

GEORGE L. PERRY

Brookings Institution

Gauging Employment: Is the Professional Wisdom Wrong?

THE BUREAU OF Labor Statistics' (BLS) monthly report on labor market developments is the government's most widely anticipated statistical release and the one that most influences markets, forecasters, and policymakers. The report provides detailed information on employment from two sources: the Current Population Survey of households, which also provides data on unemployment, and the Current Employment Statistics survey of payrolls from nonfarm business establishments and government. These two sources often produce very different estimates of the monthly change in aggregate employment, and the payroll data are widely accepted as the more reliable. They are featured on the first page of the monthly release and are the employment data most frequently discussed in the business press and other media. The annual *Economic Report of the President* has recently reaffirmed their reliability relative to the household data, as has Federal Reserve Board Chairman Alan Greenspan. Simply put, to judge by what the experts say and what the media report, most people interested in what is happening to aggregate employment rely on the payroll numbers. They use the household survey for the unemployment rate, demographic breakdowns, and various more arcane measures of labor market developments that only it provides. This paper questions whether the overwhelming preference for the payroll data as the measure of aggregate employment is justified, and concludes it is not.

Ryan Nunn provided excellent research assistance. Benjamin Friedman, Edward Gramlich, Charles Schultze, Jack Triplett, and William Wascher offered useful comments on an earlier draft.

Table 1. Explaining Changes in Treasury Yields, January 1994 to June 2004^a

<i>Independent variable</i>	<i>Regression</i>			
	<i>1-1</i>	<i>1-2</i>	<i>1-3</i>	<i>1-4</i>
	<i>Dependent variable: ten-year Treasury yield, thirty-minute window^b</i>			
Household employment shock ^c	10.8 (3.62)		1.7 (0.63)	2.5 (0.88)
Payroll employment shock ^d		54.7 (9.24)	53.0 (8.05)	52.5 (7.97)
Change in unemployment rate				5.5 (1.23)
No. of observations	125	125	125	125
Adjusted R^2	0.09	0.41	0.40	0.41
	<i>Dependent variable: two-year Treasury yield, thirty-minute window^b</i>			
Household employment shock	12.9 (3.66)		1.0 (0.32)	1.6 (0.49)
Payroll employment shock		68.6 (10.05)	67.5 (8.84)	67.0 (8.72)
Change in unemployment rate				4.7 (0.91)
No. of observations	125	125	125	125
Adjusted R^2	0.09	0.45	0.44	0.44
	<i>Dependent variable: six-month Treasury yield, twenty-four-hour window^e</i>			
Household employment shock	5.4 (2.39)		-0.5 (-0.20)	-1.1 (-0.48)
Payroll employment shock		33.9 (6.73)	34.4 (6.12)	35.1 (6.25)
Change in unemployment rate				-4.9 (-1.30)
No. of observations	122	122	122	122
Adjusted R^2	0.04	0.27	0.26	0.27

(continued)

Policy and Market Reactions

What is the evidence that financial markets and policymakers respond to the payroll employment data and ignore the household data? Table 1 analyzes the response of market interest rates to the monthly employment

Table 1. Explaining Changes in Treasury Yields, January 1994 to June 2004^a (continued)

<i>Independent variable</i>	<i>Regression</i>			
	<i>1-1</i>	<i>1-2</i>	<i>1-3</i>	<i>1-4</i>
	<i>Dependent variable: three-month Treasury yield, twenty-four-hour window^e</i>			
Household employment shock	4.7 (2.59)		1.7 (0.88)	0.8 (0.40)
Payroll employment shock		18.9 (4.19)	16.9 (3.35)	18.0 (3.57)
Change in unemployment rate				-5.6 (-1.68)
No. of observations	122	122	122	122
Adjusted R^2	0.05	0.12	0.12	0.13

Source: Author's regressions using BLS, GovPX, Federal Reserve, Bloomberg, and Money Market Services data.

a. All variables except the change in the unemployment rate are 100 times the change in logs. Numbers in parentheses are t statistics.

b. Treasury yield changes (in basis points) were calculated from five minutes before the release of the employment report to twenty-five minutes after.

c. The unrestricted model for Treasury yields is $\Delta r = a + b_1 \text{shock}_h + b_2 \text{shock}_p + b_3 \Delta UR + \epsilon$, where Δr and ΔUR are the change in the yield and in the unemployment rate, respectively. Constant terms are not reported. Shocks are defined as the actual percentage change in the variable minus the expected percentage change in the variable. Expected employment numbers are the median survey expectation as reported by Money Market Services.

d. Initially reported, unrevised changes to payrolls.

e. Treasury yield changes were calculated from the close of business on the previous day until the close of business on the day of the employment report.

release. The employment data are presented as initially reported rather than the later revisions, which reflect a much larger sample of payrolls and updated seasonal factors. Since market interest rates adjust continuously, they are modeled as responding to the shock in the monthly employment change, measured as the difference between the actual change in the payroll or household data and the expected change in employment as reported in a survey of analysts by Money Market Services.¹

In the first two panels of the table, the employment shocks are used to explain the change in ten- and two-year Treasury yields in the window from five minutes before to twenty-five minutes after the employment

1. Throughout the paper, I adjust the household data in those months when a new estimate of the population is introduced, since published changes in employment are meaningless in those months. For those months the change in the population is taken as the average of the changes in the immediately adjoining months.

data are released.² The regressions are run on data from January 1994 to June 2004, the period for which data were available. The payroll employment shock has a statistically significant and meaningful effect on yields, and the several regressions using it explain between 40 and 45 percent of the variation in interest rates in the thirty-minute window. In the last column, even the change in the unemployment rate has no significant effect when used together with the payroll employment shock. In contrast, the *R*-squared of the equation is very low when the household employment shock is used alone in the regression, and this variable is insignificant when entered together with the payroll shock. The same regressions found no consistent effects when the window over which these interest rate changes were calculated was extended to the twenty-four hours between the market close the day before and that on the day of the employment release. For these two- and ten-year maturities, the initial reaction appears to be overcome by other developments in the course of the day. Adding the change in the Standard and Poor's (S&P) stock index over the same interval, as an explanatory variable that might capture such developments, did not change this result. However, in regressions over the twenty-four-hour period for yields on shorter maturities (three- and six-month Treasury bill rates, third and fourth panels of table 1), the payroll data again show significant effects on rates and the household data do not. Data for these maturities were not available for the thirty-minute window.

Turning to the response of policymakers, table 2 reports on how policy targets have related to employment changes and changes in unemployment rates over this same period. The dependent variable is the change in the federal funds target rate, whether announced after each meeting of the Federal Open Market Committee (FOMC) or between FOMC meetings. (Announcements of no change in the rate are recorded as zeroes.) Because the timing between employment reports and announced policy changes varies, two employment reports may be issued between announcements; in such cases the average of the two is used. Since policy changes are considered at discrete intervals, actual employment changes between those

2. I am grateful to Brian Sack for providing the data from which employment shocks are calculated and the data on the change in yields over this thirty-minute window. The same expected change is used to calculate shocks for both the household and the payroll data. Regressions using raw employment changes rather than these shocks did not find significant effects, supporting the conventional view that markets had already responded to anticipated changes in employment.

Table 2. Explaining Changes in the Announced Federal Funds Target Rate, January 1994 to June 2004^a

<i>Independent variable</i>	<i>Regression</i>								
	2-1	2-2	2-3	2-4	2-5	2-6	2-7	2-8	2-9
Household employment ^b	46.5 (3.59)		31.9 (2.45)	46.1 (3.51)		28.1 (1.85)		18.3 (1.24)	18.3 (1.23)
Payroll employment ^c		90.1 (4.86)	72.6 (3.73)		90.2 (4.76)		69.1 (3.63)	64.3 (3.30)	64.4 (3.25)
Core inflation rate				5.3 (0.34)	-0.5 (-0.03)				-0.5 (-0.03)
Change in unemployment rate						-67.6 (-2.50)	-66.6 (-2.89)	-50.3 (-1.92)	-50.3 (-1.90)
Adjusted R ²	0.13	0.22	0.26	0.12	0.21	0.19	0.29	0.29	0.28

Source: Author's regressions using BLS, Federal Reserve, and Money Market Services data.

a. The dependent variable (in basis points) is the change in the target federal funds rate upon any announcement following or between FOMC meetings. All variables except the change in the unemployment rate are changes in logs. All regressions are performed on eighty-three observations. Numbers in parentheses are *t* statistics.

b. The unrestricted model for policy rates is $\Delta r = a + b_1h + b_2p_orig + b_3\Delta UR + \epsilon$. Constant terms are not reported.

c. Initially reported, unvisited changes to payroll.

intervals rather than employment shocks are used as the explanatory variable in the first panel of the table.

The payroll data also dominate the explanation of policy rates, explaining nearly twice as much of the variance in the federal funds target as the household data when no other variable is included. A payroll employment change that is one standard deviation different from the mean change is expected to change the federal funds rate by 12 basis points in the same direction. When the change in the inflation rate, measured by the core consumer price index (CPI), is added to the regressions, it is insignificant and has no effect on the employment coefficients. The same is true (results not reported) when the change in the total CPI rather than the core is used, or when the inflation rate itself (either core or total) rather than its change is used. These results on inflation almost surely reflect the absence of any worrisome inflation in this period. When the change in the unemployment rate is included in the regressions, it has a significant effect but does not change the much better fit obtained with payroll than with household employment. And when both employment series and the change in unemployment are in the regression, only the payroll data are significant.³

Comparing the Employment Series

Both the household and the payroll data provide valuable disaggregated information about labor market developments that is widely used by analysts and researchers. It thus seems odd that only the payroll data are relied on to measure short-run variations in aggregate employment. That preference, revealed both in what practitioners say and in what they apparently do, presumably rests on the lower standard errors that the BLS reports for that series, and possibly on the experience of forecasters and other analysts and their aversion to the greater monthly variability of the household series. Before presenting statistical tests of the relative usefulness of the two series, it is worth looking at some of the features of both.

3. As a check on the specification, the regressions were also run using the employment shocks in place of the employment changes; as expected, the shock variables explain almost none of the variation in the funds rate.

The BLS estimates a standard error of 212,000 for the monthly change in household employment and a much smaller standard error of about 70,000 for the monthly change in payroll employment. But these statistical measures tell less than they seem to. As the BLS notes, they measure only sampling error; other identifiable, but hard to quantify, sources of error are present in each series.⁴ Some of these are likely to be more important for monthly changes in employment, and others for trends over longer periods.

Problems in the Household Survey

The monthly household survey uses a rotating sample of about 55,000 households (60,000 are eligible for interview) and scales up the survey counts by an estimate of the population. The population estimate is benchmarked to the decennial census and projected forward using estimates of immigration and other relevant demographic information. Between census years, these population estimates are updated as new information about immigration becomes available. Nevertheless, population errors can build up over time. But, except in the month when new population estimates are introduced, population uncertainties should have only negligible effects on the reported monthly changes in employment.⁵ Another source of error is response error, always present in surveys. Here the main problem identified in the household data has been misreporting of workers as unemployed when in fact they are out of the labor force, and vice versa. Response errors are less likely to affect total employment. For example, as working arrangements at many establishments have evolved, individuals who have moved from being employees to being self-employed contractors or consultants for the same establishment may incorrectly identify themselves in the household survey as employees. This form of response error affects the difference between the payroll and the household data but not total household employment.

4. Nardone and others (2003) provide a useful discussion of a number of issues relating to the two series. This document and most other discussions of the discrepancy are mainly concerned with the longer-run divergence of the two surveys over the past several years rather than with month-to-month discrepancies. More recently the main problems were summarized in Bureau of Labor Statistics (2005).

5. As described in footnote 3, I adjust the population change in months when new population estimates are introduced.

Problems in the Payroll Survey

The monthly payroll survey is based on a sample of 400,000 establishments, which employ approximately one-third of all nonfarm payroll workers; the sample is stratified to capture both large and small establishments. The initial estimates are revised over the next two months as the number of establishments reporting grows, and annually using data on payments into states' unemployment insurance (UI) funds. Since all establishments are required to pay UI taxes on each employee, this is treated as a virtually full count of the population of employers, and the estimated standard error for monthly employment reflects the accuracy of the initial estimate relative to this presumed full count.

Several potentially important sources of error are not captured in the statistical standard error that the BLS reports for the payroll data. The most important is that establishments are born and die each year, and this process is difficult to capture either in the monthly sample or in the allocation across months of revisions that is dictated by the annual UI count. The number of such establishments is very large: an estimated 356,000 new establishments have been formed every quarter in recent years. Establishments that go out of existence typically do not report that fact and are treated as nonreporting rather than as zero employment until the BLS learns otherwise. Particularly because the UI programs are administered at the state level, reporting compliance with the UI tax, especially by small establishments, may be imperfect. But even if one assumes that the number of new establishments is captured adequately in the annual UI count, allocating their employment across months in the preceding year is a daunting task.

Over time the BLS has attempted to improve its procedures for estimating establishment births and deaths. Twenty years ago the agency determined that the payroll survey was consistently underestimating job creation among smaller establishments, and it introduced an adjustment that added about 30,000 jobs a month. In 2002 it replaced this with an adjustment for establishment births and deaths that adds an average of 68,000 jobs a month in 2004, but in a pattern that ranges widely over the months of the year, from approximately -321,000 to +270,000 in the twelve months beginning August 2003.⁶

6. These changes are discussed on the BLS website at www.bls.gov/ces/cesbdhst.pdf.

It has been understood for some time that the payroll survey double-counts jobs when employees change jobs within the survey week (the week including the 12th of the month). If such job turnover were constant, it would not affect estimates of the change in employment. But variations in turnover will distort the monthly change in employment. What evidence there is on job turnover indicates that it has both a systematic cyclical pattern and, in recent years, signs of a downward drift. There is less evidence on its monthly variability.

Differences in Volatility

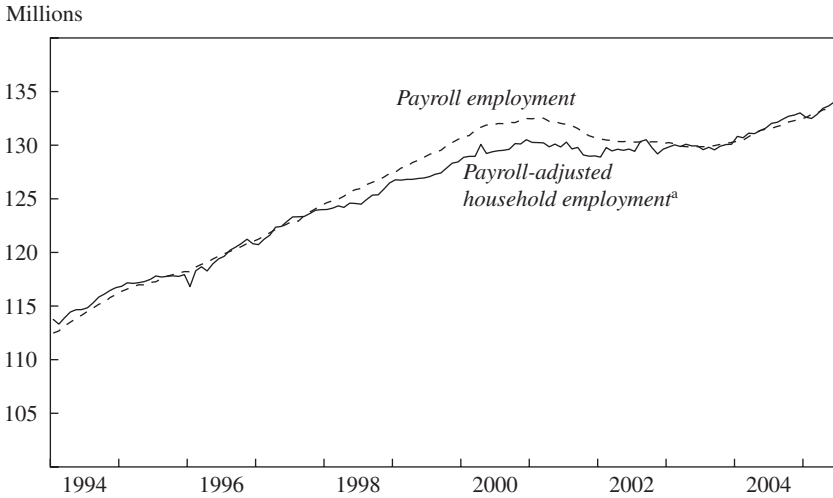
On a monthly basis, the household series has considerably greater volatility than the payroll series. The standard deviations of monthly percentage changes are, respectively, 0.0033 and 0.0025 starting in 1954, and 0.0021 and 0.0014 for the period from January 1994 to June 2005. Whether this is a reason to prefer the payroll series is another matter. For one thing, the payroll data may be smoothed excessively when the BLS processes the survey results and makes adjustments for coverage of births and deaths of establishments. For another, how much volatility should be expected in monthly changes for aggregate economic data is unclear. Volatilities for the period from January 1994 to June 2005 are higher for broad measures of consumer demand, for example, than they are for either employment series:

<i>Variable</i>	<i>Standard deviation of change in logs</i>	<i>First-order autocorrelation coefficients of change in logs</i>
Household employment	0.0021	-0.057
Payroll employment	0.0014	0.656
Real personal consumption	0.0042	-0.433
Real goods consumption	0.0093	-0.460

The BLS Research Series

For the period starting in 1994, the BLS has constructed a “research” series that adjusts household employment to the concepts of payroll employment. The main adjustments are for the omission of agricultural workers and the self-employed from the payroll data and the counting of multiple-job holders only once in the household data. Figure 1 compares this household series adjusted to payroll concepts (HP) with the payroll series (P) over the period January 1994 to June 2005. (Unless otherwise

Figure 1. Payroll Employment and Payroll-Adjusted Household Employment, January 1994 to June 2005



Source: BLS data.

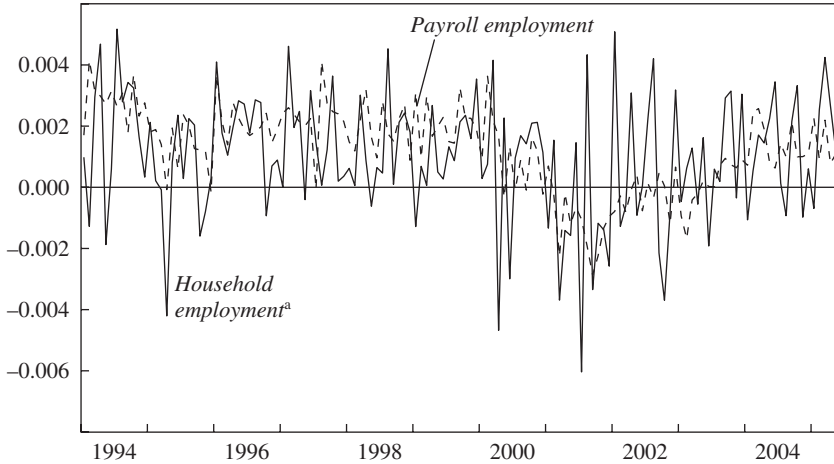
a. Research series constructed by the BLS, which adjusts the coverage of the household survey to match that of the payroll survey.

noted, the series used contain all data revisions that the BLS has made to the data.) At annual frequencies, the P and the HP series track each other closely in most years, although longer-run discrepancies remain, with payroll employment growing faster from mid-1997 to early 2000, returning to the level of adjusted household employment by early 2003, and tracking it rather closely since then. However, as figure 2 shows, striking discrepancies exist between the monthly log changes in P and HP, designated p and hp , respectively.

The period surrounding 2000 was an unusual one, during which both output and employment in the high-technology sector were subject to exceptional measurement problems. Because my purpose here is to compare the high-frequency changes in the two employment series, and to test how much this one episode of drift between them influences that comparison, I formed an additional series by subtracting a small constant (0.00053) from hp for the thirty-seven months from the time the discrepancy emerged to its peak in 2000, and adding a small constant (0.00064) to hp for the next thirty months until the discrepancy ended. The resulting

Figure 2. Change in Payroll and Household Employment, January 1993 to June 2005

Change in logs



Source: BLS data.
a. Population adjusted.

trend-adjusted series, hpt , closely tracks the low-frequency movements in p but retains the substantial differences that exist at monthly frequencies. The standard deviations of p , hp , and hpt are 0.0014, 0.0029, and 0.0030, respectively. Regressions of p on hp and p on hpt have R -squareds of 0.15 and 0.21, respectively. For the analysis of the period starting in 1994, I use all three of these “payroll concept” series, two of which come from household survey data.

Tests Using Contemporaneous Cyclical Measures

Since the known characteristics of the household and the payroll data give no decisive reason for preferring one over the other, statistical tests are needed to determine their usefulness. And since there is no definitive measure of employment against which to compare either, indirect tests are required. I start by comparing how the various employment series track other important cyclical measures. For this purpose some familiar measures are not suitable, because they are not independent of the payroll employment data: industrial production is estimated in part using payroll data, the

help wanted index measures the availability of payroll-related jobs, and insured unemployment comes from jobs lost from payrolls. GDP and private nonfarm output are two measures of real activity that are constructed independently of employment data and so are useful for these tests. So is the unemployment rate if, as seems likely, reporting errors as between unemployment and labor force nonparticipation are important, but reporting errors between employment and unemployment are not.

Tracking Output

I first run regressions to compare how the household and payroll measures track contemporaneous GDP (Q) and output in the private nonfarm business sector (QNFB). Growth in employment and output are the two main indicators of the strength of economic expansion. Policymakers and forecasters are interested in the cyclical state of the economy, and if they could see only one independent piece of data from which to estimate a quarter's output growth, they would choose growth in employment. Thus, although these contemporaneous regressions are too simple to be interpreted as a proxy either for a production function or for employment demand, finding how each employment measure correlates with output provides some evidence about its reliability as an indicator of aggregate performance. Although neither output measure is perfectly aligned with any one employment measure, one would expect, based on conceptual differences, that the household survey H, which includes agricultural and private household workers, would align better with Q, whereas series using payroll concepts would align better with QNFB.

The tracking regressions for output are reported in table 3. In these and the subsequent regressions, all variables are measured as changes in logs, designated as p, hp, hpt, p_orig (the change in payroll employment as originally reported), h, q, and qnfb. Since the output measures are quarterly, the monthly employment levels are averaged into quarters, and the quarterly log changes are used in regressions for the period 1994:1 to 2005:2. Also, all regressions are reported with Prais-Winsten corrections since results using ordinary least squares generally had poor Durbin-Watson statistics.

The results for the nonfarm business sector, shown in the first panel of table 3, are surprising in light of the profession's priors. In a comparison of *R*-squareds in the first five columns, h outperforms p_orig by a wide margin; it also outperforms p, despite the better alignment of p with the

Table 3. Tracking Changes in Nonfarm Business Output and GDP with Alternative Employment Measures, 1994:1 to 2005:2^a

<i>Independent variable</i>	<i>Regression</i>							
	3-1	3-2	3-3	3-4	3-5	3-6	3-7	3-8
	<i>Dependent variable: change in nonfarm business output</i>							
Household employment (h)	1.102 (4.15)					0.693 (1.73)		
Payroll employment, unrevised (p_orig)		0.794 (2.98)						
Payroll employment, revised (p)			0.817 (3.73)			0.414 (1.34)	0.477 (1.53)	0.264 (0.74)
Research series (hp) ^b				0.800 (3.58)			0.473 (1.53)	
Research series, trend adjusted (hpt) ^c					0.765 (4.28)			0.586 (1.95)
Adjusted R ²	0.26	0.15	0.22	0.21	0.28	0.27	0.24	0.27
	<i>Dependent variable: change in GDP</i>							
Household employment (h)	0.845 (4.10)					0.546 (1.80)		
Payroll employment, unrevised (p_orig)		0.616 (2.93)						
Payroll employment (p)			0.634 (3.60)			0.317 (1.32)	0.336 (1.41)	0.168 (0.63)
Research series (hp) ^b				0.627 (3.71)			0.414 (1.82)	
Research series, trend adjusted (hpt) ^c					0.605 (4.44)			0.494 (2.23)
Adjusted R ²	0.26	0.15	0.21	0.23	0.30	0.27	0.25	0.28

Source: Author's regressions using Bureau of Labor Statistics and Bureau of Economic Analysis data.

a. All variables are changes in logs. Regressions use the Prais-Winsten transformation. Numbers in parentheses are *t* statistics. The unrestricted model for output is $q = a + b_1x + b_2p_orig + e$, where x is the indicated variant of household employment. All regressions are performed on forty-five observations. Constant terms are not reported.

b. Household employment series adjusted by the BLS to match the payroll concept.

c. Trend widenings and narrowings between household and payroll data are removed from hp; see the text for details.

nonfarm business sector. The research series, *hp*, which adjusts household data to payroll concepts, does about as well as *p*. And the best fit of all is with *hpt*, which adjusts *hp* to match the trend in *p*. It appears that the payroll data better track the output pattern around 2000, but that the research series better tracks quarterly variations in output except for that episode.

The second panel of table 3 shows results of regressions of the same form but tracking changes in quarterly GDP rather than nonfarm business output. The unrevised payroll data, *p_orig*, again have by far the lowest *R*-squared. The three household-based series have the best fit, but the better alignment of *h* with *q* makes very little difference, indicating that the variations in GDP that matter for these comparisons come largely from variations in nonfarm business output. Regressions that add changes in unemployment to the variables shown in table 3 (results not reported) make little difference to any of these comparisons.

Table 3 also reports regressions in which pairs of employment series are entered together as right-hand-side variables. The relative ranking of the employment series, here using *t* statistics to measure their output tracking performance, is largely the same as for the single-variable regressions. However, as Christopher Sims discussed in a recent Brookings Paper,⁷ the interpretation of such regressions is unclear, especially when the independent variables all have considerable explanatory power taken alone, which is the case here.

Might seasonal adjustment help account for some of these results? Both *H* and *P* are seasonally adjusted at a disaggregated level and then aggregated to total employment. In general this procedure gives different monthly estimates than seasonally adjusting total employment itself would, and the procedure might excessively smooth the payroll data where the disaggregation is most detailed. To test this possibility, the Census X-12 ARIMA program was used to adjust total payroll employment directly. However, the resulting series was very much like the official data, and the regression results (not reported) were essentially the same.

Tracking Unemployment

I next use the same employment series to track changes in unemployment, another broad indicator of economic activity. This permits using the monthly data directly, although results with quarterly data are also reported

7. Sims (2002).

Table 4. Tracking Changes in Unemployment with Alternative Employment Measures, 1994–2005^a

<i>Independent variable</i>	<i>Regression</i>							
	4-1	4-2	4-3	4-4	4-5	4-6	4-7	4-8
	<i>Dependent variable: monthly unemployment, January 1994 to June 2005</i>							
Household employment (h)	-6.149 (-6.81)					-4.530 (4.81)		
Payroll employment, unrevised (p_orig)		-7.129 (-6.01)						
Payroll employment, revised (p)			-7.275 (-6.58)			-4.743 (-4.11)	-5.224 (-4.43)	-4.565 (-3.60)
Research series (hp) ^b				-4.057 (-5.84)			-2.717 (-3.70)	
Research series, trend adjusted (hpt) ^c					-4.229 (-6.61)			-2.801 (-3.77)
Adjusted R ²	0.25	0.21	0.24	0.20	0.24	0.34	0.31	0.32
	<i>Dependent variable: quarterly unemployment, 1994:1–2005:2</i>							
Household employment (h)	-9.918 (-7.46)					-5.455 (-3.11)		
Payroll employment, unrevised (p_orig)		-8.758 (-6.88)						
Payroll employment, revised (p)			-7.992 (-7.41)			-4.775 (-3.49)	-5.512 (-3.75)	-4.585 (-2.78)
Research series (hp) ^b				-6.367 (-5.43)			-3.506 (-2.59)	
Research series, trend adjusted (hpt) ^c					-6.275 (-6.58)			-3.681 -2.83
Adjusted R ²	0.55	0.51	0.55	0.39	0.49	0.66	0.57	0.58

Source: Author's regressions using Bureau of Labor Statistics and Bureau of Economic Analysis data.

a. All variables are changes in logs. Regressions use the Prais-Winsten transformation. Numbers in parentheses are *t* statistics. The unrestricted model for unemployment is $u = a + b_1x + b_2p_orig + \epsilon$, where x is the indicated variant of household employment. All regressions are performed on 137 observations (top panel) or 45 observations (bottom panel). Constant terms are not reported.

b. Household employment series adjusted by the BLS to match the payroll concept.

c. Trend widenings and narrowings between household and payroll data are removed from hp; see the text for details.

for comparability with the output results. The first panel of table 4 shows monthly regressions of the log change in unemployment (u) on h , p_orig , p , hp , and hpt . The unrevised payroll data again do relatively poorly, but otherwise the evidence in favor of either the household or the payroll data is less clear than in the regressions tracking output. The research series, hp , has the poorest fit, although when adjusted for the trend difference with p , it fits about as well as the payroll data. The quarterly regressions in the second panel give much the same results.

Explaining Employment

A more structural approach to comparing the employment series is to estimate aggregate employment functions. According to a standard model of employment dynamics, both average hours worked and employment adjust to initial shocks to product demand. As the shocks persist, increasingly more of the labor demand shows up in employment, and average hours moves back toward trend. An estimation model incorporating these dynamics relates current changes in employment to both changes in output and the level of actual hours relative to trend. Lagged values of output may help capture the persistence of output shocks. For the regressions, average hours was taken from the BLS series for the private nonfarm business sector, the broadest estimate available but one that relies mainly on payroll data. The trend in average hours was estimated over 1954:2 to 2004:2 using a Hodrick-Prescott filter (with a 1600 parameter, although using 400 instead changed nothing important). Using instead a moving average of average hours gave similar results (not reported). The period over which the regressions are estimated again starts in 1994.

The top panel of table 5 reports the results of this model using GDP as the output variable, and the bottom panel reports results using nonfarm business output. The results in both are much the same. Because the variations in the employment series are not too different, one can usefully compare how well the regressions explain each of them, although small differences should be ignored. The adjusted research series, hpt , usually achieves the best fit. When only current and once-lagged output, along with the average hours term, are included in the regressions, both h and hpt are better explained than in identical regressions using the payroll data. Adding two more lagged values of output improves the fit of equations explaining the payroll data relatively more, with R -squared approaching

Table 5. Explaining Changes in Employment, 1994:1–2005:2^a

<i>Independent variable</i>	<i>Employment series</i>					
	<i>Regressions using one lag of output change</i>			<i>Regressions using three lags of output change</i>		
	<i>h</i>	<i>p</i>	<i>hpt^b</i>	<i>h</i>	<i>p</i>	<i>hpt^b</i>
GDP (q)	0.255 (3.47)	0.164 (3.14)	0.441 (4.22)	0.193 (2.36)	0.187 (4.04)	0.312 (2.85)
q _{t-1}	0.236 (3.09)	0.150 (2.93)	0.297 (2.74)	0.217 (2.73)	0.226 (4.32)	0.246 (2.27)
q _{t-2}				0.149 (1.68)	0.192 (3.44)	0.297 (2.50)
q _{t-3}				0.066 (0.78)	0.169 (3.74)	0.156 (1.38)
Log deviation in hours from trend (devtrend) ^c	-0.101 (-1.17)	-0.207 (-2.49)	-0.242 (-2.02)	0.006 (0.06)	-0.136 (-1.74)	-0.023 (-0.17)
Sum of q coefficients	0.49	0.31	0.74	0.63	0.77	1.01
Adjusted R ²	0.39	0.29	0.46	0.42	0.52	0.57

(continued)

Table 5. Explaining Changes in Employment, 1994:1–2005:2^a (continued)

Independent variable	Employment series					
	Regressions using one lag of output change			Regressions using three lags of output change		
	<i>h</i>	<i>p</i>	<i>hpt</i> ^b	<i>h</i>	<i>p</i>	<i>hpt</i> ^b
Nonfarm business output (q _{nf})	0.195	0.111	0.333	0.141	0.144	0.225
	(3.51)	(2.88)	(4.16)	(2.37)	(4.41)	(2.75)
	Regressions using nonfarm business output to explain employment change					
q _{nf} _{<i>t</i>-1}	0.182	0.099	0.217	0.174	0.19	0.193
	(3.17)	(2.64)	(2.62)	(3.00)	(4.92)	(2.41)
q _{nf} _{<i>t</i>-2}				0.133	0.172	0.25
				(2.08)	(4.23)	(2.85)
q _{nf} _{<i>t</i>-3}				0.044	0.147	0.111
				(0.71)	(4.50)	(1.30)
Log deviation in hours from trend (devtrend) ^c	-0.108	-0.201	-0.255	0.001	-0.134	-0.04
	(-1.27)	(-2.39)	(-2.08)	(0.01)	(-1.80)	(-0.30)
Sum of q _{nf} coefficients	0.38	0.21	0.55	0.49	0.65	0.78
Adjusted R ²	0.39	0.26	0.43	0.43	0.58	0.54

Source: Author's regressions using Bureau of Labor Statistics and Bureau of Economic Analysis data.

a. All variables, except deviation in hours, are changes in logs. Regressions use the Prais-Winsten transformation. The model for employment is $x_t = a + b_1q_t + b_2q_{t-1} + \dots + b_3\text{devtrend}_t + e_t$, where x_t is the indicated employment series. Constant terms are not reported. All regressions are performed on forty-five observations. Numbers in parentheses are t statistics.

b. Household series adjusted to match the payroll concept (BLS calculations) and adjusted for trend by the author.

c. Trend calculated with Hodrick-Prescott filter, parameter = 1600.

those in regressions explaining hpt when GDP is the output variable and slightly exceeding them when nonfarm business is the output variable. It is unclear whether lags this long are relevant for employment decisions, but their coefficients remain significant in the payroll equations.

Results for Earlier Decades

Taken collectively, these results for the period starting in 1994 provide some evidence that the household data now provide the single best measure of aggregate employment, and no support for the widespread belief that they should be disregarded in favor of the data from the payroll survey. These inferences are especially strong when the household data are compared with the initially reported change in payroll employment. To see whether these results arise solely as a result of the 1994 improvements to the household survey, or are due to other recent developments in the economy that may have affected the reliability of the data sources, I repeated the main regressions for the four preceding ten-year intervals. These regressions use only the h and p series, since the research series is not available before 1994 and I did not resurrect the initially reported payroll data for these earlier years. For the two earliest intervals, 1954–63 and 1964–73, the professional consensus for ignoring the household measure in favor of the payroll measure is fully supported. In tracking regressions, the payroll data consistently had the better fit in these decades, and the household data had no statistical significance in any regressions for 1964–73. These strong results support the idea that the experience of data users in this early period was important in creating the strong current preference for the payroll data as the measure of aggregate employment.

The two subsequent decades produced more mixed results, which are presented in tables 6 and 7. For ease of comparison, these tables include the parallel results for the period starting with 1994. In the tracking regressions reported in table 6, for the decade ending in 1993, the household data do somewhat better than the payroll data in tracking either GDP or nonfarm business output, but not as well in tracking unemployment. For the decade ending in 1983, the household data produce slightly higher R -squareds in all regressions. Thus, whereas the results for the two early

Table 6. Tracking Changes in Output and Unemployment, Various Periods^a

Independent variable	Period								
	1994:1–2005:2		1983:4–1993:4		1973:4–1983:4				
	6-1	6-2	6-3	6-4	6-5	6-6	6-7	6-8	6-9
Household employment (h)	0.845 (4.10)		0.546 (1.80)	1.097 (6.36)		0.593 (1.87)	1.277 (6.27)		0.306 (0.66)
Payroll employment (p)		0.634 (3.60)	0.317 (1.32)		0.949 (5.97)	0.513 (1.87)		1.253 (6.13)	0.984 (2.29)
No. of observations	45	45	45	40	40	40	40	40	40
Adjusted R ²	0.26	0.21	0.27	0.50	0.47	0.51	0.50	0.48	0.48
Root mean squared error	0.0043	0.0044	0.0043	0.0040	0.0040	0.0038	0.0090	0.0086	0.0086
	<i>Dependent variable: GDP^b</i>								
Household employment (h)	1.102 (4.15)		0.693 (1.73)	1.300 (5.95)		0.856 (2.04)	1.787 (6.43)		0.556 (0.94)
Payroll employment (p)		0.817 (3.73)	0.414 (1.34)		1.074 (5.14)	0.444 (1.22)		1.799 (6.23)	1.303 (2.34)
No. of observations	45	45	45	40	40	40	40	40	40
Adjusted R ²	0.26	0.22	0.27	0.46	0.39	0.46	0.51	0.49	0.50
Root mean squared error	0.0058	0.0058	0.0057	0.0051	0.0053	0.0051	0.0116	0.0111	0.0112
	<i>Dependent variable: nonfarm business output^b</i>								
Household employment (h)	-9.918 (-7.46)		-5.455 (-3.11)	-7.015 (-6.56)		-1.669 (-1.26)	-8.509 (-10.07)		-2.159 (-1.35)
Payroll employment (p)		-7.992 (-7.41)	-4.775 (-3.49)		-7.410 (-7.99)	-6.229 (-4.83)		-8.105 (-9.11)	-6.285 (-4.03)
No. of observations	45	45	45	40	40	40	40	40	40
Adjusted R ²	0.55	0.55	0.66	0.52	0.62	0.63	0.72	0.68	0.70
Root mean squared error	0.0280	0.0272	0.0252	0.0215	0.0172	0.0171	0.0360	0.0317	0.0315
	<i>Dependent variable: unemployment^f</i>								

Source: Author's regressions using Bureau of Labor Statistics and Bureau of Economic Analysis data.

a. All variables are changes in logs. Regressions use the Prais-Winsten transformation. Numbers in parentheses are *t* statistics.

b. The unrestricted model for output is $q = a + b_1h + b_2p + e$. Constant terms are not reported.

c. The unrestricted model for unemployment is $u = a + b_1h + b_2p + e$. Constant terms are not reported.

Table 7. Explaining Changes in Employment, Various Periods^a

<i>Independent variable</i>	<i>Period and employment series</i>					
	<i>1994:1–2005:4</i>		<i>1983:4–1993:4</i>		<i>1973:4–1983:4</i>	
	<i>h</i>	<i>p</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>p</i>
GDP (q)	0.255 (3.47)	0.164 (3.14)	0.308 (4.13)	0.201 (3.90)	0.352 (6.20)	0.364 (8.03)
q _{t-1}	0.236 (3.09)	0.15 (2.93)	0.179 (2.24)	0.174 (3.13)	0.137 (2.06)	0.105 (2.13)
Log deviation in hours from trend (devtrend) ^b	-0.101 (-1.17)	-0.207 (-2.49)	-0.054 (-0.52)	-0.055 (-0.53)	-0.346 (-2.39)	-0.425 (-2.91)
Sum of q coefficients	0.49	0.31	0.49	0.38	0.49	0.47
No. of observations	45	45	40	40	40	40
Adjusted R ²	0.39	0.29	0.43	0.41	0.66	0.64
Root mean squared error	0.0024	0.0015	0.0023	0.0016	0.0040	0.0032
GDP (q)	0.193 (2.36)	0.187 (4.04)	0.330 (4.11)	0.195 (3.64)	0.338 (5.90)	0.367 (8.60)
q _{t-1}	0.217 (2.73)	0.226 (4.32)	0.124 (1.47)	0.204 (3.41)	0.151 (2.27)	0.146 (2.90)
q _{t-2}	0.149 (1.68)	0.192 (3.44)	-0.003 (-0.04)	0.072 (1.26)	0.058 (0.91)	0.083 (1.67)
q _{t-3}	0.066 (0.78)	0.169 (3.74)	0.155 (1.88)	0.003 (0.06)	0.083 (1.44)	0.113 (2.66)
Log deviation in hours from trend (devtrend) ^b	0.006 (0.06)	-0.136 (-1.74)	0.032 (0.30)	-0.051 (-0.48)	-0.191 (-1.11)	-0.298 (-1.89)

(continued)

Table 7. Explaining Changes in Employment, Various Periods^a (continued)

Independent variable	Period and employment series					
	1994:1-2005:4		1983:4-1993:4		1973:4-1983:4	
	<i>h</i>	<i>p</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>p</i>
Sum of <i>q</i> coefficients	0.63	0.77	0.61	0.47	0.63	0.71
No. of observations	45	45	40	40	40	40
Adjusted <i>R</i> ²	0.42	0.52	0.50	0.43	0.68	0.74
Root mean squared error	0.0023	0.0013	0.0023	0.0016	0.0040	0.0030
<i>Regressions using one lag of nonfarm business output</i>						
Nonfarm business output (<i>qntfb</i>)	0.195 (3.51)	0.111 (2.88)	0.246 (4.24)	0.158 (3.80)	0.269 (6.55)	0.274 (8.49)
<i>qntfb</i> _{<i>t</i>-1}	0.182 (3.17)	0.099 (2.64)	0.148 (2.44)	0.135 (3.11)	0.091 (1.95)	0.078 (2.25)
Log deviation in hours from trend (<i>devtrend</i>) ^b	-0.108 (-1.27)	-0.201 (-2.39)	-0.099 (-1.00)	-0.078 (-0.73)	-0.368 (-2.69)	-0.410 (-2.96)
Sum of <i>q</i> coefficients	0.38	0.21	0.39	0.29	0.36	0.35
No. of observations	45	45	40	40	40	40
Adjusted <i>R</i> ²	0.39	0.26	0.44	0.39	0.69	0.67
Root mean squared error	0.0024	0.0016	0.0023	0.0016	0.0038	0.0030

	<i>Regressions using three lags of nonfarm business output</i>			
Nonfarm business output	0.141	0.144	0.145	0.257
(qnf)	(2.37)	(4.41)	(3.57)	(6.18)
qnf _{t-1}	0.174	0.19	0.176	0.099
	(3.00)	(4.92)	(3.94)	(2.10)
qnf _{t-2}	0.133	0.172	0.102	0.038
	(2.08)	(4.23)	(2.55)	(0.80)
qnf _{t-3}	0.044	0.147	0.007	0.056
	(0.71)	(4.50)	(0.19)	(1.32)
Log deviation in hours from trend (devtrend) ^a	0.001	-0.134	-0.086	-0.223
	(0.01)	(-1.80)	(-0.86)	(-1.34)
Sum of q coefficients	0.49	0.65	0.43	0.45
No. of observations	45	45	40	40
Adjusted R ²	0.43	0.58	0.49	0.70
Root mean squared error	0.0023	0.0013	0.0015	0.0038
				0.73
				0.0028

Source: Author's regressions using Bureau of Labor Statistics and Bureau of Economic Analysis data.

a. The model for employment is $x = a + b_1q + b_2q_{-1} + \dots + b_t\text{devtrend} + e$, where x is alternatively h or p . All variables, except deviation in hours, are changes in logs. Regressions use the Prais-Winsten transformation. Numbers in parentheses are t statistics.

b. Trend calculated with Hodrick-Prescott filter, parameter = 1600.

decades spanning 1954 to 1973 support the preference for the payroll series, results for the decades since then do not.

The employment demand regressions reported in table 7 provide a more mixed message. With either measure of output, including the current and one lagged value of output yields a slightly better fit with the household data in both earlier decades. Adding two more lags of output to the regressions improves the fit with the household data in the 1983–93 decade and with the payroll data in the 1973–83 decade.

Conclusions: Splitting the Difference

The payroll and the household data both provide reliable information about employment and other aspects of the job market at impressive and useful levels of detail. They also track each other fairly closely over most annual intervals, the period surrounding 2000 being the recent exception. However, the series differ substantially in their estimates of monthly changes in total employment. For recent decades the statistical results reported here offer no convincing reason to reject one series in favor of the other, and in regressions the two generally had similar coefficients. This suggests the simple rule of thumb that users of the data should average the monthly change in the two.

Table 8 compares changes in the payroll and household series with this average series using both revised and initially reported payroll data. The greater volatility of the household data is apparent, but the volatility of the averaged data is little different from that of either payroll series. And, over the entire 137-month period from January 1994 to June 2005, using either revised or unrevised payroll data, the standard deviation of the averaged measures is very near that of the payroll series itself.⁸ Thus averaging the monthly changes in household and payroll employment should eliminate any reluctance to use the household series because of its greater volatility.

Finally, table 9 compares the performance of the averaged series with that of the household and payroll series in tracking output and unemploy-

8. With quarterly data, over 1994:1–2005:2 the standard deviation is lower in the household series than in the revised payroll series and about the same as in the unrevised payroll series.

Table 8. Monthly and Yearly Changes in Payroll and Household Employment, June 2003–June 2005

Thousands of workers

<i>Period</i>	<i>Household (H)</i>	<i>Originally reported payroll (P_orig)</i>	<i>Revised payroll (P)</i>	<i>Average of H and P_orig</i>	<i>Average of H and P</i>
July 2003	-254	-44	3	-149	126
August 2003	79	-93	2	-7	40
September 2003	25	57	94	41	59
October 2003	388	126	123	257	255
November 2003	418	57	96	238	257
December 2003	-46	1	83	-23	18
January 2004	406	112	117	259	261
February 2004	-142	21	94	-60	24
March 2004	71	308	320	190	196
April 2004	229	288	337	258	283
May 2004	194	248	250	221	222
June 2004	301	112	106	206	203
July 2004	464	32	83	248	273
August 2004	18	144	188	81	103
September 2004	-126	96	130	-15	2
October 2004	289	337	282	313	286
November 2004	449	112	132	281	291
December 2004	-132	157	133	12	0
January 2005	82	146	146	114	114
February 2005	-94	262	300	84	103
March 2005	344	110	122	227	233
April 2005	577	274	292	425	434
May 2005	363	78	104	220	233
June 2005	157	146	146	152	152
Standard deviation, 1994–2005	268	172	173	182	181
Year ending:					
June 1996	2,040	2,091	2,458	2,065	2,249
June 1997	2,435	2,275	2,995	2,355	2,715
June 1998	1,792	3,033	3,204	2,413	2,498
June 1999	1,905	2,340	3,004	2,122	2,454
June 2000	1,953	2,713	2,989	2,333	2,471
June 2001	-5	336	190	166	93
June 2002	-443	-1,069	-1,652	-756	1,048
June 2003	729	-512	-523	108	103
June 2004	1,668	1,193	1,625	1,431	1,647
June 2005	2,391	1,894	2,058	2,143	2,225

Source: BLS data.

Table 9. Tracking Output and Unemployment Using the Average of Changes in Household and Payroll Employment, 1994–2005^a

Independent variable	GDP ^b			Nonfarm business output ^b			Unemployment (quarterly data) ^c			Unemployment (monthly data) ^c		
	9-1	9-2	9-3	9-4	9-5	9-6	9-7	9-8	9-9	9-10	9-11	9-12
Household employment (h)	0.845 (4.10)			1.102 (4.15)			-9.918 (-7.46)			-6.149 (-6.81)		
Payroll employment (p_orig)		0.616 (2.93)			0.794 (2.98)			-8.758 (-6.88)			-7.129 (-6.01)	
Average of p_orig and h			0.832 (3.84)			1.065 (3.86)						-9.225 (-8.24)
No. of observations	45	45	45	45	45	45	45	45	45	137	137	137
Adjusted R ²	0.26	0.15	0.23	0.26	0.15	0.23	0.55	0.51	0.64	0.25	0.21	0.33
Payroll employment (p)		0.634 (3.60)										
Average of p and h			0.831 (4.26)			1.066 (4.33)						-9.243 (-8.57)
No. of observations	45	45	45	45	45	45	45	45	45	137	137	137
Adjusted R ²	0.21	0.28	0.28	0.22	0.22	0.28	0.55	0.55	0.66	0.24	0.24	0.35

Source: Author's regressions using BLS and BEA data.

a. All variables are changes in logs. Regressions use the Prais-Winsten transformation. Numbers in parentheses are *t*-statistics.b. The unrestricted model for output is $q = a + b_1x + b_2(h + p/2) + \epsilon$, where x is the indicated employment series. Constant terms are not reported.c. The unrestricted model for unemployment is $u = a + b_1x + b_2(h + p/2) + \epsilon$, where x is the indicated employment series. Constant terms are not reported.

ment in the 1994–2005 period. When the revised payroll data are used, the averaged data explain substantially more than the payroll data in all regressions, and more than the household data, although the differences are not as great in this comparison. When the originally reported payroll data are used, the superiority of the averaged series over the payroll data is even more pronounced, with the biggest improvement in the regression using monthly data. In regressions explaining employment change as in table 7 (results not shown), the average series gives the same R -squared as the payroll data when four quarters of output are used as explanatory variables, and much higher R -squareds when two quarters of output are used.

Even if users may wish to average the two series to get the most useful characterization of monthly employment growth, the BLS cannot report this average as its estimate of employment growth, because the two series measure different things. However, the BLS might consider giving the employment change from the household series the same prominence as that from the payroll series in its press releases. And it should provide a usable estimate of employment change in those months when it introduces new estimates of the population.

This paper has not directly addressed the value of the household data for longer-run issues such as aggregate productivity growth, which is calculated using labor input taken largely from the payroll data. Nonetheless, the success of the household data in this paper's short-run comparisons suggests it may be useful for analyzing such issues as well. The longer-run drift apart and back together of the two series shown in table 8 and figure 1 occurred over the same period when productivity growth accelerated, which is the most striking feature of the macroeconomy in these years and the subject of a great deal of research.⁹

It would be interesting to use the household series adjusted to the non-farm business sector, to see how it alters our understanding of aggregate productivity growth in these years. Studies that rely on data disaggregated by industry would not be possible, but the observed aggregate movements could be separated into trend and cycle components. This might provide a more plausible path for the speedup of productivity and add to the understanding of its present trend.

9. Charles Schultze has brought to my attention that the statistical discrepancy between the product and the income sides of the GDP accounts has roughly paralleled these trends in the two employment measures. These patterns are consistent with the idea that the household data better track the product-side measure of GDP at these low frequencies.

Comment and Discussion

William Wascher:¹ George Perry has written a very interesting and provocative paper aimed at anyone engaged in analyzing and forecasting the U.S. economy. As Perry notes, analysts place considerable emphasis on the Bureau of Labor Statistics' release of employment data on the first Friday of each month, essentially because these data provide one of the earliest readings on the pace of economic activity in the preceding month. One difficulty in interpreting the data, however, is that the release includes results from two independent surveys, of households and of establishments, which often give different signals about the strength of the labor market. For example, in the last month for which data were available as this volume went to press (November 2005), the household measure of employment fell by more than 50,000, whereas the payroll measure increased by more than 200,000.

Perry argues that analysts place too much weight on the monthly employment changes in the payroll survey and consequently too little on those in the household survey. In his view, analysts have misinterpreted the smoothness in the employment changes from the payroll survey as evidence of greater accuracy, because they have not taken into consideration some other shortcomings of that survey. Instead, according to Perry, forecasters would be better advised to average the monthly changes from the two surveys.

The paper's first section presents evidence on whether market participants and policymakers focus more heavily on the monthly employment changes

1. I would like to thank Jessica Williams for excellent research assistance and Stephanie Aaronson, Bruce Fallick, Andrew Figura, and Joyce Zickler for helpful discussions. The views expressed here do not necessarily reflect those of the Board of Governors of the Federal Reserve System or its staff.

from the payroll survey than on the corresponding data from the household survey. Perry's results show that federal government bond yields respond significantly to surprises in the monthly change in payroll employment immediately following the release of the data, and hardly at all to surprises in household employment. An unavoidable limitation of this analysis is that the "surprises" for both the household and the payroll employment changes are based on a single mean employment forecast gathered in a survey conducted by Money Market Services, which explicitly refers to the payroll survey. I suspect that if market participants were asked instead about the household survey, they would take into account various technical factors and time-series properties of those data, which would lead their forecasts for that employment measure to differ from that for the payroll survey. This suggests that the household employment surprise is likely measured with considerable error, which would seem to diminish the information content of Perry's regressions. Even so, the simple fact that market participants forecast only the monthly change in payroll employment, along with the existence of a derivatives auction for that portion of the release, indicates that these analysts do pay considerably greater attention to the payroll measure.

At first glance the same appears to be true of policymakers. The results in Perry's table 2 show that the announced federal funds rate target after each FOMC meeting responds more strongly to changes in payroll employment than to changes in household employment. Taken literally, the coefficients in column 2–3 of that table suggest that policymakers place about twice as much weight on the payroll survey as on the household survey. However, such a result is not necessarily evidence of a strong preference for the payroll survey. The FOMC meetings often take place well after the employment data are released, and during that period a whole host of other indicators of economic activity become available, all of which are in the FOMC's information set. In this context an alternative interpretation might be that the payroll data line up better than the household survey data do with the broad array of information used by the FOMC in making monetary policy decisions.

Nonetheless, there does seem to be a general preference among forecasters for the payroll survey, and this preference presumably reflects the view that the payroll survey's measure of employment growth has a higher signal-to-noise ratio than does the household survey measure. Perry's Figure 2 illustrates the rationale for this prevailing wisdom. With the monthly change in

household employment regularly exceeding plus or minus 500,000, one often would seem hard pressed to make heads or tails of these data.

However, Perry argues that this view of the data is too simplistic. In particular, he points to two shortcomings of the payroll survey that might tend to diminish its reliability relative to the household survey: The first is the difficulty that the payroll survey has in capturing establishment births and deaths. The second involves misreporting on the unemployment insurance (UI) tax records that serve as the benchmark for the payroll data.

Although both these critiques are valid, I think the paper overstates the extent to which they offset the benefits associated with the payroll survey's larger sample size and its adjustment to an annual benchmark—even if that benchmark is not a complete count of jobs. For one thing, the BLS has made considerable progress in addressing establishment births and deaths, for example by moving to a probability sample between 1999 and 2002 and by developing statistical models of the birth and death process, which are used to adjust the sample-based estimates of employment growth for this bias. It may be too early to judge how much these new procedures have helped. But it seems likely that any such bias will be less important going forward. In addition, although there is undoubtedly noncompliance in the UI system, BLS analysts are aware of this problem, which is why the payroll survey is benchmarked annually to the March data—when the UI data are likely to be most accurate—rather than to the quarterly UI totals.

Moreover, the household survey is also subject to significant sources of nonsampling error. The paper notes the possibility of errors in the monthly population estimates used to derive aggregate estimates of employment and unemployment from the survey-based ratios. But even the decennial census is not a perfect count of the population, and, indeed, judging from the controversy that frequently surrounds these counts, many observers would probably view noncompliance as more of a problem for the census than for the UI system. In addition, the fact that the household employment series is not benchmarked to anything at an annual frequency would seem to be a disadvantage relative to the payroll survey, even if the UI benchmark is not perfect. There are other potential sources of error as well, including proxy responses, nonresponse and undercoverage, rotation group bias, and even seasonal adjustment (most notably, problems in accounting for variation in the intervals between survey weeks). Thus, although the payroll survey indeed has its shortcomings, so does the household survey, and it is far from apparent that, from the standpoint of month-to-month changes in

employment, the problems with the payroll survey outweigh its considerable advantages.

Nonetheless, Perry correctly points out that the smoothness of a data series should not be mistaken for accuracy, and with this in mind he uses a variety of empirical tests to provide some basis for assessing the relative information content of the household and payroll surveys. This strikes me as the right way to proceed, and, indeed, the results turn out to be quite interesting.

The first test examines the contemporaneous correlation between employment and output, and Perry shows that various constructs of household survey employment perform as well as or better than payroll employment in a quarterly regression of GDP or nonfarm business output growth on employment changes. As an aside, I did not find Perry's trend-adjusted household survey measure, which takes out the divergence between the two surveys from mid-1997 to early 2003, very compelling. Gauging developments in the labor market was quite important during that period, and the divergence of household and payroll employment was a real puzzle to many analysts. More generally, from a practical standpoint, it would be difficult to know in real time whether to make such adjustments to either series. In any case, Perry's conclusions do not really rest on this construct, and in the end he recommends splitting the difference between the published changes in household and payroll employment. In the setup he uses, one can actually estimate the optimal weight that should be placed on each series by taking the fitted values for GDP growth from each equation and including them in a regression on actual GDP growth. Interestingly, the coefficients from such an exercise suggest putting a 50 percent weight on each measure, just what Perry recommends.

From the perspective of assessing the information content of alternative measures of the *monthly* change in employment, a limitation of this test is its reliance on *quarterly average* data. As Perry shows later in the paper, the use of quarterly averages yields a significant reduction in the variability of the household series, presumably because it smoothes through some negative serial correlation in the monthly estimates. This makes it look more like the payroll survey, and, indeed, on a quarterly average basis, the correlation between changes in payroll and in household employment is reasonably high, at about 0.7. Nonetheless, the results also indicate that there is independent information in the quarterly estimates from the household survey and no reason, a priori, to prefer one measure over the other.

Table 1. Regressions of GDP Growth on Alternative Measures of Employment Growth and on Predicted GDP Growth, 1994:1–2005:2*R*²s of regression except where otherwise stated

<i>Measure of employment or regression coefficient</i>	<i>Using employment data for full quarter</i>	<i>Using employment data for first month of current quarter only^a</i>
<i>Regressions of GDP growth on employment</i>		
Payroll employment	0.22	0.18
Household employment	0.22	0.13
Household employment, research series	0.25	0.12
Average of household and payroll employment	0.26	0.19
Average of household research series and payroll employment	0.29	0.17
<i>Regression of GDP growth on predicted output growth</i>		
Regression coefficient on fitted values from payroll employment regression	0.50	0.70
Regression coefficient on fitted values from household employment regression	0.50	0.30
<i>R</i> ²	0.26	0.19

Source: Perry, this volume, and author's regressions using Bureau of Labor Statistics and Bureau of Economic Analysis data.

a. Employment variables are calculated as the quarterly percent change in employment assuming that the growth rates in the second and third months of the current quarter are identical to that in the first month.

I suspect, however, that market participants would not view these results as especially relevant to their own use of the data, because they often view the employment data as providing an early indication of economic activity in the current quarter. By the time a full quarter of data becomes available, a considerable amount of other information on the current state of the economy is also available, making the quarterly average of recent employment changes less useful as a contemporaneous indicator.

To assess the information content of each monthly employment measure as a forward-looking indicator, I replicated Perry's procedure but instead assumed that only the first month of employment was known in each quarter.² The results are shown in table 1 above, where the first column repeats Perry's results with a full quarter's worth of employment data and the second reports the *R*-squareds for identical regressions using just the

2. Specifically, I assume that employment growth in the second and third months of the current quarter is the same as in the first month.

first month of employment data in each quarter. In this experiment the payroll survey proves to be a more accurate predictor of real GDP growth than the household survey, although their average does just as well. The estimated optimal weights (reported in the bottom panel) are 70 percent for the payroll survey and 30 percent for the household survey. Interestingly, these weights match up quite closely to Perry's estimate of the relative responsiveness of the federal funds rate target to each measure of employment growth (column 2–3 of his table 2).

As a second test of the relative information content of the two measures, Perry examines the correlation of employment changes with other monthly indicators of activity. Indeed, the case for the household survey would be quite compelling if there were evidence that it predicted monthly activity as well as (or better than) the payroll survey. Unfortunately, this turns out to be a frustrating exercise because of the dearth of reliable independent monthly indicators of either economic activity or labor market developments. Perry uses the unemployment rate as his indicator (which is also derived from the household survey), arguing that it can be taken as an independent indicator if reporting errors between employment and unemployment are not very important. However, given the potential for other nonsampling errors, I am skeptical of what can be learned from using another measure of the labor market from the household survey.

Instead I looked for other possible candidates that seemed reasonable to consider in this exercise. Although none is perfect, a few potentially useful indicators of general activity include the estimate of monthly real GDP growth developed by Macroeconomic Advisers, the change in the components of industrial production that are based solely on physical product data and thus not correlated with employment by construction, and the National Activity index published by the Federal Reserve Bank of Chicago, which is a summary measure of real economic activity based on the first principal component of a dataset consisting of eighty-five different indicators. In addition, I considered a couple of independently measured labor market indicators: the Conference Board's help wanted index and two series related to UI claims.

My table 2 shows the *R*-squareds from regressions that are similar to Perry's specification but replace the unemployment rate with one of these other monthly indicators. In most cases where there is any correlation at all, it tends to be higher for the unrevised and revised payroll survey measures of employment (first two columns) than for the household survey measure

Table 2. Regressions of Selected Monthly Economic Activity Indicators on Payroll and Household Employment^a, January 1994–June 1995*R*²s of regression

<i>Indicator</i>	<i>Unrevised payroll data</i>	<i>Revised payroll data</i>	<i>Household data</i>	<i>Household research series</i>	<i>Average of payroll and household data</i>	<i>Average of payroll and household research series</i>
Monthly GDP	0.00	0.01	0.00	0.00	0.00	0.00
Physical product IP	0.09	0.11	0.00	0.00	0.02	0.03
National Activity index	0.47	0.55	0.10	0.12	0.35	0.29
Help wanted index	0.07	0.07	0.06	0.06	0.09	0.08
Initial UI claims	0.31	0.44	0.07	0.05	0.03	0.07
Insured unemployment	0.08	0.12	0.03	0.05	0.09	0.09

Source: Author's regressions using data from Macroeconomic Advisers, the Board of Governors of the Federal Reserve System, the Department of Labor, the Conference Board, and the Federal Reserve Bank of Chicago.

a. All variables are percent changes except for the National Activity index and initial UI claims, which are in levels.

(third column). Of course, many of these monthly indicators are probably more naturally related to the payroll survey because they correspond more closely to payroll-related jobs than to the broader household survey definition of employment. However, this would not be the case for the BLS's research series, which adjusts the household survey data to the payroll survey job concept. Nevertheless, as shown in the fourth column, the research series also is less highly correlated with the monthly indicators than is the payroll employment series. In addition, averaging together the household and payroll employment changes (last two columns) does better than the household survey measures alone, but generally not as well as the payroll survey measures.

Perry's final test examines how each series fits in a standard model of employment fluctuations. This exercise is interesting, but it again relies on quarterly data and so speaks more to the information content of quarterly averages than to the usefulness of the monthly data.

So, is the professional wisdom wrong? My own conclusion is that, from the narrow perspective of the information content of monthly employment changes, the totality of the evidence suggests that the payroll survey is a better indicator because of its larger sample size. That said, from the stand-

point of analyzing current labor market developments, a monthly perspective is often *too* narrow, and I think Perry makes a strong case that much can be learned from the household survey over a slightly longer time frame. As his figure 1 shows, the two measures often tell the same story over longer periods. But they do not always correspond so well, and in such instances forecasters should not automatically assume that one survey is more accurate than the other. Rather, such discrepancies call for a closer inspection of the two series and an effort to ascertain their relative accuracy in the prevailing circumstances. More broadly, much can be learned from looking at a range of timely labor market indicators that extend beyond the standard measures of employment growth to include estimates of average hours in both surveys, the wide variety of other measures available from the household survey, and the relatively new Job Openings and Labor Turnover Survey (JOLTS) data now being published by the BLS.

General discussion: Gregory Mankiw said he regarded the paper as valuable and long overdue. He noted that, as household employment had risen faster than payroll employment in recent years, attempts to include the household data in assessments of the economy tended to be seen as politically motivated. He suggested constructing a measure that would adjust the payroll series to the household concept (by adding the self-employed to the payroll data) rather than vice versa, as in the research series that the paper tested. He observed that the research series performed worse than the payroll series in the regressions explaining unemployment, but not in those explaining output. Mankiw also pointed to the methodological difficulties raised by dual job holders. For example, if a worker who already has a part-time job takes a second part-time job, that worker's contribution to GDP will increase, and it is reasonable to count the job twice. On the other hand, if a worker quits two part-time jobs to take one full-time job, the contribution to GDP of that worker might not change, but the event might be counted as the loss of a job.

James Spletzer remarked that the paper set out to determine the best measure of monthly employment change, yet data limitations forced most of the statistical tests to use quarterly averages. He suggested that the usefulness of each measure might depend on its reporting frequency. Spletzer also discussed the ongoing efforts at the Bureau of Labor Statistics to understand the trend divergences between the two employment surveys and to improve their methodologies: he called attention to a Federal Economic

Statistics Advisory Committee paper and a related report forthcoming in the *Monthly Labor Review*, which document what is known about the divergences. He noted that corrections to the population controls used in the household survey, based upon the 2000 decennial census, had been found to account for a meaningful part of the divergence in late-1990s employment trends, and that these controls are reviewed and often updated annually, although on the basis of less reliable intercensal estimates. George Perry observed that population errors can have a large impact over long horizons but are of little importance for monthly or quarterly changes in employment.

Turning to the payroll survey, Spletzer pointed out that the population of establishments was highly volatile, with roughly 800,000 establishments opening and the same number closing every year on a base of 6 million. He regarded the annual tally of establishments as highly reliable.

Mankiw interpreted the results of Perry's table 1, in which the markets react only to the payroll survey, as showing that they are reacting to Federal Reserve Chairman Alan Greenspan, who has indicated that he pays close attention to that survey. Alan Blinder agreed with Mankiw and related the paper to that by Morris and Shin in the present volume: the reactions to the different employment series could be seen as the quintessential example of the Keynesian beauty contest that those authors described. He observed that the paper provided a rational expectations test, which both market participants and the Federal Reserve had failed.

References

- Bureau of Labor Statistics. 2005. "Employment from the BLS Household and Payroll Surveys: Summary of Recent Trends." Washington. (December) (bls.gov/web/ces_cps_trends.pdf).
- Nardone, Thomas, and others. 2003. "Examining the Discrepancy in Employment Growth between the CPS and the CES." Working paper prepared for the Federal Economic Statistics Advisory Committee (October 17).
- Sims, Christopher A. 2002. "The Role of Models and Probabilities in the Monetary Policy Process." *BPEA*, no. 2: 1–40.

