

Measures of Per Capita Hours and their Implications for the Technology-Hours Debate

By

Neville Francis
University of North Carolina, Chapel Hill

and

Valerie A. Ramey
University of California, San Diego
National Bureau of Economic Research

December 27, 2006

Abstract

Structural vector autoregressions give conflicting results on the effects of technology shocks on hours. The results depend crucially on the assumed data generating process for hours per capita. We show that the standard measure of hours per capita has significant low frequency movements that are the source of the conflicting results. HP filtered hours per capita produce results consistent with those obtained when hours are assumed to have a unit root. We provide alternative measures of hours per capita that adjust for low frequency movements in government and nonprofit employment, as well as the age composition of the population. When the new measures are used to determine the effect of technology shocks on hours using long-run restrictions, both the levels and the difference specifications give the same answer: hours decline in the short-run in response to a positive technology shock.

We are grateful to Shawn Sprague at the Bureau of Labor Statistics for providing the raw series used in the BLS hours index. We thank Mark Gertler, Clive Granger, James Hamilton, Mark Watson, seminar participants at the NY Fed, and two anonymous referees for their helpful comments. Valerie Ramey gratefully acknowledges financial support from National Science Foundation grant SES-0617219 through the NBER.

I. Introduction

The role of technology shocks in business cycle fluctuations has recently received considerable attention. A myriad of papers has emerged on this topic addressing the controversial conclusion reached by Galí (1999) that *technology shocks cannot be the main driving force behind cyclical movements in macroeconomic data*. This conclusion challenges the core of the long-standing Real Business Cycle (hereafter RBC) theory, thus, it comes as no surprise that so many recent papers have been written either in defense of or to challenge Galí's findings (see Galí and Rabanal (2004) for a review of the literature).

Standard RBC theory teaches that all factor inputs should rise when there is a positive technological innovation. However, recent empirical tests of the theory find that labor input falls in response to a positive shock to technology, a finding which has sparked a debate for the last five and a half years with little resolution. The crux of the debate has to do with the data generating process assumed for per capita labor input in empirical models. If one were to rely on econometrics, which fails to reject the presence of a unit root in per capita labor, one would be led to enter labor input in first differences when estimating a (structural) vector autoregression (VAR). Entered in differences the results of a typical VAR predict a fall in labor input in response to a positive shock to technology, opposite of that predicted by the standard RBC model. However, *common sense* tells us that per capita labor being a bounded series cannot have a unit root. For this reason, several papers have assumed per capita labor is stationary and, thus, should enter the VARs in levels. When entered in levels the standard result emerges that labor input rises when there is a positive innovation to technology.

In this paper, we show that there are significant (low frequency) demographic and institutional movements, over the postwar period, that are features of the commonly used measure of hours worked per capita. Our premise is that these low frequency movements have nothing to do with the kinds of *technology shocks* typically modeled in RBC theory. These low frequency movements in the standard measure distort unit root tests (which have low power to begin with), make the time series for per capita labor inconsistent over time and with RBC theory, and are the source of conflicting results in the levels versus first difference debate. We begin by showing that extraction of these low frequency movements using an HP filter produces results similar to those obtained using first-differenced hours. We then devise new measures of hours per capita that adjust for demographic and institutional changes involving government and nonprofit employment and the age structure of the population. The new series are virtually free of low frequency movements and come closer to their model-generated counterpart than the standard measure used in the literature.

The absence of significant low frequency movements in our new measures has important implications for the technology-hours debate. In particular, our new measures provide consistent implications for the role of technology shocks in business cycle fluctuations. Positive technology shocks, identified with long-run restrictions, lead to a short-run decrease in hours worked regardless of the stationary assumption made for per capita labor.

Our results also have broader implications beyond the technology-hours debate. For example, our analysis of institutional and demographic changes explains the findings of Kahn and Rich (forthcoming) and Fernandez-Villaverde and Rubio-Ramirez (forthcoming). Kahn and Rich use a regime switching dynamic factor model to detect breaks in trend productivity growth. They show that there are two important sources of low frequency movements in per capita

macroeconomic variables: technology and slow movements in labor supply that look like preference shifts. Similarly, Fernandez-Villaverde and Rubio-Ramirez estimate a DSGE model and find that low frequency movements in the preference parameter are an important part of fluctuations. Our results suggest that the sources of these movements are demographic changes in the “representative household” and institutional changes in the sectors of employment, not true preference shocks.

II. The Problem of Low Frequency Movements in the Standard Series of Hours Per Capita

Growth theory and RBC models are generally written in terms of a representative agent’s consumption, work, and leisure. To match the representative agent in empirical applications, macroeconomic variables are measured in per capita terms. Theory rarely specifies how “per capita” should be measured, yet virtually all RBC empirical applications measure “per capita” as the BLS series on the civilian noninstitutional population ages 16 and over (e.g. King, Plosser, Stock and Watson (1991), Burnside and Eichenbaum (1996)). The omission of the military, children younger than 16, and the institutionalized population is based on the desire to measure “per capita” as the available workforce rather than the entire population. To measure hours worked per capita, researchers typically use the BLS index of hours worked in private business or the narrower index of hours worked in private nonfarm business.

According to RBC models with standard preference specifications, the hours worked per capita variable should be stationary in the absence of permanent shifts in government spending, labor income taxes, and preference shifts. Yet the most widely used measure of private hours per capita shows significant low frequency movements. Figure 1 shows the behavior of private

hours divided by the civilian noninstitutional population ages 16 and over during the post-WWII period. Hours show a U-shape, with a downward trend until the mid-1970s, which partially reverses by 2006. The peak in 1973 is 13.6 percent below the peak in 1948. The low frequency movements are so pronounced that the series does not return to its mean for decades at a time.

While these low frequency features are not an issue for analyses that HP-filter the data before analyzing it, they are very problematic for structural VARs in which assumptions about stationarity are key parts of the identification. In particular, these low frequency movements in hours per capita have important implications for empirical structural VAR models that identify technology shocks using long-run restrictions. Based on the results of standard unit root tests, Francis and Ramey (2005) assume that hours per capita have a unit root, and thus enter hours in first differences in the model. They find that a positive technology shock leads to a decline in hours worked. In contrast, Christiano, Eichenbaum and Vigfusson (CEV, 2003) argue that hours per capita cannot logically have a unit root, and offer alternative empirical tests against a unit root. They enter hours in levels and find that a positive technology shock leads to a rise in hours worked.¹

To illustrate, we re-estimate the structural VAR used by Galí, Francis and Ramey, and Christiano, Eichenbaum and Vigfusson using the standard measure of hours. In the baseline bivariate case, we estimate the following system:

¹ This literature has generated a further controversy about whether these VARs can capture the results from the model. In particular, Chari, Kehoe, and McGrattan (2005) show that they can generate data from a model in which technology shocks have a positive effect on hours yet the VAR shows them to have negative effects. Christiano, Eichenbaum and Vigfusson (2006), however, show that the Chari, Kehoe, McGrattan example is an anomaly, being at odds with the data. Erceg, Guerrieri and Gust (2004) and Francis, Owyang, and Roush (FOR, 2005) also show that VARs applied to artificial data from RBC models are consistent with the simulated results of the underlying model.

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

x_t denotes the log of labor productivity, n_t denotes the log of hours per capita, ε^z denotes the technology shock, and ε^m denotes the non-technology 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.

This system applies to the case in which hours are assumed to be stationary. We also estimate a system in which hours are assumed to have a unit root:

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

We impose the same restriction, that $C^{12}(1) = 0$, to identify the technology shock. ***It is important to note that one can identify the technology shock whether hours are stationary or nonstationary.*** However, that identification of the technology shock is sensitive to the correct specification of hours, as we shall see below.

In the baseline case, we use four lags and limit our attention to a bivariate system. The data are quarterly and extend from 1948:1 through 2006:3. The standard error bands are 95 percent confidence bands based on bootstrap standard errors with 2000 replications.

Recall the previous summary of the literature. Using standard measures of hours per capita, the specification with stationary hours implies that hours increase significantly in

response to a technology shock. In contrast, the specification with a unit root in hours implies that hours fall significantly in response to a technology shock. This pattern can be seen in Figure 2 where we use the standard measure of hours per capita with the civilian population 16 and over as the population measure. The first column shows the results from the system with hours per capita in levels and the second column shows the results from the system estimated with hours per capita in first differences. The model is bivariate in the logs of labor productivity and hours, but we also show the implied effects for the log of output, since it is equal to the sum of the other two variables. The graphs display the same conflicting results from the literature.

Is the over-differencing of hours per capita leading to erroneous results or are the low-frequency movements in the level of hours per capita leading to misleading results in the levels specification? To investigate the plausibility of these explanations, we remove the very low frequency movements in hours per capita using a very conservative HP filter with a λ parameter set equal to 160,000 rather than the usual 1600 for quarterly data. Figure 3 shows the estimated trend. It displays a pronounced U-shape, with the highest part in the early part of the sample. We then use the detrended hours series in the bivariate SVAR model, both in levels and first-differenced.

Figure 4 shows the estimated impulse response functions using HP filtered hours per capita, with the results from the levels specification on the left and the results from the first-differenced specification on the right. Interestingly, both specifications imply that a positive technology shock leads to a decline in hours in the short-run, consistent with Galí's (1999) finding and Francis and Ramey's finding. Although one would suppose that the difference

specification is plagued by over-differencing when HP filtered hours are used, the results are quite similar to those when the filtered hours levels are used.²

These results support Fernald's (2006) contention that the coincidental U-shape in both productivity growth and the standard measure of hours per capita is driving CEV's finding of a positive response of hours. When Fernald removes the U-shape in productivity growth, but leaves the U-shape in the standard measure of hours per capita, he finds a negative effect of technology shocks on hours. Conversely, when we eliminate the U-shape in hours per capita by removing the low frequency component, but do not allow for structural breaks in labor productivity, we also obtain the same negative response.

One might worry, though, that HP filtering the data could distort the dynamics. Or, perhaps the HP filter is simply a crude way to correct the standard hours measure for demographic and institutional changes it does not capture. In the next section, we carefully document and control for important demographic changes in per capita hours over postwar period and find that our demographically-adjusted and HP-filtered hours results are remarkably close.

III. The Effects of Demographic and Institutional Trends

Many researchers have used private business hours per capita as their hours measure, and then appealed to the standard assumption that income and substitution effects cancel to argue that this measure should be without trend. In this section, we show that key demographic and institutional changes during the post-WWII period invalidate that assumption. First, we show the empirical importance of changes in government employment and nonprofit employment for total working hours. Second, we show how the age composition of the population can affect the

² The results are similar when an HP filter with standard parameter values is used.

average hours worked per capita. We then present a theoretical model that shows how trends in these variables can lead to trends in the standard hours per capita measure, even when leisure itself has no trend. Our aim is to develop a new measure of hours per capita that adjusts for demographic and institutional trends that are not captured by standard RBC models.

This paper builds on earlier work using annual data for the entire 20th Century (Francis and Ramey (2006)). This paper differs in three key ways. First, we present a theory that shows how to properly adjust for these trends (the procedure used in the earlier paper was not theory-consistent). Second, we use new, unpublished data on hours worked in the non-business sector in our new measure of hours per capita. Third, we investigate the effects in quarterly rather than annual data, so time aggregation is less of a problem for identification.

A. Empirical Evidence on Demographic and Institutional Trends

1. Hours worked outside of private business

As discussed above, the standard measure of hours includes only private business hours because this is the only hours series published by the BLS. The BLS focuses on this sector because the main use of its hours measure is to estimate productivity in private business. This hours series, however, neglects important trends in hours worked in two other sectors: government and nonprofit. The omission of these additional hours from the standard hours series distorts the estimate of actual hours worked and induces significant trends in the series.

To illustrate the effects, we obtained unpublished quarterly data from the BLS on all sectors of the economy from 1948 to 2006:3. Figure 5A shows the behavior of hours worked in government as a fraction of total hours worked in private business, government, the nonprofit

sector, and private households.³ Government hours rise from below 10 percent in 1948 to over 17 percent in the late 1970s and then fluctuate between 15 and 16 percent for the last thirty years. Note that the increase in government hours as a percent of total hours worked from 1948 to the 1970s mirrors the decline in private business hours per capita shown in Figure 1. One might think the standard measure compensates in part because it omits the military from both the numerator and the denominator. However, the level and trends in military hours are relatively unimportant. After the rise during the Korean War to five percent, military hours as a fraction of total hours have trended downward, with a small blip during the Vietnam War. Currently, military hours are only one percent of total hours worked.

Figure 5B shows hours in the nonprofit sector as a fraction of total hours worked. This series does not count volunteer hours, only paid hours. The fraction of hours worked in the nonprofit sector rises steadily from 2.5 percent in 1948 to 8 percent in 2006. Much of this increase is due to the growth of the medical care and education sectors, since half of the hospitals and universities are categorized as nonprofit.⁴

Figure 5C shows private business hours as a percent of total hours. Hours worked in private business represented 86 percent of all hours worked in 1948, but have fallen to only 76 percent of all hours worked. It is clear that the trends in private business hours per capita are in part due to its falling share of total hours worked.

2. Age Composition of the Population

³ Hours worked in private households are very small. They were 1.75 percent of total hours worked in 1948 and fell to 0.5 percent by 2006.

⁴ See Salaman and Solokowski (2005) for a description of the growing importance of the nonprofit sector.

The standard measure of hours uses the civilian noninstitutional population ages 16 and over. Labor supply behavior, however, differs significantly across ages within this category. Figure 6 uses Census data to show the average weekly hours worked by age group for both 1950 and 1990.⁵ The graph shows that those under age 22 and those over age 65 work much less than prime age individuals between ages 22 and 64. In 1990, individuals between the ages of 55 and 64 also worked less than those between 22 and 54.

The main reason hours worked by individuals between 16 and 21 are low is schooling. Since the returns to human capital are the greatest when investments are made while young, it is individuals in this age group who are most likely to delay market work in order to spend time in school. Moreover, over time the average number of years spent in school has increased (Francis-Ramey (2006)).

The hours of individuals ages 65 and over are low for a variety of reasons. First, a higher fraction of individuals in this age group have health conditions that make working more difficult. Second, Social Security and Medicare have substantially changed the incentives to work when old. Both programs have become more generous over time, so the incentives to work when old have become even less. Third, during a substantial part of the post-WWII sample, mandatory retirement constrained the labor supply of the older population. During the early part of the sample, mandatory retirement constraints were not as binding because a larger fraction of the population was self-employed in family businesses. As self-employment decreased, the mandatory retirement constraints increased. Because of the set of institutional incentives and constraints on individuals 65 and over, many researchers do not include older individuals in the working age population (e.g. Prescott (2004)).

⁵ We use 1990 rather than 2000 because the question regarding hours worked changed in the 2000 Census. The data appendix gives details on the data construction.

One other point of interest from the graph is that hours worked by prime age individuals were significantly higher in 1990 than in 1950. Although average hours of work for men decreased somewhat over this time period, average hours of work for women increased dramatically. Hence the average across both groups has increased.

If the age composition of the population stayed constant, these differences in hours worked by age group would not lead to trends in per capita hours. However, as Figure 7 shows, the shares of these age groups have changed over time. Figure 7A shows the fraction of the population (ages 16 and over) that is between ages 16 and 21. The effect of the baby boom is very clear. The fraction rose from eleven percent in mid 1950s to almost sixteen percent in the mid 1970s, and then fell so that it is now below eleven percent. Figure 7B shows the steady increase in the fraction of 16 and older population that is ages 65 and older. This fraction has risen from ten percent to sixteen percent during the post-WWII period. Figure 7C shows the fraction of the population ages 16 and over that is between the ages of 22 and 64. It starts above 76 percent at the start of the sample, falls to 70 percent in the mid-1970s, and is now at 73 percent. Note the similarity of this pattern to the HP filtered trend of the standard measure of hours per capita shown in Figure 3.

B. A Theoretical Model with Demographic and Institutional Trends

The last section showed several major trends that we would expect to affect the low frequency properties of the standard measure of hours per capita. We now present a simple generalization of the standard macro model to demonstrate how these trends affect private business hours per capita, as well as to guide us on how we should adjust the standard measure to eliminate these low frequency movements. The model builds on work by Finn (1998), and

Cavallo (2005), which explicitly takes into account that an important part of government spending is value added in the government sector. We model the production technologies in the government sector differently than they do, however.

Consider a representative household that maximizes the present discounted value of utility:

$$(1) \quad \sum_{t=0}^{\infty} \beta^t [\ln(c_t + \mu g_t) + \phi \ln(l_t)] N_t, \quad \text{with } l_t = 1 - h_{pt} - h_{gt} - h_{nt}$$

where c is per capita consumption, g is per capita government expenditures, l is per capita leisure, h_p is the fraction of time worked in the private business sector, h_g is the fraction of time worked in government, h_n is the fraction of time worked in nonprofit organizations, and N is the working-age population. We assume that $0 \leq \mu \leq 1$, and $\phi > 0$. If $\mu = 1$, then government expenditures are perfect substitutes for private goods in consumption.

The standard empirical implementation of the model assumes that h_p , hours worked in private business, is the only use of hours. Here we have augmented the model with government work and nonprofit work. Ramey and Francis (2006) also consider home production, which is an important source of time use. We do not include home production here because most estimates of home production are only available at ten year intervals, so we know little about the shorter-run fluctuations.

In order not to introduce complications that have little effect on our main results, we specify the rest of the economy as parsimoniously as possible. In particular, we assume that private business output is given by:

$$(2) \quad Y_{pt} = A_t H_{pt}^{\alpha} K_{pt}^{1-\alpha}$$

and nonprofit output is given by:

$$(3) \quad Y_{nt} = A_t H_{nt}^\alpha K_{nt}^{1-\alpha}$$

and government output is given by:

$$(4) \quad Y_{gt} = A_t H_{gt}^\alpha K_{gt}^{1-\alpha}$$

where $H_j = h_j N$. Note that we are assuming the same production functions for private business, nonprofit, and government output, so this is essentially a one-sector model.

The economy's resource constraint is given by:

$$(5) \quad Y_{pt} + Y_{nt} + Y_{gt} = C_t + I_{pt} + I_{gt} + I_{nt} + G_t$$

with capital accumulation equations:

$$(6a) \quad K_{pt+1} = (1 - \delta)K_{pt} + I_{pt}$$

$$(6b) \quad K_{gt+1} = (1 - \delta)K_{gt} + I_{gt}$$

$$(6c) \quad K_{nt+1} = (1 - \delta)K_{nt} + I_{nt}$$

We assume that capital is always allocated efficiently across the private and government sectors, so that the relative price of the three types of output is always unity. It is clear that the government and nonprofit sectors are paying market wages in the actual economy since individuals are freely choosing to work in those sectors (with the exception of the draft, which represents a small part of employment). Given that these sectors are constrained to pay the market wages, setting the same capital-labor ratio as the private business sector minimizes costs.

This model contains an important feature of government spending highlighted by Finn (1998) and Cavallo (2005): much of what we call “government consumption expenditures and

gross investment” is actually produced by the government. On average since 1947, government value added has constituted 60 percent of government expenditures.

Government expenditures and hours are assumed to be chosen for political reasons outside the model (wars, for instance), and hence are taken as exogenous. Since government pays the market-clearing wage, individuals are indifferent between working in the government or private sector. It is also assumed that government taxation is lump-sum, so that no distortions are introduced.⁶

We also assume that the relative size of the nonprofit sector is exogenous. However, like the government sector, it pays the market wage and market rental rate on capital so individuals are indifferent between working in the business versus nonprofit sector. Since we assume that its output is a perfect substitute for private business output, the size of this sector has no effect on the economy except for the number of employees classified as “private business” versus nonprofit.⁷

We transform the model by dividing key variables by $A_t N_t$ so that we can analyze steady-states. Let smaller case letters stand for per capita variables and smaller case variables with hats stand for the transformed variables, e.g., $k_p = \frac{K_p}{N}$, $\hat{k}_p = \frac{K_p}{A N}$. Also, let $\eta = (1+n)(1+\gamma)$, where n is the mean growth rate of the working age population and γ is the mean growth rate of technology change. The key steady-state equations for our model are:

$$(7) \quad 1 - \delta + (1 - \alpha) \left(\frac{\hat{k}_j}{h_j} \right)^{-\alpha} = \frac{\eta}{\beta} \text{ for } j = p, n, g \text{ as well as for the aggregate}$$

⁶ See Francis-Ramey (2005) for an analysis of the effect of distortionary taxes on steady-state hours.

⁷ The fact that all three sectors are important providers of medical care and education in reality suggests that this is a reasonable assumption.

$$(8) \quad \frac{Y_j}{N h_j} = A \left(\frac{\hat{k}_j}{h_j} \right)^{1-\alpha} \quad \text{for } j = p, n, g \text{ as well as for the aggregate}$$

$$(9) \quad \frac{c + \mu g}{1 - h_p - h_n - h_g} = \frac{\alpha}{\phi} \frac{y_j}{h_j} = \frac{\alpha}{\phi} \frac{y}{h_p + h_n + h_g} \quad \text{for } j = p, n, g$$

$$(10) \quad \frac{i_j}{y_j} = (\eta - 1 + \delta) \left(\frac{\hat{k}_j}{h_j} \right)^\alpha \quad \text{for } j = p, n, g \text{ as well as for the aggregate}$$

$$(11) \quad \frac{c + \mu g}{y} = 1 - \frac{i}{y} - (1 - \mu) \frac{g}{y}$$

where y is the sum of the outputs of each sector divided by population and i is the sum of the investments of each sector divided by population.

It is clear from equation (7) that the capital-labor ratio in each sector is determined by the parameters of the model, and is thus invariant to changes in the size of the sectors. Equation (8) shows that this model is consistent with the identifying assumption used in the empirical work: the only variable that has an impact on steady-state labor productivity in any of the sectors is technology, A , since the capital-labor ratio is constant. Furthermore, the result does not depend theoretically on whether hours per capita are stationary or not. Equation (9) is the marginal rate of substitution condition that sets the ratio of the marginal utility of leisure to the marginal utility of consumption to labor productivity, which is equal to the real wage. Manipulating equations (9)-(11), one can show that:

- The only variable that can affect steady-state leisure per capita is g/y , and only if $\mu < 1$. If $\mu = 1$, then variations in g/y leave steady-state leisure constant. This result has nothing to do with government employment per se. Rather, this is the standard wealth effect of

government purchases discussed by Baxter and King (1993) and Christiano and Eichenbaum (1992).

- Steady-state leisure (and hence total hours worked) per capita is invariant to changes in h_g or h_n if g/y is constant or if $\mu = 1$ (no wealth effects of government purchases).
- Steady-state private hours per capita (h_p) are lower when h_g or h_n is higher. If g/y higher at the same time as h_g and if $\mu < 1$, then the effect is weaker, since the negative wealth effect causes total hours worked to increase.

In summary, these results show that we would expect private hours per capita to be nonstationary if there are trends in government hours or nonprofit hours. Total hours worked, however, are stationary as long as there are no wealth effects from government purchases.

But what is “per capita?” The “per capita” measure is embodied in N , since we must count up the hours of the representative agents. This theory does not say exactly how N should be measured, though most researchers assume that it is the “working-age population.”⁸ Most commonly, researchers define the working-age population to be anyone age 16 and older who is not in an institution. However, many other researchers use a more narrow definition. In the empirical section, we will show that trends in hours per capita are greatly affected by small changes in the definition of “working-age population.” Since theory provides little guide, we will let the data determine how best to define “working-age population.”

⁸ Ramey and Francis (2006) argue for using the total population for N when considering all possible forms of non-leisure time, including time spent in school and home production.

IV. New Measures of Hours Per Capita

We now use the empirical and theoretical insights from the last section to construct hours per capita measures that are more consistent with theory. We begin with the standard measure and change the numerator and denominator in a couple of ways to show how important the adjustments are for the low frequency movements.

Figure 8 compares the standard measure, which is private business hours divided by the civilian noninstitutional population ages 16 and over, to total hours worked divided by the noninstitutional population ages 16 and over. Whereas the peak of the private hours series in 1973 is over 13 percent below the peak in 1948, the peak of the total hours series in 1973 is only 5 percent below the peak in 1948. Thus, a significant amount of the U-shape disappears when we use a more complete measure of hours worked. Nevertheless, even the total hours series has a period of 17 years without ever crossing its mean. One still cannot reject a unit root at standard significance levels (see Table 1).

As discussed in the theoretical section, however, even total hours worked per capita can be nonstationary if the government spending share changes and government spending is not a close substitute to private consumption. Could changes in government spending account for the decline in hours worked per capita during the 1970s and 1980s? It is unlikely. Figure 9 shows the nominal share of GDP devoted to government spending versus total hours worked divided by the population ages 16 and over. Although there seems to be a positive correlation in the short-run, such as during the Korean War and the Vietnam War, the overall correlation is -0.29. Recall that the theory predicted that this correlation should be zero or positive: if there is a negative wealth effect of government spending, then an increase in the government share should be

associated with an increase in hours per capita. Clearly, the decline in hours during the 1970s and 1980s is due to some other factor.⁹

The previous empirical discussion about demographics suggests that a prime suspect is the demographic composition of the population. If age groups known for having below-average hours of labor supply constitute a larger fraction of the population, it is likely that hours per capita will decrease.

Guided by the facts displayed in Figures 6 and 7, we consider several measures of total hours per capita based on various definitions of the “working-age population.” Figure 10A shows three such measures, one that uses the population between ages 22 and 64, another 25 to 74 and yet a third that weights the young and old groups by their hours relative to prime age workers.¹⁰ None of these measures display the U-shape so characteristic of the measures using the population ages 16 and over.

But which measure is “best?” Since our goal is to correct for demographic and institutional sources of the low frequency trend we estimated in the standard measure, we examined the following statistics: (1) a standard unit root test; (2) the fraction of the variance attributable to low frequency cycles; and (3) the correlation of the ratio of the standard measure to the new measure with the HP trend estimated earlier. The unit root test allows us to determine whether we can reject a unit root in favor of stationarity. The spectral analysis, which applies under the assumption of stationarity, allows us to gauge the importance of the lowest frequency components in the overall variance. Finally, we include the correlation with the HP trend to see

⁹ One might wonder whether labor income taxes could explain the pattern. As Francis and Ramey (2005) show, an increase in the tax on labor income should lower hours worked per capita in steady-state. However, the empirical pattern of labor income taxes does not fit these hours per capita data. Figure 1 of Eichenbaum and Fisher (2005) shows that average labor income taxes have had a steady upward trend, with a small decline at the time of the Bush tax cuts.

if there is any relationship between our demographic and institutional adjustments and the HP trend we had estimated earlier.

Table 1 shows the three statistics for various possible measures of working age population. We have highlighted the best two scores in each column. While several measures perform quite well, the one that performs the best across the board uses the population ages 22 to 64, although the measures that use the population 22 to 74 or 25 to 74 could easily be chosen instead. All three of these measures easily pass the unit root test, and have only a small fraction of their variance explained by low frequency components. The high correlation between the estimated HP trend and the ratio of the standard measure (which use only private business hours and civilian noninstitutional population ages 16 and over) to these measures (with use total hours worked and different population definitions) suggests that the HP filter was indeed isolating demographic and institutional changes in the per capita labor series. Surprisingly, the series that includes the young and old weighted by their hours seems not to do as well at removing the low frequency trends.

We will use the population ages 22 to 64 for our main results, but will also show results for the other measures. One question that arises, though, is whether we want to identify technology shocks as those shocks having long-run effects on aggregate productivity or those having long-run effects on private business productivity. To be comparable with the literature, we focus on the private business sector, but also discuss results for the total economy. We want to match private business hours with private business productivity. Thus, we use private business hours detrended by the low frequency component of the ratio of private business hours

¹⁰ To be specific, this weighted measure uses the population ages 22-64 plus the population ages 16-21, weighted by their **average** hours relative to those 22-64, plus the population ages 65+, weighted by their hours relative to those 22-64. The hours weights are based on decennial census data that are then interpolated.

to total hours, i.e., the low frequency component of the series shown in Figure 5C. To be specific, our private hours per capita measure is:

$$\text{Adjusted Private hours per capita} = \frac{\text{Private business hours}}{\theta \cdot \text{noninstitutional population ages 22 - 64}} \quad \text{where}$$

θ is the trend component (using an HP filter with $\lambda = 160,000$) of the ratio of private business hours to total hours.

Thus, this measure has the same cyclical volatility of private hours per capita, but does not have a downward trend. Figure 10B shows this series plotted over the total hours per capita series. The graph shows that the series are very similar, but the private business hours series has a little more business cycle volatility. The results reported in the next section are very similar if we instead use total hours per capita.

IV. The Effects of Technology Shocks Using the New Measure of Hours Per Capita

A. Baseline Impulse Responses

We now investigate how the use of our new measure of hours per capita changes the previous results on the effect of technology on hours. We re-estimate the structural VAR used by Galí, Francis and Ramey, and Christiano, Eichenbaum and Vigfusson using the new measure of hours. Figure 11 plots the impulse responses from the system using the new measure of hours. The levels specification is on the left and the first-difference specification is on the right. In both specifications, per capita labor hours respond negatively in the short-run to the technology shock in both the levels and first difference specifications. Hours become positive, though not significantly so, after a year or more.

Thus, in contrast to the case with the standard measure, the negative effect of technology shocks on the new hours measures is robust across specifications, whether hours are assumed stationary or not. These results also shed light on the debate concerning the results with the standard measure. CEV claim that over-differencing of hours per capita leads to different estimated effects of technology shocks on hours. This is not true with our new measure, or with the HP filtered standard measure as shown earlier. Even though the standard ADF tests reject a unit root, assuming a unit root in hours does not change the qualitative nature of the impulse response functions.

B. Robustness Checks

We tested for robustness in a variety of ways. We first investigated the effects of using various measures of population. Rather than report pages of graphs, we summarize the estimated impact effect of the estimated technology shock on hours. Table 2 shows results of the bivariate SVAR for the various measures of population listed in Table 1. In all cases that passed the unit root tests at the 5 percent level, the estimated impact effect of a technology shocks on hours is significantly negative – the responses were also significantly negative for age group 25-54 and the hours-weighted measures even though they failed to reject unit root at 5 percent. The effect becomes greater in magnitude as we drop more of the younger population from the measure.

For our baseline measure using population ages 22 to 64, we also investigated robustness in several other ways. First, we estimated a bigger system. CEV initially argued that omitted variables were the source of the Galí finding. To check the robustness, we estimate the larger system used by CEV. This system adds four variables: the federal funds rate, the rate of inflation (measured using the GDP deflator), the log of the ratio of nominal consumption to nominal GDP

(where consumption is measured as the expenditures on nondurables and services plus government expenditures), and the log of the ratio of nominal investment expenditures to nominal GDP (where investment is measured as expenditures on consumer durables and gross private investment). The $C(L)$ matrix of this system is now a block 6×6 matrix in the lag operator. If labor productivity is the first variable in the system, we identify the technology shock by imposing the restriction that $C^{1j}(1) = 0$ for $j = 2, 3, 4, 5, 6$. Because the federal funds rate is only available beginning in 1954, the model is estimated over a shorter sample.

Figure 12 shows the results when hours are specified in levels. The labor hours response continues to be significantly negative in the short run. Output and investment also dip slightly in the short-run, but they are not significantly below zero.¹¹

We also checked robustness in two more ways. Since the graph in Figure 10B looks like the mean of hours per capita for the population ages 22-64 may be slightly higher in the period after 1986, we allowed a different mean in this period.¹² In this case, the estimated impact effect was smaller in magnitude, -0.230 rather than -0.301, but still statistically significant (the standard error was 0.06). Finally, we studied the bivariate SVAR in which we used total hours per capita and total productivity (GDP divided by total hours). In this case, the impact effect was estimated to be -0.449 with a standard error of 0.108.

Cooley and Dwyer (1998) point out that the results from structural VARs may be sensitive to auxiliary assumptions with respect to lag length. Our baseline models all include four lags. To determine whether our results were due to too few lags, we re-estimated the

¹¹ Additionally, in the larger system using first differenced labor hours, the hours response is significantly negative.

¹² The change in mean specification is superior to the alternative of a linear trend. There is no significant trend once the change in mean is allowed.

bivariate system in levels and included 50 lags. The impulse responses were qualitatively similar to those from the system with only 4 lags.

Thus, when our new measure of hours per capita is used, hours always respond negatively to technology shocks. This is true for levels and first-differences specifications. It is true for specifications with more variables in the system. It is also true when we add 50 lags to the specification.

V. Conclusion

In this paper we have proposed that the conflicting results in the debate concerning the effects of technology shocks on hours stems from the low frequency movements in the standard measure of hours per capita. We first showed that removal of the low frequency movements using a very conservative HP filter produces results suggesting a positive technology shock lowers hours in the short-run. We then argued that the HP filter is capturing demographic and institutional changes, which have nothing to do with the kinds of technology shocks typically modeled in RBC theory. We modify standard theory to account for these changes and use the insights from the theory to produce a new measure of hours per capita that adjusts for these changes. Our measure, which adjusts for trends in government and nonprofit employment and the age structure of the population removes most of the low frequency trend. We find that removing these slow moving components leads to consistent results on the effects of technology on per capita hours worked, regardless of the stationary assumption assumed for per capita labor hours. That is, in contrast to results using the standard measure, our new measures produce uniformly negative effects of technology on hours, whether labor hours enter the VARs in levels or first differences.

Our findings also have direct implications for the nature of the “preference shock” identified by researchers such as Fernandez-Villaverde and Rubio-Ramirez (forthcoming) and Kahn and Rich (forthcoming). Their “preference shock” is a catchall for anything that shifts relative labor supply and is not a typical stationary *i.i.d.* stochastic process. Our findings imply that it is easily measured demographic and institutional changes rather than mysterious “preference shocks” that are the source of these important low frequency trends.

Data Appendix

Hours and Productivity:

We use published and unpublished data from the BLS on hours worked in private business, government, nonprofit, and private households. Shawn Sprague kindly provided the unpublished data. Note that government enterprises (such as the post office) are included in private business by the BLS. Private business productivity is based on the BLS series. Total productivity is constructed by dividing real GDP by total hours worked.

Population:

Annual data, including age breakdown, is from the U.S. Census, *Mini Historical Statistics*, Table HS-3, *Economic Report of the President, 2006*, Table B-34, and are updated from the Census website. Annual data are interpolated to quarterly. The quarterly civilian noninstitutional population is from the Federal Reserve Bank of St. Louis web site. The fraction of the population institutionalized is based on decennial census data that is interpolated. We distinguish the institutionalization rate of those ages 18 to 64 from those ages 65 and over.

Hours Worked By Age

We use Census data from IPUMS to calculate hours worked by age group. 1950, 1980, and 1990 use the variable "hrswork1." 1960 and 1970 use "hrswork2" with interval values calculated from overlapping "hrswork1" and "hrswork2" variables in 1950 and 1980. 2000 is constructed as follows. For 1990 and 2000, we constructed $wkswrk1 * uhrswork$ and used the implied growth rate to project hrswork1 from 1990 forward to 2000. ($wkswrk1 * uhrswork$ is actually quite close to hrswork1 in 1990.)

GDP Aggregates and Federal Funds Rate:

The GDP aggregates and deflator are from the BEA web site. The federal funds rate is from the St. Louis Federal Reserve web site.

References

- Baxter, Marianne and Robert G. King, "Fiscal Policy in General Equilibrium," *The American Economic Review*, 83 (June 1993): 315-334.
- Burnside, Craig and Martin Eichenbaum, "Factor Hoarding and the Propagation of Business Cycle Shocks," *American Economic Review*, 86 (1996): 1154-1174.
- Cavallo, Michele, "Government Employment Expenditure and the Effects of Fiscal Policy Shocks," Federal Reserve Bank of San Francisco Working Paper, 2005-16, September 2005.
- Chari, V.V., Patrick J. Kehoe, and Ellen R. McGrattan, "A Critique of Structural VARs Using Real Business Cycle Theory," May 2005 manuscript.
- Christiano, Lawrence, and Martin Eichenbaum, "Current Real Business Cycle Theories and Aggregate Labor Market Fluctuations," *The American Economic Review* 82 (June 1992): 430-450.
- Christiano, Lawrence, Martin Eichenbaum, and Robert Vigfusson, "What Happens After a Technology Shock?" 2003 NBER working paper 9819.
- , "Assessing Structural VARs," forthcoming *NBER Macroeconomics Annual* 2006.
- Cooley, Thomas F. and Mark Dwyer, "Business Cycle Analysis without Much Theory: A Look at Structural VARs," *Journal of Econometrics* 83 (March-April 1998): 57-88.
- Eichenbaum, Martin and Jonas D.M. Fisher, "Fiscal Policy in the Aftermath of 9/11," *Journal of Money, Credit and Banking*, 37 (February 2005): 2-22.
- Erceg, Christopher J., Luca Guerrieri and Christopher Gust, "Can Long Run Restrictions Identify Technology Shocks?" International Finance Discussion Papers 792. Washington: Board of Governors of the Federal Reserve System, 2004.
- Fernald, John, "Trend Breaks, Long-Run Restrictions, and the Contractionary Effects of Technology Shocks," 2006 manuscript.
- Fernández-Villaverde, Jesús and Juan F. Rubio-Ramírez, "Estimating Macroeconomic Models: A Likelihood Approach," forthcoming *Review of Economic Studies*.
- Finn, Mary G., "Cyclical Effects of Government Purchases," *International Economic Review* 39 (August 1998): 635-657.

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

-----, "The Source of Historical Economic Fluctuations: An Analysis using Long-Run Restrictions," NBER International Seminar on Macroeconomics 2004, ed. Richard Clarida, Jeffrey Frankel, Francesco Giavazzi, and Kenneth West. The MIT Press: 2006

Francis, Neville, Michael Owyang and Jennifer Roush, "A Flexible Finite-Horizon Identification of Technology Shocks," International Finance Discussion Papers 832. Washington: Board of Governors of the Federal Reserve System, 2005.

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

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

Kahn, James A. and Robert W. Rich, "Tracking the New Economy: Using Growth Theory to Detect Changes in Trend Productivity," forthcoming *Journal of Monetary Economics*.

King, Robert G., Charles I. Plosser, James H. Stock, and Mark W. Watson, "Stochastic Trends and Economic Fluctuations," *American Economic Review* 81 (1991): 819-840.

Prescott, Edward C., "Why Do Americans Work So Much More Than Europeans?" *Federal Reserve Bank of Minneapolis Quarterly Review*, 28 (July 2004): 2-13.

Ramey, Valerie A. and Neville Francis, "A Century of Work and Leisure" NBER Working Paper 12264, May 2006.

Salamon, Lester and S.Wojciech Sokolowski, "Nonprofit Organizations: New Insights from the QCEW Data," *Monthly Labor Review*, September 2005, pp. 19-26.

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

Table 1
Tests on Total Hours per Capita using Alternative Measures of Working-Age Population

Working-Age Population Definition (noninstitutional)	ADF unit root test against H_a: stationarity** (p-value in parenthesis)	Fraction of variance explained by low frequencies		Ratio correlation with HP trend in standard hours*** ($\lambda = 160,000$)
		58 year cycles	29 year cycles	
Ages 16+	-2.52(0.11)	0.48	0.56	0.87
Ages 16-64	-2.33 (0.16)	0.48	0.67	0.84
Ages 18-64	-2.74 (0.07)	0.30	0.58	0.90
Ages 20-64	-3.13 (0.02)	0.13	0.48	0.95
Ages 22-64	-3.38 (0.01)	0.06	0.41	0.97
Ages 25-64	-3.18 (0.02)	0.20	0.43	0.98
Ages 22-54	-3.47 (0.01)	0.08	0.40	0.97
Ages 22-74	-3.24 (0.02)	0.09	0.40	0.96
Ages 25-54	-2.68 (0.08)	0.43	0.62	0.95
Ages 25-74	-3.33 (0.01)	0.06	0.33	0.97
Hours- weighted*	-2.50 (0.12)	0.32	0.68	0.90

*“Hours-weighted” uses the population ages 22-64 plus the population ages 16-21, weighted by their average hours relative to those 22-64, plus the population ages 65+, weighted by their hours relative to those 22-64. Weights are based on decennial census data and interpolated.

**Tests were on the logs of the variables. Lags were chosen optimally. In every case, either 3 or 4 lags was chosen.

***This column gives the correlation of the log of the HP trend shown in Figure 3 with the log of the following ratio:

(private hours/ civilian noninst. population ages 16+) / (total hours/ working-age population)

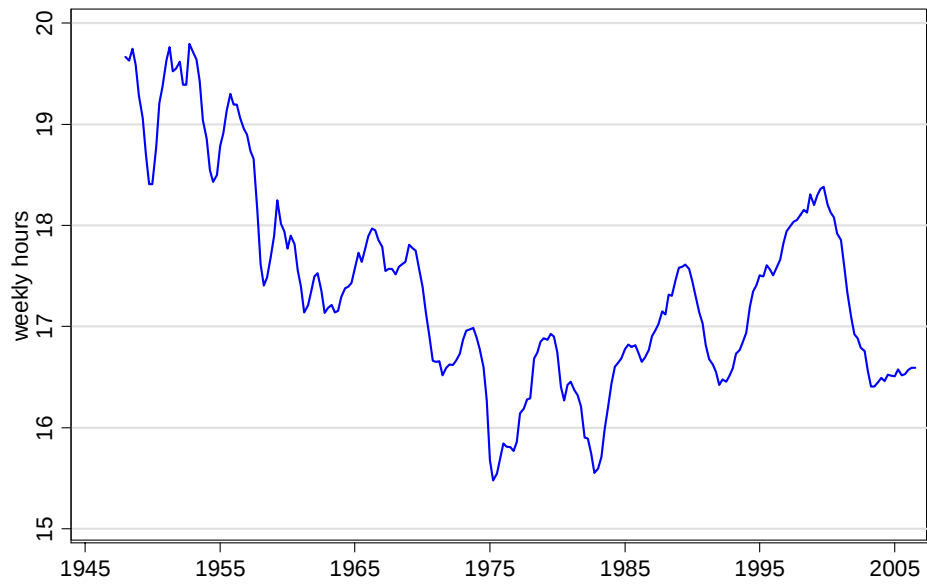
Table 2. Estimated Impact Effect of Technology Shocks on Private Hours Per Capita
(Bivariate SVAR)

Working-Age Population Definition (noninstitutional)	Estimate and Standard Error
Ages 16+	0.215 (0.067)
Ages 16-64	-0.029 (0.065)
Ages 18-64	-0.085 (0.063)
Ages 20-64	-0.170 (0.061)
Ages 22-64	-0.301 (0.062)
Ages 25-64	-0.425 (0.117)
Ages 22-54	-0.319 (0.066)
Ages 22-74	-0.153 (0.061)
Ages 25-54	-0.424 (0.155)
Ages 25-74	-0.237 (0.059)
Hours-weighted	-0.372 (0.092)

All systems include private business productivity and private business hours per capita, adjusted for the low frequency movements in the ratio of private business to total hours.

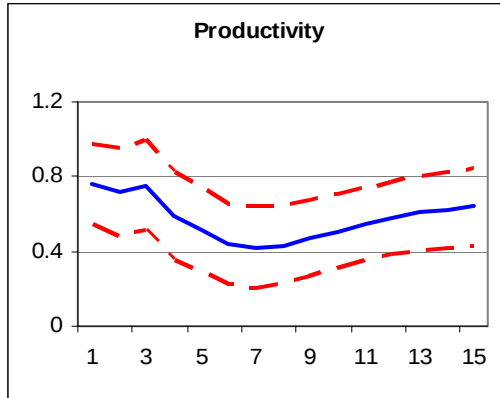
The measures that passed the unit root test in Table 1 are highlighted.

Figure 1: Private Business Hours Per Capita
(based on civilian non-institutional population ages 16 and over)



**Figure 2: Impulse Responses to a Technology Shock: Quarterly 1948-2006
(Business Hours/Civilian Population 16+, Bivariate System with 95% standard error bands)**

Hours in Levels



Hours in First Differences

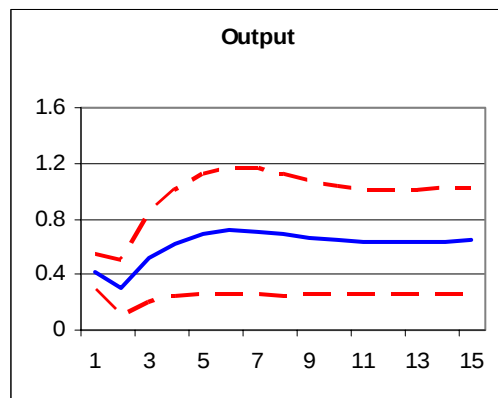
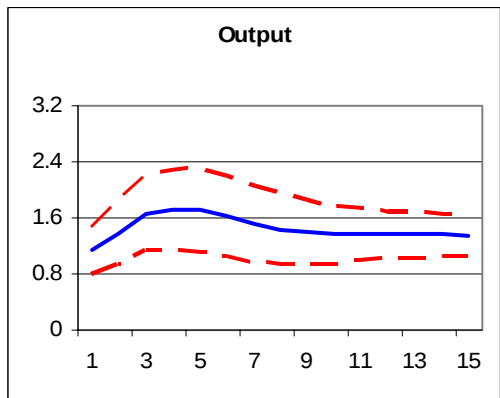
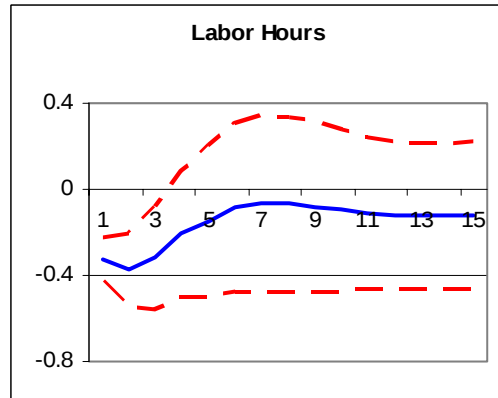
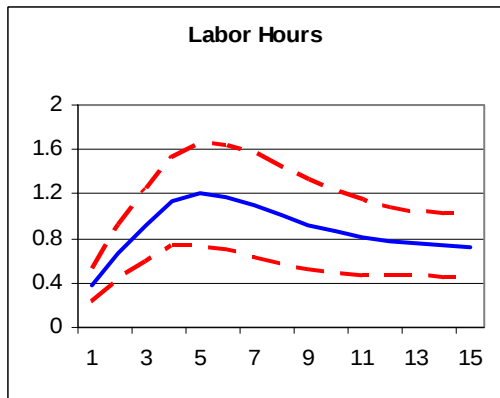
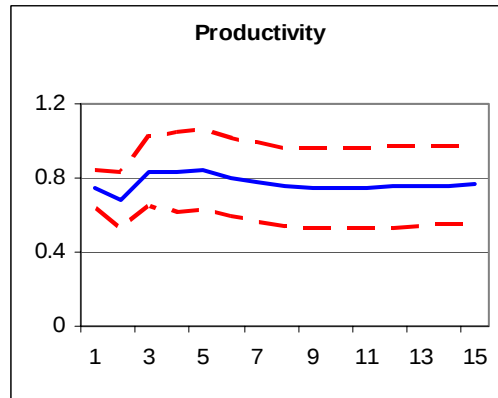
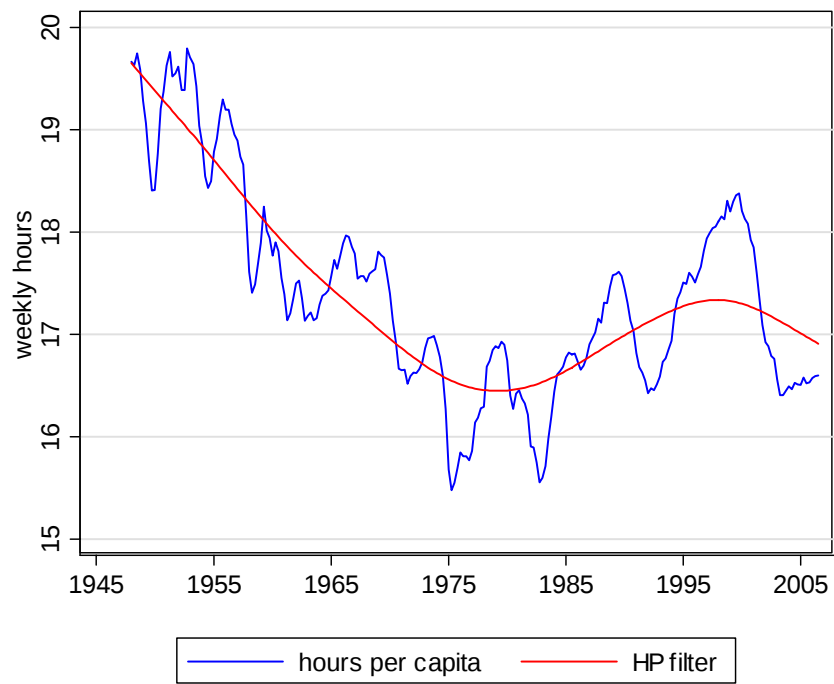
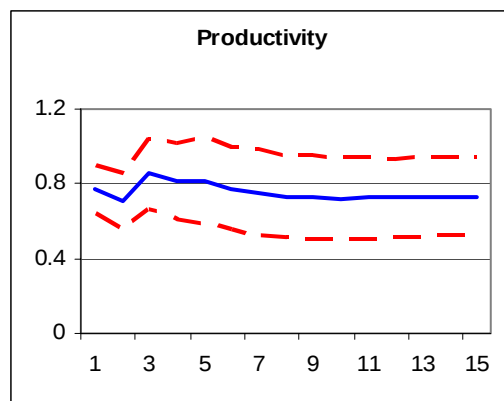


Figure 3: Plot of Hours per Capita and its HP-Filtered Trend ($\lambda = 160,000$)

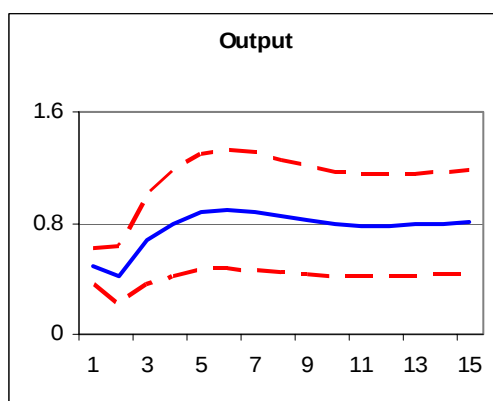
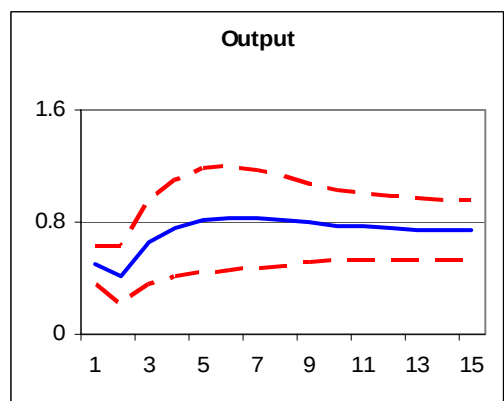
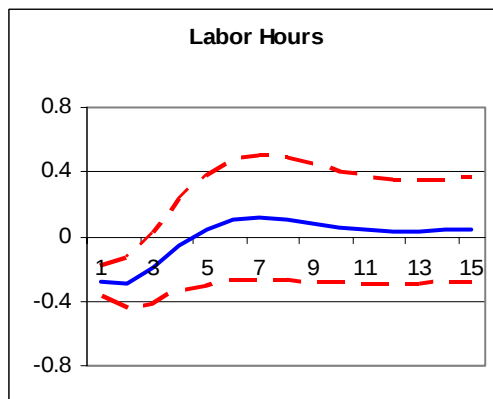
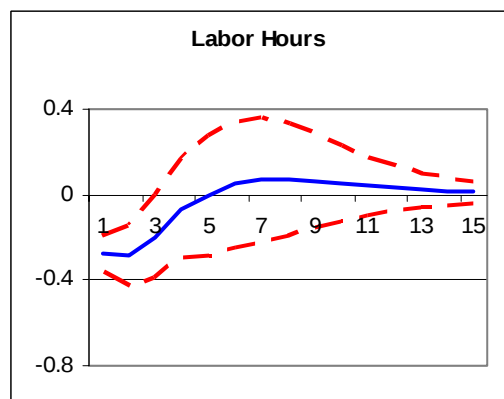
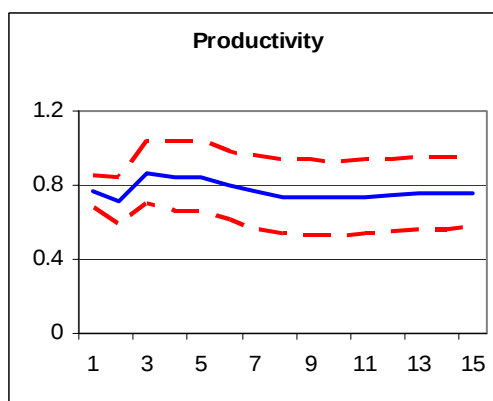


**Figure 4: Impulse Responses to a Technology Shock: Quarterly 1948-2006
(HP-Filtered Hours/Civilian Population 16+, Bivariate System with 95% error bands)**

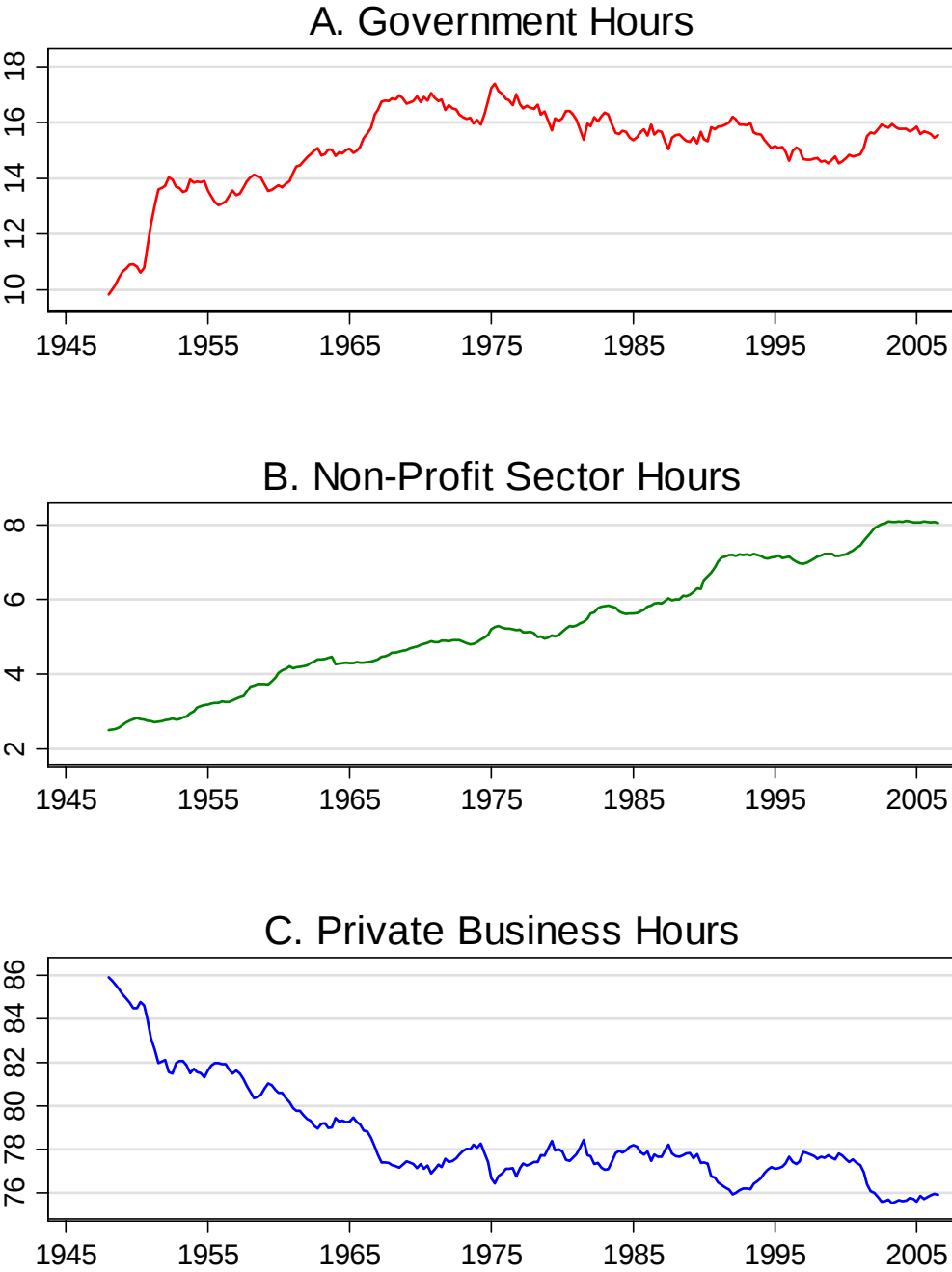
Hours in Levels



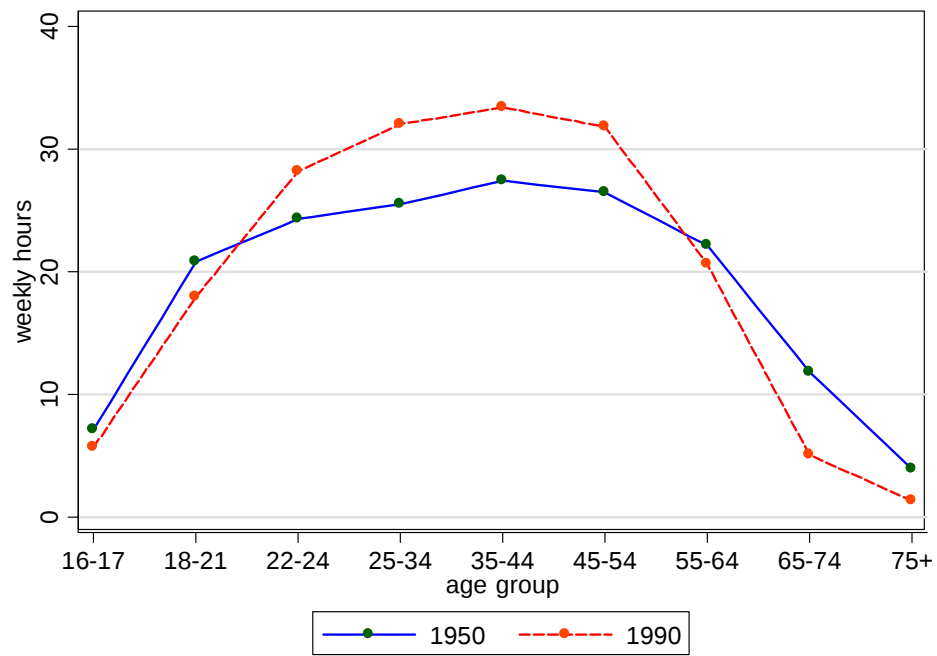
Hours in First Differences



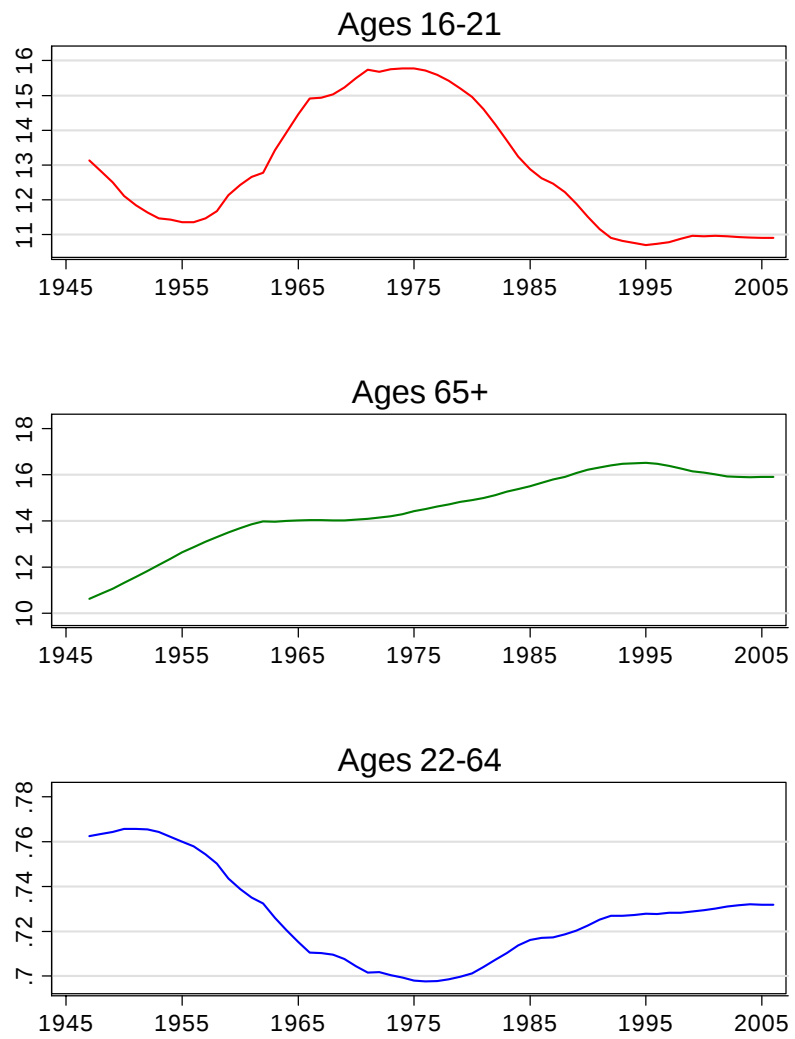
**Figure 5: Hours Worked by Sector
(percent of total)**



**Figure 6: Average Weekly Hours Worked by Age Group
(non-institutional population)**



**Figure 7. Age Composition of Population
(percent of population ages 16 and over)**



**Figure 8. Private Business vs. Total Hours Worked
(divided by population 16 and over)**

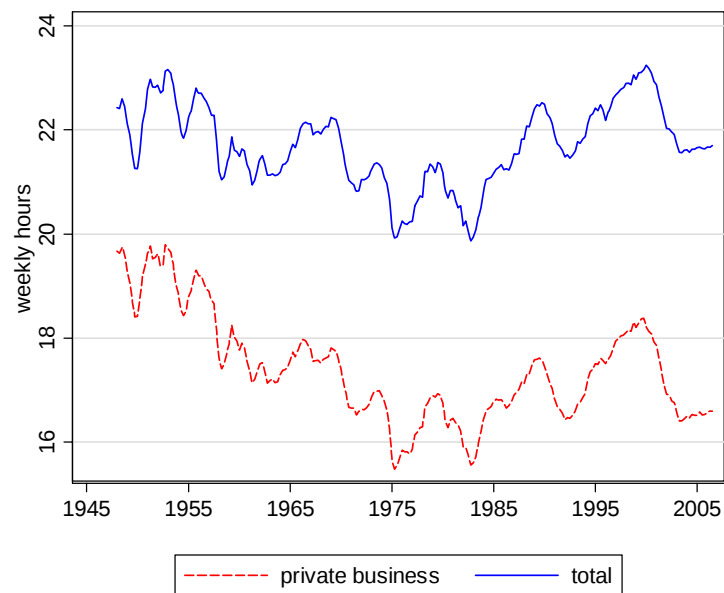
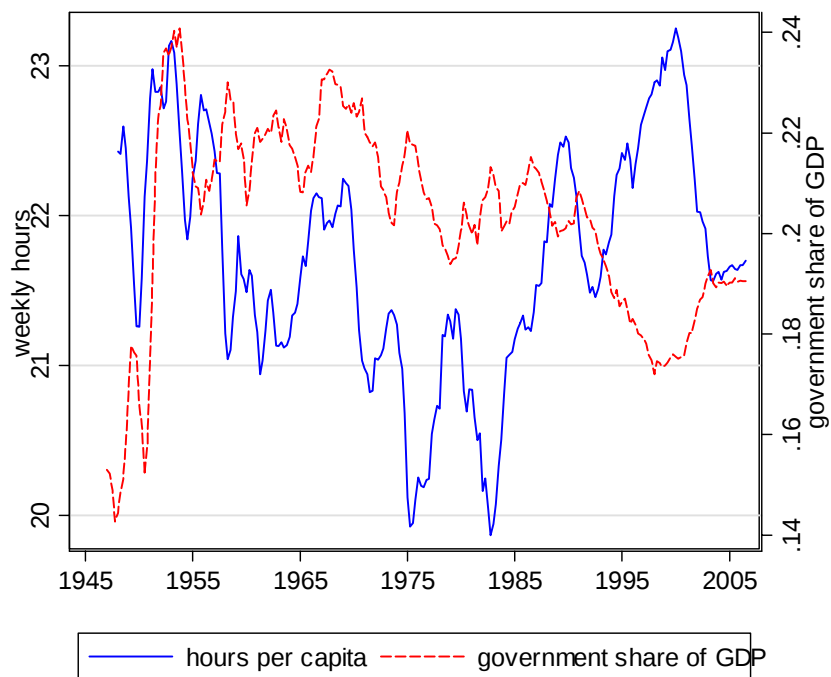


Figure 9. Government Share of GDP vs. Total Hours per Person 16 and Over



**Figure 10A. Total Hours Worked Per Capita
Effects of Alternative Measures of Working-Age Population**

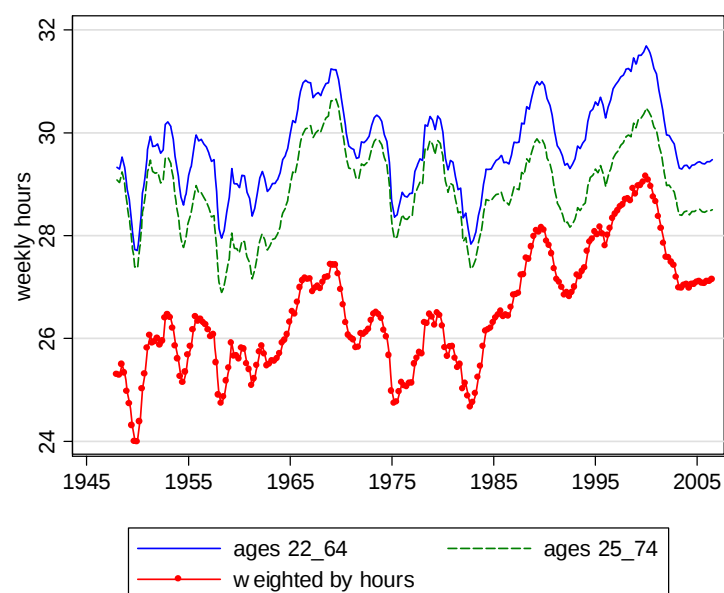
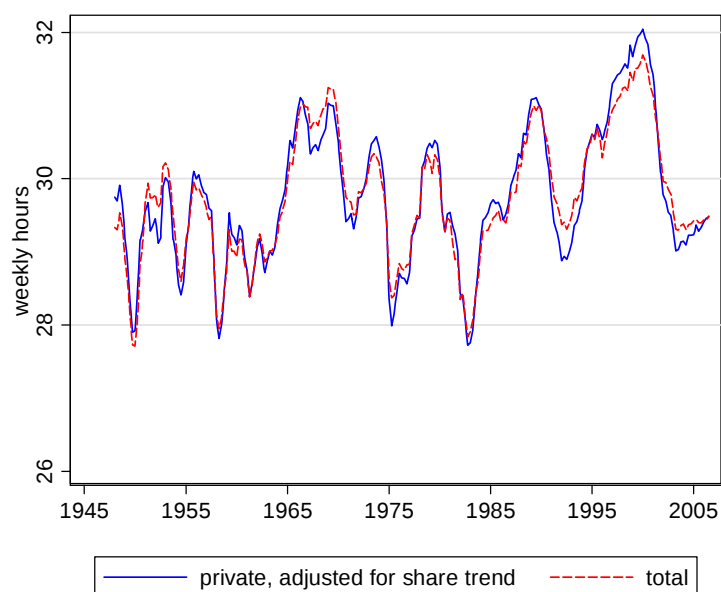
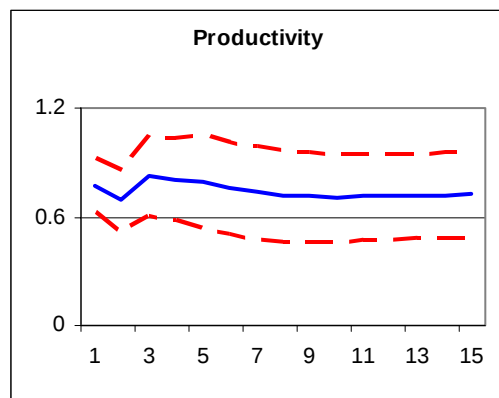


Figure 10B. Total Hours versus Adjusted Private Business Hours
(Both series are divided by noninstitutional population ages 22-64)

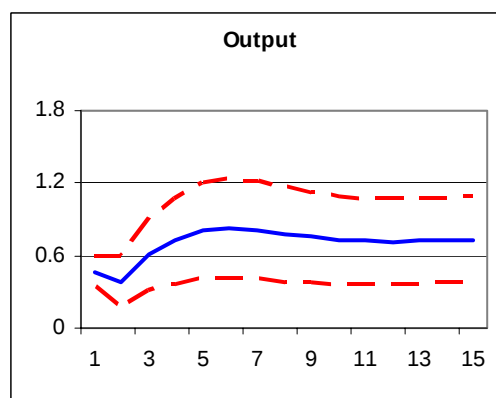
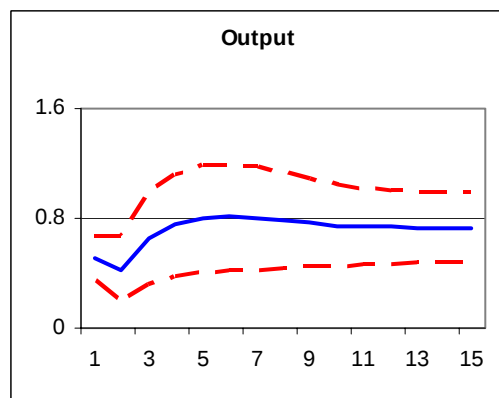
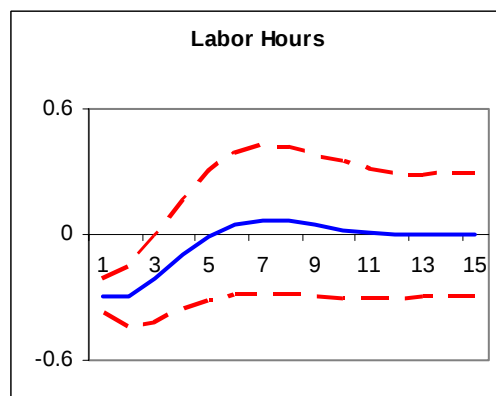
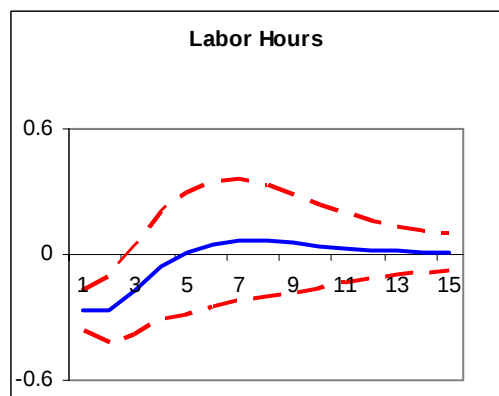
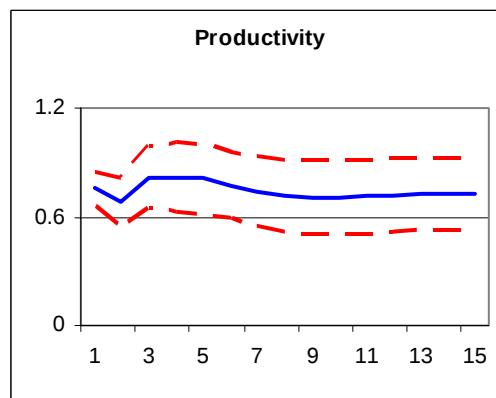


**Figure 11: Impulse Responses to a Technology Shock: Quarterly 1948-2006
(with Hours Per Capita Measure, Bivariate System with 95% standard error bands)**

Hours in Levels



Hours in First Differences



**Figure 12: Impulse Responses to a Technology Shock: Quarterly 1954-2006
(New Measure, Six-Variable VAR with 95% standard error bands, Hours in Levels)**

