Do Technology Shocks Drive Hours Up or Down? A Little Evidence From an Agnostic Procedure

Elena Pesavento
Emory University

Barbara Rossi
Duke University

November 2003

Abstract. This paper analyzes the robustness of the estimate of a positive productivity shock on hours to the presence of a possible unit root in hours. Estimations in levels or in first differences provide opposite conclusions. We rely on an agnostic procedure in which the researcher does not have to choose between a specification in levels or in first differences. We find that a positive productivity shock has a negative effect on hours, as in Francis and Ramey (2001), but the effect is much more short-lived, and disappears after two quarters. The effect becomes positive at business cycle frequencies, as in Christiano et al. (2003).

Keywords: Technology shocks, persistence, impulse response functions, Real Business Cycle Theory.

JEL Classification: C32, C12, F40.

We thank Robert Vigfusson and Neville Francis for providing codes to replicate their results, Robert Chirinko and seminar participants at Emory for comments. We also thank Pedro Duarte and Viktor Todorov for research assistance in the early stages of the project. All mistakes are ours.

Corresponding author: Elena Pesavento, Department of Economics, Emory University, Emory GA30322, USA. Phone: (404) 712 9297. E-mail: epesave@emory.edu.

1. Introduction

According to Real Business Cycles models, hours worked should rise after a positive permanent shock to technology. However, the empirical validity of this theoretical implication has been questioned in the recent literature. For example, Gali (1999) identifies technology shocks as the only shocks that have an effect on labor productivity in the long run, and estimates a persistent decline of hours in response to a positive technology shock. As Gali (1999) points out, this result is more consistent with the predictions of a New Keynesian model than those of standard Real Business Cycle models. Other papers have reached similar conclusions (see for example Shea (1999) and Francis and Ramey (2001)), which spurred a line of research aimed at developing general equilibrium models that can account for this empirical finding.

In a recent paper, Christiano Eichenbaum and Vigfusson (2003) challenge these empirical results. Using the same identifying assumption as Gali (1999), Christiano et al. (2003) find evidence that a positive technology shock drives hours worked up, not down. It seems that the estimated effects of technology shocks crucially depend on whether the empirical analysis is specified in levels or in differences. In fact, Gali (1999), Shea (1999) and Francis and Ramey (2001) specify hours in *first differences* and report that hours worked *fall* after a positive technology shock. On the other hand, Christiano et al. (2003) use hours in *levels* and report that hours worked *increase*. In Christiano et al. (2003) words: "the difference must be due to different maintained assumptions. As it turns out, a key culprit is how we treat hours worked".

Whether hours worked is a stationary or an exactly integrated process is then a key assumption in the current debate on the effects of technology shocks on business cycles. However, it is practically difficult to choose between specifications in levels or in first differences on the basis of unit root tests, because of their low power. Pesavento and Rossi (2003) show that, in the presence of a root close to unity, impulse response function estimates and confidence bands that rely on unit root pretests have bad small sample properties (in terms of median unbiasedness and coverage rates). Impulse responses based on VARs estimated in levels or first differences have bad coverage properties as well, unless the true data generating process is not persistent (in which case levels are appropriate) or it has an exact unit root (in which case first differences are appropriate).

We provide empirical evidence based on an agnostic empirical estimation procedure proposed by Pesavento and Rossi (2003). The estimation is agnostic in that it does not impose either a unit root or stationarity. These authors show that their method is robust to the presence of highly persistent processes and, thus, it is appropriate if the researcher aims at analyzing the long run effect of technology shocks on hours worked without making assumptions on the order of integration of the series. We find that a positive productivity shock has a negative impact effect on hours worked, but this effect disappears more quickly than in Francis and Ramey (after only 2 quarters), and it becomes quickly positive.

2. Empirical results

We use the same data as in Christiano et al. (2003), where per capita hours are measured as the natural logarithm of hours worked in the business sector divided by a measure of the population. Productivity is measured as the natural logarithm of output per hour in the business sector. Data are quarterly observations from 1948:1 to 2001:4 and are ultimately taken from the DRI Economics Database.[1] As in the previous literature we identify innovations to technology as the only shocks that have a permanent effect on the level of labor productivity. Figures 1 and 2 report the 90% confidence intervals and the estimated responses of per capita hours to a one standard deviation positive shock to productivity by using either a VAR in differences or a VAR in levels.[2] Results from the VAR estimated *in differences* (Figure 1) are very similar to the results in Gali' (1999) and Francis and Ramey (2001): hours worked show a negative and persistent response to a technology shock in the short run. According to point estimates, the negative effect persists for one year (4 quarters). Eventually, the effect becomes positive in the long run (although not significantly different from zero). When the VAR is estimated by using hours *in levels*, our results indicate that the initial response of hours is positive, although not significantly different from zero. The response is positive and statistically significant after one quarter, and for roughly twenty quarters.

INSERT FIGURES 1 AND 2

Table 1 shows that indeed hours are a persistent process. According to Stock (1991) method, the largest root is between 0.91 and 1.01, with a median equal to 0.96. With such a persistent process is not surprising that the Augmented Dickey Fuller test is not able to reject a unit root.

Table 1: **Unit root tests on per capita hours.**

	Hours
Augmented Dickey Fuller Test	-2.068
DF-GLS test	-0.446
Root, median estimate	0.96
Root, 95% confidence interval	(0.91, 1.01)
Number of Lags	4

[1] The mnemonics for business labor productivity, business hours and the civilian population over the age of 16 are, respectively: LBOUT, LBMN and P16. We thank Christiano et al. for the data.

[2] The IRF are multiplied by 100 so a value of 0.10 correspond to a response of 0.10%. Following the cited literature, we include a constant, but not a time trend. We focus on a bivariate VAR with hour worked and the productivity measure. As in Francis and Ramey (2001) and in Christiano et al. (2003), we do not expect our results to change if we include additional variables. We use 4 lags (chosen by the BIC criterion) in order to compare our results directly to Francis and Ramey (2001) and Christiano et al. (2003). Results are robust to different lags (e.g. 1 to 6) if we use quasi-differences to estimate the short-run dynamics.

Given that unit root tests do not strongly support the presence of a unit root, it may not be desirable to take a stand on whether the process has a unit root or not. Kilian and Chang (2000) and Pesavento and Rossi (2003) show that in the presence of large roots the coverage rates of confidence intervals for impulse response functions constructed from VARs in first differences or levels can be very bad in finite samples. The intuition is that a model that imposes a root equal to one when one of the variable is not I(1) is mis-specified. On the other hand, in small samples, a model in levels underestimates the largest root and the persistence of shocks. These apparently small mistakes and biases become extremely important at long horizons, where the difference between stationary and non-stationary processes becomes more and more important. As a result, VARs in levels and first differences have a very small probability of containing the true impulse response function, almost zero. Unit root pretests do not solve the problem, as the actual coverage of impulse response bands obtained after a pretest can be quite different from the nominal one (due to the low power of unit root tests). Furthermore, even if the tests reject a unit root, asymptotic approximations that rely on highly persistent regressors may provide better approximations in small samples. Thus we use Pesavento and Rossi (2003) "*agnostic method*" to estimate median unbiased impulse response functions and their confidence bands, which does not require the researcher to choose between the two specifications. The method proposes to construct confidence intervals for the largest root by inverting a unit root test, as originally proposed by Stock (1991). The uncertainty on the largest root is then combined with the confidence intervals for the parameters related to the short run dynamics to obtain confidence intervals for the impulse responses. Details are reported in the Appendix. Pesavento and Rossi (2003) show that this method has a coverage that is close to the nominal one at both short and long horizons in situations in which the largest root is close to one.[3]

Figure 3 reports results for the "agnostic method". It shows a negative and very short-lived impact effect. The negative effect lasts only two quarters, less than in Francis and Ramey (2001). At business cycle frequencies, the median point estimate of the impulse responses is positive, although not significantly different from zero. The confidence bands show the that effect is very likely to be positive at long horizons and at business cycle frequencies (between 6 quarters and 8 years). Comparing our median unbiased estimate of the response with that of VARs in differences, we find some evidence that the medium and long horizon effect is more positive and slightly larger in magnitude. On the other hand, the effect that we estimate is also more persistent than that obtained from VARs in levels. Finally, for comparison, Figure 4 reports

[3] Pesavento and Rossi (2003) investigate a variety of methods, all of which have good coverage. These methods build on the inversion of the following test statistics: ADF as in Stock (1991), Elliott, Rothemberg and Stock (1995), Elliott and Stock (2001), Elliott and Jansson (2001) and Elliott, Jansson and Pesavento (2003). While we report results based on ADF only, our results are qualitatively robust to the use of the other methods mentioned above.

results obtained by using Wright-style (2000) methods.[4] The results are similar, except that the confidence bands are larger (as Wright method is conservative).

INSERT FIGURES 3 TO 4

3. Conclusions

This paper analyzed the robustness of the estimate of the effect of a positive productivity shock on hours worked to the presence of a possible unit root in hours. While the literature focused on the cases in which hours are estimated either in levels or in first differences, we rely on an agnostic procedure in which the researcher does not have to choose between the two specifications. We found that a positive productivity shock has a negative impact effect on hours, as in Francis and Ramey (2001), but the effect is much more short-lived than previously found, and disappears after only two quarters. The effect then becomes positive at business cycle frequencies, as in Christiano et al. (2003).

Our empirical evidence extends the results in Christiano et al. (2003) in an important and crucial way. In their framework, the level specification implies that the first difference specification is mis-specified while the difference specification implies that the level specification is correctly specified. The latter follows from the fact that the level VAR allows for a unit root. While this is true at very short horizons, this does not need to hold at horizons that are large relative to the sample size, where the possibly downward biased estimate of the root becomes extremely important. The importance of these biases depends on the economic problem at hand and on the particular parameters that the researcher faces. Our results show that neglecting this effect may lead to very different economic results in measuring the effects of productivity shocks.

[4] The method originally proposed by Wright (2000) is univariate. We apply a method which is in spirit very much similar to his, but it is extended to a multivariate VAR with one large root.

Figure 1. Estimation in differences

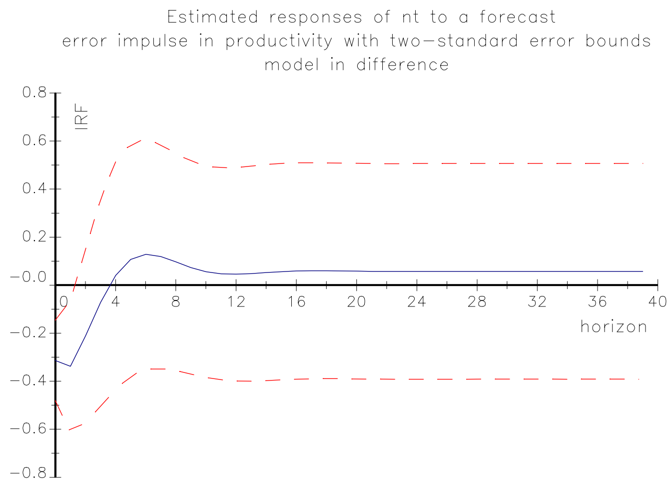
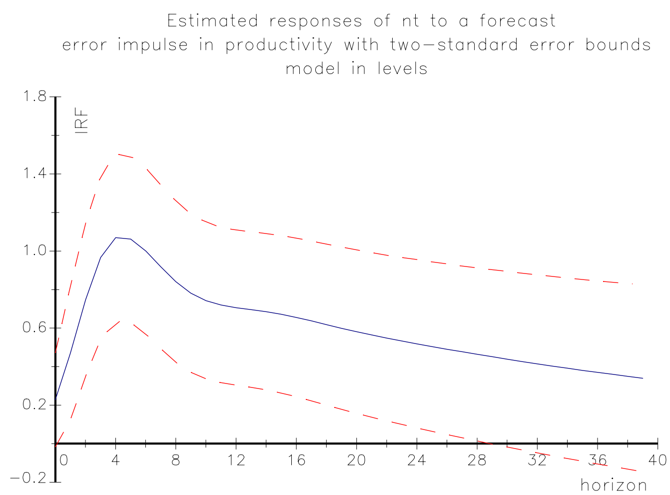


Figure 2. Estimation in levels



Figures 1 and 2 show the estimated IRF (solid line) and IRF confidence bands (dotted line) of hours worked to a one percent standard deviation increase in the productivity shock. The model is a VAR in differences in Figure 1 and in levels in Figure 2.

Figure 3. Agnostic estimation

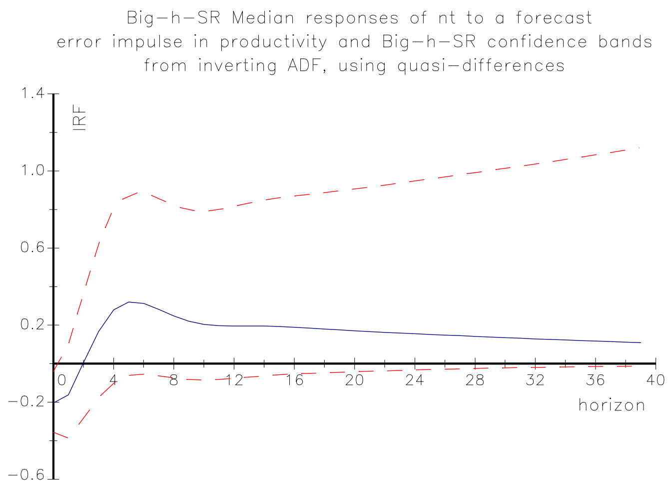
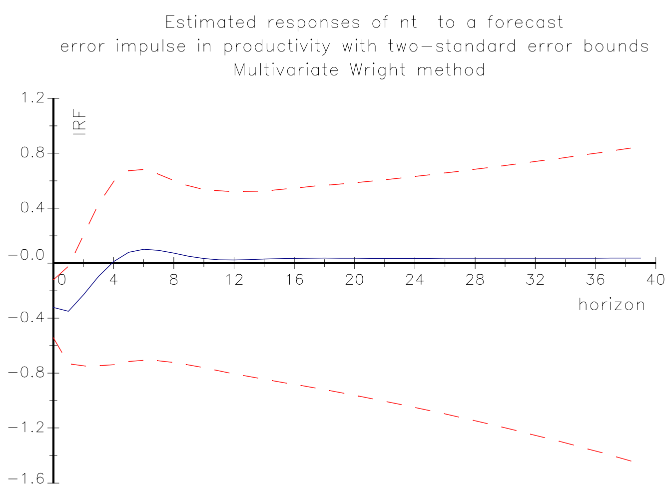


Figure 4. Wright method



Figures 3 shows the estimated IRF confidence bands of hours worked to a one percent standard deviation increase in the productivity shock. Results based on Pesavento and Rossi (2003) method robust at short horizons. Figure 4 shows the median unbiased estimate and the confidence bands that we obtain by applying a method similar to Wright (2000).

4. References

Christiano, L. M. Eichenbaum and R. Vigfusson (2003): "What Happens After a Technology Shock?", *mimeo*, Northwestern University.

Elliott, G. and M. Jansson (2003): "Testing for Unit Roots with Stationary Covariates", *Journal of Econometrics*, forthcoming.

Elliott, G. and J. H. Stock (2001): "Confidence Intervals for Autoregressive Coefficients Near One ", *Journal of Econometrics*, 103, pp. 155-181.

Elliott, G., M. Jansson and E. Pesavento (2002): "Optimal Power for Testing Potential Cointegrating Vectors with Known Parameters for Nonstationarity", *mimeo.*

Elliott, G., Thomas J. Rothenberg and J. H. Stock (1996): "Efficient Tests for an Autoregressive Unit Root", *Econometrica*, 64, pp. 813-836.

Francis, N. and V. Ramey (2001), "Is the Technology-Driven Real Business Cycle Hypothesis Dead? Shocks and Aggregate Fluctuations Revisited", *mimeo*, UCSD.

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

Kilian, Lutz, and Paolo-Li Chang, "How accurate are confidence intervals for impulse responses in large VAR models?", *Economics Letters* 69, 2000, pp. 299-307.

Ng, S. and P. Perron (2001), "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power", *Econometrica 69(6),* pp. 1519-54.

Pesavento, E. and B. Rossi (2003), "Small Sample Confidence Intervals for Multivariate Impulse Response Functions at Long Horizons", *Duke University Working Paper #03-19.*

Shea, J. (1999), "What Do Technology Shocks Do?", in B. Bernanke and J. Rotemberg, eds., *NBER Macroeconomics Annual,* Cambridge, MA: MIT Press.

Stock, J. H. (1991), "Confidence Intervals for the Largest Autoregressive Root in U.S. Macroeconomic Time Series", *Journal of Monetary Economics*, 28, pp. 435-459.

Wright, J. (2000), "Confidence Intervals for Univariate Impulse Responses with a Near Unit Root", *Journal of Business and Economic Statistics* 18(3).

Appendix

Let the data generating process (hereafter DGP) be:

$$(I - \Phi L)\, w_t = u_t \qquad t = 1, 2, ...T$$

where $w_t = \begin{bmatrix} n_t & f_t \end{bmatrix}'$ is a (2×1) vector of variables, where n_t is the log of per capita hours worked in the business sector and f_t is average labor productivity. u_t is a (2×1) stationary and ergodic moving average sequence:

$$u_t = \Theta(L)\epsilon_t \ , \tag{1}$$

ϵ_t is a martingale difference sequence with covariance matrix Σ, $\Theta(L) \equiv \sum_{i=0}^{\infty} \Theta_i L^i$, $\Theta_0 = I$, I is the (2×2) identity matrix and $\Omega^{1/2} \equiv \Theta(1)\Sigma^{1/2}$ is invertible. We assume that $\Phi = \begin{bmatrix} \rho & 0 \\ 0 & 1 \end{bmatrix}$ so that there is a unit root in f_t and estimate the VAR using the growth rate of labor productivity. We will denote by η_t the structural shocks:

$$\eta_t = A_0 \epsilon_t \tag{2}$$

such that $\eta_t = \begin{bmatrix} \eta_t^m & \eta_t^z \end{bmatrix}'$ where $\{\eta_t^z\}$ and $\{\eta_t^m\}$ denote, respectively, the sequence of technology and non-technology shocks. Following Gali (1999) we identify the technology innovation as the only shock that can have a permanent effect on productivity. This long-run identification imposes a lower triangular structure to $\Theta\,(I)\, A_0$ that allows the identification of the technology shock.

As in Pesavento and Rossi (2003), we use a local-to-unity asymptotic theory to improve the asymptotic approximation to highly persistent processes in small samples. That is, we model the real part of the (distinct) largest roots of the VAR, ρ, as local-to-unity:

$$\rho = I + \frac{1}{T} c \tag{3}$$

To obtain better asymptotic approximations to IRFs in small samples, we also assume that the lead time of the impulse response function, h, is a fixed fraction of the sample size:

$$\frac{h}{T} \underset{T \to \infty}{\to} \delta \tag{4}$$

Considering the two assumptions (3) and (4) together, we have that:

$$\rho^h \underset{T \to \infty}{\to} e^{c\delta}.$$

Pesavento Rossi (2003) show that the IRF of the effect of a technology shock, η_t^z, on n_t can be approximated by $\frac{\partial n_{t+h}}{\partial \eta_t^z} \simeq e^{c\delta} \mathbf{i}_1' \Theta\,(I)\, A_0 \mathbf{i}_2$. This provides a simple, closed-form formula for the impulse response functions at long horizons as a monotone increasing function of c. To implement the method, we need to construct a confidence interval for c. Given a confidence interval for c, confidence bands for the impulse response functions at long horizons are then obtained from equation. We construct confidence intervals for c by inverting the acceptance region of Augmented Dickey Fuller (ADF) test for a unit root in $n_t$as in Stock (1991). $\Theta(I)$ is estimated from a VAR in quasi-differences, $\left(I - \hat{\Phi}L\right) w_t$. To take correctly into account the short run dynamics at short horizons, we also add some variability to the estimation of the $\Theta' s$, as explained in Pesavento and Rossi (2003).

This paper also reports impulse response functions obtained from standard VAR using n_t both in levels and in first differences. To estimate the confidence bands in both VARs, we simulate the distribution of the impulse responses under a normality assumption with 1000 Monte Carlo replications (see Hamilton (1994) and Lutkepohl (1990) for details).