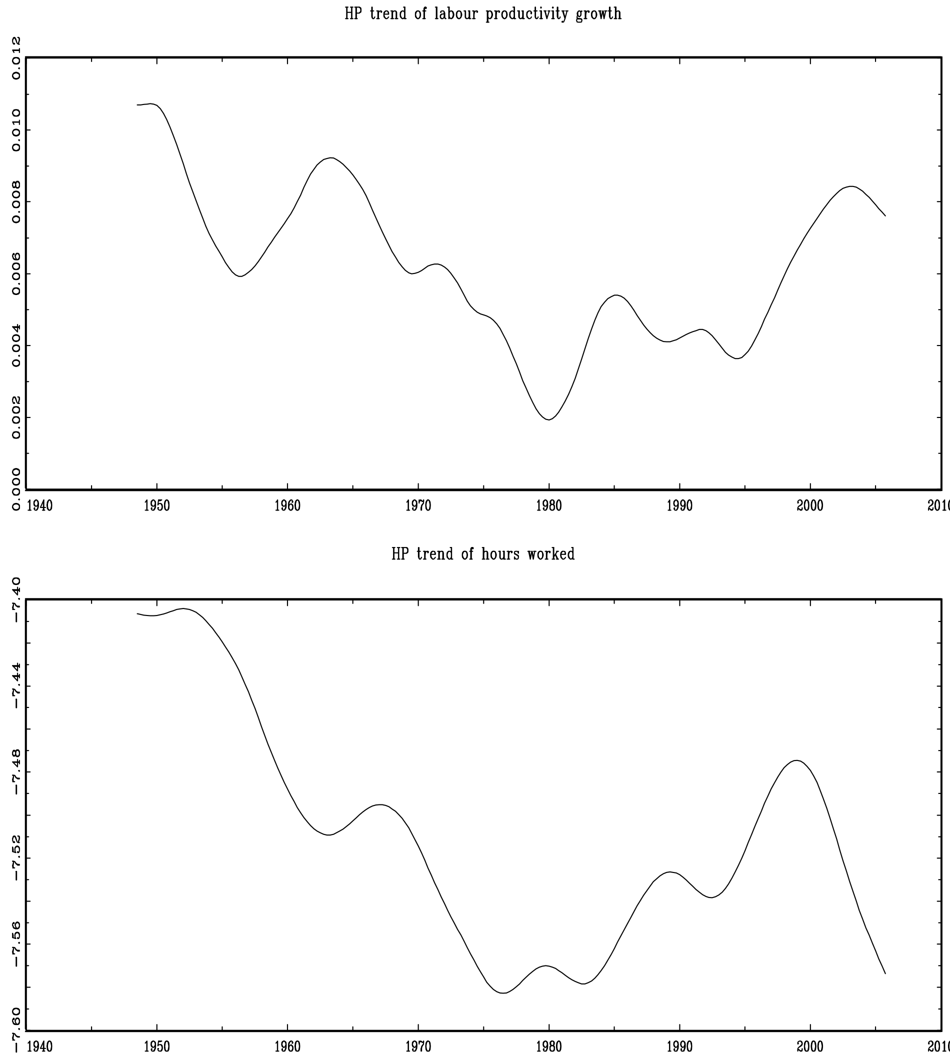

FIGURE 2. HP trends of labour productivity growth (top graph) and hours worked (bottom graph), U.S. data 1948Q2 - 2005Q3.

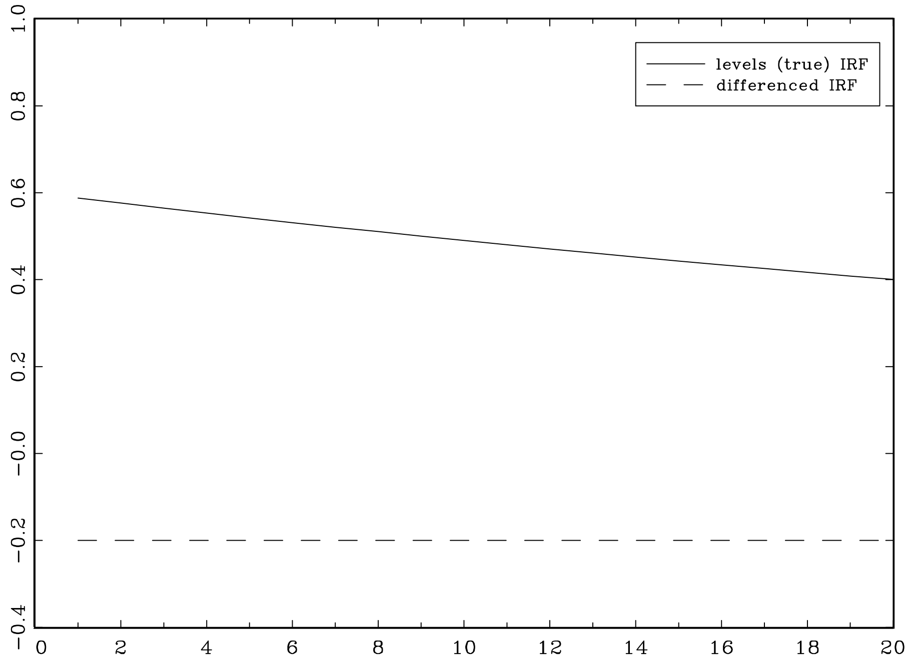

FIGURE 3. Impulse response functions computed from the levels (true) and differenced specifications in a first-order VAR with $\rho = 0.98$, $\gamma = -1$ ($\delta = 0.02$) and $\Sigma = \begin{pmatrix} 1 & -0.2 \\ -0.2 & 0.8 \end{pmatrix}$.

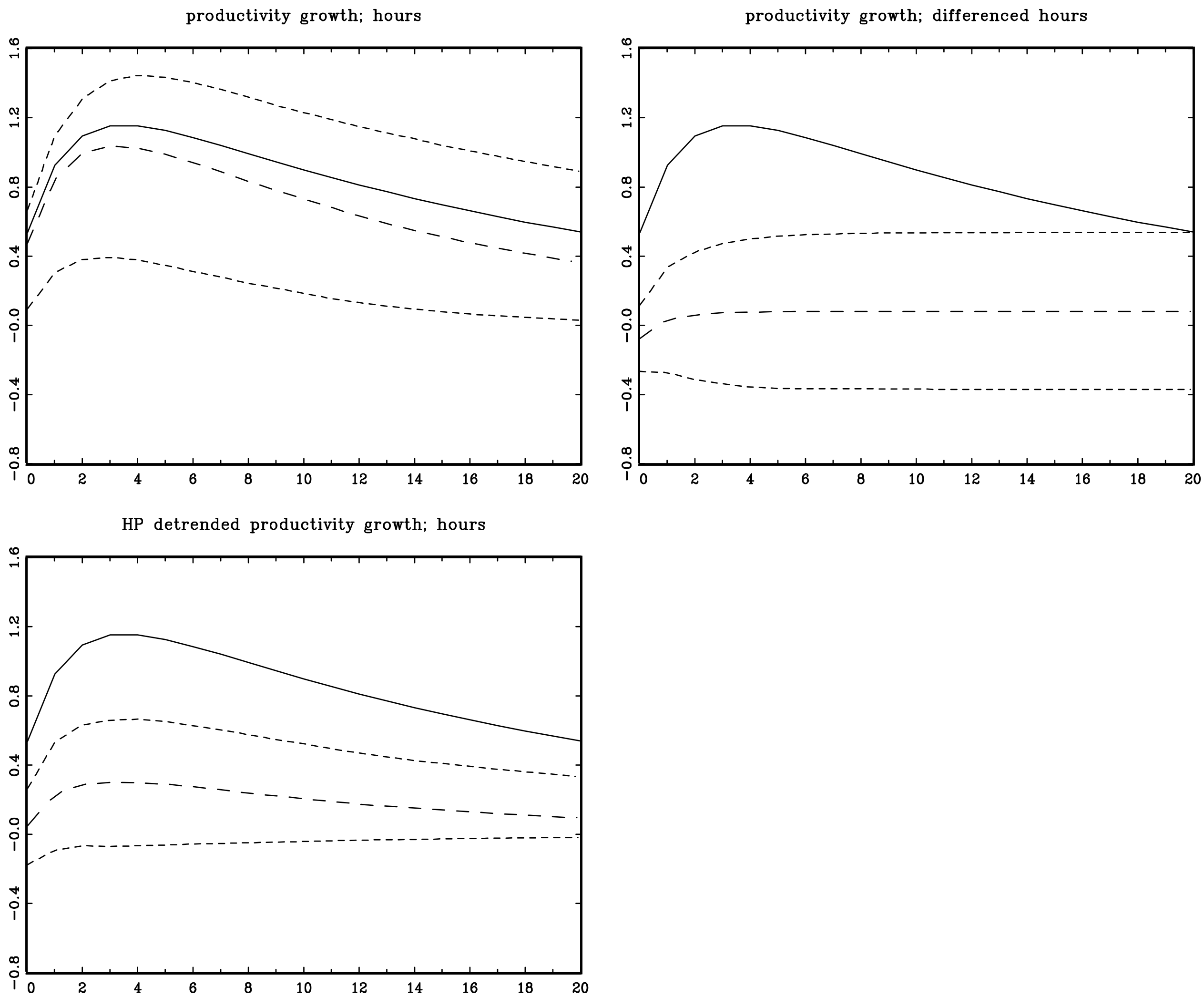

FIGURE 4. Response of hours to a positive technology shock (long-run identification) with data simulated from the model $\left[I - \begin{pmatrix} -0.05 & -0.08 \\ 0.2 & 0.55 \end{pmatrix} L\right]\left[I - \begin{pmatrix} 1 & \delta \\ 0 & \rho \end{pmatrix} L\right]\begin{pmatrix} l_t \\ h_t \end{pmatrix} = \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}$, where $\rho = 0.95$, $\delta = 0.04$ $(\gamma = -0.8)$, $(u_{1,t}, u_{2,t})' \sim iidN(0, \Sigma)$, $\Sigma = \begin{pmatrix} 0.78 & 0 \\ 0 & 0.55 \end{pmatrix}$ and $T = 250$. Solid line: true IRF; long dashes: median Monte Carlo IRF estimate; short dashes: 95% Monte Carlo confidence bands.

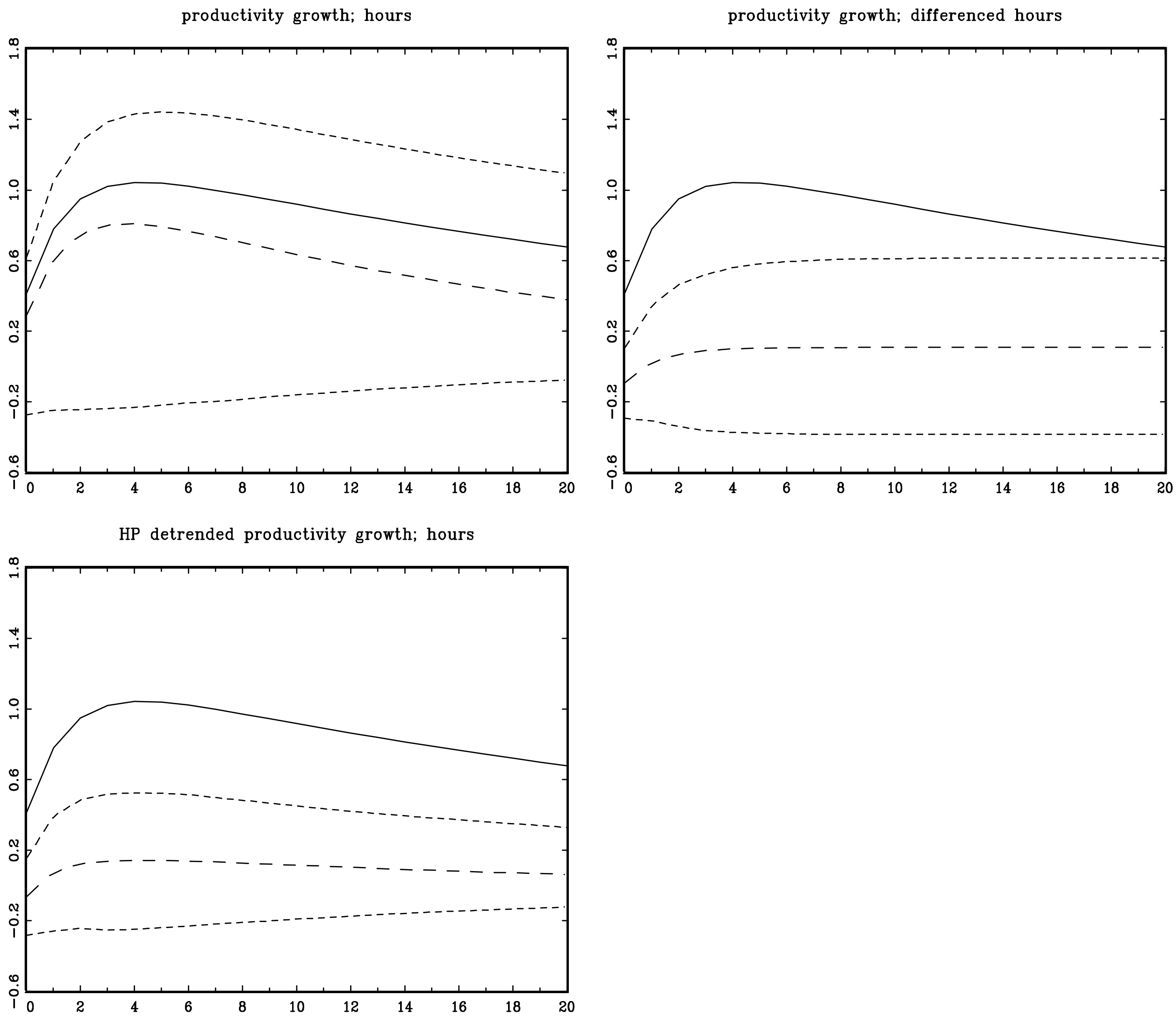

FIGURE 5. Response of hours to a positive technology shock (long-run identification) with data simulated from the model $\left[I-\begin{pmatrix}-0.05 & -0.08\\ 0.2 & 0.55\end{pmatrix}L\right]\left[I-\begin{pmatrix}1 & \delta\\ 0 & \rho\end{pmatrix}L\right]\begin{pmatrix}l_t\\ h_t\end{pmatrix}=\begin{pmatrix}u_{1,t}\\ u_{2,t}\end{pmatrix}$, where $\rho=0.97$, $\delta=0.015$ ($\gamma=-0.5$), $(u_{1,t},u_{2,t})'\sim iidN(0,\Sigma)$, $\Sigma=\begin{pmatrix}0.78 & 0\\ 0 & 0.55\end{pmatrix}$ and $T=250$. Solid line: true IRF; long dashes: median Monte Carlo IRF estimate; short dashes: 95% Monte Carlo confidence bands.

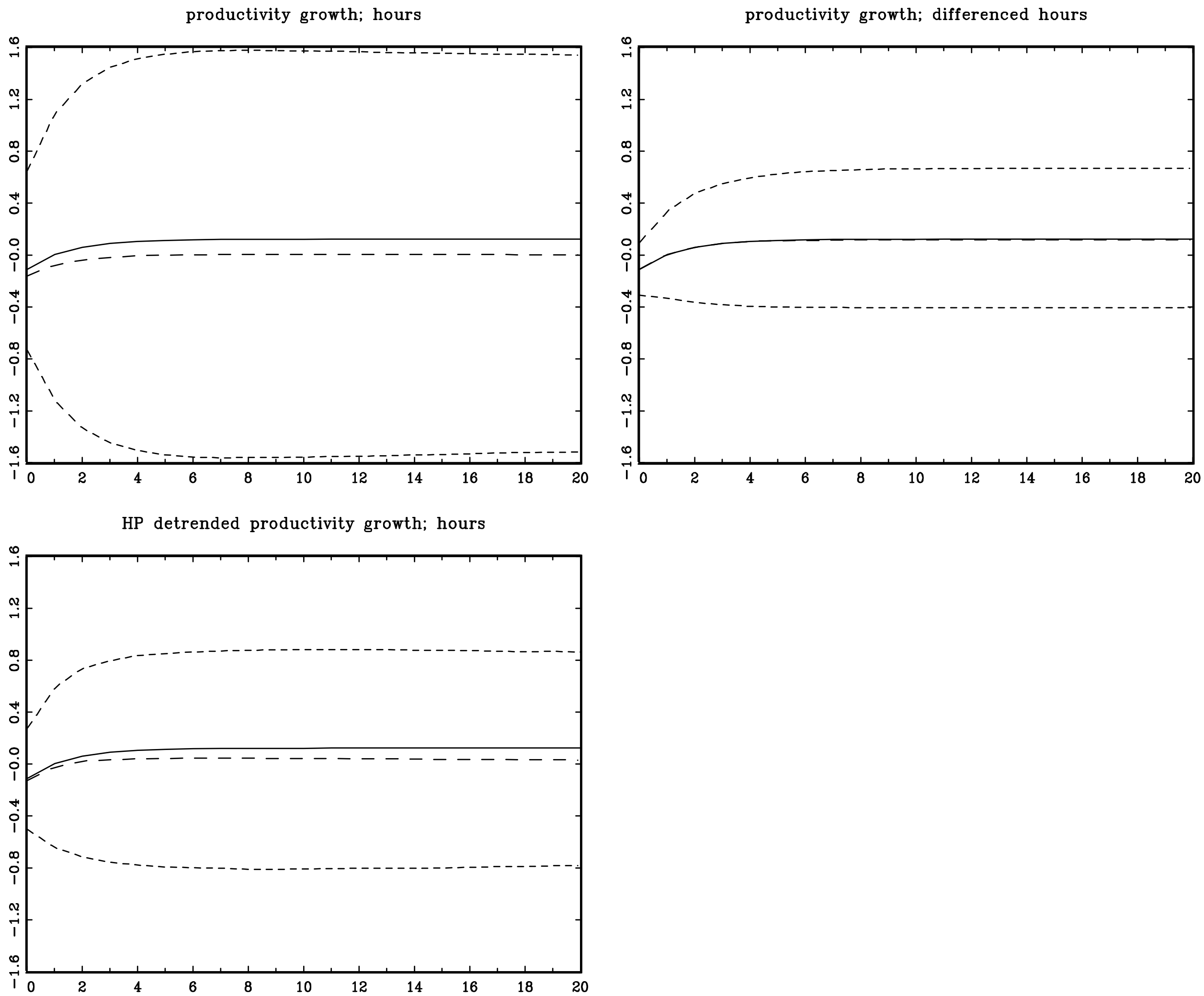

FIGURE 6. Response of hours to a positive technology shock (long-run identification) with data simulated from the model $\left[I - \begin{pmatrix} -0.05 & -0.08 \\ 0.2 & 0.55 \end{pmatrix} L\right]\left[I - \begin{pmatrix} 1 & \delta \\ 0 & \rho \end{pmatrix} L\right]\begin{pmatrix} l_t \\ h_t \end{pmatrix} = \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}$, where $\rho = 1$, $\delta = 0$ $(\gamma = 0)$, $(u_{1,t}, u_{2,t})' \sim iidN(0, \Sigma)$, $\Sigma = \begin{pmatrix} 0.78 & 0 \\ 0 & 0.55 \end{pmatrix}$ and $T = 250$. Solid line: true IRF; long dashes: median Monte Carlo IRF estimate; short dashes: 95% Monte Carlo confidence bands.

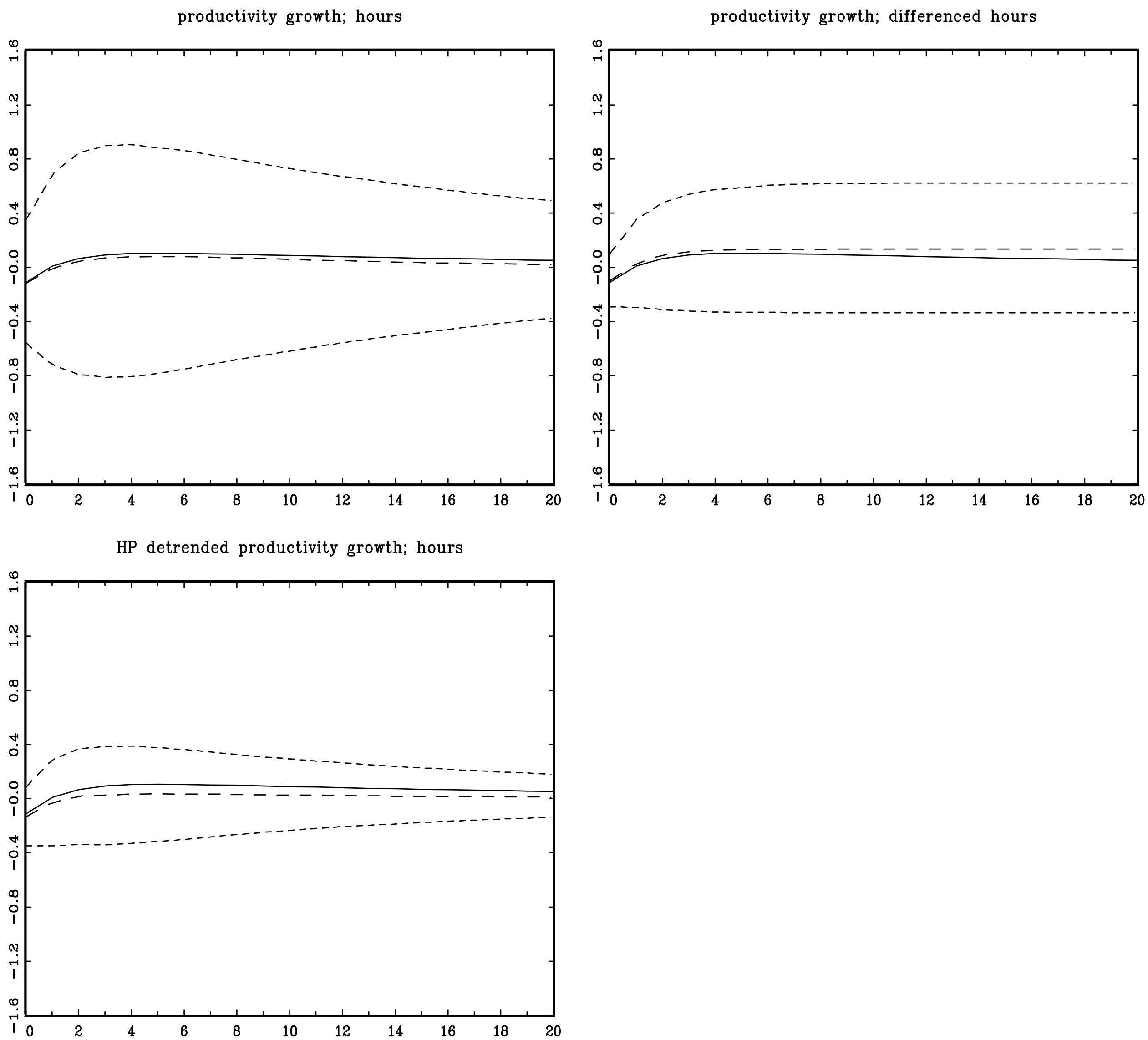

FIGURE 7. Response of hours to a positive technology shock (long-run identification) with data simulated from the model $\left[I - \begin{pmatrix} -0.05 & -0.08 \\ 0.2 & 0.55 \end{pmatrix} L\right]\left[I - \begin{pmatrix} 1 & \delta \\ 0 & \rho \end{pmatrix} L\right]\begin{pmatrix} l_t \\ h_t \end{pmatrix} = \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}$, where $\rho = 0.95$, $\delta = 0$ $(\gamma = 0)$, $(u_{1,t}, u_{2,t})' \sim iidN(0, \Sigma)$, $\Sigma = \begin{pmatrix} 0.78 & 0 \\ 0 & 0.55 \end{pmatrix}$ and $T = 250$. Solid line: true IRF; long dashes: median Monte Carlo IRF estimate; short dashes: 95% Monte Carlo confidence bands.

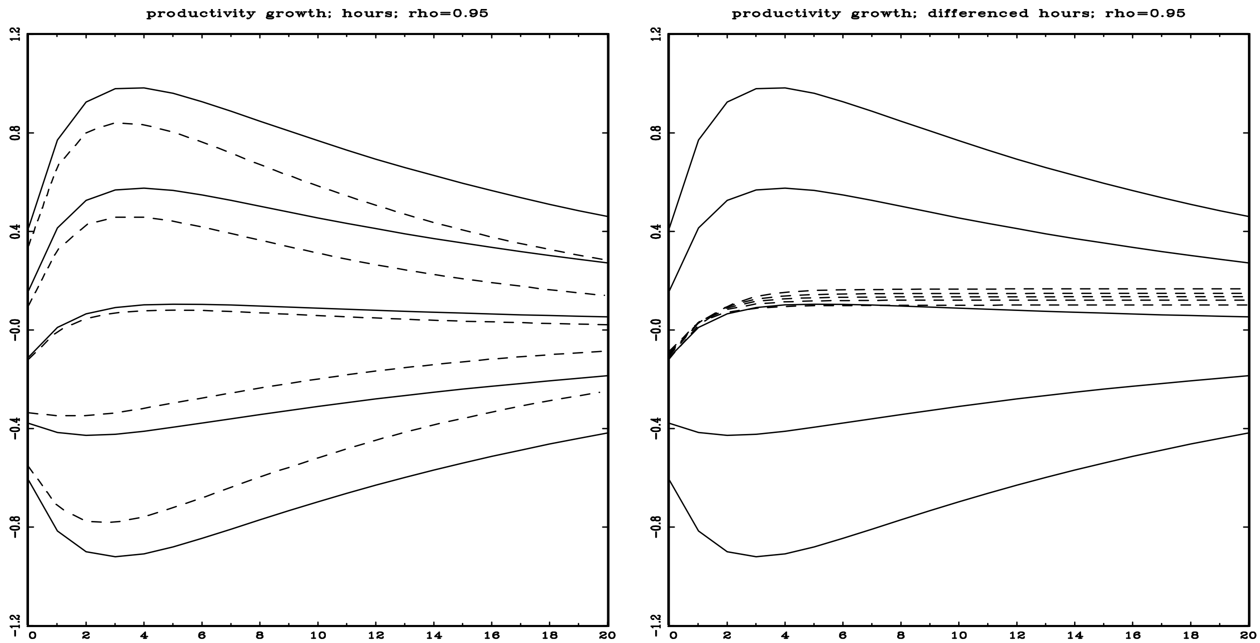

FIGURE 8. Response of hours to a positive technology shock (long-run identification) with data simulated from the model $\left[I - \begin{pmatrix} -0.05 & -0.08 \\ 0.2 & 0.55 \end{pmatrix} L\right]\left[I - \begin{pmatrix} 1 & -\gamma(1-\rho) \\ 0 & \rho \end{pmatrix} L\right]\begin{pmatrix} l_t \\ h_t \end{pmatrix} = \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}$, where $\rho = 0.95$, $\gamma = \{-0.5, -0.2, 0, 0.2, 0.5\}$, $(u_{1,t}, u_{2,t})' \sim iidN(0, \Sigma)$, $\Sigma = \begin{pmatrix} 0.78 & 0 \\ 0 & 0.55 \end{pmatrix}$ and $T = 250$. Solid line: true IRF; short dashes: median Monte Carlo IRF estimate.

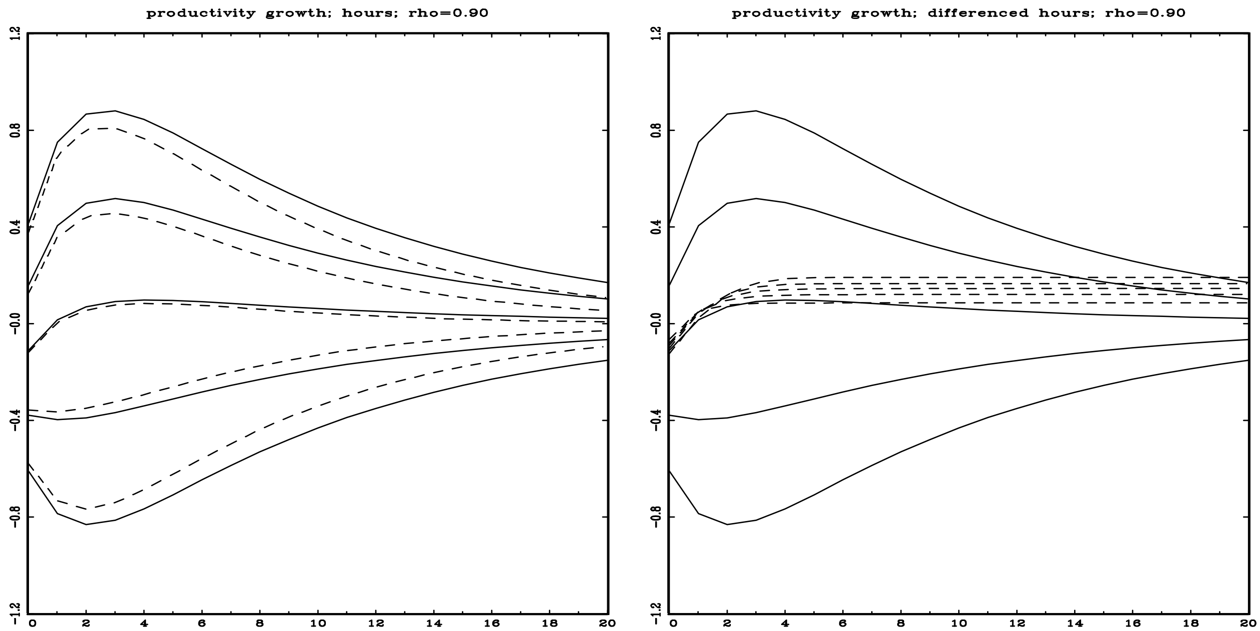

FIGURE 9. Response of hours to a positive technology shock (long-run identification) with data simulated from the model $\left[I - \begin{pmatrix} -0.05 & -0.08 \\ 0.2 & 0.55 \end{pmatrix} L\right]\left[I - \begin{pmatrix} 1 & -\gamma(1-\rho) \\ 0 & \rho \end{pmatrix} L\right]\begin{pmatrix} l_t \\ h_t \end{pmatrix} = \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}$, where $\rho = 0.90$, $\gamma = \{-0.5, -0.2, 0, 0.2, 0.5\}$, $(u_{1,t}, u_{2,t})' \sim iidN(0, \Sigma)$, $\Sigma = \begin{pmatrix} 0.78 & 0 \\ 0 & 0.55 \end{pmatrix}$ and $T = 250$. Solid line: true IRF; short dashes: median Monte Carlo IRF estimate.

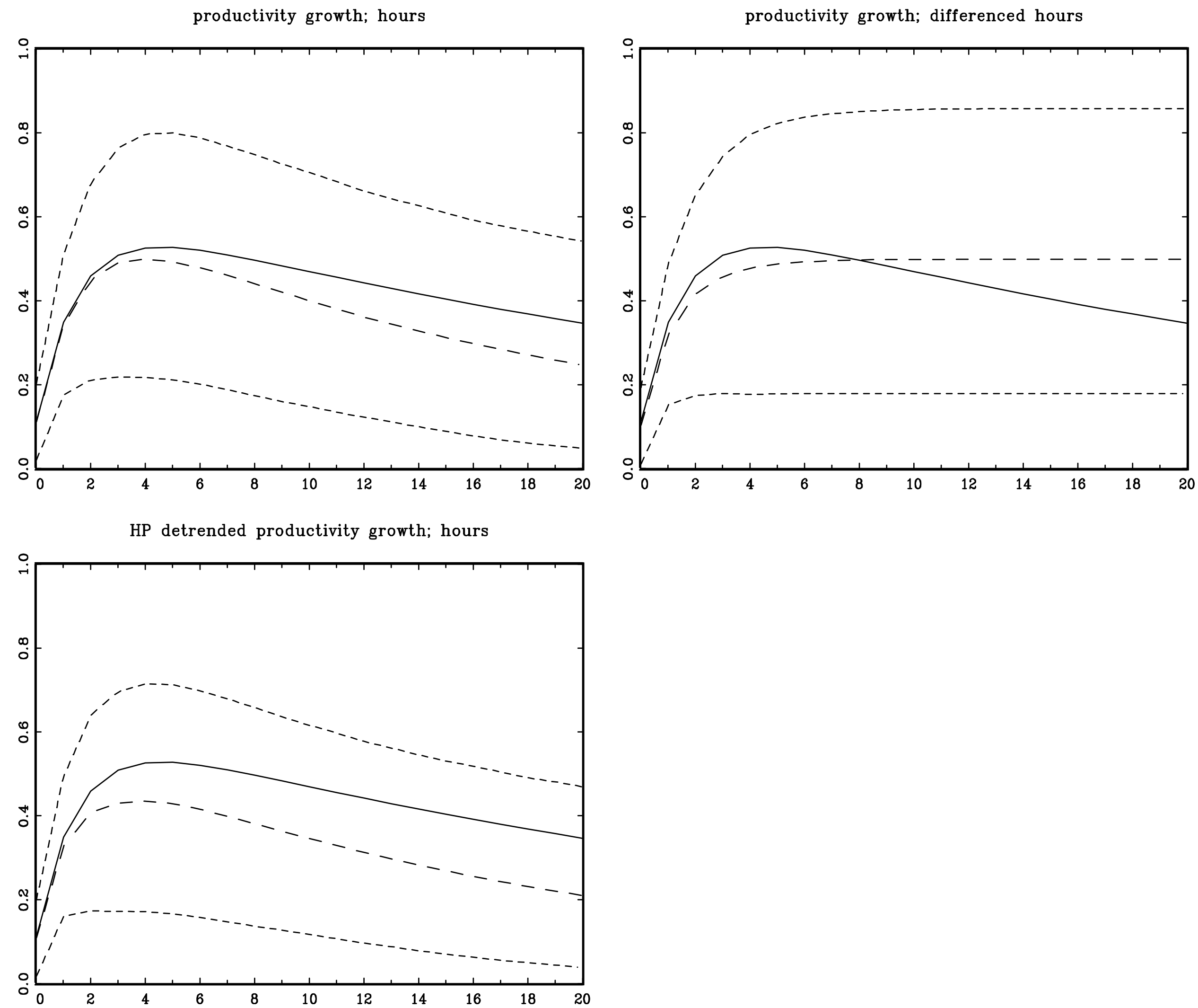

FIGURE 10. Response of hours to a positive technology shock (short-run (Choleski) identification) with data simulated from the model $\left[I - \begin{pmatrix} -0.05 & -0.08 \\ 0.2 & 0.55 \end{pmatrix} L\right] \left[I - \begin{pmatrix} 1 & \delta \\ 0 & \rho \end{pmatrix} L\right] \begin{pmatrix} l_t \\ h_t \end{pmatrix} = \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}$, where $\rho = 0.97$, $\delta = 0.015$ ($\gamma = -0.5$), $(u_{1,t}, u_{2,t})' \sim iidN(0, \Sigma)$, $\Sigma = \begin{pmatrix} 0.78 & 0.1 \\ 0.1 & 0.55 \end{pmatrix}$ and $T = 250$. Solid line: true IRF; long dashes: median Monte Carlo IRF estimate; short dashes: 95% Monte Carlo confidence bands.