NBER WORKING PAPER SERIES

THE RESPONSE OF HOURS TO A TECHNOLOGY SHOCK: EVIDENCE BASED ON DIRECT MEASURES OF TECHNOLOGY

Lawrence J. Christiano
Martin Eichenbaum
Robert Vigfusson

Working Paper 10254
http://www.nber.org/papers/w10254

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2004

The Response of Hours to a Technology Shock:
Evidence Based on Direct Measures of Technology
Lawrence J. Christiano, Martin Eichenbaum, and Robert Vigfusson
NBER Working Paper No. 10254
January 2004
JEL No. E24, E32, O3

ABSTRACT

We investigate what happens to hours worked after a positive shock to technology, using the aggregate technology series computed in Basu, Fernald and Kimball (1999). We conclude that hours worked rise after such a shock.

Lawrence J. Christiano
Department of Economics
Northwestern University
Evanston Il 60208
and NBER
l-christiano@northwestern.edu

Martin Eichenbaum
Department of Economics
Northwestern University
Evanston Il 60208
and NBER
eich@northwestern.edu

Robert Vigfusson
Board of Governors of the Federal Reserve System
20th Street and Constitution Avenue, NW
Washington, DC 20551
robert.j.vigfusson@frb.gov

The Response of Hours to a Technology Shock:

Evidence Based on Direct Measures of Technology*

Lawrence J. Christiano, Martin Eichenbaum and Robert Vigfusson†

January 16, 2004

Abstract

We investigate what happens to hours worked after a positive shock to technology, using the aggregate technology series computed in Basu, Fernald and Kimball (1999). We conclude that hours worked rise after such a shock.

Keywords: Productivity, Long-Run Identifying Assumption, Granger-causality

JEL Codes: E24, E32, O3

1 Introduction

At least since the seminal contribution of Kydland and Prescott (1982), economists have struggled to understand the role of technology shocks in aggregate fluctuations. Stimulated by the contribution of Gali (1999), there is an important strand of the literature that uses time series techniques, coupled with minimal identifying assumptions, to estimate the dynamic response of key macroeconomic variables to these shocks. These estimates are useful for assessing the source of business cycle fluctuations, and for constructing dynamic general equilibrium models.

A key issue is how to identify shocks to technology. One approach implemented by Gali (1999), Kiley (1997) and others, proceeds indirectly by exploiting the assumption that innovations to technology are the

*The first two authors are grateful for the financial support of grants from the NSF to the National Bureau of Economic Research. The authors are also grateful to John Fernald for insightful discussions, and for the hours and technology data from Basu, Fernald and Kimball (1999). The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any person associated with the Federal Reserve System.

†Christiano: Northwestern University, email l-christiano@northwestern.edu, Eichenbaum, Northwestern University email: eich@northwestern.edu, Vigfusson, Board of Governors of the Federal Reserve System, email: robert.j.vigfusson@frb.gov

only shocks that have a long-run impact on labor productivity. This assumption is satisfied by a large class of business cycle models.[1] An alternative approach, pursued by Basu, Fernald and Kimball (1999) (BFK), estimates an innovation to technology using direct measures of technology.[2] BFK's measure is arguably the state-of-the-art in the literature that builds on Solow-residual accounting.[3] An important advantage of the BFK approach is that it does not rely on the potentially questionable assumption that the only shocks with a permanent impact on labor productivity are technology shocks. For example, the presence of persistent shocks to the capital income tax rate may distort indirect estimates of the innovation to technology, but not direct estimates.

The literature on long-run identification using labor productivity reaches conflicting conclusions about whether hours worked rise or fall after a technology shock. This conflict stems from the fact that inference is sensitive to modeling details, especially details about the treatment of the low frequency component of hours worked. For example, quadratically detrending or first differencing log, per capita hours worked typically leads to the conclusion that hours fall after a positive technology shock. Quadratically detrending all variables, or modelling per capita hours as stationary in levels typically leads to the conclusion that hours rise. Christiano, Eichenbaum and Vigfusson (2003) (CEV) apply an encompassing approach for assessing the relative plausibility of these conflicting conclusions. They find that, on balance, the evidence based on long-run identifying assumptions favors the view that hours worked rise in response to a positive technology shock.

BFK develop a measure of aggregate technology based on industry-level data. They conclude that hours worked fall after a positive technology shock. So, there is a conflict between the conclusions of BFK, and those reached in CEV (2003). The purpose of this paper is to resolve this conflict.

The two key assumptions underlying BFK's analysis are as follows. First, their measure of technology is exogenous. Second, hours worked is difference stationary. We find evidence against both these assumptions. When we replace the assumptions by alternatives that are easier to defend, we find that hours worked rise after a positive technology shock. On this basis, we conclude that the approach based on long-run identification with labor productivity and the approach based on direct measures of technology shocks give rise to similar conclusions. In addition, the results help mitigate concerns alluded to above about the possibility

[1] See for example the real business cycle models in Christiano (1988), King, Plosser, Stock and Watson (1991) and Christiano and Eichenbaum (1992) which assume that technology shocks are a difference stationary process.

[2] See also Shea (1998), who assesses technological change using data on patents.

[3] See also Burnside, Eichenbaum and Rebelo (1995) and Burnside and Eichenbaum (1996), who construct Solow residual based measures of technology correcting for labor hoarding and capacity utilization, respectively.

that long run identification based on labor productivity is confounded by non-technology shocks.[4]

We now briefly summarize our argument in more detail. BFK's exogeneity assumption implies that the one-step-ahead innovation in their measure of technology coincides with the innovation to true technology and that technology is not Granger-caused by other variables.[5] We find evidence that the level of hours worked helps forecast the growth rate of technology. There are two ways to interpret this result. One is that while true technology is exogenous, BFK's measure of technology is confounded by measurement error. The presence of measurement error naturally induces Granger-causality.[6] We think it is also likely to confound the one-step-ahead forecast errors in technology. The sort of measurement errors we have in mind are the transient, high-frequency discrepancies between true and measured outputs and inputs that occur as a result of the way the economy adjusts to shocks. Examples include labor hoarding, capacity utilization and cyclical movements in the markup.

We adopt Vigfusson (2002)'s strategy for dealing with this measurement error problem. Specifically, we replace the assumption that measured technology is exogenous with the assumption that, in the long run, true innovations to technology are the only shock that affects BFK's measure of technology. In effect, we assume that the measurement distortions in the BFK technology series are only transient. Under these circumstances, we can apply Gali (1999)'s long-run identification strategy to recover an estimate of the shock to technology from BFK's measure of technology.

[4]This conclusion is reinforced by other evidence. One potentially important non-technology shock is a permanent disturbance to the capital income tax rate. Gali (2003) shows that this tax rate is not highly correlated with estimates of the innovation to technology based on long-run restrictions and labor productivity data. Moreover, estimates of the response of macroeconomic variables to the latter shock conflict in key ways from what one would expect, if these innovations were confounded in a significant way with innovations to capital income tax rates. Consider, for example, a cut in the capital income tax rate in the simple growth model. This produces a steady state fall in the rental rate of capital and a steady state rise in the wage rate. Assuming a small, or zero income effect on leisure, the latter implies a steady state rise in labor while the former and latter together imply a rise in the capital stock. So, the cut in the capital income tax rate initially leaves the economy below steady state capital. Transient dynamics in standard models have the property that labor rises immediately, and converges to the new steady state from above. This implies an initial fall in labor productivity. This conflicts with the one finding that is common across all analyses of the response of the economy to a technology shock: labor productivity increases both in the short and the long run after such a shock. (For additional discussion of the role of capital income tax rate shocks in equilibrium models, see Uhlig (2003).)

[5]We implicitly adopt the standard assumption that agents do not observe or react to advance signals on the innovation to technology. If they did do so, then the variables that react to advance signals will Granger-cause true technology. Pursuing the implications of this sort of possibility is of substantial interest, but beyond the scope of this paper.

[6]That is, suppose the past of some variable, say x_t, is sufficient for forecast purposes. If x_{t-l}, $l > 0$, is in fact measured with error, then past values of other variables might also be useful because of their correlation with x_{t-l}.

The second interpretation of the Granger-causality finding is that there is a significant endogenous component to technology. Under these circumstances, all economic shocks in principle have an impact on technology. If this impact is permanent, then the estimated innovations to technology produced by the Vigfusson (2002) strategy confound the effects of various economic shocks. Moreover, to the extent that endogeneity causes non-technology shocks to have an immediate impact on technology, they also defeat the BFK strategy of uncovering innovations to technology from the one-step-ahead forecast error in measured technology.[7] In this paper we assume that the endogenous components of technology are not important. Investigating the robustness of our results to the presence of endogeneity in technology would be of interest, but is beyond the scope of this paper.

We now turn to BFK's second key assumption, namely, that hours worked are difference stationary. CEV (2003) report that for the sample period, 1959I-2001IV, there is evidence against this assumption. They also reject the hypothesis that per capita hours worked is difference stationary. Their evidence is based on Bruce Hansen (1995)'s covariates adjusted Dickey-Fuller test. In this paper, we present additional, complementary, evidence based on an encompassing argument, in support of the view that per capita hours should not be first differenced.

When we apply long-run identification to the BFK measure of technology and work with the level of hours worked, we find that an innovation to technology leads to a rise in hours worked. This rise is comparable to the one obtained using the long-run identification strategy. Based on these findings we conclude that inference about the response of hours worked to a technology shock is robust incorporating direct measures of technology into the analysis.

The remainder of this paper is organized as follows. Section 2 describes the response of hours to a technology shock under various assumptions. Section 3 displays evidence against the Granger-causality property of the BFK model. Section 4 argues that per capita hours worked is best modeled as a stationary process. Finally, we present concluding remarks.

[7] It may be that non-technology shocks affect technology only with a lag. If so, then BFK strategy would still be appropriate, while Vigfusson (2002)'s would not. We thank John Fernald for this observation.

2 The Response of Hours Worked to a Technology Shock Under Various Assumptions

In this section we define two models and explore their implications for the response of hours worked to a technology shock. In both cases, we work with the following bivariate, two lag vector autoregression (VAR):

$$Y_t = B_1 Y_{t-1} + B_2 Y_{t-2} + Ce_t,\ CC' = V,\ Ee_t e_t' = I,$$

where e_t denote the fundamental economic shocks:

$$e_t = \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \begin{bmatrix} \text{innovation to technology}_t \\ \text{other shock} \end{bmatrix}$$

The matrices, B_1, B_2, are estimated by ordinary least squares, while V is the variance-covariance matrix of the associated regression residuals. To determine the dynamic response of the macroeconomic variables in Y_t to e_{1t} requires knowing the elements in the first column of C. At the same time, we do not have enough information to recover C. While C has four unknown elements, $CC' = V$ represents only three independent equations. Additional restrictions ('identification assumptions') are required.

The data we use are the annual hours worked and technology series covering the period 1950 to 1989, analyzed in BFK (1999). The data are for the non-farm, private-business sector. The hours worked data are converted to per capita terms by dividing by a measure of the population.[8] Throughout, the first element of Y_t is $\Delta s_t = s_t - s_{t-1}$, where s_t denotes log of technology.

The *BFK model* is defined by two assumptions. First, the second element of Y_t is Δh_t, where h_t denotes per capita hours. This corresponds to the assumption that per capita hours worked is difference stationary. Second, we impose the assumption that s_t is exogenous with respect to hours worked. This implies that the 1,2 elements of B_1 and B_2 are zero (i.e., Δh_t does not Granger-cause technology) and the $1, 2$ element of C is set to zero (i.e., the one-step-ahead forecast error in Δs_t is proportional to e_{1t}).

The *CEV model* differs from the BFK model in two ways. First, it drops the assumption that h_t is difference stationary and defines the second element of Y_t as h_t. Second, it drops the assumption that Δs_t is exogenous. In particular, we do not restrict any element of B_1 and B_2 to be zero, and we allow the $1, 2$ element of C to be non-zero. We replace the assumption of exogeneity with the restriction that e_{2t} does not

[8] The population data are from the DRI Basic Economics database with mnemonic P16.

have a long-run impact on s_t :

$$\lim_{j \to \infty} E_t s_{t+j} - E_{t-1} s_{t+j} = f(e_{1t}). \tag{1}$$

In the CEV model, e_{1t} is estimated using the instrumental variables approach in Shapiro and Watson (1988). Then, the first column of C is estimated by regressing the VAR disturbances, Ce_t, on e_{1t}.

Figure 1 reports the response of hours to a technology shock in the two models. In each case, the gray area represents a 95 percent confidence interval.[9] Panel A displays the response implied by the BFK model. Note that hours worked drops by a little over 1 percent in the year of the shock. Hours are still down in the second year, and they hover around zero in the years after that. This pattern generally reproduces the findings in BFK, even though they work with the first difference of actual hours, while we work with per capita hours. Panel B displays the response implied by the CEV model. Note that here, hours jumps by 0.5 percent in the year of the shock and the point estimates remain positive for several years thereafter. Although there is evidence of considerable sampling uncertainty in the estimated impulse response function, note that the confidence interval clearly excludes the kind of drop implied by the BFK model.

3 BFK Technology is Granger-Caused by Hours Worked

When we test the null hypothesis that the $1, 2$ elements in B_1 and B_2 are zero in the BFK VAR, we obtain an $F-$statistic of $F_{\Delta h} = 2.39$. Using conventional sampling theory, this has a $p-$value of 10 percent, indicating little evidence against the null hypothesis. However, when the test this null hypothesis in the CEV VAR, we obtain an F statistic of $F_h = 4.66$. Conventional sampling theory implies a $p-$value of 1.6 percent, so that the null hypothesis is rejected.

Which test should we believe? If we believe the one based on the first difference of hours worked, we fail to reject the no-Granger-causality null hypothesis. If we believe the one based on the level of hours worked, we reject the hypothesis. Here, we use an encompassing approach similar to the one in CEV (2003) to assess the relative plausibility of the two specifications underlying the two inferences. As in CEV (2003), we conclude that the most plausible result is the one based on the level of hours worked.

There are at least four ways to interpret the observation, $F_h = 4.66$ and $F_{\Delta h} = 2.39$. One is that the BFK VAR is specified correctly, so that the low test statistic, $F_{\Delta h}$ =2.39, is the one sending the 'correct'

[9] The confidence intervals were computed by first simulating 1000 artificial impulse response functions. Each was obtained by fitting a VAR to artificial data obtained by bootstrap simulation of the relevant VAR. The reported reported intervals are plus and minus two standard deviation intervals.

signal. A necessary condition for this conclusion to be appealing is that the BFK VAR 'explains' the low p value associated with F_h as reflecting some sort of distortion, perhaps the inappropriate application of conventional sampling theory. Another interpretation is that the CEV VAR is correctly specified, so that it is the large test statistic, $F_h = 4.66$, that is sending the 'correct' signal. For this interpretation to be appealing, the CEV VAR must be able to explain the low value of $F_{\Delta h}$ as reflecting some sort of distortion, perhaps distortions due to first differencing. Logically, there are two other possible interpretations: the BFK VAR estimated without the restriction that the 1,2 elements of B_1 and B_2 are zero, and the CEV VAR with that same restriction imposed.

For each of the four data generating mechanisms, we simulated 1000 data sets by sampling randomly from the estimated VAR residuals, Ce_t. In each artificial data set we computed $(F_h, F_{\Delta h})$ using the same method used in the actual data. For the two data generating mechanisms that involve Δh_t, we obtain artificial time series on h_t by setting an initial condition on h_t and cumulating subsequent values of Δh_t. We find that the percent of times that $F_h > 4.66$ is only 2.3 percent. Thus, the level F statistic is too large to be consistent with the null hypothesis under the maintained hypothesis of the BFK VAR. Its magnitude is grounds for rejecting that VAR. We also reject the version of the BFK VAR which allows for Granger-causality. Using this VAR, we found that F_h is greater than 4.66, only 2.0 percent of the time.

Now consider the VAR involving the level of hours, estimated subject to the constraint that h_t does not Granger-cause Δs_t. That VAR is also rejected because the percent of times that F_h exceeds 4.66 is only 1.7. The only model that can account for the observed $(F_h, F_{\Delta h})$ is the CEV VAR. When we simulated that model, we found that the observed $(F_h, F_{\Delta h})$ lies close to the center of their distribution implied by the CEV VAR. We interpret these results as indicating that there is valuable information in the level of hours, over and above what is in the first differences, for forecasting technology growth. These results reject, at conventional significance levels, the Granger-causality assumption in the BFK model.

4 Hours Worked Should Not Be Differenced

Based on the results of the previous section, we drop the restriction in the BFK model that hours do not Granger-cause technology growth. In addition, we identify the innovation to technology using the identification condition, (1). We call the resulting model, B_1, B_2 and C, the 'difference VAR'. We refer to the CEV model as the 'level VAR'. The only difference between the difference and level VAR's has to do with the treatment of hours worked.

To see what these models imply for the response of hours worked to a technology shock, consider Figure 2. The first graph in that figure displays results for the levels case. This reproduces, for convenience, the results in Panel A of Figure 1. Panel B of Figure 2 displays results for the first difference model. Note how the drop in hours worked in the difference model is even greater than it was in BFK (see Panel B, Figure 1). The drop in the second year is now statistically significant and the point estimates indicate that hours remain low for all the years displayed. Clearly, whether one works with first differences or levels of hours has a substantial impact on the outcome of the analysis.

We now apply the encompassing analysis proposed in CEV (2003), to argue that the results based on the level of hours worked are more plausible. Before turning to the quantitative analysis, we sketch some of the relevant *a priori* considerations (for a more detailed discussion, see CEV, 2003).

4.1 *A Priori* Considerations

Specification error considerations suggest that the results based on the level VAR are more plausible. However, once sampling issues are taken into account it is less clear on a priori grounds alone which result is more likely.

If the level VAR is right, then the analysis based on first differencing hours worked entails specification error.[10] For example, suppose $h_t = \rho h_{t-1} + \varepsilon_t$. Then, $\Delta h_t = \rho \Delta h_{t-1} + \varepsilon_t - \varepsilon_{t-1}$, and Δh_t does not have a finite-ordered (or even infinite-ordered!) autoregressive representation. The conventional practice of working with finite-ordered VAR's would be misspecified in this case. Now suppose the difference VAR is correct. In this case, there is no specification error in working with levels since that simply fails to impose a true restriction. Specification error considerations alone suggest an asymmetry in the assessment of the two models. If the results based on levels and difference specifications had been similar, one should be roughly indifferent between the two specifications. But, given that the results are very different, this is consistent with the notion that the difference specification is misspecified and the level specification is closer to the truth. Although this simple specification error analysis correctly anticipates the conclusion we eventually reach, it oversimplifies.

There are sampling issues to consider too. For example, suppose the level VAR encompassed the results from the difference VAR, but at the cost of predicting large serial correlation in the fitted residuals in that VAR. This would deflate our confidence in the level VAR because the fitted difference VAR, in fact, displays

[10]By specification error we mean that the true parameter values are not contained in the econometrician's parameter space.

very little serial correlation in its residuals. There are also sampling concerns related to the difference VAR. As explained in CEV (2003), if the difference specification is true, then the Shapiro and Watson (1988) instrumental variables procedure we use for estimating the innovation to technology has a weak instrument problem. Suppose the difference VAR managed to encompass the level results, but at the cost of predicting that the analyst using the level VAR should have failed to reject the weak instrument null hypothesis. This would deflate our confidence in the difference VAR. This is because a conventional statistic for detecting weak instruments in the level data in fact rejects the weak instruments hypothesis.

4.2 Quantitative Results

We begin by asking whether the level VAR can encompass the hours response estimated for the difference VAR, and vice versa. Figure 3 displays the results. Each panel reproduces the estimated response of hours worked to a technology shock. In addition, there is a mean response predicted by the indicated DGP. The gray area indicates the associated 95 percent confidence region. DGP's were simulated using a standard bootstrap procedure, by drawing randomly with replacement from the underlying fitted VAR disturbances.

Note from Panel A in Figure 3 that the level VAR easily predicts the estimated impulse response function corresponding to the difference VAR. According to the level VAR, the true sign of the response of hours worked is positive and the negative sign estimated in the difference VAR is a consequence of specification error due to first differencing. Now consider Panel B. Note the difference VAR's counterfactual prediction that the hours response in the level VAR is negative. That is, in terms of the mean, the difference VAR does not encompass the level VAR results. This is not surprising in view of the a priori considerations discussed above. At the same time, note from the width of the gray area that the difference VAR's prediction for the level VAR's hours response is very noisy. Indeed, there is so much noise that, technically, any results including the level VAR estimates are encompassed.

We quantify the implications of the results in Figure 3 as follows. Let Q denote the event that hours rise on average in the first six periods after a shock in the level VAR, and that hours fall on average over the same period in the difference VAR. Then, bootstrap simulation implies $P(Q|\text{level VAR}) = 0.84$ and $P(Q|\text{difference VAR}) = 0.41$. Assigning an equal probability to the level and difference models being true, we conclude that the odds favor the level VAR over the difference VAR by a factor of a little over 2 to 1.

The reason the difference VAR does not do worse is because of the noisiness of its prediction for the results in the level VAR. This prediction reflects the implication of the difference model that the level analysis has a weak instrument problem.[11] When we apply a standard test to determine whether the lag log, level of hours is a good instrument for the first difference of log hours, the resulting test statistic is $F = 11.50$, which exceeds the Staiger and Stock (1997) recommended value of 10. Thus, the null hypothesis that lagged hours is a weak instrument is rejected. Interestingly, this corresponds to Bruce Hansen (1995)'s covariates adjusted Dickey Fuller test for the null hypothesis that hours worked has a unit root. This weak-intrument test in effect rejects the unit root specification on classical grounds.

To integrate the weak instrument consideration into the analysis, we add the result of the weak instrument test to the event, Q, discussed above. In particular, we add the event that the weak instrument test statistic is inside the interval defined by the actual test statistic, plus and minus unity.

The weak instrument issue raises a concern about the plausibility of the difference specification. As discussed above, there is an analogous concern related to the level specification. Recall that the level specification's ability to account for the difference specification is becaues of the level VAR's implication that the first difference specification is misspecified. One might expect this specification error to manifest itself in the form of significant serial correlation in the bivariate, two-lag difference VAR's estimated in artificial data generated by the level VAR. If so, this would be a count against the level VAR. This is because the Box-Pierce q statistic for testing the null hypothesis of no serial correlation in the fitted disturbances of the difference specification is 10.73.[12] The associated $p-$value is 0.22 using conventional sampling theory, indicating little evidence of serial correlation.

To integrate this serial correlation concern into the analysis, we add the Box-Pierce q statistic to the event, Q. We add the event that the Box-Pierce q statistic lies in an interval [9.73,11.73] defined by the actual Box-Pierce statistic, plus or minus one.

[11] In particular, in applying the Shapiro-Watson method to recover the innovation to technology, the growth in hours worked is instrumented by its level. When hours worked has a unit root, the lagged level is a 'weak instrument'. To see this, note that under the unit root hypothesis the level of hours worked is heavily influenced by shocks occuring in the distant past, while the first difference of hours worked is not. As a result, there is relatively little overlap in the shocks driving the first difference of hours and the shocks driving its level. See CEV (2003) for a detailed discussion.

[12] We do tests using the multivariate Ljung–Box portmanteau (or Q) test for white noise by Hosking (1980) that is described in Johansen (1995, 22). We use four lags in the test. The resulting degrees of freedom are 8.

Let the event of interest to our analysis be the four dimensional object, Q'. Here,

$$Q' = \begin{bmatrix} Q'_1 \\ Q'_2 \\ Q'_3 \\ Q'_4 \end{bmatrix} = \begin{bmatrix} \text{average hours implied by level VAR positive} \\ \text{average hours implied by difference VAR negative} \\ \text{weak instrument test statistic, plus and minus one} \\ \text{serial correlation test statistic, plus and minus one} \end{bmatrix}$$

Table 1 displays prob(Q'_i|level VAR) and prob(Q'_i|difference VAR). It also displays the probabilities of the joint events, prob(Q'_i, Q'_{i+1}|level VAR) and prob(Q'_i, Q'_{i+1}|difference VAR), for $i = 1, 3$, as well as prob(Q'|level VAR) and prob(Q'|difference VAR). Finally, the last column provides the posterior odds, under a uniform prior, in favor of the level specification.

Table 1: Probability of Different Events and Posterior Odds

	prob(Q'_i|level VAR)	prob(Q'_i|difference VAR)	Posterior Odds
Q_1	0.88	0.41	2.14
Q_2	0.96	0.92	1.04
Q_3	0.13	0.01	13.02
Q_4	0.18	0.17	1.04
$Q_1 \cap Q_2$	0.84	0.37	2.29
$Q_3 \cap Q_4$	0.02	0.00	9.82
$Q_1 \cap Q_3$	0.11	0.01	14.56
Q'	0.02	0.00	13.00

There are several things worth noting in the table. First, prob(Q'_3|difference VAR) is very small. This reflects, in results not displayed here, that the difference VAR substantially underpredicts the weak instruments test statistic. At the same time, prob(Q'_3|level VAR) is relatively large. As a consequence, the implied posterior odds favor the level model very strongly. Second, prob(Q'_4|level VAR) and prob(Q'_4|difference VAR) are of similar magnitude, so that the posterior odds of the two models relative to that statistic are near unity. This statistic does little to move confidence one way or the other between the alternative specifications. In results not displayed here, we found that the level VAR does not predict substantial serial correlation in the fitted residuals of the difference VAR. (Of course, this result reflects the size of our data sample. We verified that if the sample had been sufficiently large, the level VAR would have predicted a sizeable amount of serial correlation in the fitted residuals.)

The bottom line in the table is the posterior probability in favor of the level VAR, given the entire joint fact, Q'. This posterior probability is very large. We conclude that the level VAR is more plausible than the difference VAR, and on these grounds we conclude that hours worked rise in response to a positive technology shock.

5 Conclusion

In CEV (2003), we argued that long-run restrictions, in conjunction with data on labor productivity, imply that hours rise in response to a technology shock. Here, we argue that the direct evidence on technology constructed by BFK contains no reason to change that conclusion.

Although this paper emphasizes some points of difference with analyses such as those of Gali (1999, 2003) and Gali, Lopez-Salido and Valles (2003), it is useful to also note the many points of common ground. For example, Altig, Christiano, Eichenbaum and Linde (2002), and CEV (2003) find that, as in Gali (1999), shocks to disembodied technical progress account for only a small component of business fluctuations. In addition, Altig, Christiano, Eichenbaum and Linde (2002) argue, as do Gali, Lopez-Salido and Valles (2002), that monetary policy has played an important role in determining the nature of the transmission of technology shocks.

References

[1] Altig, David, Lawrence J. Christiano, Martin Eichenbaum and Jesper Linde, 2002, ‘An Estimated Dynamic, General Equilibrium Model for Monetary Policy Analysis,’ Manuscript., Northwestern University Manuscript

[2] Basu, Susanto, John G. Fernald, and Miles S. Kimball. 1999. ‘Are Technology Improvements Contractionary?’ Chicago Federal Reserve Manuscript.

[3] Burnside, C., M. Eichenbaum and Sergio Rebelo, 1995, ‘Capacity Utilization and Returns to Scale’, NBER Macroeconomics Annual 1995, Bernanke, B. S. Rotemberg, J.J., eds. NBER Macroeconomics Annual 1995. Cambridge and London: MIT Press, 67-110.

[4] Burnside, C. and M. Eichenbaum, 1996, ‘Factor Hoarding and the Propagation of Business Cycle Shocks’, *American Economic Review*, 86(5), December, 1154-74.

[5] Christiano, Lawrence. J., Martin Eichenbaum and Robert Vigfusson, 2003, ‘What Happens After A Technology Shock?’ Federal Reserve Board International Finance Discussion Papers 768.

[6] Doan, Thomas 1992. *Rats Manual* Estima Evanston, IL.

[7] Francis, Neville, and Valerie A. Ramey, 2001, ‘Is the Technology-Driven Real Business Cycle Hypothesis Dead? Shocks and Aggregate Fluctuations Revisited,’ manuscript, UCSD.

[8] Gali, Jordi, 1999, ‘Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?’ *American Economic Review*, 89(1), 249-271.

[9] Gali, Jordi, 2003, ‘On the Role of Technology Shocks as A Source of Business Cycles: Some New Evidence’, Universitat Pompeu Fabra manuscript.

[10] Gali, Jordi, J. David Lopez-Salido, and Javier Valles, 2002, ‘Technology Shocks and Monetary Policy: Assessing the Fed’s Performance’, National Bureau of Economic Research Working Paper 8768.

[11] Hansen, Bruce E., 1995, ‘Rethinking the Univariate Approach to Unit Root Testing: Using Covariates to Increase Power,’ *Econometric Theory*, December, 11(5), 1148-71

[12] Hosking, J. R. M. 1980. The multivariate portmanteau statistic. *Journal of the American Statistical Association* 75: 602–08.

[13] Johansen, S. 1995.*Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. New York: Oxford University Press.

[14] Kiley, Michael 1997. “Labor Productivity in U.S. Manufacturing: Does Sectoral Comovement Reflect Technology Shocks?” Federal Reserve Board Manuscript.

[15] Kydland, Finn and Edward Prescott, 1982, ‘Time to Build and Aggregate Fluctuations,’ *Econometrica*. 50(6), 1345-70

[16] Hamilton, James B., 1994, *Time Series Analysis*, Princeton University Press, Princeton New Jersey.

[17] Shapiro, Matthew and Mark Watson,1988, ‘Sources of Business Cycle Fluctuations,’ NBER, *Macroeconomics Annual*, 111-148.

[18] Shea, John 1998, ‘What Do Technology Shocks Do?,’ National Bureau of Economic Research Working Papers 6632

[19] Staiger, Douglas, and James Stock, 1997, ‘Instrumental Variables Regression with Weak Instruments,’ *Econometrica,* 65(3), 557-586.

[20] Uhlig, Harald, 2003, ‘Do Technology Shocks lead to a Fall in Total Hours Worked,’ manuscript, Humboldt University

[21] Vigfusson, Robert J., 2002, ‘Why Does Employment Fall After A Positive Technology Shock?,’ manuscript, Board of Governors, Federal Reserve System.

Figure 1: Dynamic Response of Per Capita Hours Worked to Innovation in Technology

Panel A: BFK Model | Panel B: CEV Model

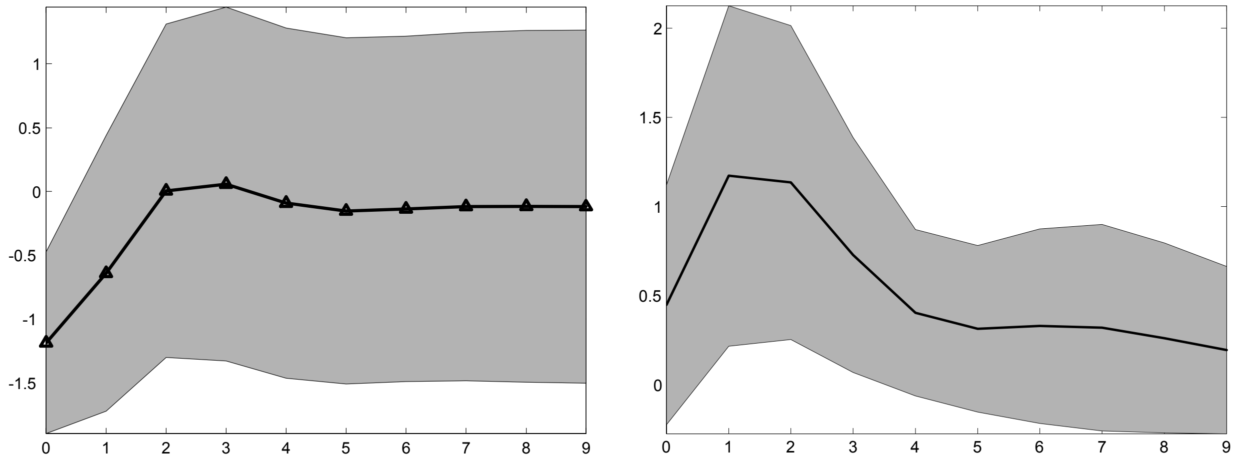


Figure 2: Response of Hours Worked to Technology: Long Run Restrictions

Panel A: Hours in Levels | Panel B: Hours in First Differences

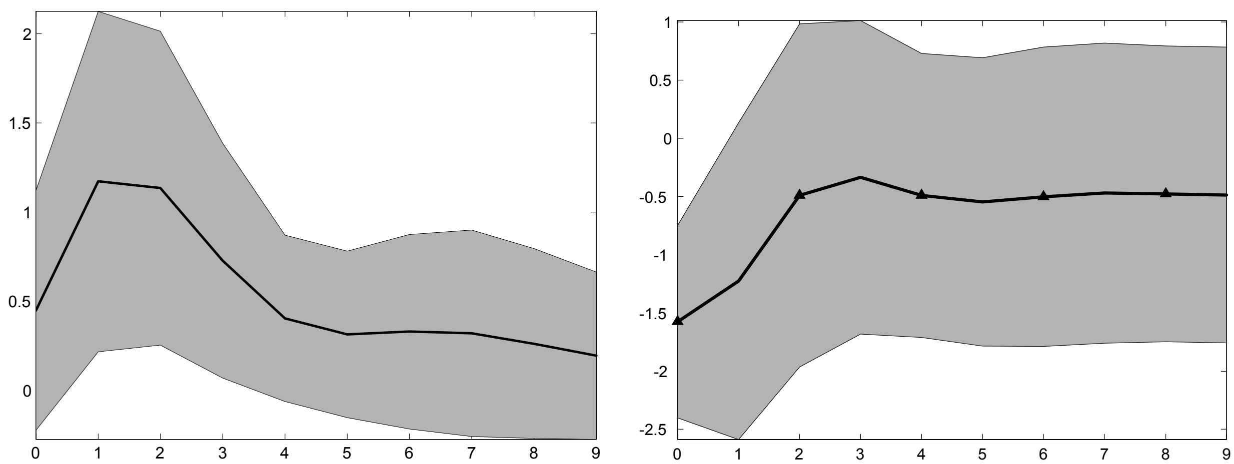
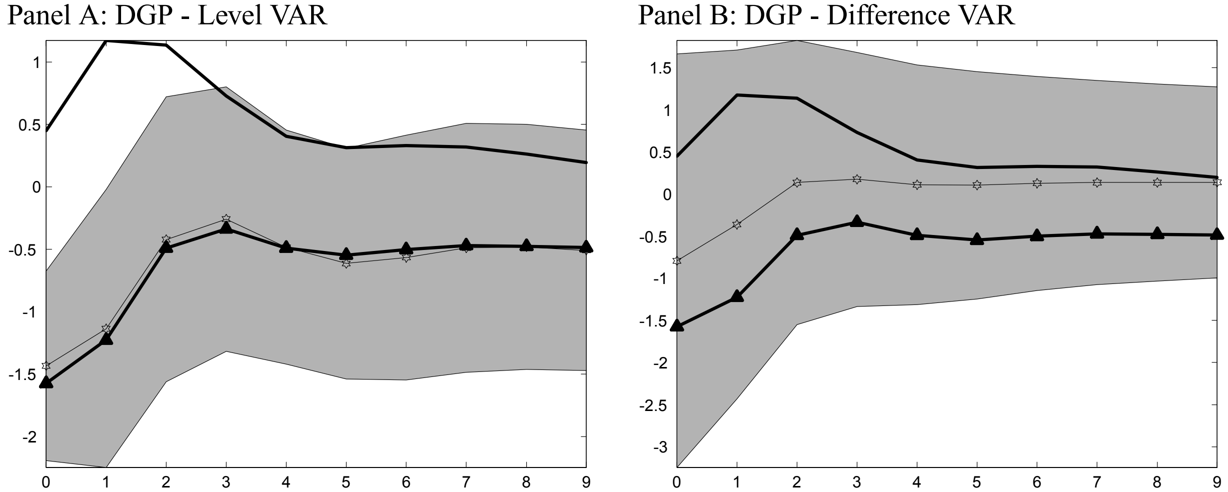


Figure 3: Evaluating the Ability of each VAR to Encompass the Hours Response of the Other

Panel A: DGP - Level VAR | Panel B: DGP - Difference VAR



Notes to Figures: Solid line - estimated response in level VAR,
Triangles - estimated response in difference VAR,
Stars - mean response implied by indicated DGP, Gray area - 95% confidence interval about mean.