Scottish Journal of Political Economy, Vol. 54, No. 5, November 2007

REAL WAGE CYCLICALITY IN THE PANEL STUDY OF INCOME DYNAMICS

*Eric T. Swanson**

Abstract

Previous studies of real wage cyclicality have made only sparing use of the micro-data detail that is available in the Panel Study of Income Dynamics (PSID). The present paper brings to bear this additional detail to investigate the robustness of the previous results and to examine whether there are important cross-sectional and demographic differences in wage cyclicality. Although real wages were procyclical across the entire distribution of workers from 1967 to 1991, the wages of lower-income, younger, and less-educated workers exhibited greater procyclicality. However, workers' straight-time hourly pay rates have been acyclical, suggesting that more variable pay margins such as bonuses, overtime, late shift premia, and commissions have played a substantial if not primary role in generating procyclicality.

I Introduction

John Maynard Keynes famously remarked that, 'for a given organisation, equipment, and technique, real wages and the volume of output (and hence of employment) are uniquely correlated, so that, in general, an increase in employment can only occur to the accompaniment of a decline in the rate of real wages' (Keynes, 1936, p. 17). Since then, the correlation between real wages and employment has been a prominent testing ground for a wide array of macroeconomic models. For example, a procyclical relationship between real wages and employment is predicted by technology-driven models of business cycles (e.g., Kydland and Prescott, 1982; Barro and King, 1984), models of strongly countercyclical markups (e.g., Rotemberg and Woodford, 1992), and models of external increasing returns to scale (e.g., Bartelsman *et al.*, 1994). By contrast, a countercyclical relationship between real wages and employment is predicted by Classical and traditional Keynesian models and by more modern DSGE models in which technology shocks play only a minor role. The goal of the present paper is to investigate the correlation between real wages and unemployment in more depth than previous studies by making fuller use of the micro-data available in the Panel Study of Income Dynamics (PSID).

*Federal Reserve Bank of San Francisco

Historically, the relationship between real wages and unemployment (or other business cycle indicators) has been explored using economy-wide measures of 'the real wage', namely aggregate wages paid divided by aggregate hours worked, with different measures of the numerator, denominator, and deflator being preferred by different authors (see Abraham and Haltiwanger, 1995, for a good survey). These studies have typically found real wages to be acyclical or slightly procyclical over the postwar period, and somewhat more procyclical since the late 1960s.

However, the highly aggregate nature of the data in these studies has led to questions about the accuracy and relevance of the results. For example, low-wage workers tend to have substantially more cyclical hours and employment than high-wage workers, so that in every recession, a large number of low-wage worker-hours are dropped from the aggregate wage statistic; this would cause the economy's aggregate wage to be countercyclical even if every individual's wage were completely fixed over the business cycle [Stockman, 1983; Solon, Barsky, and Parker (SBP), 1994]. On the other hand, highly cyclical industries such as durables manufacturing and construction also have high average wages; in every recession, then, a large number of these high-wage workers are also dropped from the economy's aggregate wage statistic, inducing a procyclical bias in aggregate measures of real wages, exactly opposite the effect outlined above (Chirinko, 1980; Solon and Barsky, 1989). Finally, even controlling for worker and industry composition change over the business cycle, the aggregate wage is an income-weighted measure, in that a 1% change in high-wage workers' earnings has a much greater impact on the aggregate wage statistic than a 1% change in low-wage workers' pay; if high-wage workers tend to experience less (or more) wage cyclicality than their low-wage counterparts, the aggregate statistic will again be a poor measure of what we might think of as the typical worker's experience.

For all of these reasons, measuring the cyclicality of the economy's aggregate wage statistic may reveal relatively little about the experience of typical workers and firms and provide relatively little insight into whether a given macroeconomic model is accurately describing the cyclical relationship between labor supply and labor demand. Micro-level panel data provide a much better medium for gaining an insight into the labor market relationship between workers and firms over the business cycle. Beginning with the availability of enough such data in the 1980s, a number of researchers have begun investigating exactly this question (e.g., Bils, 1985; Solon *et al.*, 1994). SBP in particular make it clear that the composition biases mentioned above are substantial and countercyclical on net, so that over the period covered by panel data, real wage movements of individual workers have in fact been strongly procyclical, in sharp contrast to the findings of little or no procyclicality in the aggregate statistics mentioned earlier.

However, these panel studies have left open a number of questions. Most importantly, the time period covered by the data (1966 to the present) has been one of significantly countercyclical prices, so that the findings of procyclical real wages may be due simply to rigid nominal wages and countercyclicality of the underlying price deflator. Oil shocks during this period may be playing a substantial direct role as well: a recession brought on by an exogenous increase in the price of oil may induce employers to shift away from energy and capital and lower the marginal product of labor while still reducing industry and aggregate output, generating a procyclical real wage. It is also dissatisfying that none of these panel studies attempts to present the data in a more disaggregate format – for example, SBP simply replace a regression of the *change in the average* 'real wage' mentioned earlier with a regression of the *average of the changes* in real wages experienced by the panel. Although this statistic is relatively free of the aggregation biases mentioned above, it fails to make use of an enormous amount of detail inherent in the data. These studies have also ignored the availability of local-area unemployment rates and data on straight-time hourly wages, which are available for all workers who are paid by the hour in the PSID. These additional data provide an opportunity to check for robustness in the findings mentioned above, and to examine them in greater detail.

The present paper investigates all of these issues, with a view toward relating the results to macroeconomic theories of the business cycle. The remainder of the paper proceeds as follows: Section II discusses the data and the basic empirical framework used for the analysis. Section III presents and discusses the results. Section IV concludes. Appendix A provides additional technical details about the exact data series used and empirical methods.

II Data and Methods

For data, we use the PSID, a longitudinal survey begun in 1968 that now covers some 8,000 families.[1] The two primary micro-data alternatives to the PSID are the National Longitudinal Survey (NLS) and the Current Population Survey's (CPS) Annual Demographic Supplements, but both of these have shortcomings as far as measuring wage cyclicality is concerned. For example, the NLS covers only selected segments of the US population (young men, young women, men ages 45–59, women ages 30–44, and youths), and interviews were not taken in at least five years between 1966 and 1983 due to lack of funds, so the NLS sample is lacking both in comprehensiveness and continuity. Moreover, from 1970 to 1976 the NLS did not ask respondents for hours or earnings data on their most recent job if the respondent was currently unemployed, which creates a sample selection bias in the data for those years. The CPS's Annual Demographic Supplements is a more promising alternative, having some advantages such as sample size (it covers roughly 60,000 individuals). Although the CPS data are not truly longitudinal, year-to-year changes in wages for many individuals in some years can be computed, but unfortunately these matches cannot be

[1] When families are multiplied by their 'family weights' in each year, calculated to account for differential sampling rates, mortality rates, and rates of nonresponse across demographic groups, as well as issues of family composition change, the resulting cross sections are representative of 1968 America (excluding Alaska and Hawaii) as it has evolved through the years. The PSID makes no attempt to account for immigration into the United States that has occurred since 1968.

performed for the years 1964–1968, 1971–1973, 1976–1977, and 1985–1986, which include some of the most interesting years of the sample. Moreover, before 1977 the CPS data on hours worked are for the preceding week rather than the preceding calendar year, which leads to a sample selection problem once again should an individual be unemployed in the week preceding the survey.

In the PSID, questions on income and hours worked are for the preceding calendar year, thus avoiding the sample selection problem that is present in both the NLS and CPS (Blank, 1990, confirms this empirically). Questions about job characteristics are for the job held at the time of the interview, and so they must be lagged 1 year in order to be matched to the corresponding hours and income data. Data for household heads are more detailed, accurate, and complete than for other sample members; hence, we take as our sample all 10,114 men who were ever household heads in the PSID between 1968 and 1992 (excluding the more recent Latino sample). For each year such an individual was household head, we make use of the following data: total labor income of head; wages and salaries of head; bonuses, overtime, and commissions of head; head's annual hours worked; head's race, age, and education; whether head works for government or the private sector; whether head's job is covered by a union contract; head's hourly wage if paid by the hour; unemployment rate for the head's county of residence; and PSID family weight for the head's family. Additional details regarding these series are provided in Appendix A. Data for which 'major assignments' were made by the PSID staff are omitted.

For business cycle indicators, we begin with the national unemployment rate in Figure 1. Log real GDP and its deviations from trend are presented in Figure 2 for comparison (we will also consider first differences of these series in some regressions). Later, we will consider as a cyclical indicator the unemployment rate in the respondent's county of residence, as reported in the PSID. Note that 'national unemployment rate' and 'real GDP' here refer to the civilian unemployment rate for all civilian workers, Table B–40, and GDP in 1987 dollars, Table B–2, *Economic Report of the President*, 1995.

For a measure of the price level, we focus primarily on the 1987 GDP deflator. Although both the CPI and PPI are more appropriate from a theoretical point of view, the countercyclical movements in both of these series over the period 1967–1991 substantially dominate those of the GDP deflator (see Figure 3)[2]; because previous studies have found significant real wage procyclicality over this time period, we have chosen to be conservative by taking the least countercyclical price measure. Interested readers can easily

[2] Abraham and Haltiwanger implicitly find that the PPI is actually *less* countercyclical than the CPI over the years 1970–1994. This is not due to the negligible difference in the time period covered, but rather due to differences in the cyclical indicator used: AH focus on employment and output in the manufacturing industry alone, rather than the economy as a whole. A graph of manufacturing employment over time reveals that it never fully recovers from either the 1981–1982 or 1991 recessions, so that the last 10–15 years appear essentially as one long depression, with the 1990s being particularly severe. Detrended manufacturing output suffers from the same problem to a lesser degree. Comparing these observations with Figure 3 explains the discrepancy between our results.

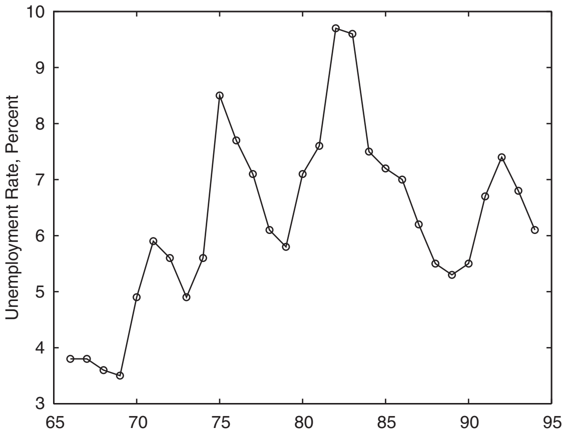


Figure 1. US unemployment rate, 1966–1994.

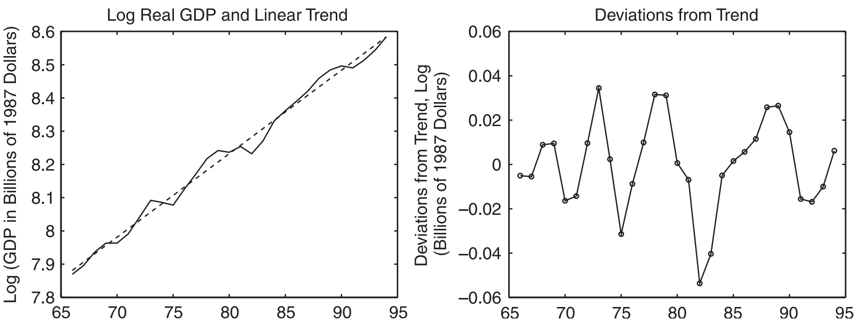
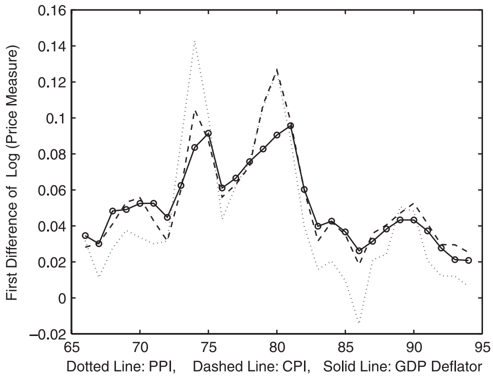


Figure 2. Log real GDP and deviations from trend.

modify the graphs of first-differenced log wages below by subtracting off the GDP deflator and adding in either the CPI or PPI as desired; the corresponding regression coefficients are likewise linear and can be similarly adjusted by making use of the coefficients in the accompanying table [although standard errors (SEs) cannot be easily adjusted for coefficients thus modified]. Note that the CPI, PPI, and GDP deflator presented here are the CPI–U for all items, Table B–59, PPI for total finished goods, Table B–64, and implicit GDP deflator, Table B–3 (*Economic Report of the President*, 1995).

Previous researchers using the PSID have typically focused on the 'total labor income' variable as their measure of wages. However, about 14% of the weighted sample earning any labor income in a given year report labor income beyond 'wages and salaries' and 'bonuses, overtime, and commissions'. About 8% of male heads who report positive labor income earn *no* wages, salaries, bonuses, overtime, or commissions – these people are primarily self-employed businessmen, professionals, farmers, and ranchers. Because they make up a nontrivial percentage of the sample, it is possible that the substantial wage



$$\Delta \log P_t = \beta_1 + \beta_2 t + \beta_3 \Delta \mathrm{Unemp}_t + \varepsilon_t,$$
$$t = 1968, \ldots, 1991.$$

Price Index	β_3
PPI	.0095 (.0073)
CPI	.0084 (.0051)
GDP deflator	.0063 (.0035)

Figure 3. Alternative measures of prices.

procyclicality found by other researchers is due partly to these nonwage sources of income. To check this, we computed wages in two ways: first using the head's 'total labor income' variable and then the 'wages and salaries' plus 'bonuses, overtime, and commissions' variables. In fact, we found essentially no differences in the cyclicality of these two measures of wages – in all of our graphs and results below, differences between using 'labor income' as compared with 'wages and salaries' plus 'bonuses, overtime, and commissions' were negligible.[3] Thus, for the remainder of the paper, we will simply present the results using annual total labor income divided by annual total hours worked as the wage measure.

To make quantitative measurements of real wage cyclicality and comparisons across demographic groups, we will run regressions of the form:

$$\Delta w_t = \beta_1 + \beta_2 t + \beta_3 \Delta \mathrm{Unemp}_t + \varepsilon_t, \tag{1}$$

where w_t is the given real wage statistic (in logs), Δw_t is its first-difference (i.e., $w_t - w_{t-1}$), t is a time trend, Unemp is the national unemployment rate, and the β_i are parameters to be estimated by ordinary least squares. When considering wage levels or deviations of wage levels from individual-specific trends, we will consider regressions of the form:

$$w_t = \beta_1 + \beta_2 t + \beta_3 t^2 + \beta_4 \mathrm{Unemp}_t + \varepsilon_t. \tag{2}$$

All regressions will be of one of these two forms unless stated otherwise. These are not, of course, structural models of real wage behavior, but simply a convenient method of computing reduced-form sample correlations between real

[3] Because data for 'bonuses, overtime, and commissions' are bracketed and hence unusable until 1975, we also compared these wage measures with 'wages and salaries' alone – the structure of the PSID questionnaire is such that this variable will often include the respondent's bonuses and overtime data anyway (and the 'bonuses and overtime' variable itself will be nil), and hence will also be a reasonable measure of total income earned from all employers. Again, we found virtually no difference in any of the results using this measure of the wage rate instead of the other two.

wages and unemployment. Because the coefficients in these regressions are not structurally interpretable, questions of 'bias' due to endogeneity or residual autocorrelation have little relevance or meaning.

Given two or more summary statistics of real wage behavior over time for different demographic groups, we will quantify differences between the two series in terms of *p*-values for a test of 'statistical significance' that the two series' coefficients on the unemployment rate are the same. For these tests, we use the Seemingly Unrelated Regressions framework, allowing for the variance of the residuals to differ between the two time series and for the residuals to be contemporaneously correlated, and test the single restriction that the coefficients on ΔUnemp (or Unemp) are identical using a standard Wald test.

III Results

Basic results

We consider first the aggregate wage statistic in the PSID, calculated using the PSID sample of male heads described above. Although there are a number of alternatives, we have chosen here total labor income of the panel divided by total hours worked (using the PSID family weights), because this is the methodology behind the BLS's Average Hourly Earnings statistic (although the BLS sample excludes government, agricultural, and nonproduction workers, while our sample here excludes those who are not male household heads, among other differences). Figure 4 presents this statistic, in (log) levels and (log) first-differences, deflated by the 1987 GDP deflator, and with a cubic trend for reference (a quadratic trend fits the series significantly more poorly, even for this brief period). A fair degree of procyclicality is evident in the diagram, and is reflected in regression coefficients on the unemployment rate: the coefficient on ΔUnemp in equation (1) is -0.0055 (SE 0.0025, $R^2 = 0.39$), while that on Unemp using model (2) is -0.0040 (SE 0.0016, $R^2 = 0.22$). Thus, aggregate 'real wage' changes of roughly 0.4–0.55% have been associated with 1% changes in unemployment over this period. These results are comparable to the findings of

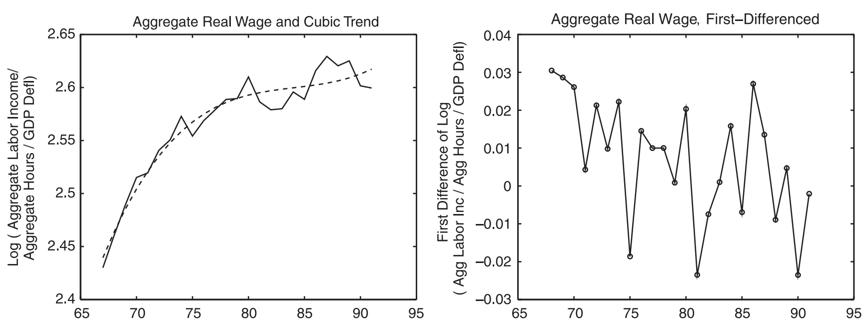


Figure 4. aggregate 'real wage' statistic.

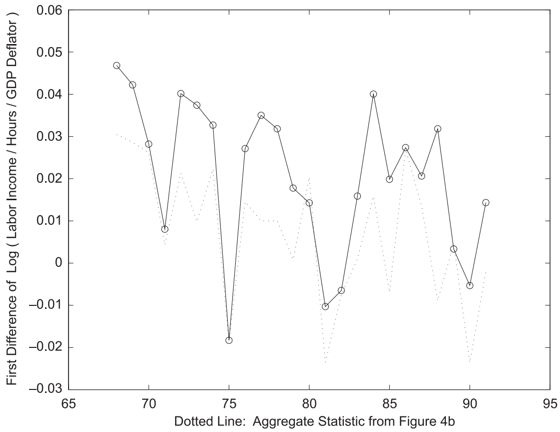


Figure 5. Average of real wage changes in the panel.

other researchers using aggregate data covering the same sample period (Abraham and Haltiwanger, 1995).

In contrast, the average year-to-year *changes* in real wages of the panel are presented in Figure 5 (along with the first-differenced aggregate statistic of Figure 4 for comparison). This average change in real wages is computed as follows: $\Delta \log w_{it}$ is calculated for each individual i with a real wage observation in both periods $t-1$ and t, and the average wage change is then taken across these individuals, weighted by the PSID family weights to make the average representative of the US population as a whole. This is the method followed by Bils (1985) and Solon *et al.* (1994), for example, and it avoids, to a very large extent, the aggregation biases mentioned earlier: workers are no longer weighted by their hours worked or their income received for the preceding year (unless they work zero hours for the year, in which case they are omitted from the sample; however, this holds for only a tiny fraction of the labor force). The result is a substantially more procyclical picture of real wages than before, as evident in the figure. The coefficient on ΔUnemp is -0.0118 (SE 0.0021, $R^2 = 0.67$), double the value of -0.0055 from Figure 4, and the difference is highly statistically significant (p-value < 0.001). This accords with the findings of SBP and others, who conclude that composition bias is a major source of error in traditional real wage studies, and that real wages at the individual level have been extremely procyclical since about 1967, typically varying by more than a full percent for each 1 percentage point change in the unemployment rate. The finding is quite striking and generally at odds with the view that workers and firms are moving along a stable aggregate labor demand curve over the course of a business cycle.

Before continuing, however, it should be noted that we fail to confirm these findings using either local-area unemployment rates in place of the national rate, or using employee's straight-time hourly pay rates instead of their annual wages divided by annual hours. We will investigate the possible reasons for this discrepancy, along with the findings themselves, in detail below.

First, however, we turn to some ways in which the above studies can be improved by making better use of the micro-data detail available in the PSID. Given the availability of data on individuals' wages, it is clearly desirable to find a more disaggregated format for its presentation. Regressions using the individual-level data have been run by many authors, with individual demographic variables and the aggregate unemployment rate as explanatory variables, but these suffer from the problem that the residuals are in general contemporaneously correlated, and hence all estimated SEs are incorrect and, in general, downward biased (Moulton, 1986, 1990).[4]

The approach adopted here is that of Figure 6. In every year between 1967 and 1991, we observe an entire distribution of real wages in the panel; Figure 6 presents a contour plot of these distributions, with contours drawn at each of the nine deciles: the middle contour plots the median real wage observed in each year, the bottom contour plots a real wage that is higher than exactly 10% of the wages observed in each particular year, etc. In order to make the diagram representative of the United States as a whole, these deciles have been computed using the PSID family weights. Three sets of axes are presented to emphasize different aspects of the distribution, ranging from its overall dispersion to finer levels of detail.

From the diagram, we can immediately see that the aggregate wage movements in Figures 4 and 5 do in fact correspond to shifts in the *entire distribution* of real wages in the economy. A pronounced degree of procyclicality is evident in all the deciles of the distribution (although noticeably more so for the bottom four, a fact that we shall return to below). This is interesting because it demonstrates that the procyclicality noted earlier has in fact been very widespread, and almost certainly experienced by a very large fraction of the population. Regressing the median contour on the unemployment rate using model (2) yields a coefficient of -0.0100 (SE 0.0019, $R^2 = 0.93$). Regression coefficients for the other contours range between -0.0075 and -0.0182.

Figures 7 and 8 present, in the same format, alternative perspectives on real wage cyclicality. In Figure 7, the distribution of real wage *changes* for each year is given [the middle contour plots the median real wage raise (or pay cut) received by the panel in each year, the bottom contour plots a real wage cut that is higher than exactly 10% of the real wage cuts experienced by the panel in each year, etc.]. This diagram is more comparable to Figure 5, and is more representative of individuals' experiences over the business cycle than was Figure 6, which details shifts in the aggregate distribution. We can see here that the procyclical shifts in wage changes are substantially smaller than was suggested by the average wage changes in Figure 5 [the regression coefficient of the median on ΔUnemp is only -0.0068 (SE 0.0015, $R^2 = 0.53$), compared with -0.0118 for the averages]. This is due to the exaggerated movements at the two tails of the distribution: a greater number of workers receive a dramatic pay cut in each recession, while at the same time the number of dramatic pay raises falls

[4] This problem can normally be corrected by a simple GLS procedure, but here the amount of data is so large that doing so is computationally intractable.

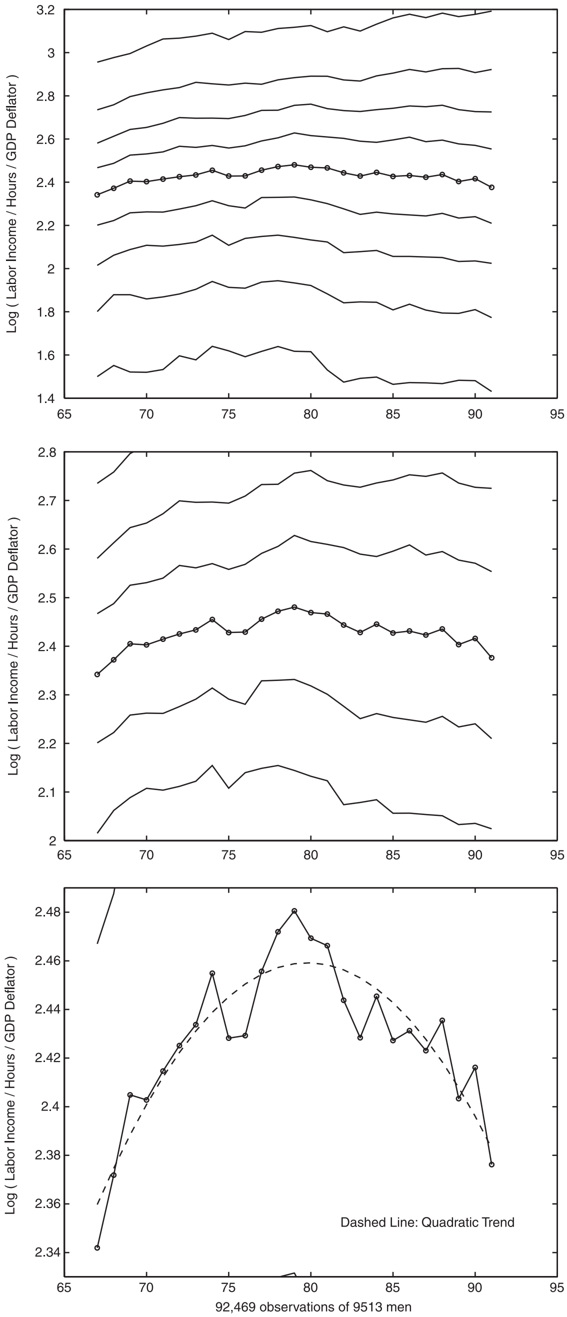


Figure 6. distribution of real wages over time.

substantially. Still, widespread procyclicality is evident in the diagram, confirming the findings in Figures 5 and 6.

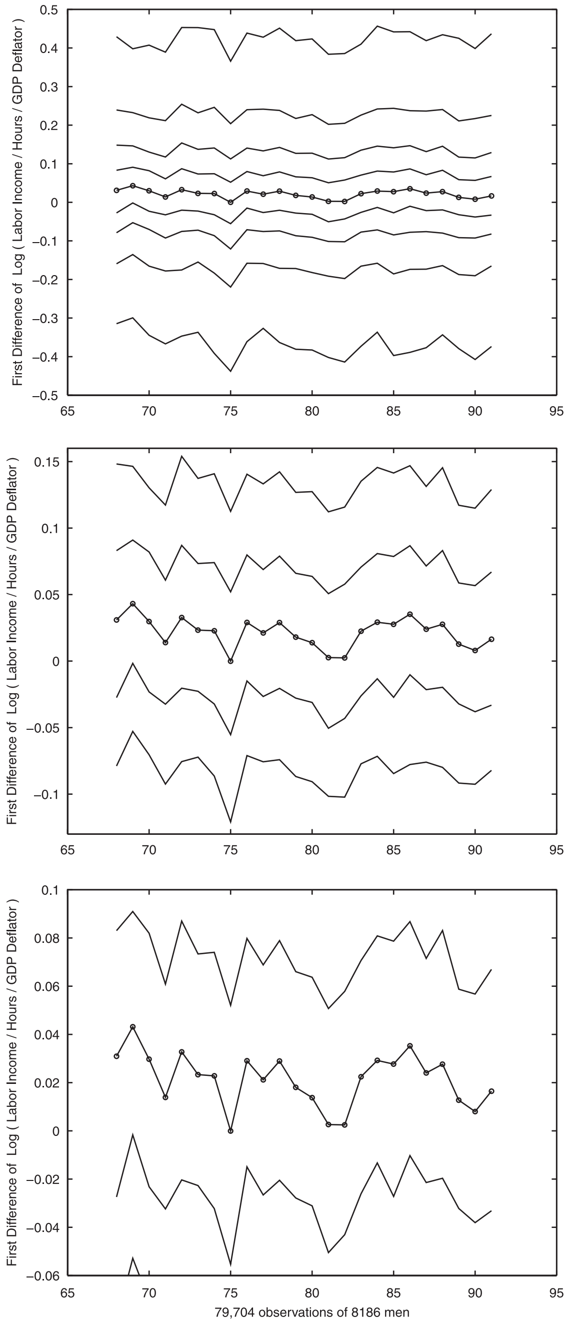


Figure 7. distribution of real wage changes over time.

Figure 8 presents a third measure of real wage cyclicality. Each of the 10,114 men in the sample can have a quadratic (or cubic or higher) trend fitted to

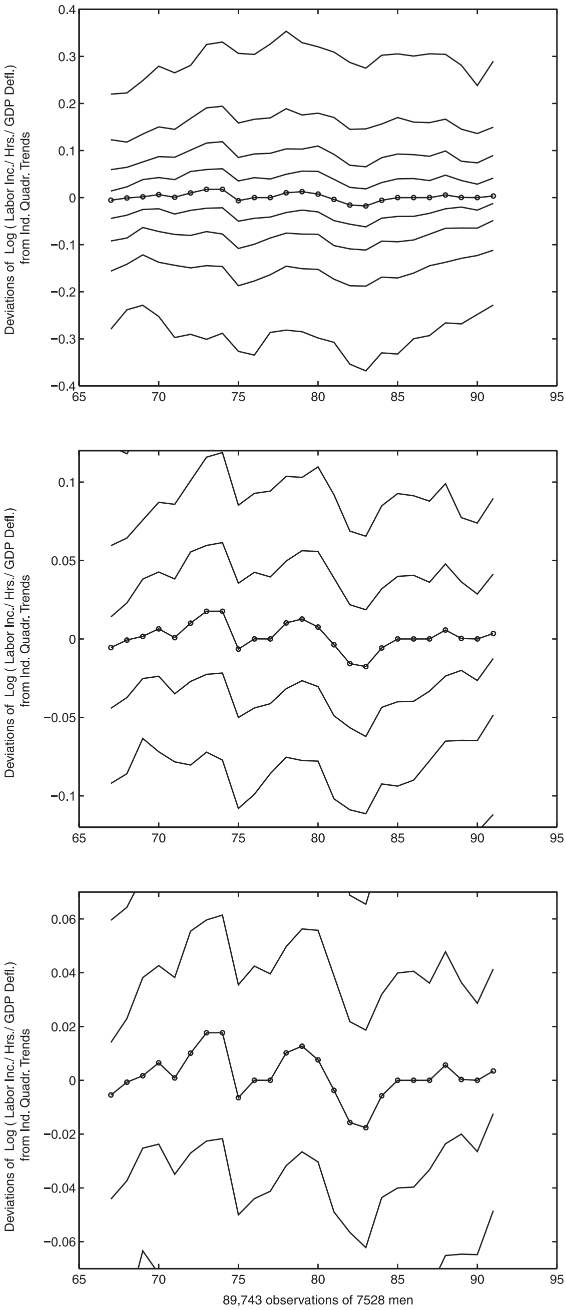


Figure 8. Real wage deviations from individual-specific trends.

his (log) wages over time. Deviations from this individual-specific trend are then computed and the distribution of these *deviations* is plotted in the figure.[5] The middle contour thus plots each year's median wage deviation from individual trend, the bottom contour plots a number that is greater than exactly 10% of the panel's wage deviations from individual trends, etc. Once again, the widespread procyclicality of Figure 6 is confirmed. Individuals are more likely to experience a decline in their real wage rate when the economy is in recession, and are more likely to receive a raise relative to trend when the economy is in a boom. The median contour's coefficient on Unemp is -0.0065 (SE 0.0011, $R^2 = 0.65$).[6]

In general, the disaggregate approach adopted here confirms other researchers' findings of strongly procyclical real wages over the period 1967–1991. In addition, we gain a sense of how consistent across recessions and how broad across the distribution of individuals these findings are. From the diagrams, it is clear that the entire distribution of wages, by several measures, is shifting downward in each recession. The magnitude of the shift appears to be between roughly -0.0065 and -0.0100, or slightly $<1\%$ for each percentage point change in the unemployment rate, in terms of the regression framework presented earlier; this is about 1.5–2 times larger than the aggregate 'real wage' would suggest, but smaller than what other researchers using panel data have concluded. These other researchers, in taking the average across the distribution, have given more weight to the greater movement in its tails than we have here.

Nominal wage rigidities and countercyclical prices?

Having noted the depth and breadth of real wage procyclicality in the micro-level data over this sample, it is natural to ask whether there is a simple explanation. For example, can the results be attributed directly to the oil price shocks of these years, or more generally to nominal wage rigidity and the countercyclicality of price movements over the sample period?

A quick look at Figure 9 suggests that oil prices are not going to be a significant explanatory variable, beyond the information that is already contained in an indicator of the business cycle. Although the oil price, represented here by the producer price of fuel (*Economic Report of the President*, Table B–64), exhibits significant spikes near each of the recessions in this period, the timing does not match the wage data nearly as well as either the unemployment rate or deviations of real GDP from trend. Moreover, the large declines in the price of oil during the 1980s are not matched well by the wage data.

These observations are borne out by a regression of the average wage change from Figure 5 on changes in the unemployment rate and the producer price of fuel. The regression equation

[5] Individuals with three or fewer wage observations over the period are excluded, because a quadratic trend would fit their data perfectly. Results using a cubic trend are very similar.

[6] For completeness, we have also tried removing individual-specific linear trends from the first-differenced data of Figure 7. The results, which are not presented here, are very similar to those in that figure.

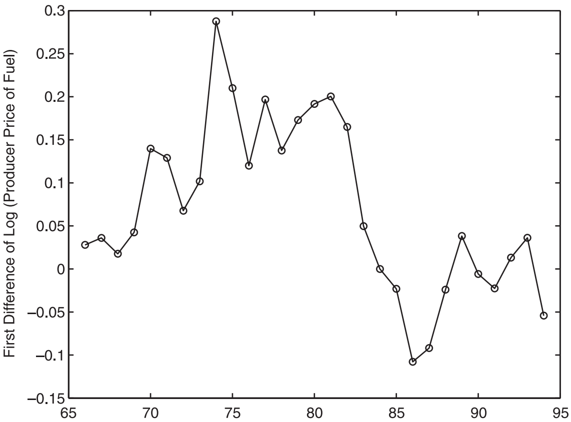


Figure 9. price of oil.

$$\Delta w_t = \beta_1 + \beta_2 t + \beta_3 \Delta \text{Unemp}_t + \beta_4 \Delta \log \text{poil}_t + \varepsilon_t, \tag{3}$$

yields a coefficient on ΔUnemp of − 0.0102 (SE 0.0022), virtually the same as in the original (nonoil) regression, while Δlog poil has a coefficient of − 0.0463 (SE 0.0271). Thus, a 1 standard deviation (1.1 percentage point) increase in the unemployment rate yields about a 1.1% decrease in the real wage, while a 1 standard deviation (10.4%) increase in the price of oil yields only about a 0.5% decrease in the real wage; moreover, the latter effect is not statistically significant (p-value = 0.104).[7] The oil shocks thus do not seem to offer a direct explanation for the behavior of real wages over this period.

Similarly, we can gain a basic sense of the importance of nominal wage rigidity by examining Figures 3 and 7. As with the oil shocks, upward movements in prices are associated with each recession over this period; again, though, the timing and magnitude of these price movements do not closely match those of the real wage changes in Figures 5 or 7. A business cycle or labor market indicator clearly yields a much better fit. This is corroborated by regression as well: the model

$$\Delta w_t = \beta_1 + \beta_2 t + \beta_3 \Delta \text{Unemp}_t + \beta_4 \Delta \log P_t + \varepsilon_t, \tag{4}$$

yields a coefficient on ΔUnemp of − 0.0102 (SE 0.0021), again virtually the same as in the original model, while the coefficient on Δlog P_t is − 0.2598 (SE 0.1180). A 1 standard deviation (2.1%) increase in the price level is associated with only a 0.5% decrease in the real wage, exactly the magnitude that was associated with the price of fuel directly. There is thus little evidence that the procyclical real wage–unemployment correlation can be explained simply by countercyclical movements in prices over the period.

[7] A regression using the median wage change from Figure 7 does yield a statistically significant coefficient on the price of oil, but the relative magnitudes of the two coefficients are unchanged.

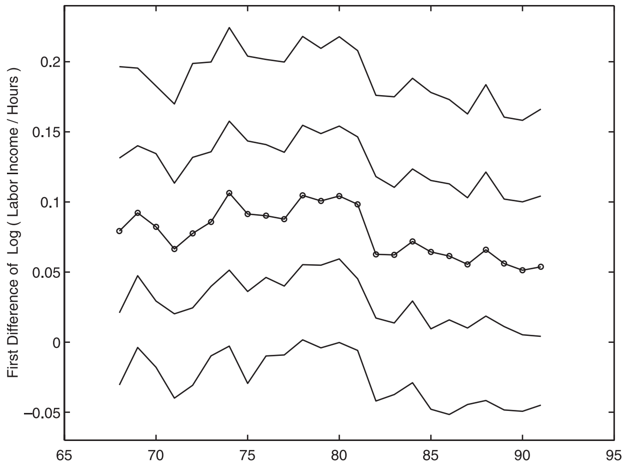


Figure 10. Nominal wage changes in the panel.

This point is further brought home by Figure 10, which depicts the undeflated, nominal wage changes corresponding to the middle panel of Figure 7. There is quite a bit of year-to-year variation in the size of nominal wage changes, and more importantly, in *every* year between 1967 and 1991, over 30% of the workers in the sample experience *nominal* wage cuts.[8] During recessions, and the 1980s and 1990s, this percentage is even greater. There is thus substantial evidence against the hypothesis that nominal wage rigidity and countercyclical prices have played any more than a modest role in the cyclical behavior of real wages over this period.

Local-area unemployment rates

Given the availability of local-area unemployment rates in the PSID, it is natural to ask whether findings of real wage procyclicality persist using these variables as well.

It is reasonable to expect that local-area unemployment rates should be related to workers' real wages, perhaps even more so than the national rate. Regional economic downturns, such as those in the Texas–Oklahoma area in 1986 or in Southern California in the early 1990s, clearly have important labor market effects. To the extent that real wage cyclicality is influenced by labor market conditions, we should expect to see a relationship at the state or local level.[9]

[8] McLaughlin (1994) notes the same phenomenon.

[9] Blanchflower and Oswald (1994) investigate this relationship in some detail, and claim to find convincing evidence of negative (i.e., procyclical in this context) relationship between real wages and local unemployment. Their use of weekly or annual wage rates rather than hourly wage rates weakens their argument, however, because weekly and annual hours will clearly be procyclical as well.

Table 1
Regression coefficients on unemployment (β_3)

Regression model	Unemployment measure: Local rate	National rate
First differences		
Basic model (5)	0.0002 (0.0008)	– 0.0114 (0.0016)
Time dummies (6)	– 0.0001 (0.0009)	N/A
Basic model with individual-specific trends	– 0.0001 (0.0007)	– 0.0093 (0.0015)
Time dummies with individual trends	– 0.0004 (0.0009)	N/A
Levels		
Basic model	– 0.0054 (0.0009)	– 0.0122 (0.0023)
Time dummies	– 0.0060 (0.0010)	N/A
Basic model with individual-specific trends	– 0.0007 (0.0004)	– 0.0112 (0.0010)
Time dummies with individual trends	– 0.0003 (0.0004)	N/A

Because these data are so disaggregate, a graphical approach similar to the previous section is infeasible. Pure regression analysis is the most convenient and informative approach. Using local-area unemployment rates, we can correct for the problem of contemporaneous residual correlation by including time dummies for each year; when using the national unemployment rate we do not have this option, because the national unemployment rate only varies in the time dimension. Unfortunately, dummies for each region cannot be used because the PSID censors the county of residence for each family, due to privacy considerations. Finally, note that the regressions here do *not* weight the observations by the corresponding family weights, because there is no reason to think that a high-weight individual has more accurate data than a low-weight individual.

The most natural regression to consider is

$$\Delta w_{it} = \beta_1 + \beta_2 t + \beta_3 \Delta \mathrm{Unemp}_{it} + \varepsilon_{it}, \tag{5}$$

without time dummies, or

$$\Delta w_{it} = \sum_{j=68}^{91} \gamma_{jt}\delta_{jt} + \beta_3 \Delta \mathrm{Unemp}_{it} + \varepsilon_{it}, \tag{6}$$

with time dummies. No matter what the model, however, the coefficient on the local unemployment rate is virtually zero. The results are presented in the upper middle column of Table 1. The third and fourth rows rerun regressions (5) and (6) after removing individual-specific linear trends from each worker's wage change data; this has the effect of controlling for *all* linear effects of individual-specific variables (e.g., race, age, education, experience, etc.) in one fell swoop. As can be seen from the table, this has little impact on the correlation with the unemployment rate.

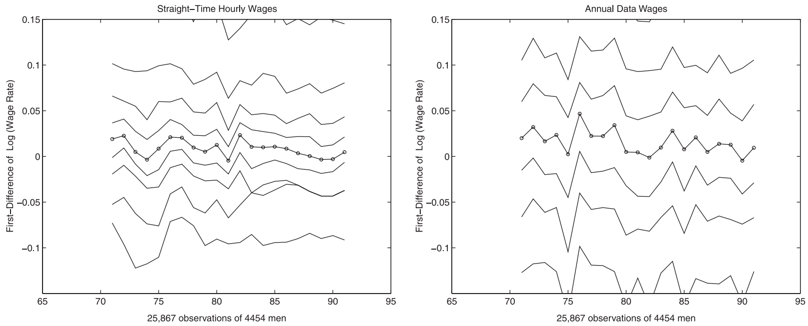


Figure 11. Straight-time hourly wages vs. annual data, for all hourly workers.

For comparison, the upper right column of Table 1 presents the results for the same regressions using changes in the national, rather than the local, unemployment rate for each individual (so $\Delta\text{Unemp}_{it} = \Delta\text{Unemp}_t$ for all i). Note that this latter set of numbers is very close to what was found in the analysis of the preceding section. Clearly, the national rate is much more closely correlated with individuals' real wages than is the local rate in this sample.

The bottom half of the table presents regression results for the corresponding levels specifications of models (5) and (6). A quadratic trend term is included in these specifications as well, both in model (5) and when removing individual-specific trends from the data. Thus the last two rows of the table control for all quadratic effects of individual-specific variables such as race, age, education, experience, etc. Using levels, the correlation with the local unemployment rate is stronger than with the first-differences, but not once individual trends have been removed from the data.

At first glance, these results are surprising. However, it is important to keep in mind that local unemployment rates are typically very poorly measured in the United States, and this is particularly true in the PSID, in which these data were binned in the years before 1981. The result of this measurement error will push all estimated coefficients on the variable toward zero, which appears to be the case here. Still, the fact that the estimated coefficients are almost exactly zero is disturbing.

Straight-time hourly wage rates

The PSID also collects data on worker's straight-time hourly pay rates, for workers who are explicitly paid by the hour on their current, primary job. In Figure 11, we plot the distribution of individual workers' straight-time hourly wage changes, along with the changes in the annual measure of wages for these exact same workers for comparison, as calculated from their reported annual earnings divided by annual hours.[10] Aside from variation in extra earnings,

[10] These data are lagged appropriately, to correspond to the year in which they were earned rather than reported, as always.

Table 2
Wage correlations with unemployment and prices

Real wage measure	Coefficient on ΔUnemp_t	Coefficient on $\Delta \log P_t$
Regression model (1)		
Annual data, all workers	– 0.0068 (0.0015)	N/A
Annual data, hourly workers	– 0.0068 (0.0017)	N/A
Straight-time wages, hourly workers	0.0000 (0.0016)	N/A
Regression model (4)		
Annual data, all workers	– 0.0051 (0.0013)	– 0.2698 (0.0743)
Annual data, hourly workers	– 0.0066 (0.0020)	– 0.0352 (0.1220)
Straight-time wages, hourly workers	0.0015 (0.0016)	– 0.2262 (0.1009)

income from extra jobs, reporting error, and mid-year job changes, these two diagrams ought to be *exactly* identical. In fact, they are considerably different. First, there is a great deal less spread in the wage-change distribution for workers' straight-time hourly pay than there is for their annual earnings divided by annual hours. Second, the cyclical movements of the former are more subdued, and slightly out of sync with the business cycle indicators in Figures 1 or 2 – note in particular the upticks in 1975 and 1982.[11] In fact, the timing of the straight-time hourly real wage movements appears to comove much more closely with the *price* changes in Figure 3, as would be the case if nominal wages were completely rigid. Regression analysis supports this observation: the relevant coefficients under models (1) and (4) are reported in Table 2. Note that changes in the price level play a much greater role for straight-time hourly wages than do changes in the business cycle indicator. This seems to indicate that nominal rigidity for straight-time wages is an important factor, although the fact that the relationship is less than one-for-one indicates that the rigidity is not perfect.

Nominal straight-time hourly wage rates are plotted in Figure 12. In comparison with Figure 10, note the greater compression (i.e., smaller cross-sectional variance) of the distribution, and the evidence of a significant point mass at zero. In sharp contrast to Figure 10, only 10–20% of workers in any given year experience a cut in their nominal straight-time hourly wage rate. The picture painted here is thus one of substantially greater nominal wage rigidity than was evident in the previous section.

There are several possible explanations for these findings. First, reporting error may be contaminating the annual income and hours data. Although Bound and Krueger (1991) and Bound *et al.* (1994) find acceptable levels of reporting error in annual earnings and annual hours, these errors are compounded when the quotient is taken to calculate average earnings per hour, and they find that this

[11] Using the PSID's 'wages and salaries' plus 'bonuses, overtime, and commissions' variables (or even 'wages and salaries' alone) rather than 'labor income' yields essentially identical results in this section, just as in the rest of the chapter.

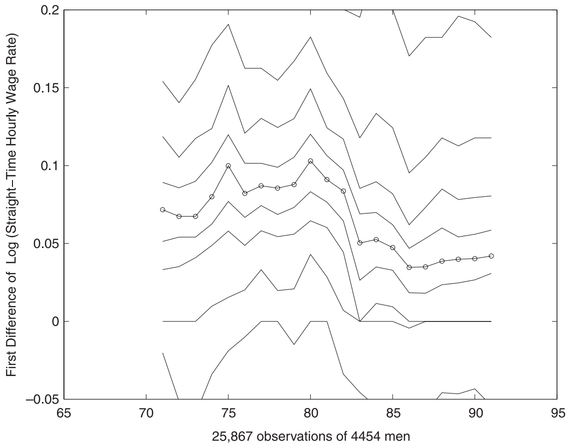


Figure 12. Nominal straight-time hourly wage changes.

results in substantial measurement error.[12] However, in the diagrams and regressions of this and the preceding sections, the measurement error is irrelevant as long as the means or medians of the distribution come out correctly each year. Only to the extent that means and medians are measured incorrectly, *and that this measurement error is correlated with the business cycle*, will the results here be affected at all. The possible impact of measurement error on our results here is thus minimized. However, it is possible that workers consistently bias their reported hours of work per week toward, say, forty, resulting in an overstatement of hours worked during recessions and an understatement during booms. If previous year's income is reported correctly, this would result in some procyclicality of the annually derived wage figures. However, it seems equally likely that workers in a downturn might exaggerate their loss of annual hours due to the recession, or unintentionally inflate the amount of overtime actually worked in a particularly robust year, and it is not clear that the first effect would dominate the second. Moreover, the fact that we are able to replicate movements in the BLS's aggregate wage statistic with a great deal of accuracy using PSID data suggests that a measurement error of this sort is not a serious problem. For all of these reasons, then, the effects of measurement error on our findings should be relatively minor.

Second, it is possible that additional income beyond straight-time hourly pay is responsible for the observed differences between the two figures. Bonuses, overtime, tips, commissions, and pay premia for evening and night shift work are all obviously very procyclical sources of income, and could potentially lead to exactly the effects documented here. Unfortunately, there are almost no direct data on these income components in the PSID. As mentioned earlier, data on extra income from bonuses, overtime, and the like are very often lumped

[12] The PSID attempts to minimize reporting errors in income by surveying its subjects shortly after income tax returns are due.

together with 'wages and salaries' in the PSID questionnaire, leaving the 'bonuses, overtime, and commissions' variable completely blank. The small amount of data that does exist in the PSID regarding these variables has no discernible effect on any of the diagrams presented so far. Evidence from the BLS's establishment survey indicates that average weekly overtime hours of production workers in manufacturing varies about 1.5 hours over the course of a business cycle. Dividing by the average, 40-hour workweek, and assuming a 50% premium for overtime yields an impact on average wages of about $((1.5)(1.5)+40)/41.5 = 1.018$, or slightly less than 2% over the course of a business cycle. From Shapiro (1996), the workweek of capital in manufacturing varies about 10% over the course of a business cycle; assuming a 25% premium for shift work (as Shapiro does) yields an additional impact of $((0.1)(1.25)+1)/1.1 = 1.023$ or 2.3% on the average wage. Unfortunately, we are not aware of any data on the importance of bonuses in US manufacturing or any other industry, especially as it varies over the business cycle. However, overtime and shift premia together are already accounting for roughly 4% movements in wages over the course of a business cycle, which is virtually as large, and certainly the same order of magnitude, as the effect we are trying to explain. Although manufacturing as an industry is very overtime- and shift work-intensive, the magnitude of movement in bonuses, tips, and commissions in other industries is likely to be of the same order of magnitude as in manufacturing.[13] Overall, then, this explanation is a feasible one.

One might think that second or even third jobs could help to explain workers' real wage procyclicality in a similar fashion. However, in order to have a procyclical rather than countercyclical effect, the additional jobs must pay higher wages than the worker's main job. Although it is plausible, we are not aware of any empirical evidence on this point. Moreover, second jobs are held by only a relatively small fraction (6%) of the work force, according to BLS statistics, and so any effects will be minor in relation to the aggregate. It is thus probably safe to dismiss this theory as a possible explanation for the findings of the present paper.

Finally, job changes over the course of a year, or job seasonality, could lead to discrepancies between a worker's reported straight-time hourly wage rate and his actual average hourly wage over the course of the year. This would be the case if, for example, a worker's straight-time wage at the time of interview was not representative of the wage he actually earned over the rest of the year. As regards job changes, both Bils (1985) and Solon *et al.* (1992) have noted the large procyclical impact of a job change on a worker's wage, but if the change occurs before the PSID interview for a given year, it will be picked up in the reported straight-time hourly wage rate for that year as well. Even if the change occurs after the time of interview for a given year, it will still be picked up in the reported hourly wage at the time of interview the following year. Thus, at worst,

[13] The bonuses, overtime, etc. explanation may also help to explain why women's wages are significantly less procyclical than men's, as documented by Solon *et al.* (1994). To the extent that men are concentrated in more overtime- and shift work-intensive industries, we would expect their wages to be more procyclical.

the effect of job changes on reported straight-time hourly wage rates will simply be to spread the change out over a 2-year period, blurring it somewhat but not hiding it altogether. This is not what we observed in Figure 11. As regards job seasonality, it is not clear that there would be any variation in this phenomenon over the course of a business cycle, which would preclude it from being a significant explanator of the discrepancy noted in Figure 11 as well. Thus, it seems that we can also eliminate both of these theories as explanations for the differing behavior of straight-time hourly wage rates and average wage rates derived from annual data over the course of the business cycle.

To summarize, then, it appears that bonuses, tips, commissions, and premia for overtime and shift work are playing a substantial role in the observations of real wage procyclicality that we made earlier. Straight-time hourly wages, by contrast, do not appear to vary significantly over the business cycle; in fact, they appear to exhibit substantial nominal rigidity,[14] and vary more (inversely) with movements in prices than with movements in a labor market indicator or indicator of the business cycle.

Demographic differences in real wage cyclicality

Having studied the behavior of real wages over the business cycle for the population as a whole, it is natural to ask to what extent observed cyclicality differs across major demographic groups. This question is interesting for two reasons: first, demographic differences in real wage cyclicality can shed light on certain macroeconomic theories; for example, in the presence of insider–outsider effects, we would expect to see the equilibrium wage of young labor market entrants vary much more than the wages of older, more established workers. Second, demographic differences in real wage cyclicality help to identify empirically the nature and magnitude of aggregation bias in the data; for example, we have already noted above how bias can arise from changes in worker composition, industry composition, and the greater weighting of high-income individuals. How large is each of these compositional effects in practice?[15]

Rather than present a full array of graphs for each demographic comparison, only the middle set of axes for the first-differences specification (corresponding to Figure 7) will be presented. Results corresponding to the methods of Figures 6 and 8 (the levels and deviations methods) are typically very similar.[16]

[14] Note that this finding is in contrast to McLaughlin (1994), who only looks at workers' annual data.

[15] Note that Solon *et al.* (1994) attribute aggregation bias in their sample to the first source, while Bils (1985) attributes it to the third. There has not been any attempt in the literature to separate out the relative sizes of these effects.

[16] The breadth of the wage level distributions of Figure 6 makes it difficult to view both cyclical variation and a reasonable fraction of the distribution at the same time, and the necessity of fitting a quadratic trend that is downward sloping in the later years of the sample is a drawback. Although one might think that the very high R^2 of the levels regressions in Figure 6 indicates a superior fit, in fact it is due primarily to the explanatory power of the quadratic trend rather than the unemployment rate; the raw sum of squared residuals for the levels regression is

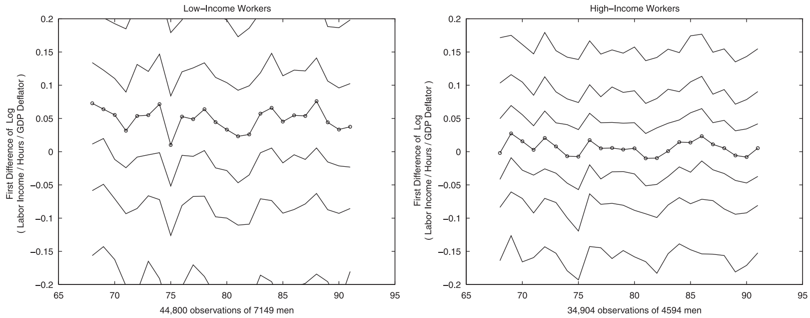


Figure 13. Wage cyclicality for low-income vs. high-income workers.

We begin with a comparison of real wage cyclicality across high- and low-income individuals. Recall that there was already suggestive evidence of such a difference in the wage-level contours of Figure 6. Moreover, Bils (1985) attributes almost all of the aggregation bias in his sample to the difference in cyclicality between these income groups, rather than to changes in sample composition over the cycle. Hence, the existence and size of a difference between these groups is important.

Note that income, rather than wages, is the appropriate basis for comparison here: for example, high-wage workers who work no hours have no impact on the average aggregate wage. Mathematically (holding hours h_{it} constant in order to isolate the income-weighting effect),

$$\Delta \log\left(\frac{\sum w_{it}h_{it}}{\sum h_{it}}\right) = \sum \frac{w_{it}h_{it}}{\sum w_{it}h_{it}} \Delta \log w_{it}. \tag{7}$$

Relative changes in wages, $\Delta\log w_{it}$, are weighted exactly by the individual's share in total labor income.[17]

We begin by calculating labor income deciles for each year. Individuals who earn zero labor income in a given year are not counted, and we use the family weights as usual. High-income individuals are defined to be those with labor income in the top five deciles, and low-income earners are those in the bottom five. In looking at first-differenced data, we apply this criterion to the first of the two years that make up a person's wage change.

The results are presented in Figure 13. It is clear that low-income workers have experienced much greater wage cyclicality than their high-income counterparts over this period; the coefficients on the medians of the figures are -0.0111 (SE 0.0023, $R^2 = 0.55$) and -0.0047 (SE 0.0018, $R^2 = 0.27$), respectively, and this difference is highly statistically significant (p-value = 0.019). Thus, low-income workers experience greater cyclicality in wages

in fact about 40% *greater* than that for the first-differences, indicating an inferior fit. Finally, the previous literature has generally focused on first-differences rather than levels, so focusing on that format here enhances comparability to previous work.

[17] The following subsection provides a more complete breakdown of aggregate wage cyclicality into its constituent components.

as well as in employment over the business cycle.[18] Income weighting may indeed be an important source of bias in the aggregate wage statistic. We will return to this question below.

We next focus on race, age, and education, looking at whites vs. blacks (and ignoring others), age groups 20–29, 30–44, and 45–54 (ignoring others), and individuals with no high school degree, those with a high school degree or GED but no years of college, and those with a 4-year college degree (again, ignoring others). The breakdown by age is of particular interest, because wages of new entrants into the labor market may be more sensitive to cyclical conditions than those already on the 'inside'.

Figure 14 presents the full array of graphs for a demographic partition along these lines. Note that the figures for blacks have been omitted due to small cell sizes that result in an enormous amount of noise and an inability to draw meaningful conclusions.[19] Even for whites, there is an issue of small sample sizes for many of the cells; still, real wage procyclicality is evident in virtually every one of the diagrams. Indeed, the estimated coefficient of the median on changes in the unemployment rate is negative for every single cell – the results are presented in Table 3.

The tremendous amount of cyclical variation in wages of the youngest and least-educated workers is the most striking feature of the diagram. Here, at least, an insider–outsider theory appears to be borne out. More generally, it does appear that wage cyclicality decreases with both age and education, at least in the first two rows of the figure. However, none of the differences in medians is statistically significant by any conventional measure, except for the youngest and least educated workers (i.e., the top left cell), who experience wage changes that are significantly more procyclical than those of any other cell (p-values all <0.01). Thus, there is suggestive evidence of decreasing cyclicality with age and education within the sample, especially for the youngest and least educated workers, but the finding has only limited statistical support on the whole.

Sources of aggregation bias

We have discussed above the possible importance of worker composition bias, industrial composition bias, and income weighting on the measured cyclicality of the aggregate real wage statistic. Researchers have differed regarding the relative importance of these factors, with Solon *et al.* (1994) coming down strongly in favor of the first source of bias, and Bils (1985) very much emphasizing the third. Moreover, with our finding of substantial differences in real wage cyclicality between high- and low-income workers in the previous section, it would be

[18] This empirical observation remains true when we control for other demographic variables such as race, age, and education as well; however, the difference between estimated medians of the corresponding figures loses its statistical significance, owing to a greater amount of noise in the data, due to the smaller number of observations.

[19] Lumping all blacks together and comparing them with all whites confirms that the wage distributions of the former fluctuate more dramatically over time, but there is not a clear difference in the wage *cyclicality* of blacks vis-à-vis whites.

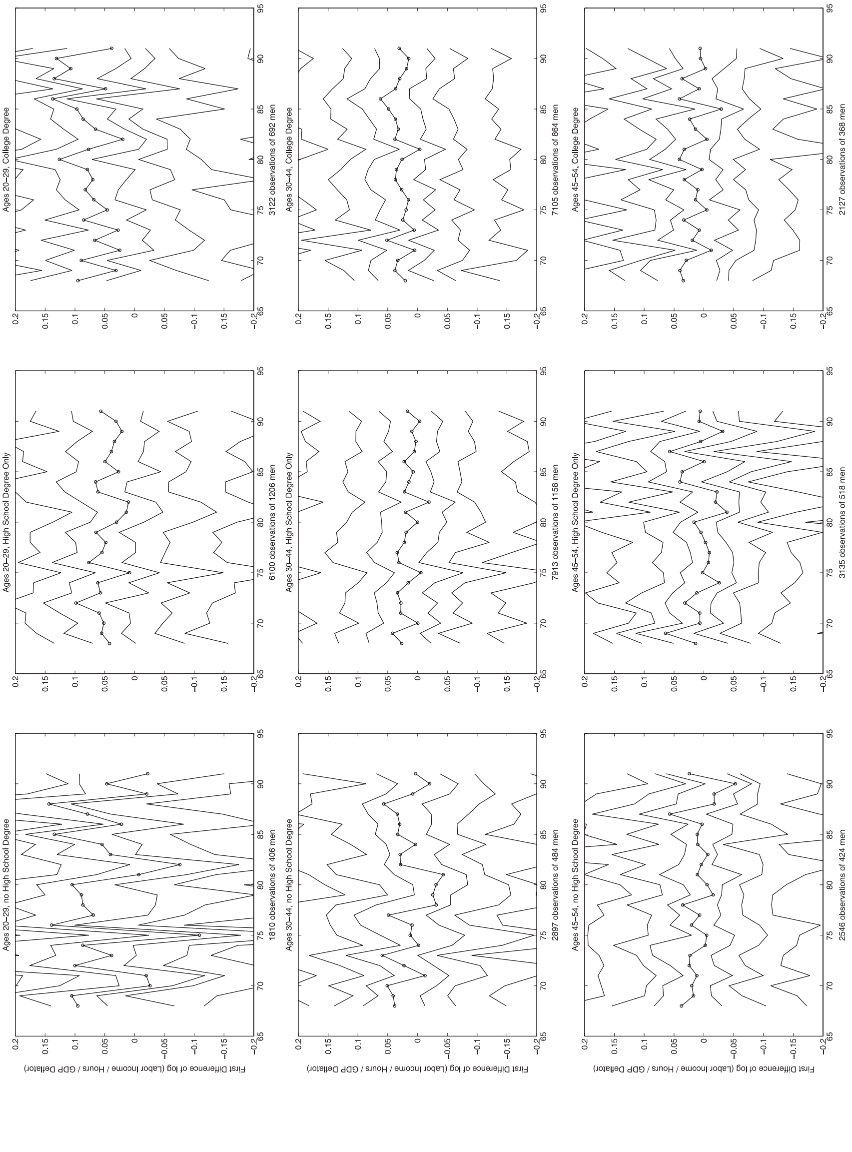


Figure 14. Wage cyclicality by demographic group, whites.

Table 3
Wage cyclicality by demographic group, whites

Age	No HS degree	HS degree, no college	College degree
20–29	– 0.0436 (0.0097), $R^2 = 0.49$	– 0.0106 (0.0034), $R^2 = 0.41$	– 0.0063 (0.0064), $R^2 = 0.18$
30–44	– 0.0045 (0.0058), $R^2 = 0.06$	– 0.0081 (0.0020), $R^2 = 0.55$	– 0.0023 (0.0030), $R^2 = 0.06$
45–54	– 0.0053 (0.0040), $R^2 = 0.21$	– 0.0076 (0.0049), $R^2 = 0.12$	– 0.0041 (0.0036), $R^2 = 0.10$

interesting to know how much this discrepancy contributes to the aggregation bias identified earlier. In this section, we provide a decomposition of aggregation bias into its constituent components, which will help shed light on these issues.

For clarity of exposition, assume that hours and wages are both perfectly correlated with the business cycle, so that

$$h_{ijt} = \bar{h}_{ij} + (\bar{h}_{ij}\gamma_{ij})u_t,$$
$$w_{ijt} = \bar{w}_{ij} + (\bar{w}_{ij}\beta_{ij})u_t,$$

where h_{ijt} and w_{ijt} are hours and wages for industry i, worker j, at time t, and u_t is an indicator of the business cycle, such as the unemployment rate (expressed as a deviation from the natural rate of unemployment).

The standard aggregate hourly wage statistic has the form

$$W_t = \sum_{i,j} \left(\frac{h_{ijt}}{\sum h_{ijt}}\right) w_{ijt}.$$

From the above, we have

$$\sum h_{ijt} = \sum \bar{h}_{ij} + \sum (\bar{h}_{ij}\gamma_{ij})u_t,$$

and because the cyclical variation in hours is small relative to total hours, the first-order Taylor series approximation,

$$\frac{1}{\sum h_{ijt}} \approx \frac{1}{\sum \bar{h}_{ij}} - \frac{\sum (\bar{h}_{ij}\gamma_{ij})}{\left(\sum \bar{h}_{ij}\right)^2} u_t,$$

is excellent. Thus,

$$\frac{h_{ijt}}{\sum h_{ijt}} \approx \frac{\bar{h}_{ij}}{\sum \bar{h}_{ij}} + \frac{\bar{h}_{ij}\gamma_{ij} \sum \bar{h}_{ij} - \bar{h}_{ij} \sum (\bar{h}_{ij}\gamma_{ij})}{\left(\sum \bar{h}_{ij}\right)^2} u_t,$$

and

$$
\begin{aligned}
W_t &= \sum\left(\frac{h_{ijt}}{\sum h_{ijt}}\right)\bar{w}_{ij} + \sum\left(\frac{h_{ijt}}{\sum h_{ijt}}\right)\bar{w}_{ij}\beta_{ij}u_t \\
&\approx \sum\frac{\bar{h}_{ij}}{\sum\bar{h}_{ij}}\bar{w}_{ij} + \sum\frac{\bar{h}_{ij}\bar{w}_{ij}\gamma_{ij}\sum\bar{h}_{ij} - \bar{h}_{ij}\bar{w}_{ij}\sum(\bar{h}_{ij}\gamma_{ij})}{\left(\sum\bar{h}_{ij}\right)^2}u_t + \sum\frac{\bar{h}_{ij}\bar{w}_{ij}}{\sum\bar{h}_{ij}}\beta_{ij}u_t \\
&= \frac{\sum\bar{h}_{ij}\bar{w}_{ij}}{\sum\bar{h}_{ij}} + \sum\frac{\bar{h}_{ij}}{\sum\bar{h}_{ij}}\bar{w}_{ij}\left[\gamma_{ij} - \sum\frac{\bar{h}_{ij}}{\sum\bar{h}_{ij}}\gamma_{ij}\right]u_t + \sum\frac{\bar{h}_{ij}\bar{w}_{ij}}{\sum\bar{h}_{ij}}\beta_{ij}u_t \\
&= \bar{W} + \sum\frac{\bar{h}_{ij}}{\sum\bar{h}_{ij}}\bar{w}_{ij}\left[\gamma_{ij} - \sum\frac{\bar{h}_{ij}}{\sum\bar{h}_{ij}}\gamma_{ij}\right]u_t + \bar{W}\sum\frac{\bar{h}_{ij}\bar{w}_{ij}}{\sum\bar{h}_{ij}\bar{w}_{ij}}\beta_{ij}u_t.
\end{aligned}
$$

This yields:

$$
\begin{aligned}
\Delta\log W_t = {}& \frac{1}{\bar{W}}\sum\frac{\bar{h}_{ij}}{\sum\bar{h}_{ij}}\bar{w}_{ij}\left[\gamma_{ij} - \sum\frac{\bar{h}_{ij}}{\sum\bar{h}_{ij}}\gamma_{ij}\right]\Delta u_t \\
&+ \sum\frac{\bar{h}_{ij}\bar{w}_{ij}}{\sum\bar{h}_{ij}\bar{w}_{ij}}\beta_{ij}\Delta u_t.
\end{aligned}
\tag{8}
$$

The second term on the right-hand side of equation (8) is the income-weighted average of the individuals' wage cyclicality coefficients, β_{ij}. The first term on the right reflects the change in sample composition over the period: it depends only on cyclical fluctuations in individuals' hours (the γ_{ij}), and not on individual wage cyclicality at all. In fact, the quantity inside the summation is exactly the sample covariance of individuals' wages and their cyclicality of hours γ_{ij}, with each individual assigned a weight equal to his average hours $\bar{h}_{ij}$ over the period. When this covariance is negative (as when low-wage workers are more likely to be laid off in a recession), a downward bias is imparted to the cyclicality of the aggregate wage statistic.

We can further separate the composition bias term into industry and worker components. Let $\bar{w}_i$ be the average wage in industry i, and define

$$\bar{w}_j \equiv \bar{w}_{ij} - \bar{w}_i$$

to be individual j's deviation from the industry average. Substituting into equation (8) gives

$$
\begin{aligned}
\Delta\log W_t = {}& \frac{1}{\bar{W}}\sum\frac{\bar{h}_{ij}}{\sum\bar{h}_{ij}}\bar{w}_i\left[\gamma_{ij} - \sum\frac{\bar{h}_{ij}}{\sum\bar{h}_{ij}}\gamma_{ij}\right]\Delta u_t \\
&+ \frac{1}{\bar{W}}\sum\frac{\bar{h}_{ij}}{\sum\bar{h}_{ij}}\bar{w}_j\left[\gamma_{ij} - \sum\frac{\bar{h}_{ij}}{\sum\bar{h}_{ij}}\gamma_{ij}\right]\Delta u_t + \sum\frac{\bar{h}_{ij}\bar{w}_{ij}}{\sum\bar{h}_{ij}\bar{w}_{ij}}\beta_{ij}\Delta u_t.
\end{aligned}
$$

There is industrial composition bias when $\bar{w}_i$ is correlated with γ_{ij}, and worker composition bias when $\bar{w}_j$ is correlated with γ_{ij}. In the United States, the first correlation is positive and the second is negative.

To determine exactly to what extent composition bias impacts the cyclicality of the aggregate wage statistic, we need only consider the term

$$\sum\left(\frac{\bar{h}_{ij}\bar{w}_{ij}}{\sum\bar{h}_{ij}\bar{w}_{ij}}\right)\beta_{ij}\Delta u_t,$$

which is just $\Delta \log W_t$ with the composition shift terms removed. Because $\Delta \log w_{ijt} = \beta_{ij}\Delta u_t$, this becomes

$$\sum\left(\frac{\bar{h}_{ij}\bar{w}_{ij}}{\sum \bar{h}_{ij}\bar{w}_{ij}}\right)\Delta \log w_{ijt}. \tag{9}$$

A regression of equation (9) on the change in the unemployment rate will then yield the desired, composition-bias free coefficient. Note, however, that this is different from the regression that other panel studies have actually run. There, it is standard practice to regress the (simple) average change in log wage,

$$\frac{1}{N}\sum \Delta \log w_{ijt} = \frac{1}{N}\sum \beta_{ij}\Delta u_t, \tag{10}$$

on the change in the unemployment rate, which purges the effects of income weighting from the cyclicality of the aggregate wage statistic as well. Regression estimates for equation (10) that yield substantially greater procyclicality than a regression of the aggregate wage statistic are thus not conclusive as to the importance of composition bias by itself. We saw in the previous section that income weighting was potentially a large source of difference between these regression estimates also.

To assess the importance of composition bias accurately, we must run regression (9) (including a constant and time trend to account for growth in average wages over time). The result is a composition-free coefficient on ΔUnemp of −0.0120 (SE 0.0027), which is essentially identical to the value of −0.0118 obtained by traditional panel studies using equation (10). Thus, despite the important differences in wage cyclicality by income group apparent in Figure 13, for all practical purposes income weighting appears to play a negligible role in the cyclicality of the aggregate wage statistic. All of the bias observed in going from the aggregate to the panel regression appears to be due to changes in sample composition. The findings here thus come down strongly in favor of the composition effects emphasized by Solon *et al.* (1994) rather than the income weighting favored by Bils (1985).

IV Discussion and Conclusions

We draw several conclusions from this analysis. First, real wages of individual workers in the United States were strongly procyclical from 1967 to 1991, much more so than previous studies of aggregated real wage data have suggested. This finding is robust across recessions as well as pervasive throughout the US economy, shifting the entire distribution of workers' wages by any of several measures. This pattern of real wage movements tracks the business cycle very closely, and cannot be explained simply by price movements and nominal wage rigidity or by the oil shocks that occurred during the period.

Second, workers' straight-time hourly pay rates vary much less over the business cycle than do their wages as measured by total annual income divided by total annual hours. Indeed, straight-time hourly pay rates seem to suffer from substantial nominal wage rigidity. A number of possible explanations were offered for this difference in cyclicality, with the evidence strongly favoring variation in bonuses, overtime, shift premia, tips, and commissions as the source of cyclical variation in total wages. The implication is that employers have some latitude with which they can easily adjust wages over the business cycle, but are more constrained when it comes to larger adjustments of nominal wages, for which changes in straight-time pay might be required.

Third, the correlation of real wages with local-area unemployment rates is small, although this may be due to the large amount of noise present in the local-area unemployment statistics, particularly as they are reported in the PSID. Future research using state-level unemployment rates could potentially resolve these difficulties, and replicate the findings of procyclicality that are apparent with respect to the national rate.

Fourth, there are demographic differences in real wage cyclicality that can provide an insight into the labor market. For example, real wage cyclicality appears to decrease with age, education, and income, supporting insider–outsider models of the labor market, in which the youngest and least experienced workers are the most susceptible to fluctuations in labor market conditions and those on the 'inside' are more insulated from these shocks. The substantially smaller cyclicality of high-income workers' wages also raises the possibility that this is a major source of the muted cyclicality that is found in the aggregate real wage statistic computed by the BLS, as suggested by Bils (1985). However, a detailed decomposition of this statistic showed that this is not the case; instead, the difference appears to be due to the changing composition of the work force over the business cycle, as maintained by Solon *et al.* (1994).

It should be emphasized, however, that just because workers' wages were significantly procylical from 1967 to 1991 does not imply that they have always been so. For example, anecdotal evidence from the Great Depression and the 1920–1921 contraction strongly suggests that real wages were countercyclical during these episodes: e.g., '[Benjamin] Strong wanted to wait until wage rates were lower. He noted that deposits had fallen off considerably, retail prices had fallen moderately, wholesale prices precipitously [56%], but wages had hardly been affected' (Friedman and Schwartz, 1963, p. 234). In fact, these observations lend support to the idea that employers can easily vary a worker's wages only to the extent that they can vary his bonuses, commissions, shift premia, and the like. Larger changes in nominal wages may be constrained by the rigidity that appears to be present in workers' straight-time hourly pay. The nominal wage declines on the order of 5% that employers managed to implement in the post-1967 PSID sample might well be swamped by larger price changes such as those that took place during the Great Depression and 1920–1921 contraction.

Finally, Swanson (2004) shows that, despite the procyclicality of individual workers' wages with respect to aggregrate price measures like the CPI and GDP deflator, workers' wages have been *counter*cyclical over both the post-War and post-1967 period when those wages are deflated by the price index of the worker's *own* two- or four-digit industry and compared with the state of economic activity in that same industry. Intuitively, a positive economic shock that impacts one sector of the economy more than others can lead to an increase in the relative price of that sector's good, a corresponding decrease in that sector's real wage deflated by its product price, and an increase in employment and the utilization of capital (and labor) in that sector. This change in capital and labor utilization is consistent with an increase in labor productivity and CPI-deflated real wages in the sector *despite* the fall in real wages deflated by the sector's product price. These effects can be demonstrated rigorously in a fully specified general equilibrium framework (Swanson, 2006). The observations in this paper and the others mentioned above suggest that this pattern may be a common feature of the post-War US economy and thus that further empirical and theoretical work along these lines might be illuminating.

Acknowledgements

This paper is based on Chapter 2 of my doctoral dissertation at Stanford. I thank John Pencavel, Jonathan Parker, Gary Solon, Michael Horvath, Robert Hall, and John Taylor for helpful comments and suggestions. The views expressed in this paper, and all errors and omissions, should be regarded as those solely of the author, and are not necessarily those of the individuals listed above, the management of the Federal Reserve Bank of San Francisco, or any other individual in the Federal Reserve System.

Appendix A: Detailed Data and Methods

The PSID data for the interview years 1968–1988 were taken from the 1988 cross-year family-individual file on cd-rom. Data for the interview years 1989–1992 were downloaded from the PSID's home page on the World Wide Web. As mentioned in section 2, my sample consists of all men who were ever household heads in the PSID, excluding the Latino sample. For each of these men, I made use of the following variables (numbers in parentheses indicate the 1992 variable number for exact reference, and years for which comparable data were available, if not the entire period): total labor income of head (V21484); wages and salaries of head (V20429, 1970–1992); bonuses, overtime, and commissions of head (V20431, 1976–1992); annual hours worked by head (V20344); head's race (V21420), age (V30736), and education (V21504, V21423); whether head is employed by the government or private sector (V20698, 1975–1992); whether head's job is covered by a union contract (V20699, 1976–1992); head's hourly wage if paid by the hour (V20707, 1970–1992); unemployment rate in head's county of residence (V21521); and family weight for the head's family (V21547). Details of these data are available in Survey Research Center (1973–1995, 1986).

A few of these variables required modification in order to remove deficiencies or make them comparable to other years' data. Labor income is top coded at \$99,999 until 1983, at which point it is top coded at \$999,999. I obtained a list of the true values for these incomes through 1988 from Gary Solon, and entered them in by hand. Between 1988 and 1992, there are only one or two people who ever reach the top-coded amount, and their data were omitted for those years. Wages and salaries are similarly top coded, and I omitted those data unless it was clear from the context that the individual's wages and salaries were equal to his labor income, in which case I assigned the labor income values from the top-coded list. Bonuses, overtime, and commissions are top coded at $99,999, and I omitted these data unless the true value could be deduced from the context as being equal to labor income minus wages and salaries. Finally, in 1992, the income of self-employed businessmen and professionals appears to have been top coded at $99,999 even within the labor income variable – this was not the case in earlier years – hence it was necessary to delete two or three observations on this basis as well.

The education variable changes formats in 1985 and again in 1992, so that in later years it becomes necessary to modify the variable for comparability, by including those with a GED in the high school bracket after 1985, and by binning the 1992 data into corresponding brackets for earlier years.

Finally, the unemployment data through 1980 are available only in 2% brackets rather than to the nearest whole percent, as in later years. I simply assigned the midpoint of the bracket to these years, which is equivalent to rounding unemployment to the nearest 2%, rather than the nearest whole percent. No attempt was made to account for heteroskedasticity in the econometric analysis. Unemployment was also lagged one year for both the 1991 and 1992 interview years, thereby replacing the 1990 value as well, due to a change in the definition of the variable.

The annual hours, labor income, and wages and salaries data were subjected to an accuracy screening in order to eliminate observations for which 'major assignments' were made by the PSID staff. This is necessary when wage cyclicality is the subject of study, because the PSID's most common assignment procedure by far is simply to give the previous year's value to the current year, creating a bias toward zero change. About 3% of the weighted observations in recent years have major assignments to labor income, and about 2.5% have major assignments to annual hours. The screening procedure for labor income is straightforward until 1976, at which point there are separate accuracy codes for wages and salaries and for 'labor income excluding wages and salaries'. I deleted the labor income observation if and only if either of these variables report major assignments. For annual hours, the procedure is straightforward until 1985, at which point there are separate accuracy codes for hours on main job, hours on extra jobs, and hours of overtime. I ignored the accuracy of overtime hours and deleted the observation if and only if hours on either the main job or extra jobs was a major assignment. Screening the wages and salaries variable is straightforward for all years.

As noted throughout the text, the data were weighted using the PSID family weights. This is superior to simply truncating the SEO subsample, because it gives us roughly 40% more observations, not to mention that it corrects for differential nonresponse across demographic groups, which truncating the SEO does not. The observations could also be weighted using the PSID individual weights. However, this is essentially equivalent to discarding nonsample spouses and doubling the weight of sample members who themselves have nonsample spouses. The resulting weighted population is the same, and so I found it preferable not to discard these observations. Finally, when dealing with wage changes, I simply assigned the family weight for the latter of the 2 years that make up the individual's wage change.

References

ABRAHAM, K. and HALTIWANGER, J. (1995). Real wages and the business cycle. *Journal of Economic Literature*, **33**, pp. 1215–65.

BARRO, R. and KING, R. (1984). Time-separable preferences and intertemporal substitution models of business cycles. *Quarterly Journal of Economics*, **99**, pp. 817–39.

BARTELSMAN, E., CABALLERO, R. and LYONS, R. (1994). Customer- and supplier-driven externalities. *American Economic Review*, **84**, pp. 1075–84.

BILS, M. (1985). Real wages over the business cycle: evidence from panel data. *Journal of Political Economy*, **93**, pp. 666–89.

BLANCHFLOWER, D. and OSWALD, A. (1994). *The Wage Curve*. Cambridge, MA: MIT Press.

BLANK, R. (1990). Why are wages cyclical in the 1970's? *Journal of Labor Economics*, **8**, pp. 16–47.

BOUND, J., BROWN, C., DUNCAN, G. J. and ROGERS, W. L. (1994). Evidence on the validity of cross-sectional and longitudinal labor market data. *Journal of Labor Economics*, **12**, pp. 345–68.

BOUND, J. and KRUEGER, A. B. (1991). The extent of measurement error in longitudinal earnings data: do two wrongs make a right? *Journal of Labor Economics*, **9**, pp. 1–24.

CHIRINKO, R. (1980). The real wage rate over the business cycle. *Review of Economics and Statistics*, **62**, pp. 459–61.

FRIEDMAN, M. and SCHWARTZ, A. J. (1963). *A Monetary History of the United States, 1867–1960*. Princeton, NJ: Princeton University Press.

KEYNES, J. M. (1936). *The General Theory of Employment, Interest, and Money*. London: MacMillan.

KYDLAND, F. and PRESCOTT, E. (1982). Time to build and aggregate fluctuations. *Econometrica*, **50**, pp. 1345–70.

MCLAUGHLIN, K. J. (1994). Rigid wages? *Journal of Monetary Economics*, **34**, pp. 383–414.

MOULTON, B. (1986). Random group effects and the precision of regression estimates. *Journal of Econometrics*, **32**, pp. 385–97.

MOULTON, B. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *Review of Economics and Statistics*, **72**, pp. 334–8.

ROTEMBERG, J. and WOODFORD, M. (1992). Oligopolistic pricing and the effects of aggregate demand on economic activity. *Journal of Political Economy*, **100**, pp. 1153–207.

SHAPIRO, M. (1996). Macroeconomic implications of variation in the workweek of capital. *Brookings Papers on Economic Activity*, **2**, pp. 79–133.

SOLON, G. and BARSKY, R. (1989). Real wages over the business cycle. NBER Working Paper 2888.

SOLON, G., BARSKY, R. and PARKER, J. A. (1992). Measuring the cyclicality of real wages: how important is composition bias? NBER Working Paper 4202.

SOLON, G., BARSKY, R. and PARKER, J. (1994). Measuring the cyclicality of real wages: how important is composition bias? *Quarterly Journal of Economics*, **109**, pp. 3–25.

STOCKMAN, A. (1983). *Aggregation Bias and the Cyclical Behavior of Real Wages*. Rochester, NY: University of Rochester.

SURVEY RESEARCH CENTER. (1973–1995). *A Panel Study of Income Dynamics: Procedures and Tape Codes*, 22 vols. Ann Arbor, MI: Institute for Social Research, University of Michigan.

SURVEY RESEARCH CENTER. (1986). *User Guide to the Panel Study of Income Dynamics*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

SWANSON, E. (2004). Measuring the cyclicality of real wages: how important is the firm's point of view? *Review of Economics and Statistics*, **86**, pp. 362–77.

SWANSON, E. (2006). The relative price and relative productivity channels for aggregate fluctuations. *Contributions to Macroeconomics*, **6**, article 10.

Date of receipt of final manuscript: 1 August 2007