The Source of Historical Economic Fluctuations: An Analysis Using Long-Run Restrictions [with Comments]
Author(s): Neville Francis, Valerie A. Ramey, Harald Uhlig, Susanto Basu
Source: *NBER International Seminar on Macroeconomics*, (2004), pp. 17-73
Published by: The University of Chicago Press
Stable URL: http://www.jstor.org/stable/40215047

1

The Source of Historical Economic Fluctuations: An Analysis using Long-Run Restrictions

Neville Francis, *University of North Carolina at Chapel Hill*
Valerie A. Ramey, *University of California, San Diego and NBER*

1. Introduction

There has been a recent surge in the number of papers studying real business cycles and the role that technology shocks plays in generating cyclical movements in macroeconomic data. Such renewed interest in technology-driven business cycles has been fueled by the finding of recent empirical studies that labor input falls, at least in the short run, in response to a positive technology shock (see Shea (1998), Galí (1999), Basu, Fernald, and Kimball (2004) and Francis and Ramey (2005)). These results have generated a good deal of discussion because they raise fundamental questions about the empirical relevance of the technology-driven real business cycle hypothesis.

The goal of this paper is to analyze the historical role played by technology shocks in the U.S. by studying the fluctuations in data extending back to the 19^{th} century. Our approach is to identify technology shocks using long-run restrictions as in Blanchard and Quah (1989) and Galí (1999). It seems particularly appropriate to use long-run restrictions for identification in truly long-run data. We carry out our analysis for the entire sample period and for two subsamples of the data, the pre- and post-WWII eras. Our subsample results are then compared to see if there have been any changes in the nature of technology shocks or in their transmission mechanism. To check the robustness of our results, we identify technology shocks using various assumptions about the source of nonstationarity in the data. The time series properties of the data are a concern at the heart of the debate concerning the validity of the standard RBC hypothesis.[1]

We construct a new series of hours per capita that adjusts for demographic and social trends. With this measure, the results for the post-WWII period are quite similar whether one assumes hours are stationary, trend stationary or nonstationary. The results differ across specifications in the early period. Only the specification that assumes a unit root in hours produces technology shocks that are not Granger-caused by monetary and government spending variables.

The preferred unit root specification gives an interesting account of the historical sources of fluctuations. Technology shocks are much more important for the forecast variance of output and hours in the early period than in the later period. The Great Depression was a time in which both types of shocks were very negative for several years. The period immediately after the Great Depression was a period of extraordinarily high positive technology shocks. Finally, the variance of both types of shocks decreased dramatically in the post-WWII period.

2. Overview of the U.S. Historical Data

This section presents an overview of the U.S. data as a preliminary step to estimation of the structural VAR. The data are annual for the time period 1889–2002. The principal variables studied are labor productivity and hours for the private business sector. Private output is constructed from these two variables. In augmented models and in additional tests, data on consumption, investment, government spending, the price level and money are also used.

Data for the early part of the sample come from John W. Kendrick's *Productivity Trends in the United States* (1961), *Historical Statistics,* Balke and Gordon (1989), and Anderson (2003). Data for the later part of the sample are obtained from the Bureau of Labor Statistics (BLS), the BEA, the *Economic Report of the President,* and the Federal Reserve. The appendix provides a detailed description of all the data and their sources.

Figure 1A shows a graph of the logarithm of output per hour in the private sector. It is difficult to distinguish the cyclical movements in output per hour because the overall upward trend is so strong. A slowdown in the rate of growth beginning in the early 1970s is apparent from the graph, however. Figure 1B shows the growth rate of output per hour. The most noteworthy feature of this graph is the difference in volatility between the pre- and post-WWII period.

The only assumption required for our identification technique is that output per hour have a unit root. As Table 1, Panel A shows, one cannot reject a unit root in labor productivity against either of the three alternative hypotheses. This result also holds in the pre- and post-WWII

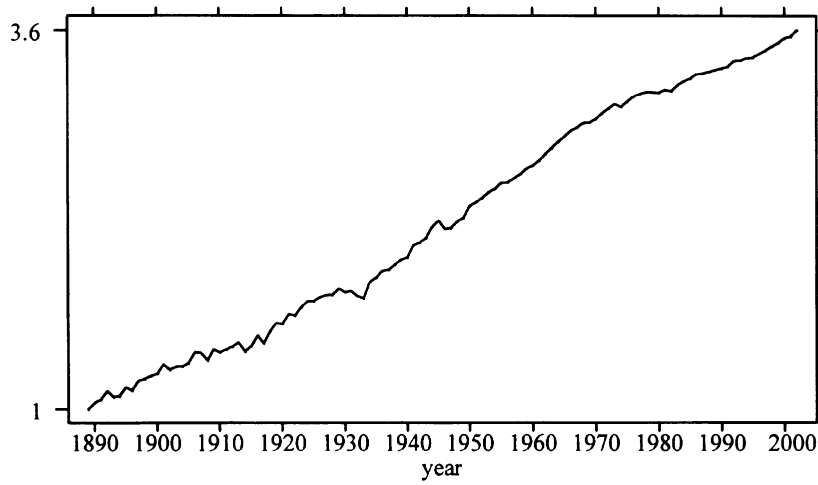


Figure 1A
Labor productivity in the private sector

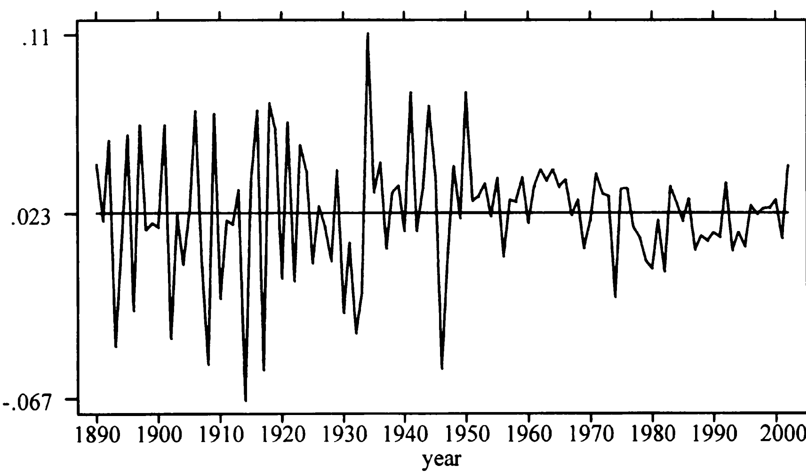


Figure 1B
Growth rate of labor productivity in the private sector

Table 1
ADF unit root tests
Private business sector

(Logarithms, lags in parenthesis were chosen optimally up to max=4)
P-values for null hypothesis of a unit root

A. Labor Productivity

	Alternative hypothesis		
Time Period	No trend	Linear trend	Quadratic trend
1889–2002	0.964 (4)	0.760 (4)	0.970 (4)
1889–1940	0.982 (4)	0.578 (2)	0.617 (2)
1948–2002	0.461 (2)	0.760 (2)	0.994 (2)

B. Private Hours Divided by Population 16+ Years

	Alternative hypothesis		
Time Period	No trend	Linear trend	Quadratic trend
1889–2002	0.625 (3)	0.512 (3)	0.353 (3)
1889–1940	0.875 (2)	0.677 (2)	0.009 (4)
1948–2002	0.183 (4)	0.825 (4)	0.270 (3)

C. Private Hours Divided by Adjusted Population

	Alternative hypothesis		
Time Period	No trend	Linear trend	Quadratic trend
1889–2002	0.057 (3)	0.127 (3)	0.108 (3)
1889–1940	0.575 (2)	0.498 (2)	0.040 (4)
1948–2002	0.420 (4)	0.266 (3)	0.460 (3)

sub-periods. On the other hand, further ADF tests (not shown) overwhelmingly reject a second unit root. Thus, the data support the key identifying assumption.

Consider now the time series properties of hours per capita. The nature of this series in the post-WWII period has been the source of much recent controversy because different assumptions can lead to different results in structural VARs (e.g., Christiano, Eichenbaum, and Vigfusson (2003), Francis and Ramey (2005), Fernald (2005) and Galí 2004)). Figure 2A plots the standard measure used—it divides total private hours by the population 16 and older. This measure of hours per

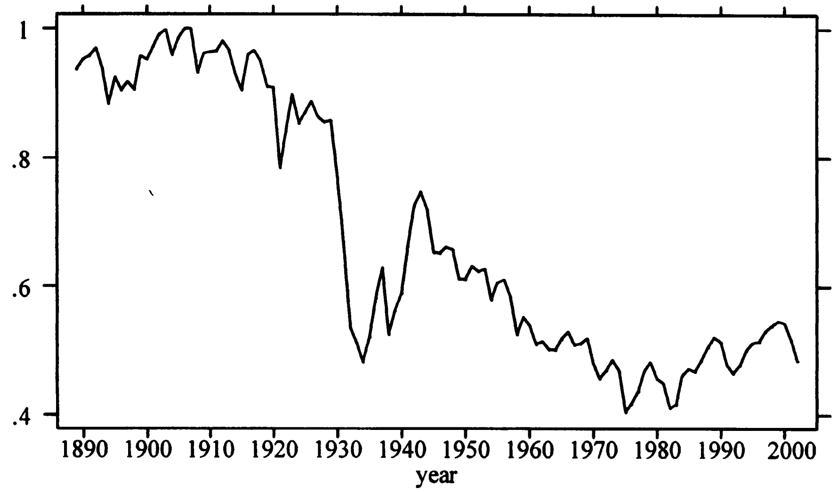


Figure 2A
Private hours divided by population 16+ years

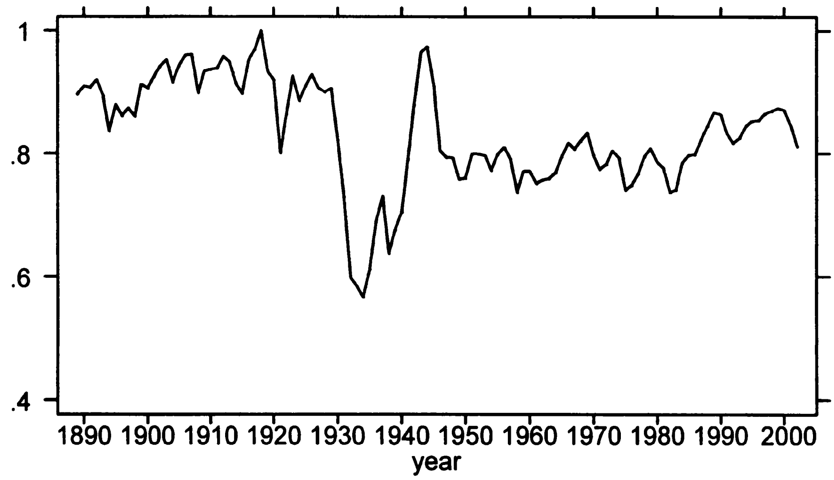


Figure 2B
Private hours divided by demographically-adjusted population

capita clearly has a trend during the 114 years of the sample, falling by 37 percent from 1889 to 2002. Table 1, Panel B shows that the only case in which one can reject a unit root is in the early sample, in favor of the alternative with a quadratic trend.

Our investigations reveal that the main source of the trend in this measure is not taxes and regulations, as some have argued (e.g., Mulligan 2002). Rather, most of the trend is due to changing demographics, school enrollments and the growth of government employment. Consider an alternative series, shown in Figure 2B. This series adjusts for three important trends over the sample: (1) the growing fraction of workers in government; (2) the changing fraction of population that is too young to work or is in school; and (3) the growing fraction of retirement age population, 65 and over. This series consists of dividing total private hours by total population minus (1) the full-time equivalent number of employees working in government; (2) the population aged 0-4; (3) the number of individuals enrolled in school; and (4) the population aged 65 and over.[2]

The demographically adjusted series in Figure 2B paints a very different picture. Average hours were still higher at the start of the 20th century than at the end, but the difference is much smaller. Hours per capita in the year 2000 were only 3.5 percent lower than they were in 1900. Thus, most of the trend in the standard series stems from slow-moving demographic, government employment, and educational trends.

Figure 3 shows graphs of the various series that were used to adjust population. The fraction of the population employed by the government shows spikes during the two world wars, but also is higher in the post-WWII period than in the pre-WWI period. The fraction of the population age 65 and older shows a steady-upward trend.[3] Both graphs on the right hand side show the effects of the baby boom, which started in the 1950s. After the baby boom graduates, the fraction of the population in school falls, but remains significantly higher in the late century than in the early century. Years spent in school have increased significantly, both because of increased government subsidies and because of the rising returns to education. The low frequency movements in these series lead to low frequency movements in unadjusted hours per capita.

Panel C of Table 1 shows the results of unit root tests on the demographically adjusted hours series. The p-value of a unit root test against the alternative of stationarity is 0.057 for the entire period. On the other

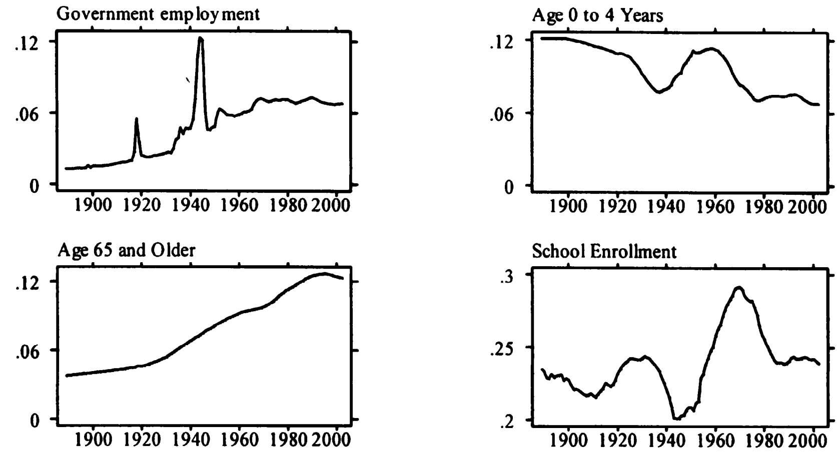


Figure 3
Sub-groups as a fraction of total population

hand, in the subperiods, one can reject a unit root only against a quadratic trend in the pre-WWII subperiod.

Since the unit root results do not give a clear picture, we initially consider three possible specifications: hours in levels, deterministic trend in hours, and a unit root in hours. The reason for looking at trend specifications is that there are clearly still some low frequency movements in our adjusted hours per capita. For example, hours per capita in the 41 years from 1889–1929 never cross the mean of hours per capita in the years 56 years 1947–2002 and vice versa. It is likely that other demographic influences that we have not measured account for the remaining low frequency movements as well. To account for the unmeasured slow-moving demographic forces, we subtract a quartic trend from our measure of hours. Since the purpose of the trend is to capture slow-moving demographic forces rather than business cycle forces, we estimate the trend omitting the large outliers from 1930–1946. Figures 4A and 4B show the trend and its effect. The fitted trend is very similar to what one would obtain using a Hodrick-Prescott filter with a very high value of the λ parameter (the parameter that penalizes changes in the estimated trend), such as 5000 rather than the standard 100 for annual data. The implied cyclical movements, shown in Figure 4B, line up well with official NBER dates. Hours are generally below trend during official recessions and above trend during expansions.

A. Quartic Trend

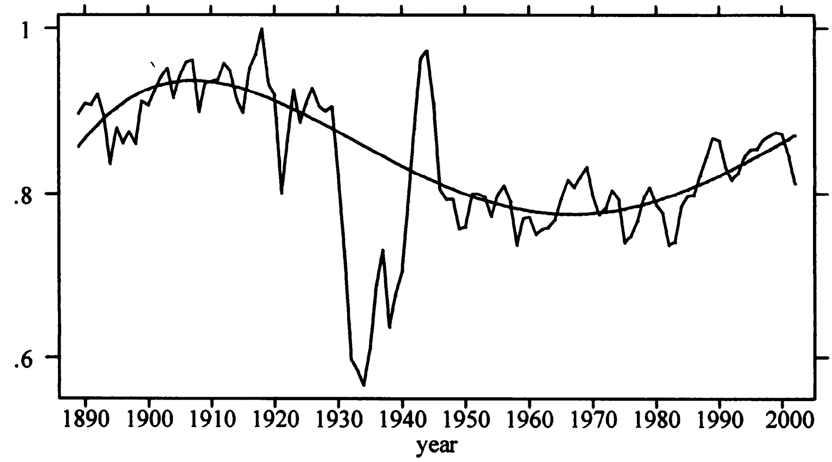


Figure 4A
Quartic detrending of demographically adjusted hours

B. Detrended Hours

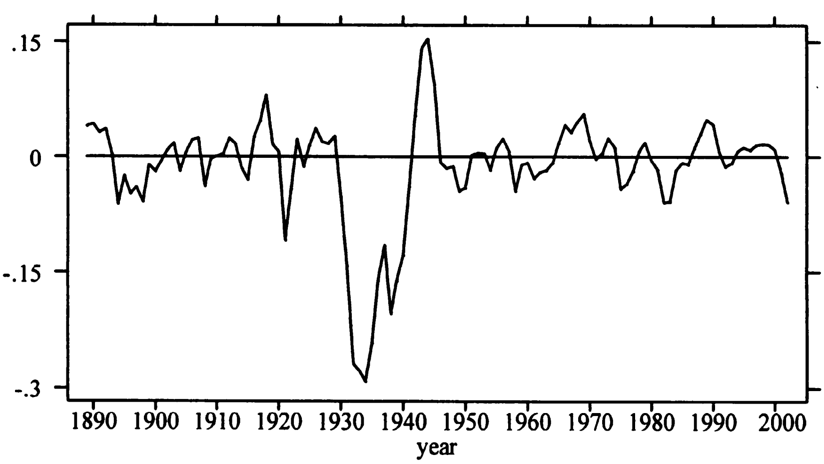


Figure 4B
Quartic detrending of demographically adjusted hours

3. Estimated Responses to a Technology Shock

3.1 Econometric Methodology

The baseline specification is a bivariate model of labor productivity and labor input similar to the benchmark models of Galí (1999) and Francis and Ramey (2005). Under this specification technology is identified as the only shock that can have permanent effects on labor productivity. This assumption is less restrictive than Blanchard and Quah's (1989) identification assumption since it allows nontechnology shocks, such as changes in government spending, to have permanent effects on output. On the other hand, if changes in distortionary taxes affect the capital-output ratio, and hence labor productivity, this identification scheme classifies them as technology shocks. For example, a cut in capital tax rates that permanently raised labor productivity would be called a "technology shock" in our model.[4]

Consider the system in which hours are assumed to be stationary or trend stationary:

$$\begin{bmatrix} \Delta x_t \\ n_t \end{bmatrix} = \begin{bmatrix} C^{11}(L) & C^{12}(L) \\ C^{21}(L) & C^{22}(L) \end{bmatrix} \begin{bmatrix} \varepsilon_t^z \\ \varepsilon_t^m \end{bmatrix}$$

where x_t denotes the log of labor productivity, n_t denotes the log of labor input (or deterministically detrended log of labor input), ε^z denotes the technology shock, and ε^m denotes the nontechnology shock. $C(L)$ is a polynomial in the lag operator. We maintain the usual assumption that ε^z and ε^m are orthogonal. Our assumption identifying the technology shock implies that $C^{12}(1) = 0$, which restricts the unit root in productivity to originate solely in the technology shock.

Another way to think about this restriction is through the estimation method suggested by Shapiro and Watson (1988). Consider the following system of equations:

$$\Delta x_t = \sum_{j=1}^{p} \beta_{xx,j} \Delta x_{t-j} + \sum_{j=0}^{p-1} \beta_{xn,j} \Delta n_{t-j} + \varepsilon_t^z. \tag{1a}$$

$$n_t = \sum_{j=1}^{p} \beta_{nn,j} n_{t-j} + \sum_{j=1}^{p} \beta_{nx,j} \Delta x_{t-j} + \theta \varepsilon_t^z + \varepsilon_t^m. \tag{1b}$$

As discussed by Shapiro and Watson (1988), imposing the long-run restriction is equivalent to restricting the other variables to enter the equation in differences. Because the current value of Δn_t will be correlated with ε_t^z in the first equation, instrumental variables must be used to estimate the equation. Using lags one through p of Δx_t and n_t as instruments for the first equation yields estimates that are identical to those obtained using matrix methods.[5] The baseline specification uses two annual lags of each variable.

The second shock to the system, ε_t^m, is identified by including the estimated residual from the first equation in the second equation, along with the standard lags of the variables, as shown in equation (1b). The estimated residual from this equation, ε_t^m, is identified as the "nontechnology" shock.

If, instead, hours have a unit root, the system to be estimated is as follows:

$$\Delta x_t = \sum_{j=1}^{p} \beta_{xx,j} \Delta x_{t-j} + \sum_{j=0}^{p-1} \beta_{xn,j} \Delta^2 n_{t-j} + \varepsilon_t^z. \tag{2a}$$

$$\Delta n_t = \sum_{j=1}^{p} \beta_{nn,j} \Delta n_{t-j} + \sum_{j=1}^{p} \beta_{nx,j} \Delta x_{t-j} + \theta \varepsilon_t^z + \varepsilon_t^m. \tag{2b}$$

In this system, the hours must be double-difference in the first equation in order to impose the long-run restrictions. The second equation is run in the first-difference of hours.

For all specifications, both equations are estimated jointly using GMM. The estimated variance-covariance matrix takes into account the fact that the technology shock that appears as a regressor in the hours equation is estimated from the first equation. It is also robust to heteroscedasticity, an important feature in light of the evidence presented below that the variance of the shocks changes over time. The standard error bands for the impulse response functions are derived by generating random vectors from a multivariate normal distribution with mean equal to the coefficient estimates and variance-covariance matrix equal to the estimated one, and then calculating the impulse response functions.

3.2 *Impulse Response Functions for Three Difference Specifications*

The results from estimating the model with hours in levels, detrended with a quartic trend, and with a unit root are shown in Figure 5. Recall that productivity is assumed to have a unit root in all specifications.

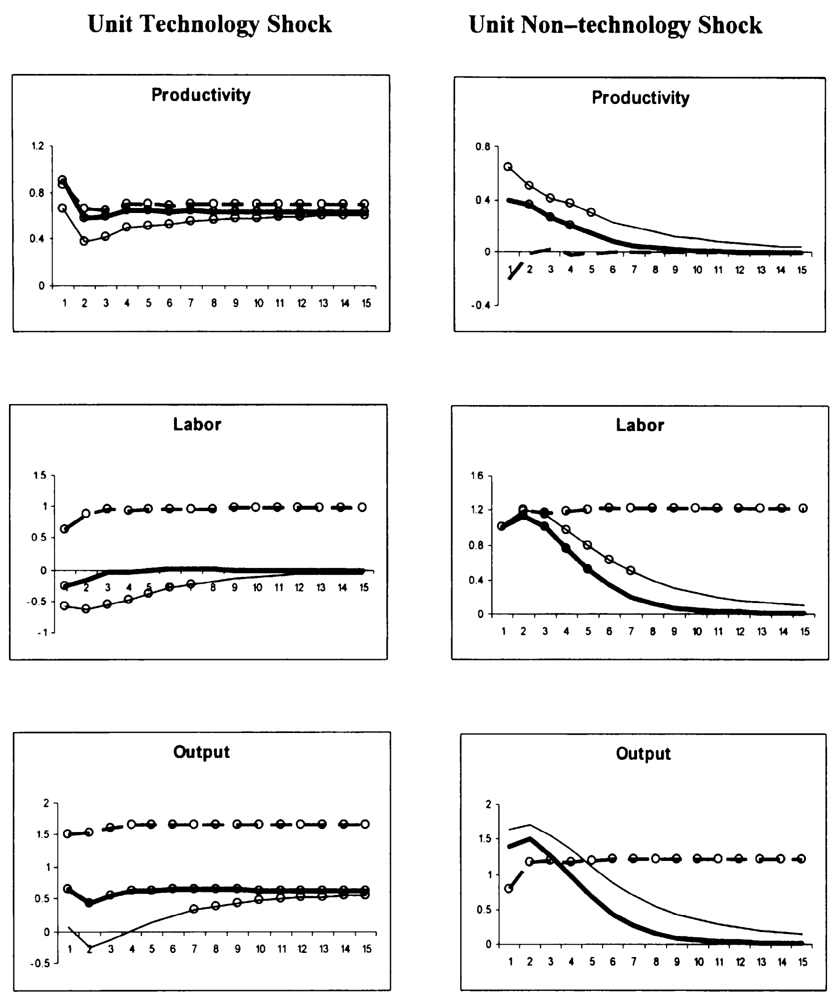


Figure 5
Impulse response functions: 1892–2002
(thin line: hours in levels; thick line: detrended; dashed line: unit root)
(Circles indicate significance at 10 percent level)

The first column of Figure 5 shows the effect of a unit technology shock and the second column shows the effect of a unit "non-technology" shock.

Consider first the results in the first column. In all three cases, a positive technology shock leads to a permanent increase in productivity, with some initial overshooting. The unit root and deterministic trend specifications have very similar effects, while the response under the levels specification is somewhat smaller. The estimates are all significantly different from zero at the 10 percent level.

The response of hours differs significantly across specifications. In the levels specification and the deterministic trend specification, hours fall temporarily in response to a positive technology shock. This result is consistent with the findings of Galí (1999) and Francis and Ramey (2005) and others for the post-World War II period. Christiano, Eichenbaum, and Vigfusson (2003) find the opposite result for the post-war period when they assume that hours per capita are stationary. Some have argued that their measure of hours per capita, which divides by the total population age 16 and over, is not stationary. Our measure, which demographically adjusts the population, produces a negative response. In contrast, the unit root specification that uses our measure predicts a significant increase in hours. As we shall see below, this result does not hold over the entire sample.

All three specifications show a permanent rise in output in response to the technology shock, although the levels specification shows output actually declining for a few years before becoming positive. This initial negative response of output also appears in Basu, Fernald, and Kimball (2004), who use annual data as well.[6]

The second column of Figure 5 shows the responses to a "non-technology" shock. This shock includes any shock that does not have a permanent effect on labor productivity. The levels and deterministic trend specifications show this type of shock raising productivity temporarily, while the unit root specification shows a very transitory fall in productivity.

Hours and output behave quite similarly across specifications in the first few years after the nontechnology shock. The key difference is that the effects on hours and output are permanent in the unit root specification whereas the effects are transitory in the two other specifications.

How stable are these estimates across time and to what extent are they accounted for by the dramatic movements of World War II? To answer these questions, we re-estimate the model for the subperiods 1892–1940 and 1948–2002. Breaking the sample in this way is supported by structural break tests. When we estimate a model with an unknown break that affects all coefficients and the variance, the log likelihood reaches a maximum for both the productivity and hours equations in the late 1940s, typically between 1948 and 1949.

Figures 6A and 6B show the results for the various specifications across the two samples. Figure 6A shows the effects of a unit technology shock. Consider first the graphs in the second column, which show the results for the models estimated from 1949–2002. The effects of a technology shock are quite similar across specifications. This result stands in contrast to the results from the literature that use hours divided by population age 16 and over. There, the levels specification gives very different results from the unit root and deterministic trend specifications. With our measure, all specifications give very similar results. In particular, all three measures show hours declining for at least one year in response to a technology shock.

The story is different in the early period, shown in the first column of Figure 6A. The effects of a technology shock differ across specifications in the early period, with the most positive effects coming from the unit root specification. Hours fall temporarily in the levels and trend specifications, while they rise permanently in the unit root specification. The results are in fact similar to those estimated over the entire sample, which is not surprising since the early period has more dramatic movements in the variables. The levels specification implies that a positive technology shock depresses output for a number of years in the early period, though the estimates are not significant.

Within each specification, the most notable change in a response across the pre- and post-WWII periods is the response of hours in the unit root specification. The response to a technology shock is strongly positive in the early period, but temporarily negative in the later period. The potential source for this structural break is discussed in Section 3.4 below.

Figure 6B shows the effects of a nontechnology shock across periods. The levels and deterministic trend specifications suggest that the effects of nontechnology shocks are somewhat less persistent in the later period than in the earlier period. The first-difference specification shows that the effects on hours and output are permanent in both periods.

Table 2 compares the estimated technology shocks across specifications and time periods. The correlation between the technology shocks estimated under the assumption of a unit root and the other specifications is quite low in the pre-WWII period, between 0.12 and 0.42. In contrast, the technology shocks from the levels and deterministic trend

1892 – 1940 1949 – 2002

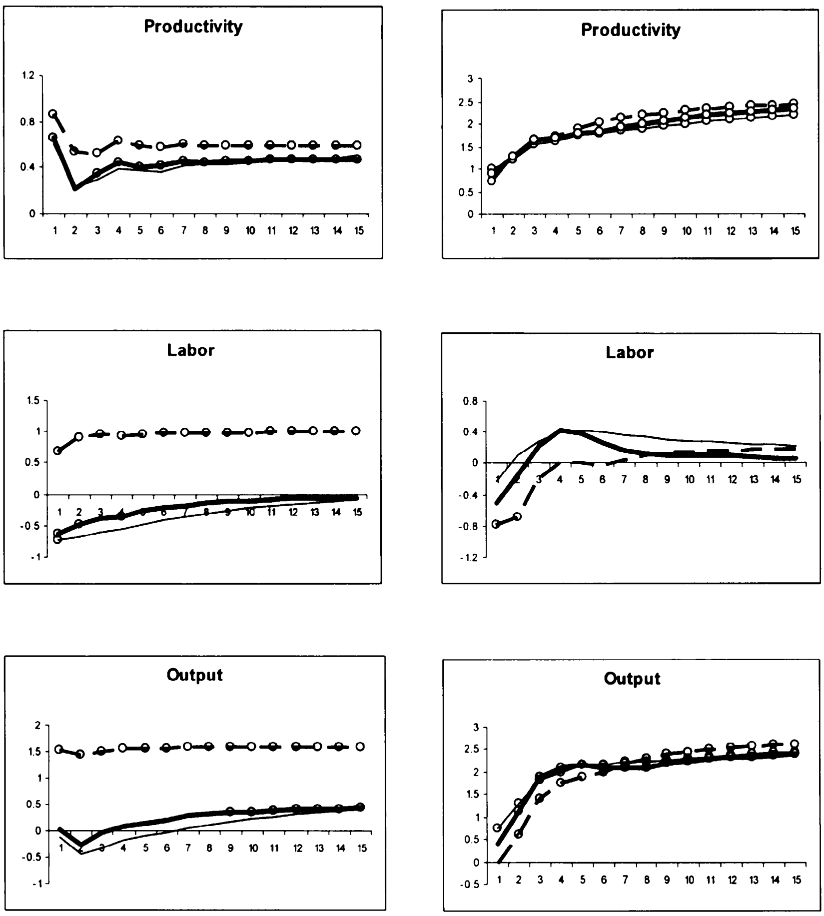


Figure 6A
Impulse response functions: Effects of a technology shock
(thin line: hours in levels; thick line: detrended; dashed line: unit root)
(Circles indicate significance at 10 percent level)

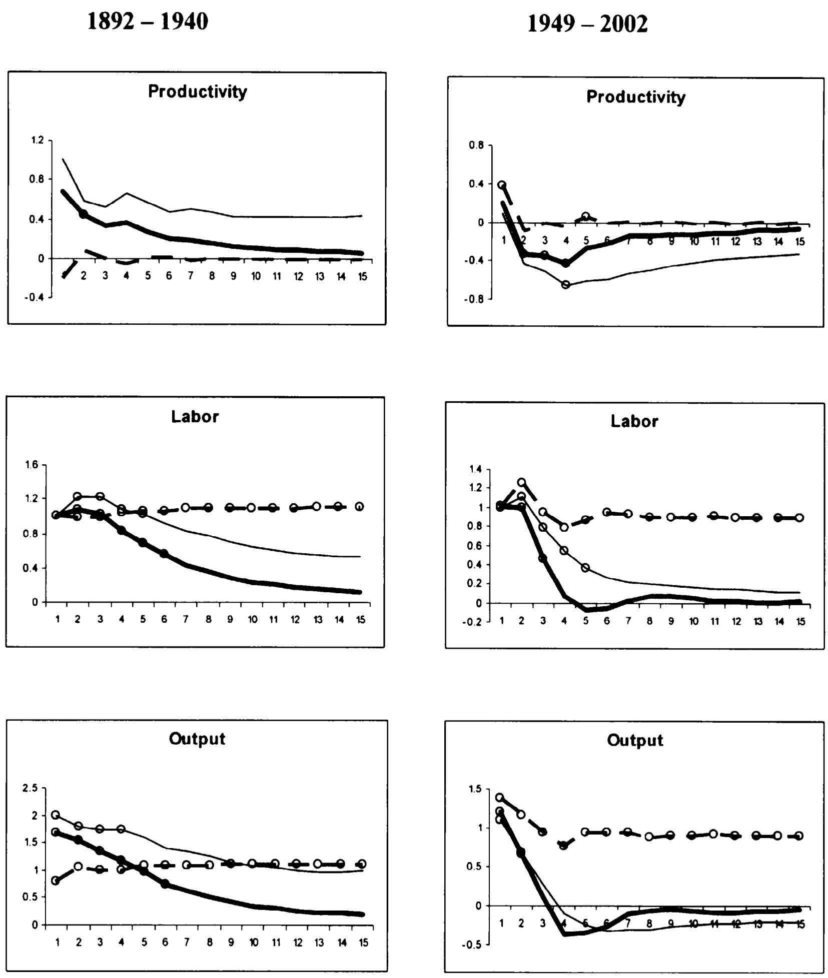


Figure 6B
Impulse response functions: Effects of a nontechnology shock
(thin line: hours in levels; thick line: detrended; dashed line: unit root)
(Circles indicate significance at 10 percent level)

Table 2
Correlation of technology shocks across specifications

Pre-WWII: 1892–1940			
Specification	**Levels**	**Deterministic trend**	**Unit root**
Levels	1.00		
Deterministic Trend	0.95	1.00	
Unit root	0.12	0.42	1.00
Post-WWII: 1949–2002			
Specification	**Levels**	**Deterministic trend**	**Unit root**
Levels	1.00		
Deterministic Trend	0.98	1.00	
Unit root	0.87	0.92	1.00
Addendum: Standard deviation of the technology shock relative to the nontechnology shock			
Specification	**Levels**	**Deterministic trend**	**Unit root**
1892–1940	1.6	0.96	1.0
1949–2002	0.57	0.69	0.73

specifications are highly correlated with each other in the early period. In the post-WWII period, the technology shocks have a very high correlation across specifications, above 0.87. The correlations for the nontechnology shock (which are not shown) have a very similar pattern across specifications for both time periods.

The bottom panel of Table 2 shows the relative standard deviations of the technology and nontechnology shocks. The levels specification suggests that the standard deviation of the technology shock is 60 percent higher than that of the nontechnology shock in the early period, whereas the quartic and unit root specifications suggest that they are about equal. In the late period, all three specifications indicate that the standard deviation of the technology shock is between 60 and 70 percent of the nontechnology shock.

3.3 *Which Specification Is Most Valid?*

For the purposes of analyzing the source of shocks for the post-WWII period, all three specifications give similar answers to most questions.

The impulse response functions are similar at short-horizons and the shocks identified are highly correlated across specifications. The only difference is the permanent effect of nontechnology shocks on output and hours in the unit root specification. In contrast, the answers for the pre-WWII period depend very much on which specification one uses.

We assess the validity of each specification by subjecting the estimated shocks to an Evans test. As argued by Evans (1992), technology shocks should not be Granger-caused by nontechnology variables such as government spending and monetary variables (Granger 1969). Evans cast doubt on the use of the Solow residual as a measure of technology shocks by showing that monetary variables and government spending Granger-caused the Solow residual. Thus, an additional means to test whether the identified shocks are really technology shocks is to test whether they are Granger-caused by these types of variables.

For each subperiod, we regressed the estimated shocks on two lags each of the growth rates of per capita government spending, money, and prices. (See the data appendix for details on these variables.) We then tested whether these variables Granger-caused the technology shock. The results are reported in Table 3.

Table 3
Granger causality tests

A. Dependent Variable: Identified Technology Shocks

	Prewar: 1892–1940		Postwar: 1949–2002	
Model	P-value on F-test	R-squared	P-value on F-test	R-squared
Levels	0.001	0.392	0.105	0.193
Detrended	0.005	0.342	0.119	0.187
Unit root	0.210	0.174	0.174	0.168

B. Dependent Variable: Identified Nontechnology Shocks

	Prewar: 1892–1940		Postwar: 1949–2002	
Model	P-value on F-test	R-squared	P-value on F-test	R-squared
Levels	0.181	0.183	0.085	0.203
Detrended	0.093	0.219	0.162	0.172
Unit root	0.002	0.384	0.139	0.179

The tests are based on a regression of the shock on a constant and two lags each of the growth rates of per capita government spending, per capita money, and GDP deflator. The null hypothesis is that all coefficients on these variables (excluding the constant) are zero.

The results indicate that only the technology shock estimated under the assumption of a unit root passes the test. In neither time period do the nominal and government variables Granger-cause the technology shock as estimated by the unit root specification. In contrast, the p-values are very low for both the levels and detrended specifications in the early period. In fact, government spending and the nominal variables explain between 34 and 40 percent of the variance of the "technology" shocks estimated with these specifications. The variables have more explanatory power for the technology shock than the nontechnology shock. In contrast, the variables explain little of the technology shock estimated with the unit root specification, but explain 38 percent of variance of the nontechnology shocks from this specification. All p-values for the Granger causality tests on the technology shocks are above 0.1 in the late period. Recall that this is the period where all three specifications gave similar results.

To summarize, the Evans-type tests favor the unit root specification because it is the only specification whose estimated technology shocks are not Granger-caused by monetary and government spending variables in either period.

Another reason to favor the unit root specification is provided by Fernald (2005). In post-WWII data, Fernald finds that the results of the levels specification are sensitive to low frequency movements in the hours and mean productivity growth. In particular, Christiano, Eichenbaum, and Vigfusson's (2003) finding that technology shocks raise hours reverses signs if one allows for statistically-supported breaks in mean productivity growth. On the other hand, the unit root specification is robust to these low frequency movements.

Based on the Granger-causality results above and Fernald's findings, we will use the unit root specification for the analysis in the remainder of the paper. The appendix shows some of the results with the quartically detrended data. Because of the evidence for a break in the late 1940s, we will continue to estimate the model separately over the two samples.

We also subject the unit root specification to two more robustness checks. First, we increase the number of lags to four years and re-estimate the models. The additional lags are not significant and the results from this specification are similar to those from the model with only two annual lags. Second, we augment the system with consumption and investment. Both consumption and investment appear to have unit roots, and do not exhibit cointegration with each other or with productivity over the sample. Thus, both appear as double-differences in the productivity equation. This system gives similar results for the effects of the shocks on productivity, hours, and output. Consumption and investment rise permanently in response to both technology and nontechnology shocks in both sample periods. In the post-WWII period, the initial impact on investment is negative and then becomes positive.

3.4 Discussion of the Impulse Response Functions from the Unit Root Specification

Before proceeding to an analysis of the shocks, we discuss two aspects of the impulse response functions for the unit root specification that raise questions. Both concern the effect of a technology shock, shown in Figure 6A.

The first question raised by these estimates is why the response of hours to a technology shock changes so dramatically from the pre-WWII period to the post-WWII period. We begin by investigating whether the Great Depression or WWI is the key source of this behavior. When the model is estimated from 1892 to 1929, omitting the Great Depression, the patterns are similar, though muted. In particular, whereas hours climb to a permanent plateau around unity for the 1892–1940 sample, in the 1892–1929 sample hours rise to 0.27 on impact then fall to a plateau of 0.17. None of the movements is significantly different from zero. When we estimate the model omitting the WWI years 1917–1920 (with the extra years omitted because of the lags in the regression), we find very similar results to those for the period 1892–1940.

Thus, it appears that there was structural change in the economy that was not just limited to the Great Depression or WWI. Why do hours rise in the early period but fall in the later period in response to a technology shock? Francis and Ramey (2005) show that real rigidities such as adjustment cost on investment and habit formation in consumption can produce a temporary negative response of hours to a technology shock. King and Wolman (1996) and Galí and Rabanal (2004) show how price and wage rigidities (with suitable monetary policy rules) can produce a negative response to hours. If one were to apply one of these explanations to our results, one would have to identify structural changes in real rigidities, price and wage rigidities, or monetary policy.

A simpler explanation comes from applying the insights of Manuelli (2000), Rotemberg (2003) and Lindé (2003). They show that if technology diffuses slowly, so that technology growth is slightly persistent, hours will fall temporarily in response to a positive technology shock. After the shock, agents expect productivity to grow even more, implying that wages will be higher in the future than they are now. Agents thus decide to work less now. On the other hand, if there is no persistence in technology growth, hours will rise temporarily in response to a technology shock.

The sets of graphs for the unit root specification across the two time periods in Figure 6A look quite similar to the two standard RBC simulations Lindé shows in Figure 2 of his paper.[7] In particular, the early period looks like the case with no positive persistence in the growth rate of technology. Productivity and output actually overshoot their new levels in our data, hours rise, and output rises to close to its new level. In contrast, the later period looks like Lindé's simulation in which technology growth has positive persistence. Productivity and output rise more slowly and hours decline temporarily. Hours eventually become positive in Lindé's simulation, though not in our empirical results. Thus, the change in the impulse response for hours across the two periods can be explained using a very standard RBC model in which the productivity growth process displays the changes shown in our impulse response functions.

The second question raised by these graphs is the apparent permanent effect of a technology shock on hours in the early period. While not inconsistent with a general RBC model, this result is inconsistent with the standard specification of preferences in an RBC model. Because hours per capita have changed little relative to real wages over the century, most RBC models specify a utility function in which the wealth and substitution effects of a technology shock exactly cancel. The impulse response functions for the later period clearly suggest a transitory effect on hours, but not for the early period. Whatever shocks are leading to permanent increases in productivity in the early period are also leading to permanent increases in hours.

4. The Source of Historical Fluctuations

We are now ready to assess the roles of technology and nontechnology shocks in historical fluctuations based on our preferred specification, which assumes a unit root in hours. We begin by analyzing the importance of each type of shock in the overall variance of productivity, output and hours. We then analyze the nature of the shocks over specific historical periods.

4.1. Variance Decomposition

To determine which type of shock is important for the variance of the key variables, we perform a forecast error variance decomposition. The results for each period are shown in Table 4. According to the unit root specification, technology shocks are the main source of the variance of productivity and output at all horizons in the early period. Technology shocks account for 33 percent of the variance of the forecast error of hours at the one-year horizon, rising to 46 percent by the 20 year horizon.

The story is different for the later period for hours and output. While technology shocks continue to account for the bulk of the variance of the forecast error of productivity, accounting for two-thirds of the variance of productivity at the one-year horizon and becoming more important with each year, they account for much less of the other two variables. Technology shocks account for 25 percent of the one-year forecast of hours, but then decline in importance. Of course, as the impulse response functions show, the technology-induced movements in hours in the later period are negatively correlated with output. The final column shows that technology shocks are unimportant for output at business cycle horizons. These results for the post-WWII period are consistent with those of other researchers who use unit root specifications, such as Galí and Rabanal (2004).[8]

Table 4
Variance decomposition: Unit root specification

Percent of Forecast Variance Explained by Technology Shocks

Horizon (in years)	1892–1940			1949–2002		
	Productivity	Hours	Output	Productivity	Hours	Output
1	95	33	80	65	25	0
2	96	40	73	88	18	5
3	96	44	72	94	15	22
4	97	44	72	96	12	37
5	98	45	72	97	11	45
10	99	46	71	99	7	64
20	99	46	71	100	4	72

4.2. *Patterns of Shocks*

It is also interesting to study the historical pattern of the two types of estimated shocks. For completeness, we use our estimates from the period 1892–1940 to produce shocks for the WWII period as well. The shocks are virtually the same as if we used the coefficients estimated from 1892–1948. The shocks from 1949 to 2002 are derived from the model estimated over the 1949–2002 sample.

While studying the shocks, it is important to keep two points of interpretation in mind. First, what we call a "technology shock" is any shock that has a permanent effect on labor productivity. While true technology shocks fit this definition, other shocks such as government policies that subsidize education also fit this definition. "Non-technology shocks" are any shocks that do not have permanent effects on labor productivity. Examples of these types of shocks are monetary policy shocks and military spending shocks. Second, what we call a "negative technology shock" is any technology shock that is lower than average. Since technology growth is generally positive, what we call a negative technology shock can in some cases simply be a positive shock that is lower than average.

Figure 7 shows the historical pattern of technology shocks and non-technology shocks estimated with the unit root specification. The recession dates identified by the National Bureau of Economic Research (NBER) are shown in the shaded areas. Note that the technology shocks tend to be negative around recession dates during the early period, but not so much during the later period. The reverse is true of the non-technology shocks, which tend to be negative around recessions in the post-WWII period.

The three years during the early period with the most negative technology shocks are 1908, 1914, and 1932. It is interesting to note that all three of these dates are associated with problems in the financial system. A banking panic occurred in October 1907. 1914 marked the outbreak of WWI, which brought some financial difficulties. For example, the New York Stock Exchange had to be closed for a day (Friedman and Schwartz 1963). Finally, the banking crises of the early 1930s are well-researched. We will discuss the behavior of shocks during the Great Depression in more detail below.

During the post-WWII period, the three years with the most negative shocks are 1959, 1974, and 1987. 1959 is associated with the steel strike, 1974 with the first oil crisis and the collapse of the exchange rate system,

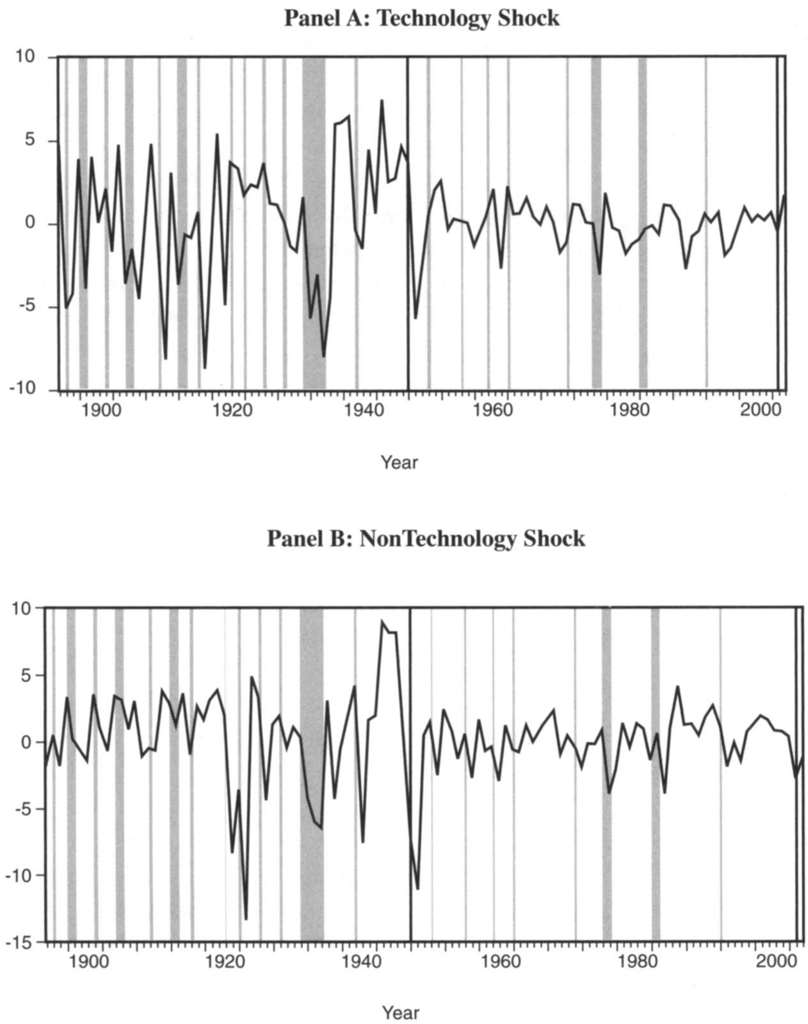


Figure 7
Plots of implied shocks from unit root specification
(Based on sub-sample estimates; shaded areas represent NBER recession dates)

and 1987 with the stock market crash (see Eckstein and Sinai 1986 for a chronology of the post-war events). It is not clear, though, that there was a causal link between these events and the estimated shocks.

On the positive side, the estimates suggest that the period from the late teens to the mid-1920s was characterized by a long string of positive technology shocks. The most positive technology shocks occurred in 1934, 1935, 1936, and 1941 in the early period. During the later period, the most positive technology shocks occurred in 1950, 1958, and 1960. The 1960s were characterized by a long string of positive technology shocks. We will discuss several of these periods in more detail below.

The four most negative nontechnology shocks occur in 1919, 1921, 1938, and 1946. According to the estimates, there was also a series of large negative nontechnology shocks during the Great Depression. During the postwar period, the three most negative nontechnology shocks occurred in 1974, 1982, and 2001. The most positive nontechnology shocks during the early period were 1941–1943. During the later period, the most positive nontechnology shocks occurred in 1950, 1984, and 1989.

4.3 Historical Chronology

We now study the pattern of shocks during prominent periods of history to ascertain the key driving forces for fluctuations. We continue to refer to the graphs of shocks in Figure 7.

4.3.1 The Early 1900s

During the period 1892 to WWI, it is clear that much of the volatility in the economy was due to the volatility of technology shocks rather than nontechnology shocks. During this period there were a number of very positive technology shocks as well as a number of very negative technology shocks. As mentioned before, the two most negative technology shocks occurred during banking crises.

4.3.2 The 1920s

According to our estimates, the 1920s began with a series of large negative nontechnology shocks in 1919–1921. These shocks were no doubt linked to the wind-down from WWI and the conduct of monetary policy (see Friedman and Schwartz's 1963 account of this period). On the other hand, from 1918 to 1926, every year saw positive technology shocks, suggesting that this was a period of steady technological progress.

4.3.3 The Great Depression

The graphs in Figure 7 show that both types of shocks contributed to the Great Depression. Both technology and nontechnology shocks were slightly positive in 1929, and then became very negative in the early 1930s. 1933 finally saw a positive nontechnology shock, perhaps owing to the increase in government spending, but the technology shock remained negative. Comparing this period to the early 1920s, we see that had the technology shocks not been negative, the early 1930s would have been more similar in magnitude to the recession of the early 1920s.[9]

4.3.4 1934–1940

Particularly interesting is the series of high estimated positive technology shocks in the second half of the 1930s. The notion of large positive technology shocks during the 1930s is at odds with the conventional wisdom. Recent work by Field (2003), however, argues that the 1930s were the most technologically progressive decade. He shows that productivity growth between 1929 and 1941 was higher on average than the period 1919–1929 and cites several micro studies on innovations.

In his study, Field does not distinguish 1929–1933 from 1934–1941 because he wants to compare years with similar unemployment rates. What the average hides is how impressive productivity growth was during 1934–1941. Our calculations show that labor productivity growth was –1.61 per year from 1929–1933 but was 4.55 percent per year from 1933 to 1941. This rate is substantially higher than the 2.36 percent per year rate of growth from 1919 to 1929. Thus, productivity growth during the second period was exceptionally high.

Microeconomic studies of innovations support the notion of this period as a period of high innovative activity.[10] Mensch's (1979) listing of basic innovations for the first half of the twentieth century (Table 4-4, pages 127-128) shows that fully 24 percent of them occurred in the three years from 1934–1936, which our estimates show to be years with high positive technology shocks. For example, in 1934 alone the innovations included the diesel locomotive, fluorescent lighting and radar. Kleinknecht (1987) summarizes the results of several studies on major innovations (Table 3.2, page 70) by five-year periods. All of them show a huge burst of innovation in the period 1935–1939. Kleinknecht's (1987) tabulation shows that the five-year period 1935–1939 was rivaled only by 1960–64 in the number of radically new products and improvement and process innovations. According to his classification, each of these five-year periods had 14 innovations.[11] Thus, our time series estimates of large positive technology shocks in the period 1934–36 is consistent with the microeconomic evidence on innovative activity.

4.3.5 World War II

The studies of innovation discussed above also showed the period covering WWII to be a period of relatively high innovative activity, so it is not surprising that there were many positive technology shocks during this period as well. Of course, dramatic increases in government spending were also important, and these show up as sustained highly positive series of nontechnology shocks.

4.3.6 The Post-World War II Period

The year 1946 had very negative technology and nontechnology shocks. Both were probably related to the end of WWII. The start of the Korean War in 1950 appears to have been associated with both positive technology and nontechnology shocks. The period 1960–1967 experienced a sustained string of positive technology shocks. Kleinknecht's (1979) study also shows high innovative activity during this period, with 14 innovations from 1960–1964 and 11 innovations from 1965–1969. The nontechnology shocks were also mostly positive during this period, with the exception of 1967 and 1969.

1974 was a year with very negative (by post-WWII standards) technology and nontechnology shocks. This was the only recession since 1946 when both the technology and the nontechnology shocks were negative.

The second half of the 1990s, which has attracted attention for its high productivity growth rates, was not marked by particularly positive technology shocks. Rather, it experienced a series of small positive technology shocks without any intervening negative technology shocks. The nontechnology shocks were also uniformly positive during this period.

4.4 *The Changing Volatility of Shocks*

Finally, a noticeable feature of both the technology and nontechnology shocks in Figure 7 is that both series appear to have become far less volatile in the postwar era. The reduced volatility of the postwar recessions has been documented by many, including Zarnowitz and Moore (1986), Taylor (1986), and DeLong and Summers (1986).[12] We carry out an F-test of equal variance between the variances of the prewar and postwar technology shocks. Given the variance of the prewar technology of 15.59 with sample size 49, and similar figures for the postwar technology of 1.59 and 54 respectively, the value of the F-statistic is 9.81 ($15.59 \div 1.59$). We compare this to a critical F-value with 49 numerator degrees of freedom and 54 denominator degrees of freedom. We reject the null of equal variance for all conventional values of the F-statistics which implies that the postwar technology is indeed (significantly) less volatile than the prewar technology.

The results are similar for the nontechnology shock, whose variance falls from 14.36 in the prewar period to 3.01 in the postwar period. An F-test of equal variance yields an F-statistic of 4.77. With the same sample sizes as above we reject the null of equal variances of the nontechnology shock at all conventional levels of significance.

Thus, according to these estimates both types of shocks became significantly less volatile in the postwar period. A comparison of the numbers, however, indicates that the volatility of the technology shocks fell by even more than the volatility of the nontechnology shocks. Thus, our results suggest that it was not improved policy alone that stabilized the U.S. economy in the post-WWII period.

5. Conclusions

This paper has presented estimates of models with long-run restrictions on historical U.S. data in order to study the nature and consequences of the shocks moving labor productivity, hours and output. Following Galí (1999), we identify the technology shock to be the only shock that can have a permanent effect on labor productivity.

We developed a new demographically-adjusted measure of hours per capita and estimated the model under a variety of assumptions about the nature of hours. We compared results from models that assumed stationary hours, a quartic trend in hours and a unit root in hours. All three specifications gave similar results for the post-WWII period, but gave different results for the pre-WWII period. Granger-causality tests on the shocks led us to conclude that the unit root specification led to the most reasonable results. According to this specification, a positive technology shock leads labor to rise in the period from 1889–1940. In contrast, the same type of shock leads labor to fall in the period from 1949–2002.

We then investigated some of the characteristics of the shocks and their role in fluctuations. Technology shocks were much more important for fluctuations in the pre-WWII period than in the post-WWII period. The periods with the most notable series of positive technology shocks were the late teens to the mid-20s, 1934–1936 and the early 1960s. The Great Depression was a period characterized by very negative technology and nontechnology shocks. Finally, both types of shocks are responsible for the reduction in the variance of output in the post-WWII period, suggesting that better policy is not the sole cause of the reduction of GDP volatility in the post-WWII era.

Notes

Valerie Ramey gratefully acknowledges support from National Science Foundation grant # 0213089. We are grateful to Susanto Basu, Garey Ramey, and Harald Uhlig for helpful comments.

1. Several economists have expressed concern that this new finding could be a direct result of the stationarity assumptions made of the time series used in the structural vector autoregressions; see Christiano, Eichenbaum, and Vigfusson (2003).

2. We also created a series that added total government hours to the numerator rather than subtracting government employment from the denominator. This series is very close to the one presented.

3. Moreover, the labor force participation rate of this age segment has decreased over time relative to the population age 16 and older. Adjusting for changes in the differential rate of labor force participation does not noticeably change in the series in Figure 2B.

4. Francis and Ramey (2005) and Gali (2004) rule out capital income taxes as an explanation for the technology shock results. Francis and Ramey found that their results were robust to the inclusion of capital income tax in their estimations while Gali found that innovations to such taxes were uncorrelated with his identified technology shocks.

5. See the appendix of Francis, Owyang, and Theodorou (2003) for an explicit derivation of the equivalence of the Shapiro-Watson and matrix methods.

6. Analyses using annual data often find an initial decline in variables such as output and investment, whereas those using quarterly data typically do not.

7. Rotemberg (2003) shows similar results for the case of slow diffusion in Figure 5 of his paper.

8. As Table A1 shows, the quartically detrended specification gives very different results relative to the unit root specification in the early period. In the detrended specification, technology shocks only account for 40 percent of the forecast error variance of productivity at the one-year horizon and 62 percent at the ten year horizon. The hours and output numbers in the early period are similar to those for the unit root specification in the later period.

9. For a very different accounting of the shocks during the Great Depression, see Figure A1, which shows the shocks estimated from the specification with the quartic trend. According to these estimates, there were mostly positive technology shocks during the Great Depression.

10. The date of innovation is defined by these authors as the time when a newly discovered material or technique is first produced on a regular basis or when a market for a new product is first formed.

11. These numbers were calculated as the sum of columns (2) + (4) from Kleinknecht's Table 3.2, page 70. The table covers the period 1900–1969.

12. More recently papers such as McConnell and Perez-Quiros (2000), Blanchard and Simon (2001) and Stock and Watson (2002) have documented a decline in output volatility post 1984.

References

Anderson, Richard. 2003. "Some Tables of Historical US Currency and Monetary Aggregates Data." Federal Reserve Bank of St. Louis Working Paper no. 2003-006A.

Balke, Nathan S., and Robert J. Gordon. 1989. "The Estimation of Prewar Gross National Product: Methodology and New Evidence." *Journal of Political Economy* 97: 38–92.

Basu, Susanto, John Fernald, and Miles Kimball. 2004. "Are Technology Improvements Contractionary?" NBER Working Paper no. 10592. Cambridge, MA: National Bureau of Economic Research.

Blanchard, Olivier, and Danny Quah. 1989. "The Dynamics Effects of Aggregate Demand and Supply Disturbances." *American Economic Review* 79: 654–673.

Blanchard, O.J., and Simon, J. 2001. "The Long and Large Decline in U.S. Output Volatility." *Brookings Papers on Economic Activity* 1: 135–164.

Christiano, Lawrence, Martin Eichenbaum, and Robert Vigfusson. 2003. "What Happens After a Technology Shock?" Working Paper no. 9819. Cambridge, MA: National Bureau of Economic Research.

Delong, J. Bradford, and Lawrence H. Summers. 1986. "The Changing Cyclical Variability of Economic Activity in the United States." In Robert Gordon, ed., *The American Business Cycle: Continuity and Change.* Chicago: NBER and University of Chicago Press.

Eckstein, Otto, and Allen Sinai. 1986. "The Mechanisms of the Business Cycle in the Postwar Era." In Robert Gordon, ed., *The American Business Cycle: Continuity and Change.* Chicago: NBER and University of Chicago Press.

Evans, Charles L. 1992. "Productivity Shocks and Real Business Cycles." *Journal of Monetary Economics* 29: 191–208.

Fernald, John. 2005. "Trend Breaks, Long-Run Restrictions, and the Contractionary Effects of Technology Shocks." Federal Reserve Bank of San Francisco Working Paper no. 2005-21.

Field, Alexander. 2003. "The Most Technologically Progressive Decade of the Century." *American Economic Review* 93: 1399–1413.

Francis, Neville, and Valerie A. Ramey. 2005. "Is the Technology-Driven Real Business Cycle Hypothesis Dead? Shocks and Aggregate Fluctuations Revisited." *Journal of Monetary Economics* 52(8): 1379–1399.

Francis, Neville R., Michael T. Owyang, and Athena T. Theodorou. 2003. "The Use of Long-Run Restrictions in Identifying Technology Shocks." Federal Reserve Bank of St. Louis *Review* 85 (6): 53–66.

Friedman, Milton, and Anna Schwartz. 1963. *A Monetary History of the United States, 1867–1960.* Princeton: Princeton University Press.

Galí, Jordi. 1999. "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations." *American Economic Review* 89: 249–271.

Galí, Jordi. 2004. "On the Role of Technology Shocks as a Source of Business Cycles: Some new Evidence." *Journal of European Economic Association* 2-3: 372–380.

Galí, Jordi, and Pau Rabanal. 2004. "Technology Shocks and Aggregate Fluctuations: How Well does the RBC Model Fit Postwar US Data? *NBER Macroeconomics Annual.*

Goldin, Claudia. 1999. "A Brief History of Education in the U.S." Working Paper no. H0119. Cambridge, MA: National Bureau of Economic Research.

Granger, C.W.J. 1969. "Investigating Causal Relations by Econometric Models and Cross-Spectral Models." *Econometrica* 37: 424–438.

Kendrick, John W. 1961. *Productivity Trends in the United States.* Princeton: NBER and Princeton University Press.

King, Robert G., and Alexander Wolman. 1996. "Inflation Targeting in a St. Louis Model of the 21st Century." *Federal Reserve Bank of St. Louis Review* 78: 83–107.

Kleinknecht, Alfred. 1987. *Innovation Patterns in Crisis and Prosperity.* London: The MacMillan Press Ltd.

Lindé, Jesper. 2003. "The Effects of Permanent Technology Shocks on Labor Productivity and Hours in the RBC Model." Centre for Economic Policy Research Discussion Papers no. 4827.

Manuelli, Rodolfo E. 2000. "Technological Change, the Labor Market and the Stock Market." Working Paper no. 8022. Cambridge, MA: National Bureau of Economic Research.

McConnell, M.M., and Perez-Quiros, G. 2000. "Output Fluctuations in the United States: What has Changed since the Early 1980s?" *American Economic Review* 90 (5): 1464–1476.

Mensch, Gerhard. 1979. *Stalemate in Technology: Innovations Overcome the Depression.* Cambridge, MA: Ballinger Publishing Company.

Mulligan, Casey. 2002. "A Century of Labor-Leisure Distortions." Working Paper no. W8774. Cambridge, MA: National Bureau of Economic Research.

Rotemberg, Julio J. 2003. "Stochastic Technical Progress, Smooth Trends, and Nearly Distinct Business Cycles." *American Economic Review* 93: 1543–1559.

Shapiro, Matthew D., and Mark Watson. 1988. "Sources of Business Cycle Fluctuations." *NBER Macroeconomics Annual,* 111–148.

Shea, John. 1998. "What Do Technology Shocks Do?" *NBER Macroeconomics Annual,* pp. 275–310.

Stock, James H., and Watson, Mark. 2002. "Has the Business Cycle Changed and Why?" *NBER Macroeconomics Annual.*

Taylor, John B. 1986. "Improvements in Macroeconomic Stability: The Role of Wages and Prices." In Robert Gordon, ed., *The American Business Cycle: Continuity and Change.* Chicago: NBER and University of Chicago Press.

Zarnowitz, Victor, and Geoffrey H. Moore. 1986. "Major Changes in Cyclical Behavior." In Robert Gordon, ed., *The American Business Cycle: Continuity and Change.* Chicago: NBER and University of Chicago Press.

Data Appendix

Population

Data Sources: 1900–2002 data, including age breakdown, is from the U.S. Census, *Statistical Abstract of the United States: 2003,* Table HS-3 and *Economic Report of the President, 2003,* Table B-34. 1889–1899 on the total resident population is from *Historical Statistics,* Table A-119.

Series Creation: Only the resident population was available before 1939. To obtain a better estimate of the total population, we added the number of armed forces overseas during WWI. Before 1900, population by age was only available in 1890. We interpolated by multiplying the ratio of resident population in a particular age group in 1900 to resident population (all ages) in 1900. 1900–1938: Resident population age 16+ plus armed forces overseas during WWI.

School Enrollment

The school enrollment numbers were obtained by combining information from the *Digest of Education Statistics, 2002, Historical Statistics* Table H442, and Claudia Goldin "A Brief History of Education in the U.S." August 1999, NBER working paper H0119. The *Digest of Education Statistics* contained total enrollment figures annually from 1964–2002, and every ten years before that. We used Goldin's and the *Historical Statistics* enrollment numbers for K–12 to interpolate the total enrollment numbers.

Government Employment

1889–1929 data are from Kendrick *Productivity Trends in the United States*, 1961, Table A-VI. Data from 1929–2002 were from BEA NIPA Tables 6.8A-D. The data were spliced using overlap data at 1929. Employment is full-time workers.

Real GDP, GDP Deflator, Consumption, and Investment

Data Sources: Real GNP and deflator 1889–1928 from Balke and Gordon, *Journal of Political Economy*, 1989. Real consumption expenditures and gross private investment 1889–1928: John Kendrick, *Productivity Trends in the United States*, 1961, Table A-IIa. Chain-weighted GDP, consumption and investment 1929–2002: BEA NIPA from www.bea.gov.

Series Creation: The pre-1928 data were multiplied by the ratio of the BEA data in 1929 to the historical data in 1929.

Productivity, Hours, and Output in Private Business

Data Sources: 1889-1946: John Kendrick, *Productivity Trends in the United States*, 1961, Tables A-X, A-XXII, A-XXIII. 1947-2002: BLS Productivity data from www.bls.gov.

Series Creation: 1889–1946 data were multiplied by the ratio of the BLS data in 1947 to the historical data in 1947.

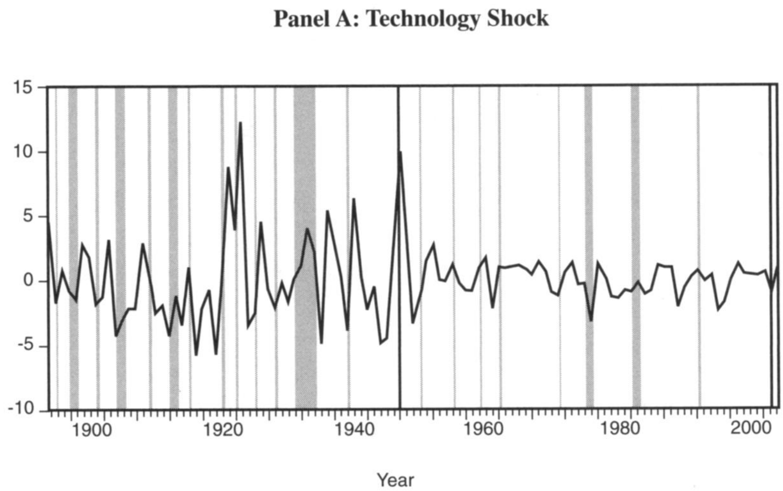
Money

M2: For the period 1959–2002, we used M2 from the Board of Governors of the Federal Reserve. The earlier series are from Richard Anderson, "Some Tables of Historical US Currency and Monetary Aggregates Data," April 2003 working paper. For 1947–1958, we use Rasche's M2 series. Because Anderson argues that Friedman and Schwartz M4 series is most comparable for the early period to M2 for the later period, we use M4 where possible, and otherwise M3.

Table A1
Variance decomposition: Quartic trend specification

Percent of Forecast Variance Explained by Technology Shocks

Horizon (in years)	1892–1940 Productivity	1892–1940 Hours	1892–1940 Output	1949–2002 Productivity	1949–2002 Hours	1949–2002 Output
1	40	26	0	90	10	5
2	33	21	2	88	6	26
3	36	18	1	90	6	55
4	40	18	1	89	10	68
5	44	18	1	91	12	75
10	62	17	5	96	14	88
20	79	17	20	98	14	95



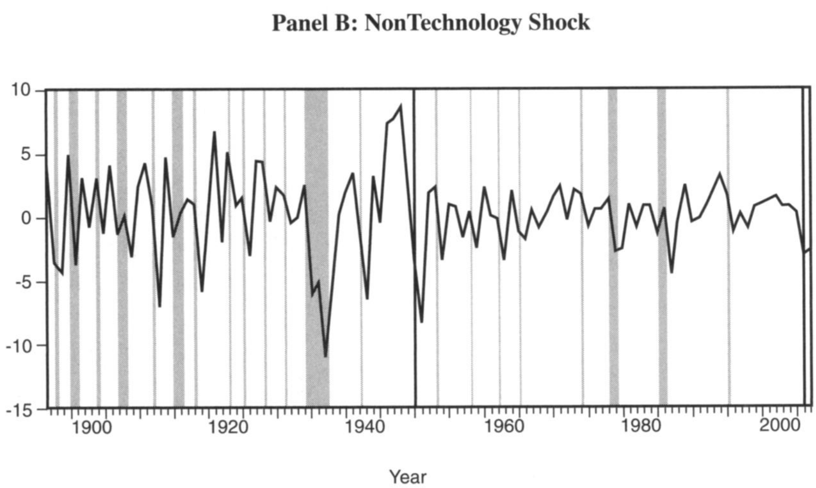


Figure A1
Plots of implied shocks from quartic trend specification
(shaded areas represent NBER recession dates)

Comment

***Harald Uhlig**, Humboldt Universität zu Berlin, Tilburg University, Bundesbank, and CEPR*

1. The Issue

What is the consequence of technological progress? To the lay person, the answer often seems obvious: technological progress is both a virtue and a vice. It is a virtue, because it has made our lives more comfortable. But it is also a vice because people are loosing jobs. Factories, where humans used to work together in the past to create the products to be sold, are now instead filled with machines and the occasional operator: the other workers are out of a job. The latter aspect is viewed as a negative aspect of technological progress and often dominates public debates.

The economist gives a dramatically different assessment. According to the data presented by Francis and Ramey, labor productivity in the private sector has increased by the factor 13.5 between 1889 and 2002. Francis and Ramey show that hours worked in the private sector per capita really has not changed all that much, once one takes certain things into account, e.g., trends in schooling or government employment. Thus, using a standard Solow growth accounting exercise, i.e., postulating that capital and output are on the steady state growth path and labor is constant, and postulating a Cobb-Douglas production function with a capital share of one-third, one finds that total factor productivity has increased by a factor of 5.7 between 1889 and 2002. Therefore, even if the (quantity of the) capital stock per worker had not changed at all, a worker now would need only 10 minutes to produce what his great-great-grandfather would have taken an hour to do. With the additional capital accumulated in the meantime, that number shrinks to less than five minutes. The welfare gains from this dramatic increase in productivity are obvious and self-evident. The virtuous aspects of technological progress dominate by far.

But what about the vice of potential lost jobs? Take again the evidence produced by Francis and Ramey, that hours worked in the private sector have not changed much over the last century. Clearly then, over the long haul, technological progress has not led to machines replacing workers: rather, the desires of humans for consumption relative to enjoying leisure have risen with the technological progress (and therefore the opportunity costs of leisure) we have made. We need to work just as hard as more than 100 years ago to keep up. And surely, even if one does not buy into the evidence produced by Francis and Ramey and rather believes that hours have shown a secular decline, few economists would interpret this as evidence that technological progress is the culprit for the high unemployment rates observed in some of the modern welfare states. Instead, the secular technological progress is viewed as freeing up time for people to enjoy as leisure. This macroeconomic perspective may be complicated by redistributional issues—some workers may benefit more from "replacement" by technological progress than others, and some may even indeed suffer from it—but overall the conclusion remains: technological progress is a virtue.

So, why should we care what technological progress—or better, surprise changes in the rates of technological—do to the labor market in the short run? After all, this is the question that Francis and Ramey seek to answer, extending the research agenda of Blanchard and Quah (1989) and Galí (1999). That research is challenging the bold claim initially made by Kydland and Prescott (1982) and others, that random fluctuations in technological progress are the cause of business cycles. (As an aside, let it be pointed out that "technology" in the Arrow-Debreu sense simply refers to the possibility of turning inputs into output, so that the term "technology shock" simply refers to any changing production function. Instead, the recent literature as well as the Francis-Ramey paper has focussed on the common language interpretation of the word "technology," i.e., patents, engines and microchips. I shall follow along for the purpose of this discussion.). That research helped to explain the challenging observation that labor productivity is procyclical. Demand-driven theories need to be worked pretty hard to cough up this and other key business cycle facts.

Some of these business cycle facts are listed in Table 1, using the Francis-Ramey data from 1889 to 2002. Note the positive correlations between output on the one hand and labor as well as labor productivity on the other, no matter how and when one measures it. However, also note the low correlations between labor and labor productivity: for

Table 1
Correlations of private sector output, private sector hours, adjusted hours (see Francis-Ramey), private sector productivity, consumption, investment and government spending. The numbers above the diagonal are for the first differences of the logs of the series, whereas the numbers below the diagonal are for the HP-filtered series, using $\lambda = 7$, see Ravn and Uhlig (2002). All data are from Francis and Ramey. Note the low correlation between labor and labor productivity, possibly explaining

Part A: Correlations for 1889–2002

	y	n	n adj.	prod.	c	x	g
y		0.82	0.83	0.71	0.73	0.59	0.20
n	0.85		0.92	0.18	0.56	0.55	0.11
n adj.	0.85	0.92		0.28	0.43	0.38	0.39
prod.	0.72	0.25	0.36		0.56	0.34	0.20
c	0.69	0.57	0.40	0.53		0.61	−0.27
x	0.55	0.52	0.31	0.34	0.70		−0.46
g	0.29	0.21	0.50	0.25	−0.29	−0.45	

Part B: Correlations for 1950–2002

	y	n	n adj.	prod.	c	x	g
y		0.85	0.83	0.53	0.85	0.89	0.18
n	0.90		0.95	0.00	0.64	0.80	0.11
n adj.	0.86	0.98		0.03	0.60	0.72	0.23
prod.	0.50	0.07	0.01		0.59	0.40	0.18
c	0.87	0.73	0.67	0.54		0.75	−0.07
x	0.91	0.81	0.75	0.46	0.84		−0.15
g	0.12	0.11	0.15	0.06	−0.19	−0.20	

Part C: Correlations for 1889–1940

	y	n	n adj.	prod.	c	x	g
y		0.83	0.83	0.74	0.84	0.79	−0.02
n	0.87		0.97	0.23	0.65	0.75	−0.05
n adj.	0.86	0.98		0.26	0.61	0.72	0.11
prod.	0.77	0.35	0.36		0.67	0.47	0.04
c	0.85	0.69	0.64	0.71		0.57	−0.17
x	0.82	0.76	0.72	0.56	0.65		−0.29
g	0.10	0.11	0.27	0.05	−0.13	−0.17	

1950 to 2002, that correlation appears to be near zero. This correlation is low and lower than the numbers usually given in the literature (see e.g., Cooley 1995). This simple statistic already sheds a lot of light on the results by Francis and Ramey. Clearly, if there are three statistics—output, labor and productivity—that are not simultaneously highly correlated with each other, two rather than one source of randomness are needed to explain most of it. Since it is labor and productivity, which show low correlation, whatever explains the movements in productivity won't explain the movements in labor and vice versa (which is essentially also what Francis and Ramey find in their VAR estimates). So, this alone merits much deeper investigation: is it really true that labor productivity and hours worked show low correlations at business cycle frequencies? This is a bold claim, which is only implicit in this paper, and which could lead to a change in our thinking about business cycles, if it holds up to scrutiny.

In light of these tables, the key business cycle question is: what explains the high labor-output and productivity-output correlations, while generating low correlation between labor and labor productivity? If one buys into the findings by Galí and now by Francis-Ramey, the explanation cannot be business cycle theories driven by technology shocks. Thus, perhaps rather than the welfare question as to whether technological progress is a good thing (it is, in practically all reasonable models), the issue at stake is: what explains business cycles.

There is a long literature criticizing the claim of the real business cycle paradigm already, though. Even if technology shocks were to lead to initial increases in labor, they may contribute little to the variance of output (see e.g. Christiano, Eichenbaum, and Vigfusson 2004; or Altig et al. 2002). Variance decompositions rather than impulse responses might be the most interesting object of investigation here: Francis and Ramey provide it in their interesting Table 4. More importantly, it is all too easy to rule out explanations of business cycles—instead, we need good, convincing theories explaining them! I keep on being surprised how easily the model by Hansen (1985)—which nowadays should be considered as a strikingly simple model—explains many of the key business cycle facts. That model sets a standard, and every graduate student in economics should learn it well. Is there a similarly beautiful, alternative explanation in sight, which works even better? I do not think there is, and even if there were, it does not seem that the profession has decided to raise that one on its shield yet as the new key paradigm of business cycles. Perhaps, too much effort has gone into shooting down a model that works ok. More work should instead go into providing a model that works better.

The issue thus cannot be as to whether technology shocks explain business cycles or not. Instead, the issue is whether one can properly identify technology shocks, using long run restrictions, and what they imply about labor movement in the short run. Let me thus turn to the contribution of Francis and Ramey. It is two-fold. First, they carefully put together a long-run data set, which is an interesting object of investigation in its own right. Second, they use a variety of VAR specifications to analyze the question at hand, thereby addressing some of the issues raised in the recent debates.

2. Long-Run Data

The first contribution of Francis and Ramey is to put together a set of long time series, and to carefully account for the hours worked in the private sector, addressing the issue of changes in education and the role of the government. This is meant to achieve three goals. First, it should provide more data to study the question at hand: clearly, this is always a good idea. The alternative (which should be pursued!) is to investigate more countries rather than longer time series, but certainly, having longer time series cannot hurt.

The second goal is to provide a more balanced view of the changes in labor input, leading to some surprising results (and as for the third goal, see the next section). Francis and Ramey make the bold claim that hours per capita in the private sector have not changed very much over more than a century; see e.g., their Figure 2B. The average for 1950 to 2002 is not even seven percent below the average for 1889 to 1940. Their adjustment of the standard labor force accounting comes from three sources.

First, rather than just considering everyone above age 16, they implicitly account for child labor by subtracting only children in school from the population above age four. But one could go further. Probably, one ought to e.g., also account for changes in the death rates in early childhood due to a variety of diseases. Figure 1 shows the decline in children death rates: unfortunately, I only had data for ages one to four rather than five to 16 at hand. Nonetheless, that figure shows a dramatic fall in death rates by a factor of 60: while two percent of all children between age one and four died in 1900, only 0.03 percent of them did so in 2000, see also the right-top panel of Figure 3 in the Francis-Ramey paper. It is plausible that there also was a dramatic decline in death rates for

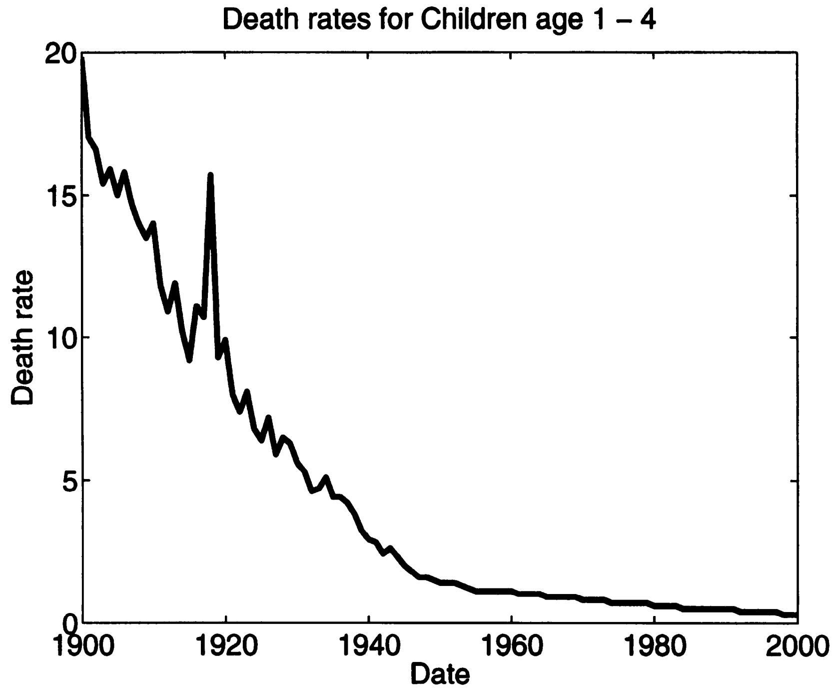


Figure 1
Death rates for children per 1000, age 1 to 4
Source: AmeriStat, Population Reference Bureau

children between the ages five and 16. Certainly, the demographics in 1889 were much more heavily tilted towards younger people and children than it was in 2000, so that one is dividing by a larger number in 1889 than in 2000. This is ok, if indeed all children above age four not in school were indeed working. But I tend to think that this is a somewhat extreme view.

Figures 2 through 4 show that this could be the case. In Figure 2, the fraction of children between the ages of five and 16 compared to the total population is plotted: one can see the large swings between nearly 25 percent at the top and slightly above 15 percent at the bottom. The numbers are very high towards the beginning of the sample and during the "baby boom" years. The numbers become more dramatic, once one subtracts out the population above 65 and government employees in the denominator: see Figure 3. Now, the numbers swing between 21 percent and 30 percent. These swings would matter a lot, if all children between age five and 16 were counted as part of the work force.

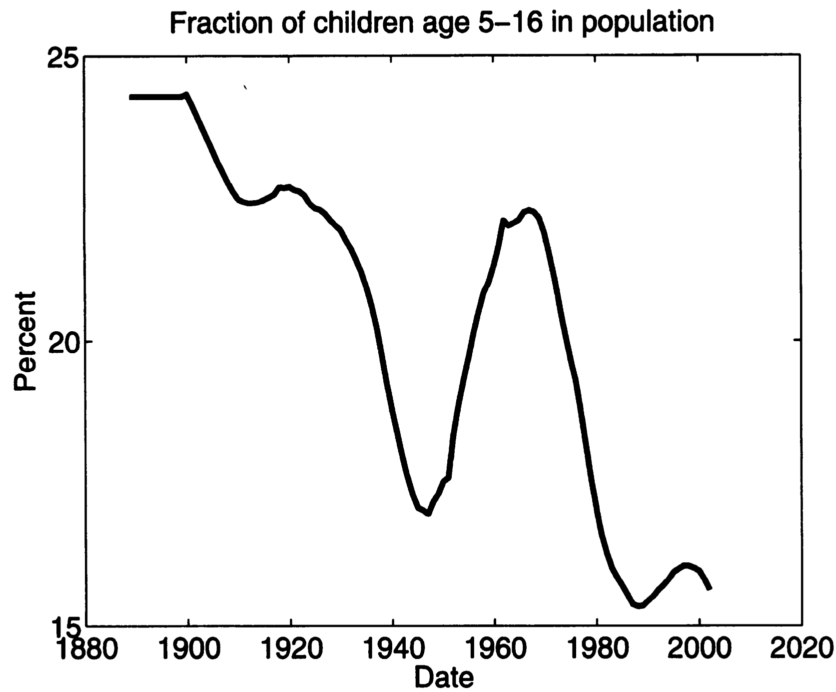


Figure 2
Fraction of children age 5 to 16 in total total population

Thus, Francis and Ramey rightly subtract school enrollment. We do this too both in the numerator and denominator in Figure 4. Now, the numbers turn negative, though: from a peak near 3 percent to the lowest point at near –18 percent. One can already in Figure 3 of the Francis-Ramey paper that this must be so: according to their numbers, nearly a quarter of the total population was enrolled in school during the last 20 years. Clearly then, this must include people who are not "too young to work." Francis and Ramey effectively make the extreme assumption, that everyone enrolled in school is not working. Given the large fractions of the population enrolled in school, one may have some serious doubts here. Finally, much of the apparent rise in hours worked in the recent 20 to 30 years is due to female labor force participation. To me, it is plausible, though, that a sizeable share of women have worked hard in the private sector previously as well, but outside official government statistics at the turn of the century. I can only hope that the original data used by Francis and Ramey has taken this properly into account, but I am skeptical.

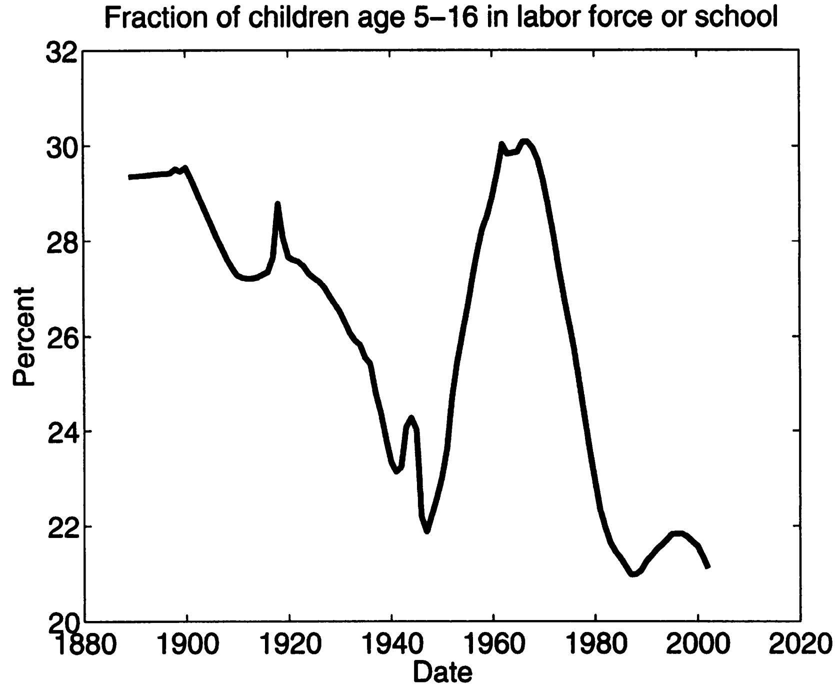


Figure 3
Fraction of children age 5 to 16 in labor force, not accounting for school enrollment, i.e., (pop - pop16 - pop4)/(pop - pop4 - pop65 - govemp)

In sum, one has to wonder whether the calculations by Francis and Ramey really are the final word on measuring the actual work force. Probably, they are not. Nonetheless, their findings are fascinating. According to conventional wisdom, hours worked per capita have declined over the last 100 years or so. Francis and Ramey have taken a first and very useful stab at that issue, and challenged that wisdom. This really calls out for an intensive analysis of the data to settle this issue.

3. VAR Estimates

The third goal of providing a longer data set is to take more confidence in imposing long-run restrictions. But here, I believe, one should be careful. Classical econometric time series analysis makes one believe that there is a big difference between an exact unit root and a root smaller than one. But this difference concerns the hypothetical exercise of considering ever longer data sets, governed by some given stochastic law

Figure 4
Fraction of children age 5 to 16 in labor force, accounting for school enrollment, i.e., (pop - pop16 - pop4 - school)/(pop - pop4 - pop65 - govemp - school)

of motion. However, a different view is more sensible. Most of the time, the data is of given length, and one has to make inferences about the various competing possibilities for the roots governing the process (see e.g., Sims-Uhlig 1991). Or—as is argued here—one is indeed provided with longer time series, but then one has a hard time believing that the stochastic properties do not change. Simple econometric models are plausibly understood to be parsimonious descriptions of key features of the data, rather than true, genuine data generating mechanisms that remain unchanged across several centuries. Given the difficulties described above in truly calculating the work force and thus in truly calculating labor productivity, the impression that more than a century of data really helps in separating the unit root parts from the non-unit-root parts may thus be just a misleading illusion. More data helps, yes. But with more data, one ought to consider a richer set of econometric possibilities. In particular, slow variations in the regression coefficients would invalidate the inference made here.

My interpretation of the evidence presented is therefore different. Evidence for a unit root in the data should be interpreted as evidence for some very persistent feature. Whether the effect gradually dies out after 100 years or whether it does not, is probably not of relevance. The "technology shock" identified in this paper is then simply that shock which leads to the most persistent changes in labor productivity among all the shocks one can consider. This is interesting.

I am concerned that it should make a difference as to whether hours are regarded as stationary or not. Since OLS provides consistent estimates of VAR coefficients, regardless of whether the variables are stationary or not, one can identify this most persistent shock certainly also in a VAR in levels, even if hours have a unit root. Conversely, if hours do not have a unit root, then a VAR which uses hours in first differences is misspecified. Christiano, Eichenbaum, and Vigfusson (2004) have nicely demonstrated how the level results encompass the results for first-differences, and that one therefore ought to trust the level results rather than the results for first-differences, regardless of what the unit root tests for labor show. I find their argument in favor of the level specification more plausible than the arguments given in the Francis-Ramey paper in favour of the first-difference specification. To put it succinctly: VARs should typically be estimated in levels, unless there are very good theoretical reasons not to. And one should use Bayesian methods, which can deal with the uncertainty regarding the presence of unit roots in a very natural and practical manner (see Uhlig 1994). Certainly, detrending with a quartic trend strikes me as something that may potentially be rather misleading. Are we allowed to extrapolate that trend towards infinity? And what does it do to inference about cause and effects in VARs, if current data is "cleansed" from a trend, which in turn is estimated with the help of future data?

The level results look particularly damaging to the benchmark real business cycle perspective, though see Francis and Ramey's Figure 5. Apparently, the permanent shocks to labor productivity—which Francis and Ramey call and identify as technology shocks—are only a minor cause of output fluctuations (see the bottom row in Figure 5) and thus, it is no surprise that they also do not do much to labor and even lead to a short, initial decline (see the middle panel). Interestingly, the sample here matters a lot. Francis and Ramey's Figure 6A shows that labor shows practically no initial response to a permanent productivity shock for postwar data, while it shows a large negative response in prewar data. The first of these findings is one of the key points in Christiano, Eichenbaum and Vigfusson (2004): the level specification overturns the findings of Galí (1999) for postwar data. So, Francis and Ramey come to Galí's rescue by showing that the level specification makes things "worse" for the real business cycle school, once one takes that level specification to prewar data. This point has also been made by myself in Uhlig (2004), using a previous version of the Francis-Ramey data set.

There are three potential conclusions one can draw: at this point, a reader should feel free to choose any one of them. First, the econometric model does not remain stable over time—which seems to me to further invalidate the whole idea of using long-run identification. Second, policy and in particular labor market regulations have changed over time, leading to different behavior (see also Galí, López-Salido and Vallés 2003). The particular change here presents a bit of a challenge. Armchair reasoning would suggest that labor markets were more flexible in prewar years than in postwar years. Furthermore, financial markets may have been less efficient back then. If so, can we think of models explaining the different responses, documented by Francis and Ramey? That strikes me as an interesting research agenda.

Third, perhaps the stochastic properties are fairly stable, but inference based on long-run identification simply is too fragile. In Uhlig (2004), I argue for a variety of reasons to rely on medium-run identification instead. The technical details are in that paper, but the idea is this: While long-run identification finds that shock (or shock direction) that explains as much as possible of the variance of the revision of the long-run (more precisely, the infinite-horizon) forecast in productivity, medium-run identification seeks that shock that explains as much as possible of that variance for some medium-run forecast revision, say, three years out. The results are in Figure 5. Now, the impulse responses for labor remain a lot more stable, which makes me want to trust these results more than the results from long-run identification. According to these results, labor does not fall much in the prewar years either in response to a technology shock. Actually, labor does not react much at all—which is effectively a restatement of the low labor-productivity correlations of Table 1.

A careful reader might also note the larger numbers. According to the data I received from Francis and Ramey, labor productivity growth has an annual standard deviation of 3.1 percent for the entire sample. The MLE standard deviation of the one-step ahead prediction error for productivity is 2.7 percent, when using four lags and a constant

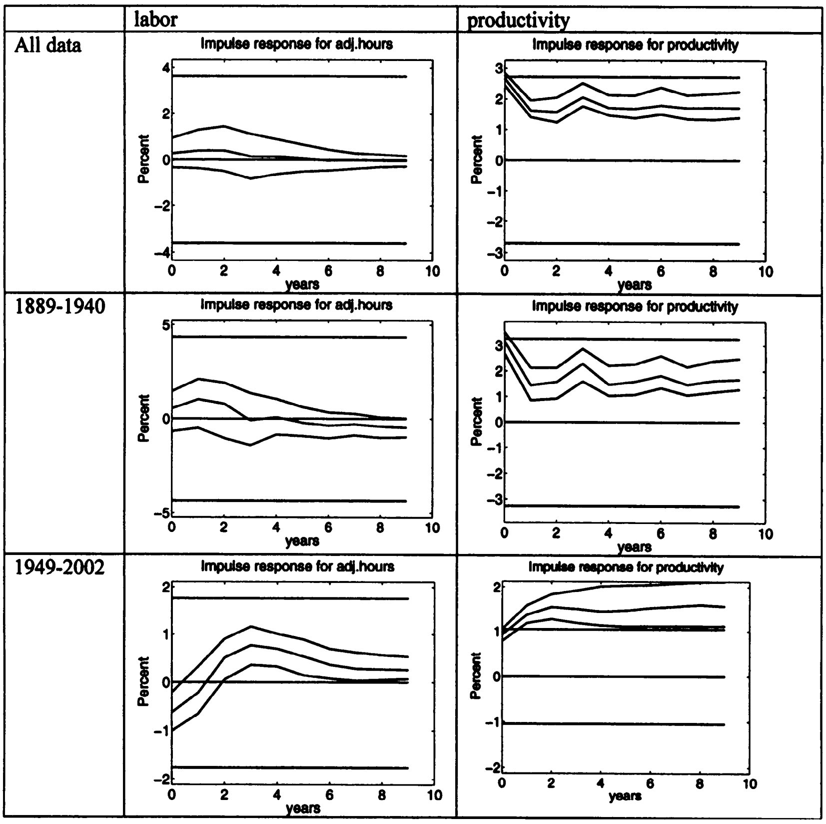


Figure 5
Results from medium-run identification, i.e., for the shock that explains as much as possible of the three-year ahead revision of the forecast for productivity. The horizontal line is the MLE of the one standard deviation of the one-step ahead prediction error for comparison

in a bivariate VAR and logs of the data. The squares of the impulse-responses of productivity at date zero should add up to this number, if one uses shocks one standard deviation in size, so I suspect that some rescaling somewhere is the cause of the difference between the results here and in Francis and Ramey. More importantly, it should be noted that this standard deviation is about three times as large in prewar data as in postwar data, pointing to changing stochastic properties across the sample. This again casts doubt on long-run identification. It also says that the Francis-Ramey results are dominated by the prewar sample, since that part of the sample contains a lot more variance than the postwar sample. Given that the data for the prewar years is probably not nearly as good as the data for the postwar years, caution may be advised in putting too much weight on these results.

Francis and Ramey do not provide error bands for their impulse responses. When I tried to recalculate their results, using my Bayesian methodology with a "very long" medium-run identification, I found a very wide error band for the prewar sample. Again, this may be more evidence for fragility of the long-run approach.

Finally, it is laudable that Francis and Ramey check that their technology shocks are not caused by government spending or M2. There is a simpler way of achieving an identification guaranteeing that: one can simply add government spending and M2 to the VAR. It would have been interesting to see the results from that exercise.

4. Theory

Neoclassical growth theory and its cousins ask one to focus on total factor productivity rather than labor productivity alone, and they ask one to consider the accumulation of capital as a key component in the low-frequency movements of labor productivity. Indeed, Chari, Kehoe, and McGrattan (2004) have recently shown how leaving away capital in the sort of empirical exercise provided by Francis and Ramey can lead to serious problems from misspecification and to potentially very misleading results, in particular, when using differenced data for hours. They show how a standard real business cycle model (where technology shocks lead to rises in hours worked) would generate data, which would deliver the results found by Francis and Ramey, that hours fall in response to a technology shock identified in a bivariate VAR from long-run restrictions. Chari, Kehoe, and McGrattan even go so far as to suggest that one should be generally skeptical about using SVARs. But one can also give their results a constructive reading and add capital into the picture: so let me pursue that.

To construct capital, I start from steady state capital in 1889, calculated private output in 1889 times the average investment-output ratio divided by $(1 - (1 - \text{delta})/g)$, where g is the average growth factor of output, i.e., the average of $y(t)/y(t-1)$, and delta is the annual depreciation rate, set to 10 percent. I calculated the investment-output ratio to be 0.25 from NIPA postwar data, where I included durable consumption with gross private domestic investment, and excluded government spending from output. For the same reason, I rescaled the raw investment series from 1889 to 2002 by the factor 1.44 in order to roughly capture the otherwise unmeasured investment in consumer durables (these rescalings do not make much of a difference in the end, since we use the logs of all series in our estimations anyways). I then calculated capital via $k(t) = (1 - \text{delta})\, k(t - 1) + \text{inv}(t)$, and used a Cobb-Douglas production function with a capital share of one-third to calculate total factor productivity. In order to also account for potential policy influences, I now estimate a VAR with six variables, i.e., TFP, adjusted hours, capital, government spending, M2 and the dividend tax rate provided by McGrattan, see Figure 8 and the comment below.

When using the entire sample and medium-run identification, seeking the shock which explains as much as possible of the variance of total factor productivity three years after the shock (see Uhlig 2004 for details), the results look remarkably reasonable (see Figure 6). In particular, hours worked move up substantially, following a technology shock identified in this manner, although the initial response is somewhat muted, and the peak response is about two years after the shock. Nonetheless, this looks very much to be in accord with standard real business cycle theory. One item that looks a little puzzling is the rather uncertain response of capital, though.

The response of capital looks much worse, when using long-run identification, though (implemented as medium-run identification applied to the 20-year ahead revision of the forecast for total factor productivity). Now, not only labor, but also capital falls after an initial increase in total factor productivity. Indeed, this is already true in the simple, bivariate VAR of labor productivity and adjusted hours, when adding the capital series: the results are in Figure 7. I suspect that Francis and Ramey would find the same, if they included capital in their econometric exercise.

This finding seems rather odd at first. It is hard to think of reasonable theories, according to which technological progress would lead to disinvestment. What then, might be going on?

A hint is already in Figure 5, possibly partly explaining even the muted response of capital in that figure: the capital income tax rate rises substantially in response to a technology shock identified from medium-run restrictions. Indeed, Figure 8 juxtaposes the output-to-capital ratio with that tax rate series, where we added 50 percent to the tax rate just in order to show both series on the same scale. First of all, that tax rate shows large and very persistent movements. Second, the output-to-capital ratio comoves with this series remarkably closely. Whether it is appropriate to interpret that tax rate as the marginal

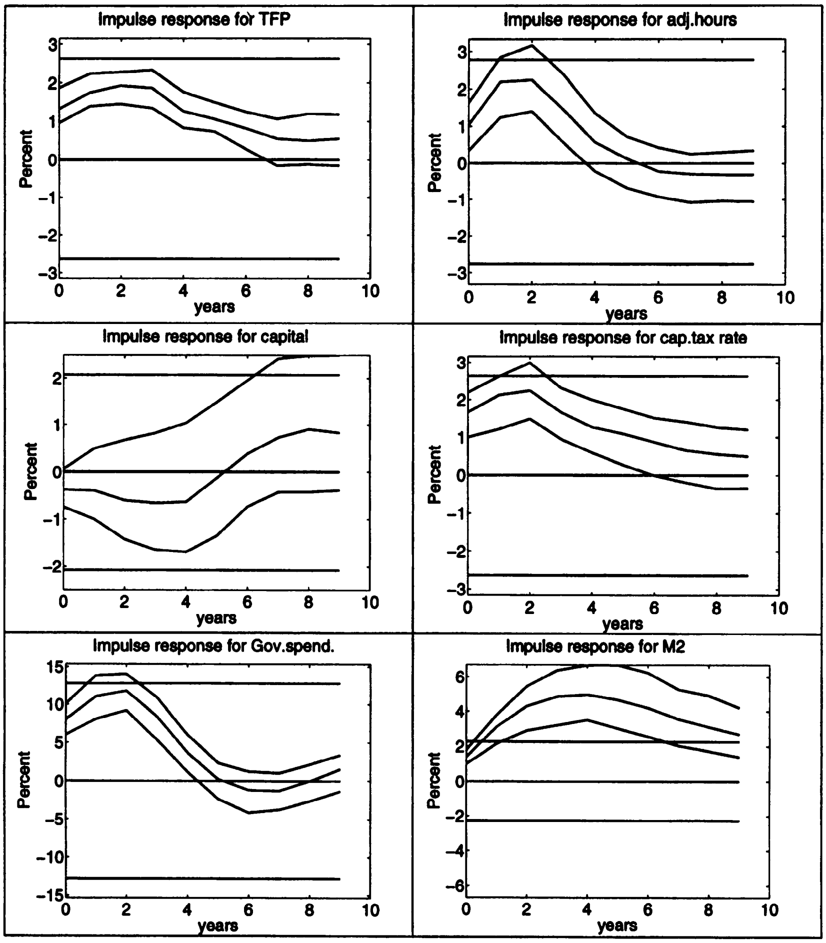


Figure 6
Impulse responses to a technology shock identified from medium run identification applied to the revision of the three-year forecast of total factor productivity, and using a six-variable VAR, including government spending, M2 and capital income tax rates

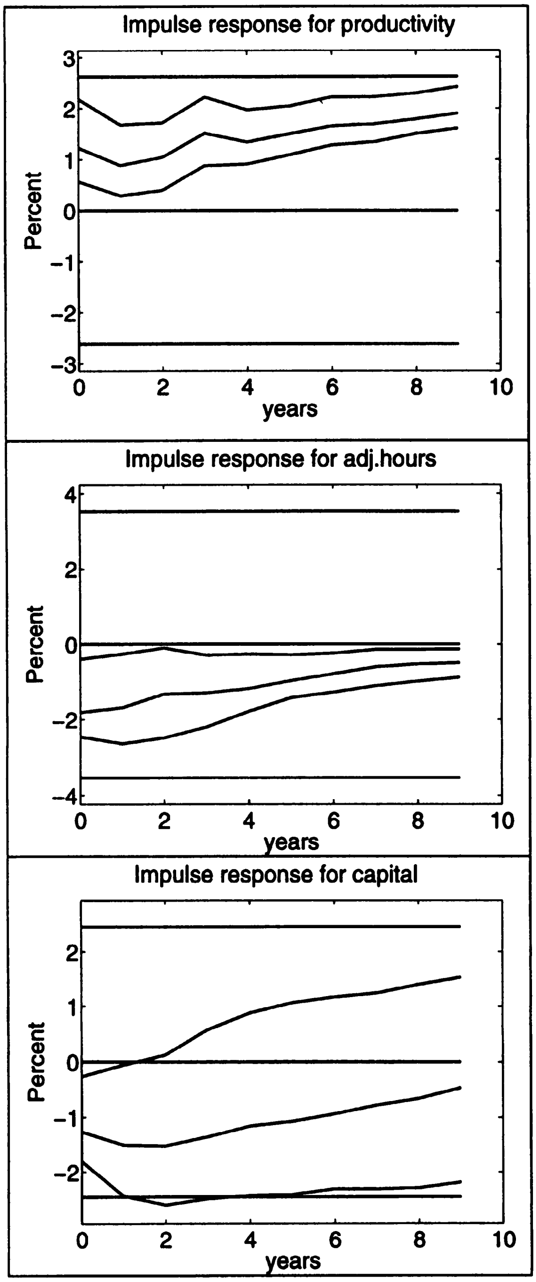


Figure 7
Results from long-run identification in a VAR with labor productivity, adjusted hours (in levels) and capital. Note, that capital falls, following a "technology shock" identified this way

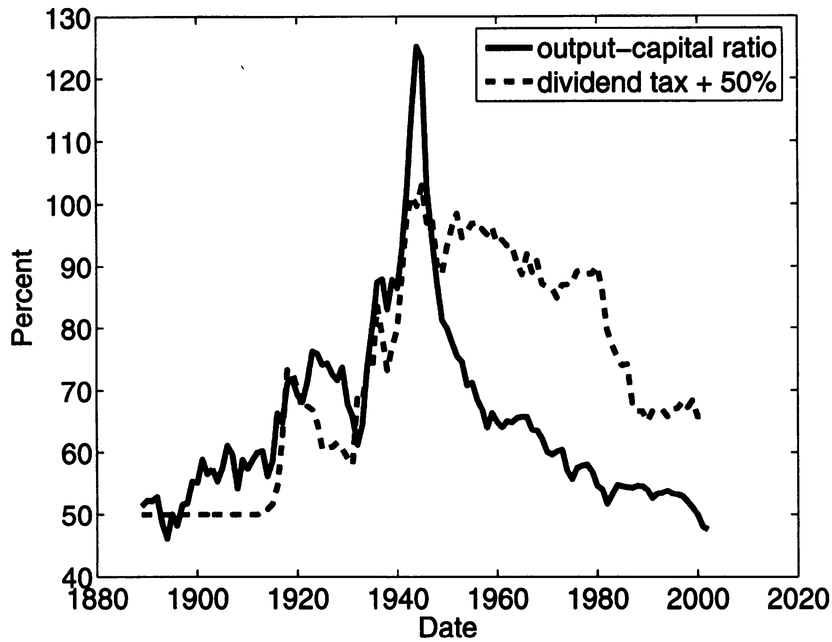


Figure 8
Juxtaposing the annual output-capital ratio to capital income tax rates (plus 50 percent in order to show both on the same scale)

dividend tax rate faced by entrepreneurs or not, does not even matter much. It is certainly plausible that this tax rate indicates the general investment climate that entrepreneurs are facing. Governments which do not hesitate to impose a high dividend-tax rate probably do not hesitate to expropriate entrepreneurs in other ways as well. So, Figure 8 is very much what one would expect, following some simple neoclassical reasoning. And persistent changes in capital income tax rates can surely lead to persistent changes in the capital-labor ratio and thus to persistent changes in labor productivity, which have nothing whatsoever to do with technology.

Francis and Ramey point to a variety of other references to rule out changes in the tax treatment of dividends as a source of their identified technology shocks. Perhaps, they are right. But given the rather dramatic shape of Figure 8, caution may be advised. Mountford and Uhlig, for example, argue that tax rate changes often are debated for a long time in the public and in parliament before they are actually implemented: so the statistical finding that technology shocks are not

Granger-caused by past changes in the capital-income tax rate may provide a very misleading picture. Case studies and additional research on this topic would help.

But capital tax rate changes may not be the only additional cause of long-run movements in labor productivity. Uhlig (2004) argues that gradual and persistent changes in the social attitude towards the work place lead to similarly persistent distortions in the way labor input is measured and thus to persistent changes in labor productivity. For many white-collar workers, the work place nowadays has become a central part of their social life, with possibilities to surf the Internet or to meet marriage partners. And finally, any endogenous growth theory makes any shock have persistent effects on labor productivity.

Clearly, the main feature of labor productivity is the fact that it is trending upwards, see the figure in Francis and Ramey. There is little doubt that this is due to technological progress. But whether the persistent random fluctuations around this trend are due to random fluctuations in technological progress also should be much in doubt. Only the simplest of theories would lead one to believe that this is so.

5. Conclusions

The paper by Francis and Ramey is an excellent and careful contribution to a growing literature, investigating whether technology shocks lead to a fall or a rise in hours worked. They provide a long-run data set which is a genuine gift to the profession and which I expect to be used a lot for a variety of purposes: certainly, I have already made ample use of it. Equipped with that data set, they show that hours worked per capita really have not changed all that much during the last 100 or so years: a bold and fascinating claim, which needs to be subjected to further research and scrutiny. Likewise, their data implies that hours and labor productivity are at most mildly correlated over the cycle, which might explain some of their findings and which could lead to a substantial rethinking about business cycles. Next, Francis and Ramey provide a careful investigation of the data, using a variety of specifications for their VAR and taking a variety of recent suggestions and criticisms on board. This is a fine and informative piece, which pushes the frontier forward.

But I do not buy into their conclusions yet. First, their historical accounting of the labor force is interesting and much more careful than what one usually sees. But for a variety of reasons, it does not seem to go far enough. And mismeasurements here imply mismeasurements in the long-run movements of labor productivity, which their methodology seeks to identify.

Second, the long-run restrictions seem to be too fragile an identification device to provide convincing conclusions. The very fact that the first and the second half of the sample look rather different already sheds substantial doubt on any strategy seeking to identify anything from long-run behavior. Medium-run identification provides more robust results, and may therefore provide a more sensible alternative.

Finally, theory suggests that one should take capital accumulation and the policy variables related to it into account. There is strong comovement in the data over the last 100-something years between the output-capital ratio and the dividend tax rate: this alone might lead to serious distortions when identifying technology shocks from long-run movements in labor productivity. Furthermore, Chari, Kehoe, and McGrattan (2004) have shown that leaving out capital in the sort of exercise performed by Francis and Ramey can lead to serious problems from misspecification. Finally, theory suggests many reasons why labor productivity fluctuates in the long-run, aside from technology. If the medium-run fluctuations in total factor productivity are more strongly dominated by technology shocks than the long-run, then medium-run identification is more informative about the impact of technology shocks. Figure 6 shows that technology shocks identified from medium-run identification seem to lead to rising rather than falling hours worked.

What we need are good theories, explaining the key business cycle facts, like the comovements between labor productivity and output, and the relative volatilities of consumption, output and investment over the cycle. The original real business cycle model has been attacked a lot, and perhaps it is false. But what should replace it? That question still remains unanswered. The evidence provided by Francis and Ramey tells us that we need to think even harder about the answer, even if that evidence is not conclusive enough either.

Note

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649, "Economic Risk." I am grateful to Neville Francis and Valerie A. Ramey for sharing their data with me.

References

Altig, David, Lawrence J. Christiano, Martin Eichenbaum and Jesper Linde. 2002. "An Estimated Dynamic General Equilibrium Model for Monetary Policy Analysis." Working Paper. Northwestern University.

Blanchard, Olivier J., and Danny Quah. 1989. "The Dynamic Effects of Aggregate Demand and Supply Disturbances." *American Economic Review* 79: 655–673.

Chari, V.V., Patrick Kehoe, and Ellen R. McGrattan. 2004. "Are Structural VARs Useful Guides for Developing Business Cycle Theories?" Draft, Federal Reserve Bank of Minneapolis.

Cooley, Thomas F., ed. 1995. *Frontiers of Business Cycle Research*. Princeton, NJ: Princeton University Press.

Christiano, Lawrence J., Martin Eichenbaum and Robert Vigfusson. 2004. "The Response of Hours to a Technology Shock: Evidence based on Direct Measures of Technology." *Journal of the European Economic Association* 2(2–3): 381–395.

Galí, Jordi, David López-Salido and Javier Vallés. 2003. "Technology and Monetary Policy: Assessing the Fed's Performance." *Journal of Monetary Economics* 50: 723–743.

Hansen, Gary D. 1985. "Indivisble Labor and the Business Cycle." *Journal of Monetary Economics* 16: 309–327.

Kydland, Finn E., and Edward C. Prescott. 1982. "Time to Build and Aggregate Fluctuations." *Econometrica* 50(6): 1345–1371.

Mountford, Andrew, and Harald Uhlig. 2004. "What are the Effects of Fiscal Policy Shocks?" Draft, Humboldt Universität zu Berlin.

Ravn, Morten, and Harald Uhlig. 2002. "On Adjusting the HP-Filter for the Frequency of Observations." *Review of Economics and Statistics* 84(2): 371–376.

Sims, Christopher A., and Harald Uhlig. 1991. "Understanding Unit Rooters: A Helicopter Tour." *Econometrica* 59(6): 1591–1599.

Uhlig, Harald. 1994. "What Macroeconomists Should Know about Unit Roots: A Bayesian Perspective." *Econometric Theory* 10(3/4): 645–671.

Uhlig, Harald. 2004. "Do Technology Shocks lead to a Fall in Total Hours Worked? " *Journal of the European Economic Association* 2(2–3): 361–371.

Comment

Susanto Basu, *Boston College and NBER*

Francis and Ramey have written a very interesting paper that covers a lot of ground in disarmingly straightforward fashion. In fact, this paper makes three important contributions to the large and fiercely contentious literature that tries to estimate the short-run effects of technology shocks. First, it contributes a new time series on hours per worker, based on a new measure of the population of "potential workers." This series will find many applications beyond the current paper. Second, as advertised, it uses data that cover a long sample period, and also reports results for some natural sub-periods. Third, it discusses new models interpreting the finding that hours worked fall temporarily following a positive technology shock. Finally, this reviewer, at least, appreciates the very clear discussion of estimation method and identification. It is a refreshing change from the discussion found in much of the preceding literature, which is opaque at best.

The demographic adjustment to population is an excellent idea. It is important for a variety of macroeconomic debates—for example, the discussion over whether hours per worker should be modeled as constant in the long run, as required by the famous King-Plosser-Rebelo (1988) utility function. However, the graphs in Figure 3 suggest that the education adjustment is most important for reaching the conclusion that per-capita hours worked have not declined since 1960. But this adjustment seems less defensible than the others—if working for pay is a feasible choice but people choose to get education instead, it is not clear why they should be excluded from the population that is eligible for work. On the other hand, excluding these workers may be the right choice for cyclical purposes if entering and leaving the workforce has significant fixed costs. In that case, those who are committed to receiving a full-time education may not stop to work temporarily, and then are not "employment-eligible" if there is only a small, transitory shock to wages. This particular adjustment is intriguing, but merits more investigation.

The second contribution of the paper is to use the demographically-adjusted series to revisit the debate on the short-run effect of technology shocks. Here Francis and Ramey uncover an interesting puzzle. As the paper notes, previous research using the same methodology on postwar U.S. data finds that hours worked fall after a positive technology shock if hours per worker are entered in levels, but rise if hours per worker are entered in differences. Using their full sample of annual data from 1892–2002, Francis and Ramey find just the opposite result! Moreover, they find that this result is not stable over sub-samples. In the pre-war data, the result matches the finding for the full sample. But in the post-war data, all transformations of hours per worker imply that hours fall after a technology improvement (although only the results for the unit-root specification are significant).

Before trying hard to interpret these results, should we believe that they represent the reaction of the economy to actual technology shocks? Various arguments have been advanced to show that the structural VAR (SVAR) method may confound technology shocks with other permanent shocks to labor productivity. For example, Uhlig (this volume) suggests that changes in capital tax rates would have the same effects as shocks to technology; Francis and Ramey note that the same is true of changes in government subsidies to education.

There is some reason to believe that impulse responses estimated in the postwar sample are the responses to true technology shocks. The reason is that the estimates are quite similar to those obtained by Basu, Fernald and Kimball (2004) [BFK] using a completely different identification strategy. BFK use a method that is less elegant and far more data-intensive, but perhaps more direct: They estimate sectoral Solow residuals, correcting for increasing returns to scale and variable utilization of capital and labor. The aggregate technology shock is the average of the sectoral residuals. These two methods are complementary, because the weaknesses of one are the strengths of the other. For example, the Solow residual method is robust to changes in capital taxes and education policy, since the capital stock and educational composition of the labor force are entered directly as controls in the regressions. On the other hand, long-run identification might be better able to control for classical measurement error in inputs, which is a problem with the Solow residual method.

Due to the lack of detailed data for constructing sectoral Solow residuals, there is no such easy cross-check for the prewar period. And there is a well-known issue which raises a real concern that the prewar "technology shocks" are in fact something else. It is well known that over the period of the late 19th century to World War II, there was a major, permanent outflow of workers from agriculture to manufacturing. Suppose it is the case that hours worked in agriculture were not fully reported, and neither was much of agricultural output, which was often grown for home consumption. Both are quite likely to be true in an agrarian economy full of family farms. Suppose furthermore that the measurement problem is more severe for agricultural output than for agricultural hours—that is, that labor productivity in agriculture is under-measured. Then a movement of workers from agriculture to formal employment with no change in technology would be measured as a permanent increase in labor productivity, which the VAR procedure would misinterpret as a positive technology shock. Furthermore, it would appear that this "technology shock" was leading to a permanent *increase* in hours worked—just as Francis and Ramey find for the prewar period in their preferred specification, but not for the postwar period. This hypothesis is sufficiently intriguing that it might be interesting to do a cross-check of the estimated technology shocks with migration data in the prewar period.

Whether or not the hypothesis I've just proposed is correct, the apparent permanent response of hours to a technology shock in the prewar period deserves more investigation. Francis and Ramey concentrate on the differences in the short-run response, and suggest that changes in the autocorrelation of technology shocks between the periods might account for the difference. But the short-run dynamics of the technology process should not influence the long-run response of hours to a permanent change in technology. Thus, Francis and Ramey's preferred interpretation does not account for all the facts. In a representative-worker framework, accounting for the differences between the two periods also requires assuming a change in preferences. Furthermore, it requires assuming that preferences in the prewar period were such that the substitution effect of higher real wages from a technology improvement dominated the income effect. Of course, if that were true then work hours should have risen significantly from 1890–1930, when real wages rose steadily. But Francis and Ramey's own Figure 2B shows that this was not the case—they find that average hours worked were roughly the same on the eve of the Great Depression as they were 40 years earlier.

In sum, Francis and Ramey have contributed a valuable paper and a useful adjustment to a widely-used data series. Their results for the post-war period are quite consistent with a variety of previous work. But their prewar results are a puzzle worthy of further investigation.

Reference

Basu, Susanto, Fernald, John G., and Kimball, Miles S. 2004. "Are Technology Improvements Contractionary?" Working Paper no. 10592. Cambridge, MA: National Bureau of Economic Research.