DISCUSSION PAPER SERIES

IZA DP No. 16699

The Mismeasurement of Work Time: Implications for Wage Discrimination and Inequality

George J. Borjas
Daniel S. Hamermesh

DECEMBER 2023

DISCUSSION PAPER SERIES

IZA DP No. 16699

The Mismeasurement of Work Time: Implications for Wage Discrimination and Inequality

George J. Borjas
Harvard University, IZA and NBER

Daniel S. Hamermesh
University of Texas at Austin, IZA and NBER

DECEMBER 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Mismeasurement of Work Time: Implications for Wage Discrimination and Inequality*

A comparison of measures of work time in the CPS-ASEC data file (based on recall) with contemporaneous measures reveals many logical inconsistencies and probable errors. About 8 percent of ASEC respondents report weeks worked last year that contradict their work histories in the Basic monthly interviews; the error rate is over 50 percent among workers who move in and out of the workforce across their monthly interviews. Over 20 percent give contradictory information about whether they usually work a full-time weekly schedule (35 or more hours per week). A small part of the inconsistency arises because an increasing fraction of observations in the ASEC (over 20 percent by 2018) consists of people whose record was fully imputed. The errors and imputations are not random: The levels and trends differ by gender and race, and they affect the calculation of wage differentials between 1978 to 2018. After adjusting for the measurement errors and excluding imputations, we find that gender wage gaps among all workers narrowed by 4 log points more than is commonly reported, and that residual wage inequality decreased by 6 log points more. The biases also exist in measures of wage gaps and residual inequality among full-time year-round workers. Using a more carefully defined sample of such workers shows that gender and racial wage differentials have narrowed slightly less than previously estimated using ASEC data, but much more than indicated by commonly used estimates from CPS Outgoing Rotation Groups.

JEL Classification: J22, J71, D63

Keywords: work time, measurement, wage discrimination, wage inequality

Corresponding author:
Daniel S. Hamermesh
University of Texas at Austin
110 Inner Campus Drive
Austin, TX 78705
USA

E-mail: hamermes@eco.utexas.edu

* We thank Francine Blau, Charles Brown, Anthony Edo, Jennifer Hunt, Theodore Joyce, Matthias Krapf, Kevin Lang, Bruce Meyer, Andrés Mira, participants at a seminar at the NBER—New York, and especially Jay Stewart, for helpful comments.

There are two things you should not watch being made: Sausages and data.
—Attributed to Zvi Griliches

I. Introduction

Immense literatures in economics have charted the time paths of measures of wage differentials and of inequality over the past few decades. The studies typically compute these measures after adjusting for differences in the amount of work time (such as weekly hours of work or annual weeks worked) that generated the earnings that survey participants report. The self-reported measures of work time likely contain substantial measurement error, both classical and non-classical, as has been shown for weekly work hours by Borjas (1980), Rodgers *et al*. (1993), and Bound *et al*. (1994). Similar errors may arise from misreports of weeks worked. If these misreports are correlated with workers' demographic characteristics, and if those correlations change over time, estimates of the levels and even the directions of the time paths of wages and wage differentials will be biased.

We examine possible misreports of weeks and hours worked in the monthly Current Population Surveys (CPS), the data source that informs much of our understanding of the evolution of wages in the United States. The March CPS in year *t* includes the Annual Social and Economic Supplement (ASEC), in which respondents are asked to report their labor earnings in year *t*-1, as well as the number of weeks and usual weekly hours worked that year. Due to the sampling scheme used by the CPS, many ASEC respondents appear in the Basic monthly CPS surveys in year *t*-1. For this subsample we have contemporaneous measures of whether and how much the respondents worked in each of their (up to four) appearances in the monthly surveys.

The overlap of the ASEC and the Basic CPS allows us to measure the extent of logically inconsistent reports of weeks worked, i.e., retrospective reports in the ASEC contradicting what is recorded in the Basic monthly interviews. Additionally, it allows us to spot possible errors (that are not necessarily logical inconsistencies) when reports of work time in the ASEC differ from the picture the CPS respondent gave in the Basic CPS. We conjecture that retrospective reports of work time in year *t*-1 in the ASEC contain more errors than contemporaneous reports in the monthly CPS interviews, as suggested by evidence on recalled work time over much shorter periods (Barrett and Hamermesh, 2019, and the huge literature on recall error, including Sudman and Bradburn, 1973, and Celhay *et al*., 2021). A key question is whether the differences in the alternative reports of work time correlate with demographic variables in ways that change over time and could thus alter inferences about the evolution of wage differentials and inequality.

A potential difficulty in determining whether the information provided by the ASEC and Basic CPS files is internally consistent is that a sizable (and increasing) fraction of the observations in the ASEC consist of "replicants"—people whose *entire* ASEC record has been imputed using a "hot-deck" procedure based on the responses of observationally similar participants.[1] We calculate the magnitude of these logical inconsistencies, probable errors, and "full-record imputations" and examine their time paths and demographic correlates. We show that these considerations change some generally accepted conclusions about the history of wage outcomes in the U.S. labor market. To the extent that wage trends in other industrialized economies are based on data that are similarly ambiguous or error-ridden, this approach can be applied elsewhere.[2]

The next section demonstrates the existence of unambiguous logical inconsistencies and probable errors between reports of weeks worked in the Basic monthly interviews and the ASEC. Section III uses the data to create a moving panel of workers and constructs alternative measures of weeks worked, one based on retrospective ASEC reports, the other based on contemporaneous work behavior in the Basic interviews. We demonstrate how the logical inconsistences in reported weeks worked and the differences between the alternative measures of weeks worked have changed over time. Section IV shows how these recall errors can bias the estimated coefficients of regressors in earnings functions and contaminate measures of residual wage inequality. Section V stresses the diversity and near-total lack of uniformity in the samples researchers choose from the CPS and the ASEC to analyze these labor-market outcomes.

[1] This designation is based on the term used in the 1982 movie *Blade Runner*. Replicants are androids who are virtually identical to human beings.

[2] Some of the problems of measuring cyclical changes in hours of work are discussed by Burda *et al*. (2013). Bick *et al*. (2019) discuss the difficulty of making international comparisons of work hours.

Section VI re-examines the evolution of gender and racial wage differentials among all workers, as well as residual wage inequality, from 1978 to 2018, and illustrates their sensitivity to the measurement issues that are the foci of our analysis. Section VII revisits the same questions, concentrating on the sample of full-time year-round workers that is sometimes used in the literature and expanding our analysis of measurement issues to include the misreporting of both weeks worked per year and hours worked per week.[3]

II. Measurement Error in Weeks Worked

The typical respondent in the CPS is interviewed eight times over a period spanning 16 months, assuming s/he resides at the same address over this time. Upon entering the sample, the person is interviewed for four consecutive months as part of the Basic monthly survey. That respondent then goes on an "interview hiatus" for the next 8 months, and is then interviewed again for four consecutive months before departing the CPS.

Each of the Basic monthly surveys asks respondents about their employment status in the (reference) week preceding the interview. People who are interviewed in March are asked many additional questions, which form the Annual Social and Economic Supplement (ASEC). The additional information includes the (recalled) number of weeks worked in the previous calendar year and the (recalled) weekly hours usually worked that year. The panel rotation sampling implies that some of the respondents in the March survey in year t appear in several Basic monthly surveys in year t-1. The panel aspect of the CPS thus provides a way of determining if the recalled numbers of weeks and hours worked in the ASEC are consistent with the contemporaneously reported work history in the Basic monthly surveys.

Table 1 illustrates the possible rotations for the subset of interviewees in the CPS whose appearance in the March survey in year t occurs in the last half of the interview cycle (i.e., who have already been interviewed at least four times).[4] To illustrate, consider respondents who happen to be in their seventh

[3] Our analysis only examines measurement issues in household data, the CPS and its components, and ignores the well-known differences between household- and establishment-based reports of work time (Frazis and Stewart, 2010).

[4] Those whose ASEC interview is the fourth (out of the potential eight) entered the CPS in December of year t-1. We ignore members of this rotation group, as their reported work experience in year t-1 consists of only a single month,

interview in March of year *t* and are asked to report how many weeks they worked in year *t*-1. As the table shows, this subsample must have entered the CPS in January of year *t*-1, been interviewed as part of the Basic survey from January through April that year, and then reentered the CPS in January of year *t* for their final four interviews. Similarly, people in their final CPS interview in March of year *t* must have entered the CPS in December of year *t*-2 and been part of the Basic monthly interviews in January, February, March, and December of year *t*-1.

Note that for the subsample whose appearance in the ASEC falls during the second half of the CPS interview cycle, we observe *four* contemporaneous reports of time worked in year *t*-1. The key implication is that for a large portion of the ASEC sample we can check if the recalled measure of weeks worked is consistent with the observed work history of the respondent in year *t*-1. The comparison also allows checking for obvious discrepancies in reported usual weekly hours. Any discrepancies would reveal measurement error in the widely used measures of weeks and hours worked available in the ASEC, potentially contaminating estimates of the determinants of (and trends in) earnings.

To calculate the extent of measurement error in weeks worked as recollected in the ASEC, we rely on a key assumption: If someone in the Basic monthly survey reports being employed in the reference week, that person worked the entire month. Conversely, if the person reports not working during the reference week, we assume that s/he was not working at all that month. This assumption makes it easy to catch the most obvious recall errors made by ASEC respondents. We show below that it likely describes actual work schedules accurately and that the frequency of logical inconsistencies is almost as great if the response on work status in the Basic monthly interview applied only to the reference week.

To illustrate, suppose that someone reports being employed in each of the four monthly Basic interviews in year *t*-1. That person, therefore, *must have worked at least 16 weeks* in that year. A clear recall error exists if the person reports in the ASEC having worked fewer than 16 weeks. Similarly, consider a

making an extrapolation of their labor force behavior to the entire year unreliable. Respondents whose ASEC interview is between the first and the third months were not included in any Basic monthly surveys in year *t*-1 and thus cannot be used in an analysis that compares recall and contemporaneous measures of work time for those years.

person who reported working in three of the four Basic monthly interviews in year *t*-1. There is a clear discrepancy if the respondent reports in the ASEC having worked fewer than 12 weeks. There is *also* a clear discrepancy if the ASEC response is more than 48 weeks in year *t*-1; after all, we know from the contemporaneous work history that the person did not work in *at least* one month that year. This final example illustrates that our calculation *underestimates* the actual measurement error in the ASEC measure of weeks worked. It is possible that this individual worked only three months in year *t* (coincidentally, three of the four months that s/he participated in the Basic monthly surveys), but we are only considering ASEC responses above 48 weeks as being unambiguously incorrect.

We calculate the frequency of these logical inconsistencies using a merged sample of Basic and ASEC data between 1978 and 2018, creating essentially a series of one-year panels of CPS respondents.[5] The sample consists of people who are in the last half of the CPS interview cycle in year *t*. We link the ASEC and Basic CPS data files using the personal identifier (IPUMS variable *cpsidp*).[6] We exclude the 2.2 percent of linked respondents whose gender indicator changed between the Basic and ASEC files, or whose reported age declined, or whose reported age rose by more than two years. We include persons ages 25-59, thus avoiding most of those who may still be in school or who may already be cutting work time as they near retirement.[7]

As Panel A of Table 2 demonstrates, 18 percent of the respondents included in the dataset reported not working in any of their four Basic CPS interviews; 9 percent reported working in 1-3 of those

[5] All the Basic and ASEC data files were downloaded from the IPUMS archive. We exclude the years 1985 and 1995 from the analysis because of coding difficulties in linking respondents between the Basic and ASEC files in those years. See https://cps.ipums.org/cps/cps_linking_documentation.shtml.

[6] The IPUMS now provides an additional linking variable (*cpsidv*) that checks for matching inconsistencies based on age, sex, *and* race (see https://cps.ipums.org/cps-action/variables/CPSIDV#description_section). Using that variable instead of *cpsidp* to create the panels more than doubles the number of unmatchable observations, no doubt because of the major changes in the coding of the race variable over time. Using this overly restrictive definition, however, hardly alters any of the results reported in this study.

[7] We do not consider possible trends in the amounts of actual work effort when someone reports having worked in either the ASEC or the Basic CPS, nor how these affect the wage comparisons that we make. The issue of effort per hour worked goes back at least to Florence (1924). We know that there are cyclical changes in actual work time per hour on the job (Burda *et al*., 2020), but know nothing about trends.

interviews; and 73 percent stated that they worked in all four interviews. Over twice as many women as men answered that they did not work in any of the 4 months, and men are about 20 percentage points more likely to have worked in all four interviews.

Panel B of Table 2 reports the estimated error rates in weeks worked, defined as the (unweighted) fraction of sample respondents whose Basic work history unambiguously contradicts their recall in the ASEC. The overall error rate is 7.5 percent, with women having a higher error rate than men (8.4 versus 6.4 percent). The table also shows a striking relationship between a person's work history in year *t*-1 and the error rate. The error rate is lowest among those who report being continuously employed in the Basic interviews in year *t*-1: Men (women) who work all four months have an error rate of only 1.2 (2.4) percent.[8]

The error rate, however, can be much higher for other work histories. Among people who did not work in any of the four Basic monthly surveys the error rate, the fraction of ASEC respondents who claim to have worked more than 36 weeks in year *t*-1, is 12.2 percent among men and 3.9 percent among women. The error rate is even higher among respondents who moved in and out of the workforce in year *t*-1 (i.e., who report being employed in one, two, or three of the four Basic interviews). In this "mobile" group the error rate uniformly exceeds 50 percent, with slightly higher rates among women.

Note that we calculated these error rates by extrapolating a respondent's work status in the reference week of the Basic monthly survey to the entire interview month. Although it is not possible to test the validity of this assumption within the CPS, the work time data in the Survey of Income and Program Participation (SIPP) suggest that the assumption is very reasonable. In particular, the SIPP provides information on how many weeks the respondent worked within a given month. In the 2014 SIPP (Wave 4), for example, 85 percent of men and 73 percent of women aged 25-59 worked in a survey month. In this subsample of workers, only 0.67 percent of men and 0.73 percent of women report working fewer than four weeks in that month.

[8] The Basic monthly interviews for the persons included in our short panels occurred between January and June or in December of year *t*-1. There is no seasonality in the errors: The size of the measurement errors is independent of which months the ASEC respondent worked in the Basic CPS (or, alternatively, in what survey month from 5 to 8 the person was part of the ASEC).

Even if we do not assume that employment in a given month generally implies employment throughout the entire month, our conclusion about the extent of measurement error in the ASEC would not be greatly affected. Suppose, for example, that a person works three out of four times in the Basic interviews. There would then be a logical inconsistency only if the ASEC variable indicates that the person reported working fewer than 3 weeks or working 52 weeks in the previous year. As Panel C of Table 2 shows, this alternative and clearly incorrect assumption leads to only a slight drop in the estimated error rates, to 5.4 percent among men and 7.0 percent among women. The error rates remain near 50 percent in the sample of "mobile" workers. This similarity, of course, implies that most errors are at the extremes—people incorrectly reporting 52 weeks worked (or nearly 52) or zero weeks worked (or nearly zero) in the ASEC. Henceforth, our analysis will therefore extrapolate work status from the reference week to the month in which the Basic monthly interview took place.

The merged CPS panel indicates that error rates have been increasing over time. The top panel of Figure 1 shows that the overall error rate rose between 1978 and 2018, particularly among men. The male error rate was below 4 percent in 1978 and had doubled to more than 8 percent by 2018. The bottom panel of the figure illustrates the trends in the frequency of logical errors found in the subsample of people who moved in and out of the workforce in year t-1. The error rate for these "mobile" workers increased substantially among both men and women, and stood at remarkably high levels in 2018, with nearly 70 percent of these ASEC respondents unambiguously contradicting their contemporaneously recorded work histories in the Basic interviews.[9]

Although we have interpreted contradictions in work time reports between the Basic and ASEC files as recall errors, there are two unrelated sampling issues that might also create such conflicts. The first is that a fraction of responses about work time in both the ASEC and Basic files are "allocated"—imputed

[9] The potential inconsistencies between responses about employment status in these surveys were noted by Akerlof and Yellen (1985), who showed that the monthly Basic CPS unemployment rate was higher than that implied by retrospective reports from the ASEC in all but two years during the period 1960-81. Bound *et al*. (2001) discuss the issue of measuring unemployment in the different surveys in more detail.

for non-respondents using information from observationally similar participants who responded fully to the CPS. Some of these allocations are for *item* non-response, for example, on weeks worked in the ASEC or employment status in the Basic CPS. This CPS hot-decking procedure then imputes the specific missing information based on the person's characteristics.

Most of the allocations that could affect our calculations, however, are full-record imputations, where *every* variable for a particular observation in the ASEC file has been imputed using the hot-decking procedure. Stewart (2007, p. 184) provides a rare discussion of this potentially important issue:

> …If a large number of variables need to be imputed, then the *entire* record is allocated using a hot-deck procedure…The resulting records are internally consistent within the Income Supplement [ASEC] variables, although *not* between the Income Supplement and the monthly CPS. [Italics added].

A very high (and increasing) number of ASEC observations are replicants. Whether including them generates misleading estimates no doubt depends on the specific context being considered (see Heffetz and Reeves, 2019; Bee and Rothbaum, 2023). Figure 2a shows the trend in allocation rates that could affect any comparison of recalled weeks worked in the ASEC file with the contemporaneous work history in the Basic interviews.[10] A small number of people (fewer than 2 percent) did not report weeks worked in the ASEC, and that specific variable was allocated. An equally small number of observations did not report work status in at least one of the four months in the Basic monthly interviews, and that work status was also allocated. In contrast, the rate of full-record imputations hovered around 10 percent before 2010 and increased to over 20 percent by 2018.

[10] It is important to note that the quality flags that are typically used in the IPUMS to report imputation for a particular variable, such as the flag *qwkswork* for the recalled measure of weeks worked (*wkswork1*), do not provide any information about the full-record imputation issue. The full-record imputation flag is instead available in two separate and less familiar IPUMS variables (*uh_suprec_a1* for 1989 through 2006, and *uh_suprec_a2* since 2007). These variables correspond to the variables *FL-665* (before 2007) and *FL_665* (since 2007) in the original ASEC documentation published by the Bureau of the Census. Beginning in 2019, the item-specific imputation information reported in the IPUMS quality flag differs from the way the Census documentation reports the allocation information. For example, the Census allocation flag for the recalled weeks worked variable now distinguishes three possibilities: "no change," "allocated," or "full-record imputation," where the full-record imputation code turns on for specific values of the *FL_665* variable.

In principle, the very high frequency of full-record imputations can produce substantial discrepancies when comparing weeks worked recalled in the ASEC to contemporaneous work records in the Basic interviews, since the imputation need not lead to an ASEC record that is internally consistent with the Basic information. To determine whether recall error persists even after adjusting for the high number of replicants in the ASEC data, we recalculated the error rates after excluding all the observations that have any type of work time imputation in the Basic or ASEC files. As Panel D of Table 2 shows, the high error rates are not produced by the large number of replicants. The overall error rate remains around 7 percent, and the error rate among mobile workers is still over 50 percent.

The other potential problem with isolating recall error is that a large number of CPS respondents are "proxies"—the information about the person's work time in the CPS is provided by someone else in her/his household. The frequency of proxies is quite high: Responses to the ASEC question on weeks worked are provided by a proxy over half the time in our sample period. Our panel data set consists of five separate interviews, four interviews in the Basic Monthly files in year t-1 and the ASEC interview in year t. We calculate the fraction of observations where the respondent self-responded all five times, so that no proxies were ever used, and no measurement error could be introduced by different people providing responses about the individual's work time. The total absence of proxies for those included in our panel data is somewhat rare, particularly at the beginning of the period. Figure 2b, for example, shows that in 1990 only about 16 percent of men self-responded to all five questionnaires that we use, but this fraction increased to 36.5 percent by 2018. Among women, the rate of self-response in all five questionnaires hovered around 35 percent for much of the period. Despite the high frequency of proxy use, Panel E of Table 2 shows that the rates of recall error remain high even after we exclude replicants *and* proxy respondents. Given the large fraction of observations where proxies were used, we ignore this issue in most of our work, but it is one more caveat to consider when evaluating what we know about trends in wage differentials and inequality.

III. The Determinants of Errors in Recalled Work Time

Given the presence of recall errors in the ASEC weeks worked variable, it is worth constructing an alternative measure of weeks worked to assess the impact of the errors on the determinants of and trends in weekly earnings. We continue to assume that people who are employed during the reference week in a Basic monthly survey are employed during the entire interview month (and, conversely, that those who are not employed then are not employed for the entire month). This assumption allows us to calculate the total number of weeks that a person works during the four months in the Basic monthly survey. If we further assume that the work history of a respondent during the four months in the Basic monthly surveys in year *t*-1 is representative of the person's work history during the calendar year, we can then multiply this quantity by three to produce an alternative measure of weeks worked in year *t*-1.[11]

Note that our imputed measure of weeks worked also contains measurement errors, as a four-month segment of a person's work history provides incomplete information about the person's work time over a full year. Nevertheless, given that concurrent reports of work are likely to be closer to the truth than retrospective reports of work time one year later, it is instructive to determine if this partial "fix" affects our understanding of the trends in weekly earnings.

We thus have two alternative measures of weeks worked: S^A, the recall measure from the ASEC, and S^B, the imputation from the Basic monthly interviews. Define:

$$\Delta S = S^A - S^B, \quad (1)$$

which gives the discrepancy between recalled and contemporaneous reports of weeks worked. We have no expectations about the correlates of these errors/biases—our purpose is simply to measure them and use them to examine alternative calculations of earnings differentials.

[11] In constructing the alternative measure of weeks worked, we account for the different numbers of weeks in each month and for the changing number of weeks in February because of leap years. For example, a report of employment in a February Basic survey counts as 4 weeks worked (except in leap years, when it counts as 4.143 weeks); a report of employment in April counts as 4.286 weeks; and one of employment in January counts as 4.429 weeks.

Table 3 presents probit regressions relating the probability of a logical inconsistency (as defined in the previous section) to various demographic covariates, to the number of months reported working in the Basic CPS, and to a time trend. The table also presents OLS regressions relating the error gap ΔS to those same variables. A key implication of the regressions is that recall errors are not random. For example, the probit derivatives reported in the first two columns show that logical inconsistencies are trending up and that there are sharp demographic differences in them. Women are more likely than men to make these errors, and the same is true for each racial/ethnic group compared to the baseline group of non-Hispanic whites. Error rates fall with additional education and as workers age. The differences by race/ethnicity and age are not small, implying an excess of about 20 percent in the error rate over the base group. Because the probit in column (2) holds constant the number of months reported working in the Basic CPS, we can conclude that these partial effects are not driven by demographic differences in labor-force attachment.

The last two columns of Table 3 report the coefficients from regressions relating the error gap ΔS to the same regressors. The average value of ΔS is negative (and equal to -0.73), indicating that people in the ASEC under-report how many weeks they actually worked in the previous year by almost a week. Note, however, that the size of this error gap is trending down (becoming less negative) over the period. On average, women grossly understate their weeks worked, by about 2.2 weeks relative to men (after holding constant the intensity of labor force attachment). Racial/ethnic differences in the error gap are small, as are differences by age (except that workers in their late 50s greatly understate their previous year's weeks worked). The other large demographic difference in ΔS is across education groups. College graduates overstate their weeks worked by much more than workers with some college, who in turn overstate it more than high-school graduates, who overstate it more than high-school dropouts.[12]

[12] The panel data used in the regressions include the matched respondents in the Basic CPS in year *t*-1 and the ASEC in year *t* who are in the second half of the CPS interview cycle. If we calculate S^B for all Basic CPS respondents in year *t*-1 and calculate S^A for all ASEC respondents in year *t*, the average value of the gap ΔS equals -0.66 from 1978-2018. In short, the average difference between the two measures of weeks worked in the panel data closely approximates the difference found in the entire sample. The 1994 CPS re-design might have altered the time paths of the error gap ΔS. To evaluate this possibility, we re-estimated all the regressions reported in Table 3 by adding an indicator for years 1994 and later. None of the estimates showed a qualitatively important difference from those listed in the table.

IV. The Impact of Recall Error on Wage Differentials and Inequality

An important reason for imputing the different estimate of work time given by S^B—an estimate that adjusts for logical inconsistencies and probable errors in the ASEC recall of weeks worked—is to discover how these measurement problems affect calculations of weekly earnings and modify our understanding of the determinants and evolution of individual earnings. The discussion suggests two possible ways to define the weekly wage:

$$w^A = \frac{Y}{S^A} \qquad \text{and} \qquad w^B = \frac{Y}{S^B}, \tag{2}$$

where Y is the worker's annual wage and salary income in the ASEC (IPUMS variable *incwage*); w^A gives the weekly wage implied by the recall measure of weeks worked reported in the ASEC (S^A); and w^B gives the weekly wage implied by the measure of work time imputed from the Basic work history (S^B).

It is instructive to illustrate the bias induced by recall error on the coefficients of the regressors in a typical (weekly) earnings function. Consider the alternative regression models:

$$\log w^A = \beta^A x + \epsilon^A, \tag{3a}$$

$$\log w^B = \beta^B x + \epsilon^B. \tag{3b}$$

Equation (3a) gives the earnings function that is widely estimated in the literature using the ASEC recall data, while (3b) gives the analogous model using the weekly wage implied by the work history in the Basic monthly interviews. All variables in the regression models are expressed as deviations from their means.

We have shown that the ASEC measure of weeks worked is contaminated by error. Suppose the relationship between the two measures of weeks worked can be expressed as:

$$\log S^A = \log S^B + v, \tag{4}$$

where we are implicitly assuming that the imputed measure of weeks worked implied by the Basic interviews is the "true" measure of weeks worked and the random variable v is the recall error. In the simplest one-regressor framework, the OLS coefficients of the regression models that estimate the impact of x on the alternative measures of weeks worked are given by:

$$\hat{\beta}^A = \frac{\Sigma\, x \log w^A}{\Sigma\, x^2} \qquad \text{and} \qquad \hat{\beta}^B = \frac{\Sigma\, x \log w^B}{\Sigma\, x^2}. \tag{5}$$

It follows from (4) that there is a linear relationship between the two alternative measures of the (log) weekly wage: $\log w^A = \log w^B - v$. Equation (5) then implies a linear (and mechanical) relationship between the two alternative estimators of the impact of x on the weekly wage:

$$\hat{\beta}^A = \hat{\beta}^B - \hat{\beta}_{vx}, \tag{6}$$

where $\hat{\beta}_{vx}$ is the OLS coefficient from a regression that relates the (log) difference in the two alternative measures of weeks worked to the regressor x (i.e., $\hat{\beta}_{vx} = \Sigma\, xv/\Sigma\, x^2$). The analysis presented in Section III showed that the classical assumption that measurement error in the dependent variable is uncorrelated with the regressors in the model is demonstrably false in our context. In other words, $\hat{\beta}_{vx} \neq 0$, and recall error biases the coefficients of earnings functions estimated using the ASEC measure of the weekly wage. Equation (6) shows that this type of non-classical measurement error in the dependent variable yields a regression coefficient that subtracts the impact of x on measurement error from the true effect of the regressor x.

The direction of the bias depends on the sign of the correlation between x and recall error. For example, the regressions presented in Table 3 indicate that highly educated workers tend disproportionately to overstate the number of weeks they worked when recalling their work time in the past year (so that $\hat{\beta}_{vx} > 0$). Equation (6) then implies that an ASEC-based estimate of the rate of return to school would be biased downwards. Although equation (6) was derived in the case of a single regressor, the algebraic relationship between the alternative coefficients generalizes to a multivariate framework.[13]

[13] The vector of estimated coefficients in a multivariate regression that uses the ASEC measure of the weekly wage is $\hat{\beta}^A = (X'X)^{-1}X'z^A$, where $z^A = \log w^A$. Analogously, the estimated coefficient vector using the Basic measure is $\hat{\beta}^B = (X'X)^{-1}X'z^B$. Note, however, that $z^A = z^B - v$, which implies that $\hat{\beta}^A = \hat{\beta}^B - \hat{\beta}^v$, where $\hat{\beta}^v = (X'X)^{-1}X'v$.

The model can also be used to show how recall errors bias estimates of residual wage inequality. Consider the case of a single regressor x. Let $\log \hat{w}^k$ ($k = A, B$) be the predicted (log) weekly wage given by estimates of the coefficients in either (3a) or (3b), so that the residual from the regression is $\hat{\epsilon}^k = \log w^k - \log \hat{w}^k$. Equation (6) implies that the residual from the (log) earnings function that uses the ASEC weekly wage as the dependent variable is:

$$\hat{\epsilon}^A = \log w^A - \log \hat{w}^A = (\log w^B - v) - (\hat{\beta}^B - \hat{\beta}_{vx})x = (\log w^B - \hat{\beta}^B x) - (v - \hat{\beta}_{vx}x), \quad (7)$$

$$\hat{\epsilon}^A = \hat{\epsilon}^B - \hat{\epsilon}^v, \quad (8)$$

where $\hat{\epsilon}^v$ is the residual from a regression of the difference in the log of weeks worked in the ASEC and the Basic CPS on the regressor x. Equation (8) implies that the relation between the two alternative measures of residual wage residual inequality is given by:

$$Var(\hat{\epsilon}^A) = Var(\hat{\epsilon}^B) + Var(\hat{\epsilon}^v) - 2Cov(\hat{\epsilon}^B, \hat{\epsilon}^v). \quad (9)$$

Equation (9) shows that two distinct factors will create a difference between the amount of residual wage inequality estimated using the ASEC or the Basic weekly wage. First, the variance in recall error will necessarily produce greater residual wage inequality in regressions that use the ASEC measure of the weekly wage (w^A) as the dependent variable. Residual wage inequality measured in the ASEC, however, is also contaminated by the correlation between the adjusted recall errors ($\hat{\epsilon}^v$) and unobserved "skills" (as measured by $\hat{\epsilon}^B$). If this correlation is positive, so that workers with relatively high levels of unobserved skills tend disproportionately to overstate how many weeks they worked, the ASEC measure of residual wage inequality will underestimate the true measure.[14]

[14] The relationship between the alternative measures of residual wage inequality in (9) also holds when comparing estimates of the unadjusted variance of alternative measures of the log weekly wage. The predicted wage in a model

Equation (9) implies not only that the level of residual wage inequality depends on which measure of the weekly wage is used, but also that the different measures may yield different trends. For example, a secular reduction in the variance of recall errors or an increase in the positive correlation between unobserved skills and recall errors will lead unambiguously to a decline in measured wage inequality in the ASEC relative to the analogous measure in the Basic data. An increasing variance of recall errors and a declining positive correlation will produce the opposite effect on the gap between the two measures. We show below that both the variance in $\hat{\epsilon}^v$ and the correlation between $\hat{\epsilon}^B$ and $\hat{\epsilon}^v$ have trended down, creating different trends in inequality using the different measures of the weekly wage.

V. A Diversity of Data

Studies of female-male wage differentials, black-white wage differentials, and wage inequality in the U.S. typically use data drawn from the ASEC and/or the Outgoing Rotation Groups (ORG) of the Basic monthly CPS (which provides a measure of usual weekly earnings at the time of the survey). We focus on these widely used data sets, combined with the work information in the individual monthly interviews, in our examination of measurement issues.[15] We concentrate on the mismeasurement of work time and only tangentially consider how our estimates are affected by imputations of labor income in the ASEC (but see Bound and Krueger, 1991).

We would ideally take a particular sample of workers that has been widely used in the literature, replicate its results, and then determine if the measurement errors that we have identified change our understanding of the underlying determinants and trends in wages. This simple exercise, however, cannot be done, because there is huge variation among studies in the samples of workers being analyzed. Table 4 describes the specific samples used by various studies of female-male and black-white wage differentials,

---

where the only regressor is a constant is simply the mean (log) wage in the sample, implying that $Var(\log w^A) = Var(\log w^B) + Var(v) - 2Cov(\log w^B, v)$.

[15]Other studies (e.g., Thompson, 2021; and Hirsch and Winters, 2014) have used various National Longitudinal Surveys or data from the decennial Census (before 2000) or the American Community Survey (after 2000) to examine wage differentials.

and of wage inequality. Each study makes different choices about the age group analyzed and whether to use all workers or somehow define a subset of persons who might arguably be viewed as full-time year-round (FTYR) workers. It is evident that there is little agreement among studies as to which is the "preferred" sample. Regrettably, there is no equivalent of a genetically pure "lab rat" on which to study the impacts of choices about measuring labor inputs on inferences about wage differentials or inequality.

VI. The Evolution of Earnings Differentials Among All Workers

We use a sample of workers in the ASEC, modify the income variables to account for measurement errors in work time, and then compare the results across alternative specifications to isolate the bias produced by measurement error. We focus on three measures of the returns to work: The female/male wage gap, the black/white wage gap among men, and residual wage inequality. This section discusses the impact of measurement error on the level and trend in each of these outcomes in the sample of all workers. The next section discusses the impact of the errors in the restricted sample of full-time year-round workers.

We estimate the regression models using our two alternative definitions of weekly earnings, w^A and w^B. As shown earlier, some of the conflicts between the alternative measures of work time provided by the ASEC and Basic interviews arise because many respondents have imputed information. We also estimate a third specification that uses w^B as the measure of weekly earnings, but the sample includes only those respondents who did not have any imputed information on work time.[16] (Note that the excluded persons in this specification consist almost entirely of replicants with full-record imputations in the ASEC file).

We illustrate the impact of measurement error by reporting figures that show the different trends in the economic outcome of interest using the three alternative specifications. The difference between the ASEC- and Basic-based estimates thus isolates the bias produced by the measurement error issues we

[16] Throughout the study the sample in the earnings regressions is trimmed to exclude the top and bottom 1 percent of the weekly earnings distributions. All top-coded observations of earnings are multiplied by 1.5.

examine. As equation (6) shows, the (negative of the) difference in the coefficients between the ASEC and the preferred specification estimates the impact of the regressor of interest on the error v. All regressions are weighted by the March CPS sampling weights and include vectors of covariates indicating the worker's age, education, race/ethnicity, state of residence, and industry and occupation (at the three-digit level).[17]

A. The Female/Male Wage Gap

Figure 3a presents the trend in the (adjusted) female/male wage gap among all workers in the 1978-2018 period using the matched Basic/ASEC panels. As in Blau and Kahn (2017), regardless of the measure of weekly earnings used, the relative female wage rises nearly steadily throughout the sample period. Note that the relative wage of women is higher when using the ASEC measure of weeks worked (shown in Figure 3a by the **black solid line**) than when adjusting for the recall errors by using the imputed measure based on the person's contemporaneous work history and excluding replicants (shown by the thicker **red dashed line**), particularly in the early part of the sample period. Equation (6) then implies that the regression coefficient $\hat{\beta}_{vx}$ must be negative. In other words, the variable measuring the size of the recall error ($v = \log S^A - \log S^B$) is more negative among women (as suggested by Table 3). Recall error leads to an understatement of the size of the gender wage gap.

The estimated gender wage gap drops by 26.9 log-points over the sample period when using the ASEC measure of weekly earnings (comparing the average in the first three years, 1978-1980, to the average in the last three, 2016-2018). Using the Basic measure without replicants shows an even larger decline, 31.4 log-points (although replicants are only excluded beginning in 1990, since information on such allocations is not available for earlier years). Figure 3a thus shows that the gender wage gap among all workers narrowed substantially more over the past 40 years, by 4.5 log points, than is suggested by

[17] The general patterns shown in this section differ little if we use a smaller set of covariates or if we use the raw wage differentials. The earnings functions estimated in this section (and the next) use the same basic set of covariates as the probits on the error rate and the regressions on Δ reported in Table 3, but also include the worker's industry and occupation in the previous calendar year. The industry and occupation fixed effects were not included in the regressions estimated in the previous section because those variables are only available for ASEC respondents who recall working at some point in the previous calendar year.

simply considering earnings and retrospective reports of work time in the ASEC.[18] The exclusion of replicants in the analysis affects these conclusions in the most recent decade, perhaps because of the magnitude and the sharp increase in the fraction of full-record imputations.

That the difference between the ASEC and Basic estimates is narrowing over time implies that $\hat{\beta}_{vx}$, although negative, is trending toward zero. Put differently, gender is a weaker predictor of the recall error in the latter part of the period. In fact, the estimates imply that the $\hat{\beta}_{vx}$ is around -0.06 at the beginning of the sample period, but only about -0.01 by 2018. The declining correlation between gender and the size of the error gap is likely produced by the rapid growth in the fraction of women with a strong attachment to the labor force (i.e., working continuously through the year) during the period, and that workers with strong labor force attachments are, as we showed, far less likely than others to make recall errors.[19]

B. The Black/White Wage Gap

Because the annual samples of black working men in our merged Basic/ASEC CPS are relatively small, annual measures are noisy, making it difficult to infer the extent of any differences in the time paths of the racial wage gap using the alternative measures of work time. To mitigate the problem, we estimate the underlying earnings functions by pooling three years of data in our panels so that, in a sense, the relative wage trends are "averages" over each three-year period.[20] Figure 3b shows these estimates of the black/white wage gap among men (in log-points), fully adjusted for all covariates, and the differences between the racial gaps measured using the ASEC data and the alternative specifications.

Regardless of which measure of the weekly wage is used, Figure 3b depicts a general upward trend in the relative wage of black men, consistent with the evidence presented in Lang and Lehmann (2012) through 2010. Despite the differences in their levels, the ASEC measure and our preferred measure tell the

[18] This result is essentially unchanged if we estimate the gender wage gap in a sample of white non-Hispanic workers.

[19] The fraction of women who worked in all four Basic monthly interviews rose from 48.8 to 68.3 percent during the sample period, as compared to a decline from 87.9 to 79.6 percent among men.

[20] Because of the missing data for 1985 and 1995, some of the pooled data points cover four calendar years.

same stories about the trend in the male black/white wage gap: Both imply a reduction in the black/white wage gap of about 4.5 log points, although our preferred specification suggests that there was a slight reversal of this trend in the 2010s.

Figure 3b shows that the difference in the coefficients between the ASEC and the two alternative specifications diverged after 2000. This divergence is likely related to the fact that the ASEC records of a large (and growing) fraction of black workers are fully imputed. In the early 1990s, about 9 percent of white men had fully imputed records compared to about 14 percent of black workers. By 2018, almost 21 percent of white workers had fully imputed records, but the black fraction had risen to 25 percent. The high frequency of full imputation for black workers can easily contaminate many studies of wage trends in this population to the extent that the growing population of replicants differs increasingly from the black workers included in the surveys.

C. *Residual Wage Inequality*

We measure inequality as the standard deviation of the residual of the log weekly wage, after adjusting for all the covariates. Figure 3c presents trends in the various measures of residual wage inequality among all workers. It clearly shows that the *level* of residual inequality among all workers depends greatly on the measure of the weekly wage being used in the analysis. Residual inequality tends to be much lower (by between 5 and 10 log points) with the ASEC measure of the weekly wage than with the Basic measures.

Figure 3c also demonstrates that the alternative wage measures produce sharply different trends in residual inequality. Using the ASEC data alone, for example, suggests a slight *rise* in inequality of about 1 log-point over the period, while using the weekly earnings measure w^B that excludes replicants suggests a *drop* of about 5 log-points. The figure also shows that failing to remove replicants attenuates the decline in wage inequality. The crucial point is that the ASEC measure implies essentially no change in inequality, while our preferred measure shows a substantial decline.[21]

[21] Haltiwanger *et al*. (2023) offer a different perspective on alternative measures of earnings inequality, one not based on the measurement of labor inputs.

Equation (9) shows that the difference in the levels (and trends) in residual wage inequality depends on the (adjusted) variance in recall error and on the correlation between recall error and unobserved skills. It implies that a necessary condition for the ASEC measure of residual wage inequality to be smaller than the Basic measure is that there be a positive correlation between recall errors and skills (holding all covariates constant). We examine the empirical counterparts to the terms in Equation (9) in Figures 4a and 4b. There was a noticeable decline in the estimated dispersion of recall error, but this decline cannot produce the slight rise in the ASEC measure of residual wage inequality. There was also a sizable decline in the correlation between unobservable skills and recall error, from 0.59 to 0.44, and this decline could account for the relative rise in the ASEC measure of inequality and for the sharp difference between it and our preferred measure.[22]

VII. Wage Differentials among Full-Time, Year-Round Workers

The previous section examined how measurement errors in weeks worked affect wage trends in the sample of all workers. Some of the attention in empirical studies, however, has been devoted to the subsample of workers who work full-time year-round (FTYR) somehow defined. By inference, this literature is attempting to measure differences in pay per unit of time worked among workers who are implicitly assumed to be working the same amount.[23] As Table 4 showed, however, there is a great deal of

[22] One might be concerned that the age range 25-59 is too narrow and produces misleading estimates of the outcomes considered here. Re-estimating the models in this section including all individuals in the panel who are ages 16-64, as is common in much of the literature, produces nearly identical results on the sizes of the trends in each of the measures shown in Figures 3a-3c and of the differences in the trends. Another possible concern is the extent to which the inclusion of proxy responses, whose incidence is ubiquitous and changing, as Figure 2b showed, alters our inferences about these outcomes. Excluding proxy responses changes the levels of the curves in Figures 3, but the differences in the levels of the three measures and in their trends remains extremely close to those shown.

[23] This inference is clearly incorrect: FTYR workers may differ in the number of hours they work, and Kuhn and Lozano (2008) show that the excess above 35 hours rose over much of the period we examine. To account for these changes, in unreported estimates we include weekly workhours as an additional regressor among the covariates used to adjust the raw wage differentials in the ASEC sample. Doing so does not alter any of the conclusions reported in the text, since the changes in hours and the response of labor income to additional hours beyond 35 are similar by gender and by race. It is also worth noting that the literature ignores the selection bias produced by estimating the determinants of earnings in the highly non-random sample of persons that compose the FTYR workforce.

dispersion in the literature on how much work time, especially along the dimension of weeks of work, actually defines FTYR status.

Recall error produces a measure of weeks that can differ substantially from the number reported contemporaneously; but it will also contaminate measures of weekly hours of work. Unlike weeks worked, however, we cannot easily demonstrate that there are logical inconsistencies in reports of "usual" weekly hours in the Basic and the ASEC surveys. For instance, a person who reports usually working full-time (35+ hours) in all four Basic monthly surveys, but states in the ASEC that last year's usual workweek was less than 35 hours, is not necessarily logically inconsistent (since s/he could have been working full-time in the other eight months), although the report is probably in error. Similarly, reporting in the ASEC that the usual workweek last year was at least 35 hours, but stating in all four Basic CPS interviews that it was less than full-time, is also probably incorrect, but again not logically inconsistent.

The identification of unambiguous errors in work hours is complicated by the fact that the Basic monthly surveys do not provide a measure of (total) usual hours worked weekly throughout the 1978-2018 sample period. The variable reporting the total number of usual hours worked weekly is available only beginning with the 1994 redesign of the CPS. Throughout the entire sample period, however, the Census Bureau uses all the information contained in the array of work time questions to construct a "full-time status" variable (IPUMS variable *wkstat*) for each respondent in the Basic monthly surveys.[24] This variable specifically indicates if the person's work schedule is "usually full time."

We first use the post-1994 surveys to demonstrate how the recalled number of hours in the ASEC differs from the contemporaneous measures of usual hours worked reported in the Basic monthly interviews, and how that difference contaminates the classification of FTYR status. Specifically, we define a person to be a full-time (FT) worker in the ASEC if s/he recalls a usual workweek of at least 35 hours per week in calendar year t-1. Our merged Basic/ASEC data file provides information on how many usual hours

[24] The coding of the *wkstat* variable also changed with the 1994 CPS redesign. Before 1994 a single code (*wkstat* = 10) indicates if the worker is in a "full-time schedule." After 1993 the usual full-time schedule is captured by various codes of the variable (*wkstat* = 11, 12, 13, or 21).

that person actually worked in four months of that year. We can thus use the (post-1994) Basic data to create an alternative FT status variable based on contemporaneous reports of work time. This alternative variable indicates if the worker usually worked more than 35 hours per week in *all* four Basic survey appearances in year *t*-1.

Panel A of Table 5 reports the error rate in the FT classification, defined as the fraction of workers for whom the two alternative variables provide contradictory information about FT status. The error rate is shockingly high: There is conflicting information on full-time status for 22.5 percent of all workers in the 1994-2018 data. Note, for instance, that 11 percent of those reporting that they did not work in *any* of their four Basic appearances nonetheless still insist in the ASEC that they usually worked full-time hours in the previous year. If we exclude replicants (and other work-time related imputations), the error rates on FT work are only slightly reduced (Panel B of Table 5).[25]

The definition of FTYR status uses information on both weeks worked and hours worked, and Panel C shows the error rate in alternative definitions of this status (for the entire 1978-2018 sample period and thus not comparable to the statistics in Panels A and B). Our ASEC definition of FTYR status requires a person to have worked 52 weeks in the previous year with a usual workweek of at least 35 hours.[26] The Basic definition of FTYR status requires that a person worked in all four monthly interviews *and* that the Census Bureau classified the person as having a full-time schedule in each of those interviews. We then define the error rate as the frequency of contradictions between the alternative FTYR classifications. As Table 5 shows, the error rate in FTYR classification is about 14 percent, about the same for men and women.

Figure 5 illustrates trends in this error rate. Among women the error rate increased substantially during the sample period, from about 12 percent in 1978 to nearly 15 percent by 2018, while the rate among

[25] Specifically, the error rates reported in Panel B exclude replicants, respondents whose recalled usual hours of work in the ASEC is allocated, and respondents whose total usual hours of work in the Basic CPS are allocated.

[26] Some of the studies of FTYR workers listed in Table 4 base FT work on 50+ weeks in year *t*-1, slightly less restrictive than our definition. Unsurprisingly our conclusions about differences between the series and their changes over time differ only minutely between the two definitions. We prefer using 52 weeks, as this definition is more likely to generate better measures of the wage rate—the monetary return per unit of work time. Moreover, ASEC respondents are instructed to include paid vacation time as part of their annual weeks worked.

men varied less, averaging around 15 percent. The key implications of Figure 5 are that there are significant measurement problems inherent in constructing a subsample of full-time, year-round workers; that the potential misclassifications into FTYR status are not random; and that both the levels and trends of important economic variables may be affected as a result.

To assess the sensitivity of wage levels and trends to the composition of the sample of FTYR workers, we construct three alternative samples that can conceivably represent this workforce. The first simply uses the ASEC definition and includes those who claim to have worked 52 weeks and who report a usual workweek consisting of 35+ hours in year *t*-1. Because the work time information provided by the ASEC and Basic CPS files often differs, a second definition of the FTYR workforce includes those workers who can be classified as FTYR based on the (recalled) work time reported in the ASEC *and* who also report working a full-time workweek in each of their four appearances in the Basic monthly surveys. In other words, this alternative sample, which we denote as the ASEC/Basic FTYR sample, combines the contemporaneous year *t*-1 work history from the Basic interviews with the recalled information from the ASEC to create a subsample of workers who are most likely to have truly worked full-time each month in year *t*-1.

Finally, some studies in the literature examine wage trends in the FTYR workforce using the weekly earnings data available in the ORG (that is, workers in their 4th or 8th monthly interview). We restrict the analysis to workers who were interviewed in the Basic monthly files four separate times in the calendar year of their ORG appearance. We then combine the work time information provided across all four interviews to create the FTYR sample of ORG workers. In particular, the ORG FTYR sample consists of workers in the outgoing rotation group (where weekly earnings are reported) who worked at least 35 hours in each of their four appearances in the Basic monthly files in that year. Note that this is a more restrictive definition of FTYR status than is typically used in studies of ORG wage trends, but it is more consistent with the definition of FTYR status in studies that use ASEC data.[27]

[27] Over the 40-year period, the ASEC FTYR definition classifies 75.2 percent of workers as FTYR, the ASEC/Basic definition classifies 68.1 percent, but the ORG definition classifies 79.2 percent.

Workers in the ASEC and the ASEC/Basic FTYR samples both work 52 weeks a year, so that differences in the levels and trends of annual earnings across groups arise only because of the changing sample composition, since the ASEC measure of annual earnings is the numerator for both. Calculations using the ORG sample may differ both because of the differing definition of FTYR status and because the wage measure differs.

We cannot be sure a priori about how the characteristics of the samples differ. In fact, there are important demographic differences. Females comprise 41.4, 42.5, and 39.8 percent of the ASEC, ORG, and ASEC/Basic FTYR groups respectively; blacks comprise 9.9, 10.6, and 9.5 percent; and the fractions non-Hispanic white are 74.2, 72.5, and 75.3. In short, the demographic characteristics of the ORG FTYR sample differ slightly along observable dimensions, especially that of ethnicity, from that of both the ASEC and ASEC/Basic; and the directions of the differences suggests that the strictest definition of the FTYR workforce selects workers with higher earnings, at least along observable dimensions.

These differences in observable demographic characteristics among the samples suggest that they may differ in unobservable skills, and these unobservable differences may produce different wage trends when comparing the ASEC and the ASEC/Basic FTYR samples because the latter is a nonrandom subsample of the former. The generic earnings function that estimates a wage gap is:

$$\log w = \beta x + \epsilon \,. \tag{10}$$

The ASEC/Basic FTYR subsample truncates the ASEC sample. To retrieve the coefficient β from a wage regression estimated in this subsample, we would need to account for the non-random selection:

$$E[\log w \mid \text{ASEC/Basic FTYR}] = \beta x + \gamma\lambda \,, \tag{11}$$

where we have assumed that the errors are jointly normally distributed and λ is the inverse Mills ratio (Heckman, 1979). If we included the inverse Mills ratio in the earnings function estimated using the ASEC/Basic FTYR subsample, we would retrieve the same estimate of β as if equation (10) were estimated using the ASEC FTYR sample. This insight provides valuable information about how the measured impact of x varies when the regression is estimated over a subsample of workers who are more likely to represent the FTYR workforce correctly.

Suppose $\gamma > 0$, so that the ASEC/Basic FTYR subsample disproportionately contains those workers who have higher earnings, consistent with what is suggested by the differences in observable characteristics. In other words, only the most unobservably productive workers are selected into the sample with the strictest definition of FTYR status. A regression of (log) annual earnings on x in the ASEC/Basic FTYR subsample that ignores the selection problem is mis-specified, and:

$$\text{plim}\,\hat{\beta} = \beta + \gamma\, Corr(x, \lambda). \tag{12}$$

The assumption that workers who are most likely to be in FTYR status are positively selected implies that the direction of the bias in the estimate of β depends on the sign of the correlation between x and the inverse Mills ratio. Note, however, that λ is a negative function of being selected into the ASEC/Basic FTYR subsample. For example, if the unobservable characteristics of women are less productive than those of men, making them less likely to have true FTYR status, then $Corr(x, \lambda) > 0$. This means that our estimate of the gender wage gap in the restricted sample will be more positive (will be smaller) than the estimate using the ASEC definition.

That the ASEC/Basic FTYR subsample truncates the ASEC FTYR subsample also has implications for measures of residual wage inequality. In particular, the one-sided truncation implied by positive selection will necessarily reduce measured residual wage inequality in the truncated sample (Burdett, 1996). In short, the measurement issues apparent in properly defining FTYR status can lead to biased inferences about both the levels and trends in important economic variables.

A. Gender and Racial Wage Differentials

As in the previous section, the estimated gender and racial wage gaps adjust for the array of covariates (age, education, state of residence, industry, and occupation). Figure 6a indicates that the time path of the gender wage differential using the preferred ASEC/Basic FTYR subsample shows a narrowing of 18 log-points, two log-points less than the narrowing implied by the standard definition of FTYR status using the ASEC information. By the end of the sample period, however, there was essentially no difference among the three measures of this wage differential. The major difference depicted in Figure 6a is that the preferred measure of the gender wage gap shows more narrowing from 1982-84 (the first years for which the ORG measure is available) than the ORG measure, 2 log-points. The implied narrowing is even greater when the ASEC FTYR definition is used. The crucial point here is that even among workers who can be presumably classified as FTYR, the different data sets, and the different restrictions on work time, yield somewhat different insights about levels and trends.

The comparisons of the calculations of the black/white earnings differential differ only slightly in the ASEC and our preferred sample, as Figure 6b (again based on three-year pooled data) shows. Both show a substantial reduction in the wage disadvantage of FTYR black working men over this nearly 40-year period, with the gains of the 1980s and 1990s remaining constant after 2000. The preferred sample does, however, show a slightly smaller net reduction in the wage gap over the 40 years.

The ORG data tell a very different story: Over the 36-year period for which all three series are available, while the ASEC and ASEC/Basic samples show *reductions* in the differential of 5.4 and 4.0 log-points respectively, the ORG data show an *increase* in the differential of 5.9 log-points. This huge discrepancy demonstrates the difficulty in drawing conclusions about this essential labor market outcome. We cannot tell whether it arises because of differences in the definition of FTYR or differences in the measures of earnings per time period.[28]

[28] These figures include workers whose wages were imputed using the hot-deck procedure, which has been shown to matter in other contexts (Lillard *et al*., 1986; Hirsch and Schumacher, 2004; Bollinger *et al*., 2019). We estimated various additional models which exclude workers with wage imputations. Deleting such workers changes nothing about the impact of mismeasured work time on the comparisons of wages, nor does it alter the comparisons to the

B. Residual Wage Inequality

As shown in Figure 6c, both the ASEC sample and our preferred more restricted sample show increases in wage inequality among FTYR workers of around 4.5 percentage points over the entire period. The ORG data, however, suggest a much larger increase in inequality over the 36 years, 9.3 percentage points. Indeed, as Figure 6c shows, less inequality was implied in the ORG data when they initially became available than in the other series, but by 2018 those data suggested greater inequality than our preferred sample (or the ASEC).

The astute reader will note that the differences between the ASEC and the ASEC/Basic FTYR measures in this section are much smaller than those calculated in the previous section over the sample of all workers. The reason is clear: FTYR workers work 52 weeks; and, as we showed in Section II, the measurement errors arise disproportionately among those workers who respond that they are not working in all four Basic CPS interviews.

VIII. Conclusions and Implications: The Role of Errors

There are many logical inconsistencies in the sample of people included in the Annual Social and Economic Supplement (ASEC) of the CPS when we compare their recall of the previous year's work time with their contemporaneous reports of work time during their earlier appearances in the Basic CPS monthly interviews. These inconsistencies (and the additional recall errors made by respondents whose recall and concurrent work histories are not necessarily inconsistent) suggest that care needs to be taken in measuring levels and trends in earnings per unit of time worked in the data file that serves as the main source of wage and employment information in the United States. We showed that these measurement errors are sizable and non-random both in the sample of all workers and in the subsample of workers that is typically classified

ORG measures, which themselves change little if we exclude workers whose employment status or earnings are imputed. The difference does not stem from issues in the imputations of earnings in either the ASEC or the ORG data The discrepancy that we have demonstrated results from differences in the measurement of weeks and hours of work.

as working full-time, year-round. We also demonstrated that the measurement problem affects trends in estimates of such key labor market outcomes as the gender wage gap and wage inequality.

We have shown substantially different results for all workers and for full-time year-round (FTYR) workers across various sample restrictions, with the recall errors mattering much more in comparisons among all workers than among samples of FTYR workers defined very strictly. Our preferred definition of FTYR status is appropriately more stringent than that used in any previous study, thus providing a closer approximation to the ideal calculation of a "wage rate" that measures pay per unit of work time. Indeed, one might argue that some studies examining FTYR workers define the concept so broadly as to approximate samples of all workers.

Table 6 brings together the results reported and discussed in Sections VI and VII. It shows the 1978-2018 change (1982-2018 for FTYR workers) in the measure of a particular economic outcome (a wage differential or residual wage inequality) using both the measure of work time recalled in the ASEC and the preferred measures of work time. The changes and levels in the variables are all listed in log-points, and averages are based on three years of data.

Accounting for the inconsistencies and errors that we documented in the ASEC work time data substantially affects what we know about trends and levels of wage discrimination and inequality among all workers. In particular, the gender wage gap among all workers narrowed by more than was previously thought. Similarly, the use of contemporaneously measured work histories suggests that wage inequality was characterized by a decline rather than the slight increase suggested by the ASEC recall data. Also, the extent of measured gender discrimination among FTYR workers was less over the last 40 years than previously thought (although today it is the same); and wage inequality among them was and is less than previously measured.

These observations are specific to the U.S. labor market and the time period we consider. They arise from changing labor force behavior, particularly the sharp increase in women's attachment to the labor force. The general point, however, applies to any economy and any time when the labor-force behavior of various demographic groups is changing and when economists measure wage differentials and inequality using data that are likely to be affected by recall errors and biases.

It is important to note that, while our analysis only addresses measurement issues in the CPS-ASEC data, similar problems exist with other data files that are widely used to draw inferences about labor market dynamics in the United States. The micro files of the decennial Census of Population (through 2000) and the annual American Community Surveys (after 2000) collect the same type of recall information on weeks and hours worked last year as the ASEC, and these measures of work time have been used in numerous studies as divisors to calculate a person's weekly earnings or hourly wage rate. When using the Census-type data, however, there is no way of correcting for the recall error, as there is no independent information on a person's prior work history that can be used to "adjust" the recall measures of weeks worked (although one could use the extraneous information generated in this study to adjust the findings). Our results suggest that much of the reported evidence that relies on recall measures of work time to infer wages per time period may need to be re-examined, and that researchers studying the U.S. labor market need to think more carefully about errors in measuring work time when generating their research results.

REFERENCES

Akerlof, George, and Janet Yellen, 1985. Unemployment through the Filter of Memory. *Quarterly Journal of Economics* 100 (August): 747-73.

Autor, David, Lawrence Katz, and Melissa Kearney, 2008. Trends in U.S. Wage Inequality: Revising the Revisionists. *Review of Economics and Statistics* 90 (May): 300-23.

Barrett, Gary, and Daniel Hamermesh. 2019. Labor Supply Elasticities: Overcoming Nonclassical Measurement Error Using More Accurate Hours Data. *Journal of Human Resources* 54 (Winter): 255-65.

Bee, Adam, and Jonathan Rothbaum, 2023. Using administrative data to evaluate nonresponse bias in the 2023 Current Population Survey Annual Social and Economic Supplement. Unpublished Memo, U.S. Bureau of the Census, September 12.

Bick, Alexander, Bettina Brüggemann, and Nicola Fuchs-Schündeln. 2020. Hours Worked in Europe and the United States: New Data, New Answers. *Scandinavian Journal of Economics* 121 (October): 1381-1416.

Blau, Francine, and Lawrence Kahn. 2017. The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature* 17 (September): 789-865.

Bollinger, Christopher, Barry Hirsch, Charles Hokayem, and James Ziliak. 2019. Trouble in the Tails? What We Know about Earnings Nonresponse 30 Years after Lillard, Smith, and Welch. *Journal of Political Economy* 127 (October): 2143-85.

Borjas, George. 1980. The Relationship between Wages and Weekly Hours of Work: The Role of Division Bias. *Journal of Human Resources* 15 (Summer): 409-23.

Bound, John, and Alan Krueger. 1991. The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right? *Journal of Labor Economics* 9 (July): 1-24.

Bound, John, Charles Brown, Greg Duncan, and Willard Rodgers. 1994. Evidence on the Validity of Cross-sectional and Longitudinal Labor Market Data on Earnings. *Journal of Labor Economics* 12 (July): 345-68.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. Measurement Error in Survey Data. In *Handbook of Labor Economics, Vol. 5*, Orley Ashenfelter and Richard Layard, eds., 3705-3843.

Burda, Michael, Daniel Hamermesh, and Jay Stewart. 2013. Cyclical Variation in Labor Hours and Productivity Using the ATUS. *American Economic Review, Papers and Proceedings* 2013 103 (May): 99-104.

Burda, Michael, Katie Genadek, and Daniel Hamermesh. 2020. Unemployment and Effort at Work. *Economica*, 87 (July): 662-81.

Burdett, Kenneth. 1996. Truncated Means and Variances. *Economics Letters* 52 (September): 263-267.

Burnette, Jeffrey. 2017. Inequality in the Labor Market for Native American Women and the Great Recession. American Economic Association, *Papers and Proceedings*, 107 (May): 425-9.

Celhay, Pablo, Bruce Meyer, and Nikolas Mittag, 2021. Errors in Reporting and Imputation of Government Benefits and their Implications. NBER Working Paper No. 29184.

Charles, Kerwin, and Jonathan Guryan. 2008. Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination. *Journal of Political Economy* 116 (October): 773-809.

Florence, P. Sargant. 1924. *Economics of Fatigue and Unrest*. New York, Henry Holt.

Frazis, Harley, and Jay Stewart. 2010. Why Do BLS Hours Series Tell Different Stories About Trends in Hours Worked? In Katharine Abraham and James Spletzer, eds., *Labor in the New Economy*. Chicago, University of Chicago Press, pp. 343-72.

Haltiwanger, John, Henry Hyatt, and James Spletzer. 2023. Increasing Earnings Inequality: Reconciling Evidence from Survey and Administrative Data. *Journal of Labor Economics* 41 (October): S61-S93.

Heckman, James. 1979. Sample Selection Bias as a Specification Error. *Econometrica* 47 (January): 153-61.

Heffetz, Ori, and Daniel Reeves. 2019. Difficulty of Reaching Respondents and Nonresponse Bias: Evidence from Large Government Surveys. *Review of Economics and Statistics* 101 (March): 176-91.

Hirsch, Barry, and Edward Schumacher. 2004. Match Bias in Wage Gap Estimates Due to Earnings Imputation. *Journal of Labor Economics* 22 (July): 689-722.

Hirsch, Barry, and John Winters. 2014. An Anatomy of Racial and Ethnic Trends in Male Earnings in the U.S. *Review of Income and Wealth* 60 (December): 930-47.

Hoffmann, Florian, David Lee, and Thomas Lemieux. 2020. Growing Income Inequality in the United States and Other Advanced Economies. *Journal of Economic Perspectives* 34 (Fall): 52-78.

Hubbard, William. 2011. The Phantom Gender Difference in the College Wage Premium. *Journal of Human Resources* 46 (Summer): 568-86.

Juhn, Chinhui. 2003. Labor Market Dropouts and Trends in the Wages of Black and White Men. *Industrial and Labor Relations Review* 56 (July): 643-62.

---------------, Kevin Murphy, and Brooks Pierce. 1993. Wage Inequality and the Rise in Returns to Skill. *Journal of Political Economy* 101 (June): 410-42.

Kuhn, Peter, and Fernando Lozano, 2008. The Expanding Workweek? Understanding Trends in Long Work Hours among U.S. Men, 1979–2006. *Journal of Labor Economics* 26 (April): 311-43.

Lang, Kevin, and Jee-Yeon K. Lehmann. 2011. Racial Discrimination in the Labor Market: Theory and Empirics. *Journal of Economic Literature* 50 (December): 959-1006.

Lemieux, Thomas. 2006. Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill? *American Economic Review* 96 (June): 461-98.

Lillard, Lee, James Smith, and Finis Welch. 1986. What Do We Really Know about Wages? The Importance of Nonreporting and Census Imputation. *Journal of Political Economy*, 86 (June): 489-506.

Mulligan, Casey, and Yona Rubinstein. 2008. Selection, Investment, and Women's Relative Wages over Time. *Quarterly Journal of Economics* 123 (August 2008): 1061-1110.

Rodgers, Willard, Charles Brown, and Greg Duncan. 1993. Errors in Survey Reports of Earnings, Hours Worked, and Hourly Wages. *Journal of the American Statistical Association* 88 (December): 1208-18.

Stewart, Jay. 2007. Using March CPS Data to Analyze Labor Market Transitions. *Journal of Economic and Social Measurement* 32 (February): 177-197.

Sudman, Seymour, and Norman Bradburn. 1973. “Effects of Time and Memory Factors on Response in Surveys.” *Journal of the American Statistical Association* 68 (344): 805–15.

Thompson, Owen, 2021. Human Capital and Black-White Earnings Gaps, 1966-2017. NBER Working Paper No. 28586.

Western, Bruce, and Becky Pettit. 2005. Black-White Wage Inequality, Employment Rates, and Incarceration. *American Journal of Sociology* 111 (September): 553-78.

Wilson, Valerie, and William Rodgers. 2016. Black-White Wage Gaps Expand with Rising Wage Inequality. Employment Policy Institute.

Table 1. CPS Rotation Groups and the Timing of Interviews
(workers in second half of CPS interview cycle)

	Interview Number in ASEC and Year								
	5		**6**		**7**		**8**		
Month	*t*-1	*t*	*t*-1	*T*	*t*-1	*T*	*t*-2	*t*-1	*t*
January					✓	✓		✓	✓
February			✓	✓	✓	✓		✓	✓
March[a]	✓	**X**	✓	**X**	✓	**X**		✓	**X**
April	✓		✓		✓				
May	✓		✓						
June	✓								
July									
August									
September									
October									
November									
December							✓	✓	

[a] ASEC appearance in year *t* denoted by a bold **X**.

Table 2. Incidence of Recall Error in ASEC Measure of Weeks Worked (%)

	Months reported as working in Basic CPS					
	0	1	2	3	4	0-4 months
A. Percent distribution of respondents:						
All	17.6	2.9	2.3	4.3	72.9	100.0
Male	9.5	2.2	2.1	4.0	82.3	100.0
Female	25.0	3.5	2.5	4.6	64.4	100.0
B. Assumes employed in entire interview month						
All persons	6.1	55.1	52.7	54.5	1.8	7.5
Male	12.2	51.8	50.2	54.4	1.2	6.4
Female	3.9	57.0	54.6	54.6	2.4	8.4
C. Assumes employed only in reference week						
All persons	5.1	49.3	43.5	47.8	1.2	6.2
Male	10.6	45.3	41.1	48.2	0.8	5.4
Female	3.1	51.6	45.3	47.4	1.7	7.0
D. Excludes replicants						
All persons	5.4	53.7	52.0	54.4	1.3	6.5
Male	11.5	49.3	49.2	53.9	0.9	5.9
Female	2.8	56.4	54.3	54.8	1.7	7.0
E. Excludes replicants and proxy responses						
All persons	2.3	48.2	46.3	48.8	1.1	5.2
Male	2.8	42.6	45.9	49.2	0.7	4.4
Female	2.1	50.7	46.5	48.6	1.3	5.8

Notes: Recall errors are clear contradictions between a person's concurrent work history in the four Basic monthly interviews and the ASEC (recall) measure of weeks worked in the previous calendar year. Panel A assumes that the work status of the respondent in the Basic reference week applies to the entire interview month. There is a contradiction if a person who did not work at all in any of the four Basic interviews claims in the ASEC to have worked more than 36 weeks; if a person who worked 1 month claims to have worked fewer than 4 or more than 48 weeks; if a person who worked 2 months claims to have worked fewer than 8 or more than 44 weeks; if a person who worked 3 months claims to have worked fewer than 12 or more than 48 weeks; or if a person who worked all 4 months claims to have worked fewer than 16 weeks. Panel D assumes that the work status of the respondent in the reference week applies only to that week. The error rates reported in Panels B and C use the 1990-2018 sample period (i.e., neither the replicant variable nor the proxy response information are available in the panel data prior to 1990). The "replicants" excluded in Panel C are people whose ASEC record was fully allocated, as well as persons whose weeks worked variables in the ASEC was allocated or whose employment status in the Basic was allocated. The observations with a proxy response are those where someone in the household other than the respondent answered the questions in at least one of the CPS questionnaires used in the analysis.

Table 3. Demographic Determinants of Differences in the Error Rate and ΔS, 1978-2018

Variable:	Error Rate		$\Delta S = S^A - S^B$	
	(1)	(2)	(3)	(4)
Year	0.0010	0.0009	0.010	0.006
	(0.0001)	(0.0002)	(0.001)	(0.001)
Female	0.023	0.010	-0.733	-1.993
	(0.001)	(0.001)	(0.028)	(0.029)
Race: Black	0.033	0.017	0.324	-0.030
	(0.001)	(0.001)	(0.059)	(0.056)
Asian	0.021	0.014	0.338	0.065
	(0.003)	(0.002)	(0.112)	(0.109)
Hispanic	0.020	0.009	0.243	0.194
	(0.001)	(0.001)	(0.058)	(0.056)
Other Race	0.021	0.012	0.547	0.164
	(0.002)	(0.002)	(0.095)	(0.091)
Education: High School	-0.021	-0.002	0.052	1.224
	(0.001)	(0.001)	(0.054)	(0.053)
Some College	-0.035	-0.007	0.070	1.670
	(0.001)	(0.001)	(0.058)	(0.056)
>=College	-0.053	-0.013	-0.201	1.874
	(0.001)	(0.001)	(0.055)	(0.055)
Age: 30-34	-0.015	-0.005	-0.155	-0.041
	(0.002)	(0.001)	(0.066)	(0.063)
35-39	-0.024	-0.007	-0.091	0.106
	(0.002)	(0.001)	(0.063)	(0.064)
40-44	-0.026	-0.006	-0.142	0.141
	(0.002)	(0.001)	(0.063)	(0.060)
45-49	-0.030	-0.006	-0.187	0.201
	(0.002)	(0.001)	(0.063)	(0.060)
50-54	-0.029	-0.006	-0.312	-0.171
	(0.002)	(0.001)	(0.064)	(0.061)
55-59	-0.030	-0.006	-0.500	-.861
	(0.002)	(0.001)	(0.065)	(0.063)
Months worked in Basic: 0 months	---	0.032	---	7.436
		(0.001)		(0.043)
1-3 months	---	0.194	---	4.762
		(0.001)		(0.083)
Pseudo- or Adj. R^2	0.023	0.390	0.002	0.070
Mean	0.075	0.075	-0.731	-0.731

* Robust standard errors in parentheses below the parameter estimates. The regressions are weighted by the March CPS sampling weight. The coefficients of the determinants of error rates are probit derivatives. The coefficients of the determinants of the error gap ΔS are from least-squares regressions. All regressions include state of residence fixed effects and have 803,771 observations.

Table 4. Sample Definitions Used in Studies of U.S. Wage Discrimination and Inequality

Study	Age Range	Sample
A. Female/Male Wage Differential		
Blau-Kahn (2017)	25-64	ASEC, 26+ weeks, 35+ usual hours
Mulligan-Rubenstein (2008)	25-54	ASEC, 40+ weeks, 35+ usual hours
Hubbard (2011)	18-65	ASEC, 50+ weeks, 35+ usual hours
B. Black/White Wage Differential		
Lang-Lehmann (2012)	20+	ASEC, 50+ weeks, 35+ hours
Juhn (2003)	1-30 years of experience	ASEC, all
Western-Pettit (2005)	22-64	ORG, all
Charles-Guryan (2008)	16-64	ORG, all, weighted by usual hours
Wilson-Rodgers (2016)	All	ORG, all
Burnette (2017)	16+	ORG, all
C. Residual Wage Inequality		
Juhn *et al*. (1993)	18-65	ASEC, all, 35+ divide by weeks×usual hours
Lemieux (2006)	16-64	ASEC, all weighted by weeks×usual hours
		ORG, all, weighted by usual hours
Autor et al. (2009)	16-64	ASEC, 40+ weeks, 35+ usual hours
	16-64	ORG, all, weighted by usual hours
Hoffmann *et al*. (2020)	25-64	ASEC, 40+ weeks, 40+ usual hours

Table 5. Misclassifications of FT and FTYR Status, by Work History in the Basic CPS

	A. Months worked in Basic CPS					
	0	1	2	3	4	0-4 months
A. Contradictory FT Status in ASEC and Basic (%), 1994-2018						
All persons	11.0	41.9	56.2	65.0	20.9	22.5
Male	20.1	58.0	72.4	79.1	22.4	26.2
Female	6.9	31.4	42.9	53.1	19.3	19.0
B. Same as Panel A, but excludes replicants						
All persons	9.1	39.9	54.9	64.4	17.2	19.0
Male	17.5	57.0	72.2	79.7	18.5	22.5
Female	5.4	29.3	41.1	51.9	15.8	17.2
C. Contradictory FTYR Status in ASEC and Basic (%), 1978-2018						
All persons	3.9	11.7	18.9	29.2	15.8	14.2
Male	9.6	17.2	24.8	36.6	13.8	14.6
Female	1.8	8.6	14.4	23.4	18.1	13.9

* Panels A and B: Full-time (FT) status in the ASEC requires usually working at least 35 hours per week in the previous calendar year; FT status in the Basic requires usually working at least 35 hours per week in each of the four appearances in the monthly interviews. The calculations use the 1994-2018 period. The "replicants" excluded in Panel B include persons whose ASEC record was fully allocated, as well as persons whose hours worked weekly information was allocated in either the ASEC or Basic files (specifically, usual hours worked weekly last year in the ASEC and total usual hours worked weekly in the Basic).

* Panel C: A worker is classified as full-time year-round (FTYR) in the ASEC if s/he reports working 52 weeks in the previous calendar year and a usual workweek of at least 35 hours that year. A person is classified as FTYR in the Basic if the person worked in all four monthly interviews and is classified as having a full-time schedule by the Census Bureau in each of the interviews. The calculations use the entire 1978-2018 sample.

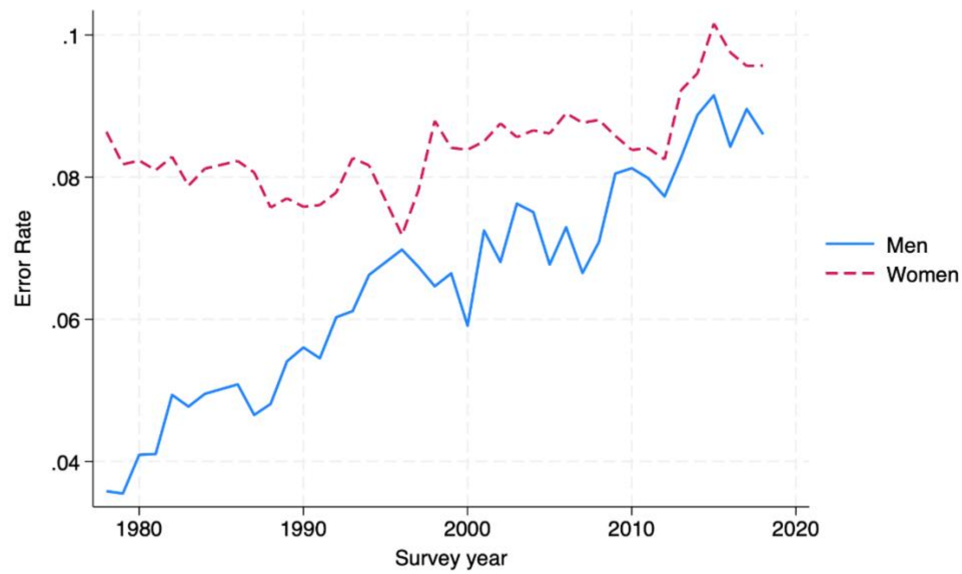
Table 6. Summary of Results on Wage Differentials and Inequality

	All workers, 1978-80 –2016-18		FTYR workers, 1982-84 – 2016-18		
	ASEC	Basic Non-allocated	ASEC	ASEC/Basic	ORG
Female/Male Log Wage Gap					
Early	-0.509	-0.562	-0.323	-0.307	-0.276
2016-2018	-0.240	-0.248	-0.193	-0.194	-0.180
Change	+0.269	+0.314	+0.130	+0.113	+0.096
Black/White Log Wage Gap, Men					
Early	-0.171	-0.185	-0.148	-0.140	-0.118
2016-2018	-0.126	-0.138	-0.094	-0.100	-0.177
Change	+0.045	+0.047	+0.054	+0.040	-0.059
Residual Wage Inequality					
Early	0.501	0.603	0.402	0.387	0.373
2016-2018	0.514	0.551	0.444	0.434	0.466
Change	+0.013	-0.052	+0.042	+0.047	+0.093

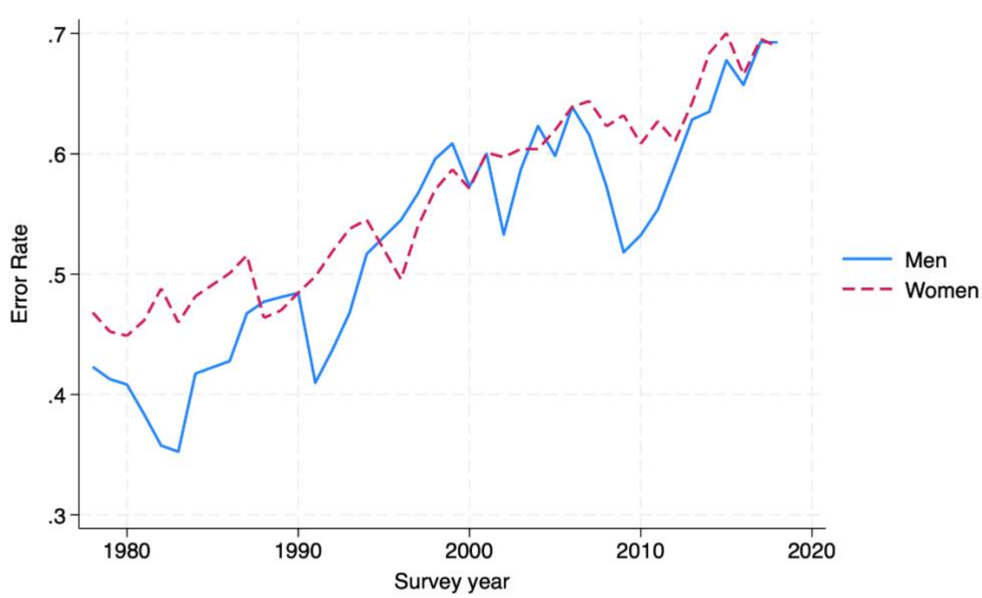
Source: Computed from the coefficients underlying Figures 3a-3c and 6a-6c. Among all workers the changes are calculated from 1978-80 to 2016-18. Among FTYR workers the change is from 1982-84 to 2016-18.

Figure 1. Trends in the Error Rate in Weeks Worked, 1978-2018, by Gender

a. All Workers



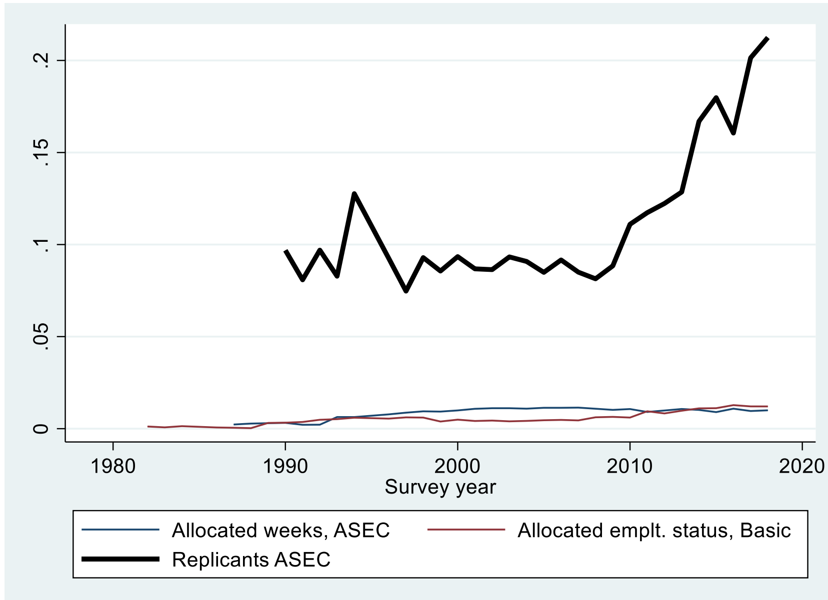
b. Worked 1 to 3 Months in Year *t*-1 in the Basic CPS



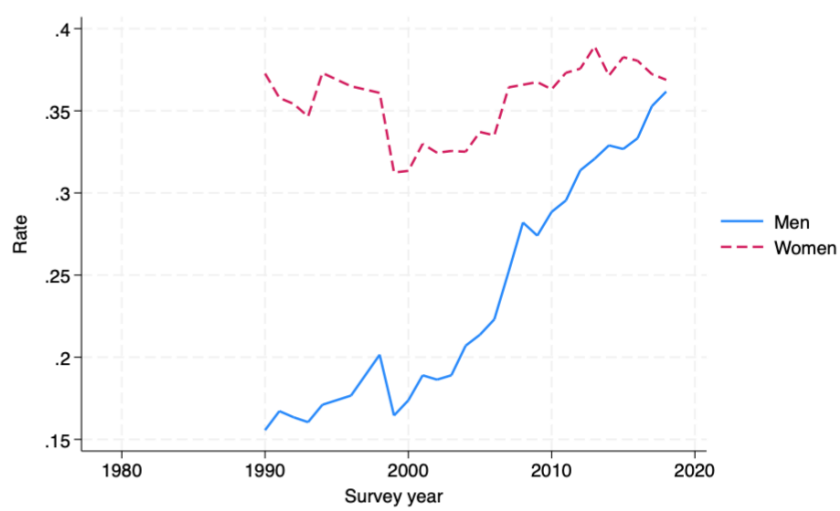
Note: The error rate measures the relative number of logical inconsistencies between the recall measure of weeks worked in the ASEC and the contemporaneously recorded work history of the person in the Basic monthly interviews; see Table 2 for details.

Figure 2. Frequency of Allocation and Proxy Responses, by Gender

a. Frequency of allocated responses



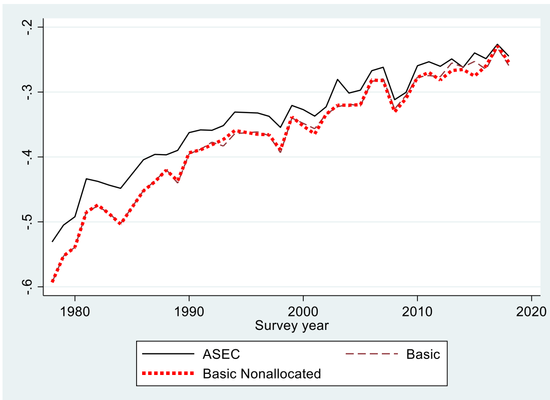
b. Frequency of Non-Proxy Responses



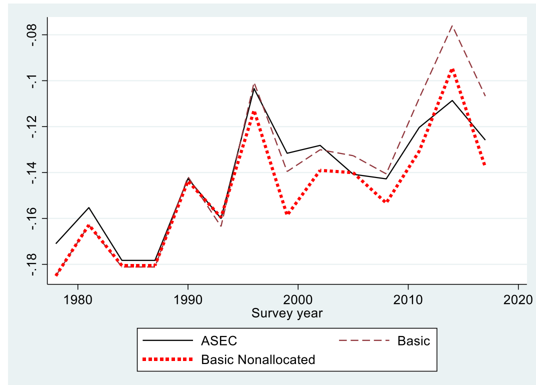
Note: Panel A gives the rate of allocated responses in the ASEC and Basic merged data panel. Panel B gives the fraction of observations where the respondent self-answered the questionnaire in all five interviews (i.e., the four Basic monthly interviews and the ASEC interview.

Figure 3. Wages with Alternative Measures of Weeks Worked as Divisors, All Workers (Adjusted for Covariates)

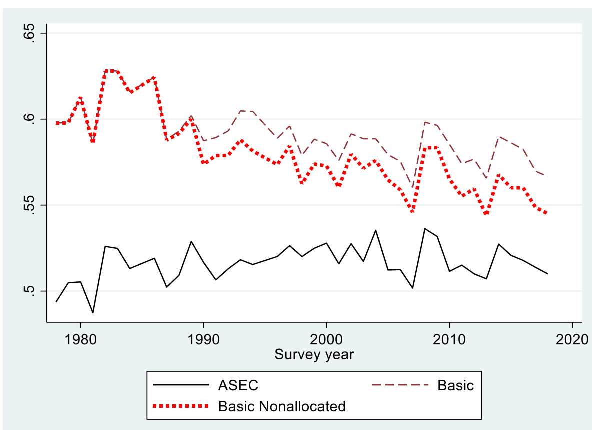
a. log (Female wage/Male wage)



b. log Male (Black wage/White wage)



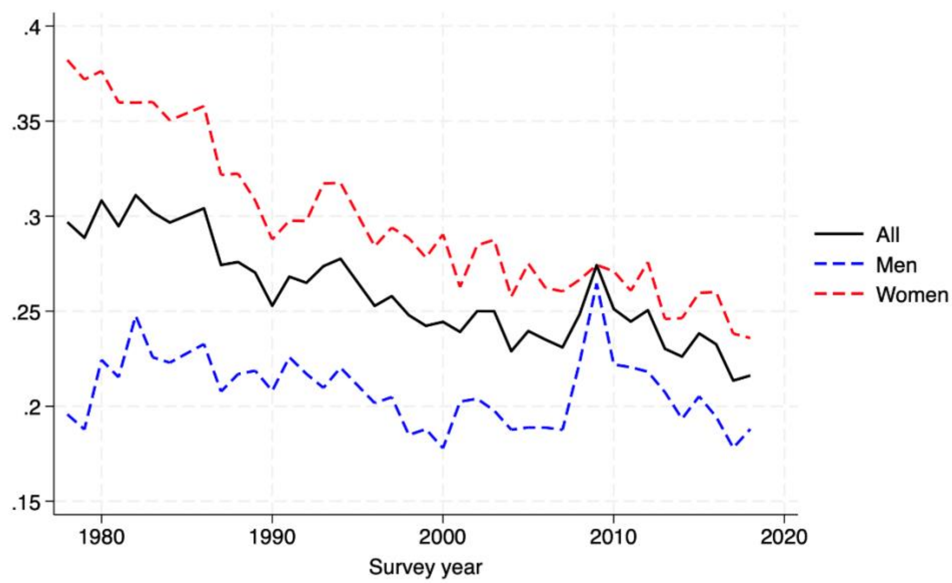
c. Standard Deviation log (Weekly Wage)



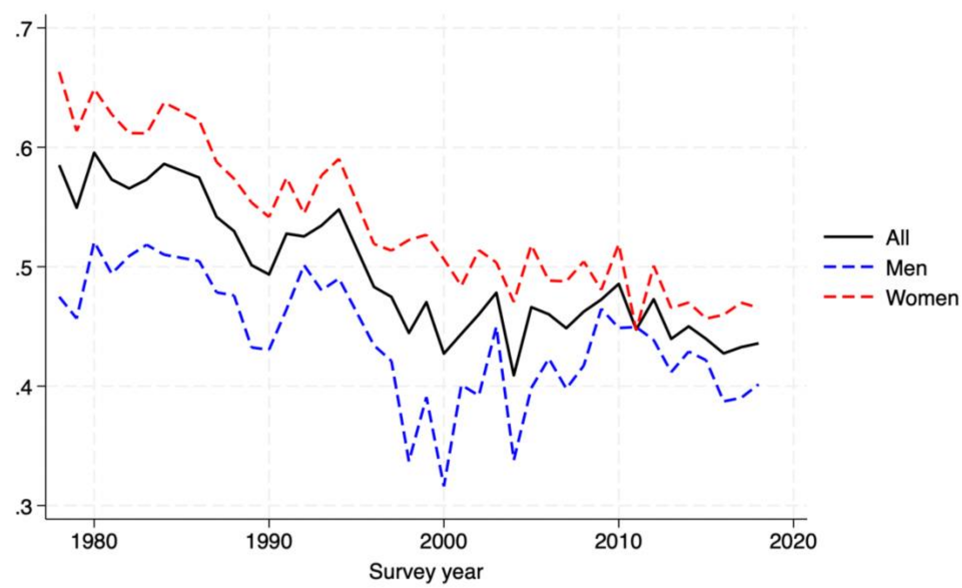
Notes: The ASEC measure of the weekly wage is the ratio of wage income to the number of (recalled) weeks worked. The divisor in the Basic measure is the imputed number of weeks implied by the Basic work history. The "Basic Non-allocated" measure is the same, but excluding all CPS respondents whose information was allocated.

Figure 4. Correlation between Recall Error and Unobserved Skills, and the Standard Deviation of Recall Error (Adjusted for Covariates), By Gender

a. Trends in Standard Deviation of $\hat{\epsilon}^{v}$

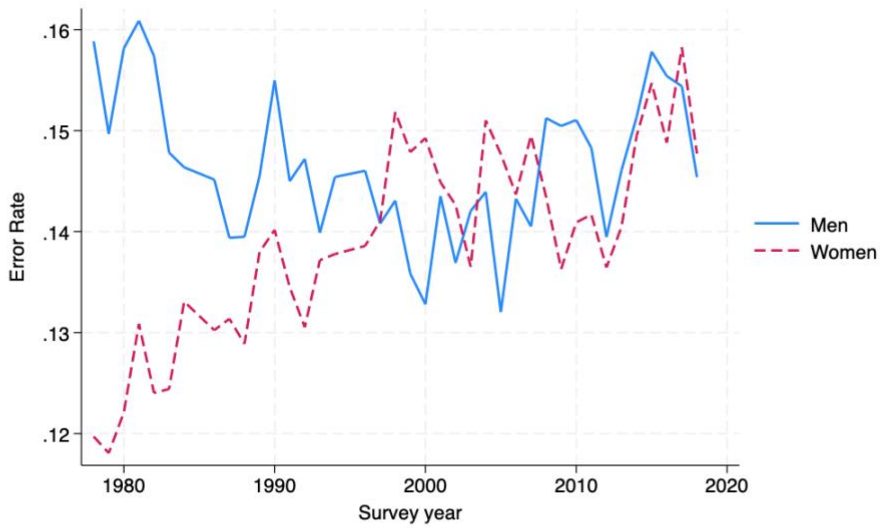


b. Trends in $Corr(\hat{\epsilon}^{B}, \hat{\epsilon}^{v})$



Notes: $\hat{\epsilon}^{B}$ is the residual from a regression of the Basic measure of the (log) weekly wage on all covariates. $\hat{\epsilon}^{v}$ is the residual from a regression of log weeks worked in the ASEC on imputed log weeks worked in the Basic (also including all covariates). The regressions are weighted by the March CPS sampling weights.

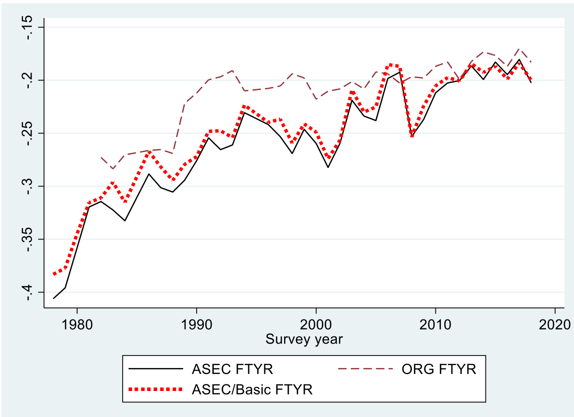
Figure 5. Trends in the Error Rate in Assigning Full-time Year-Round Status, 1978-2018, by Gender



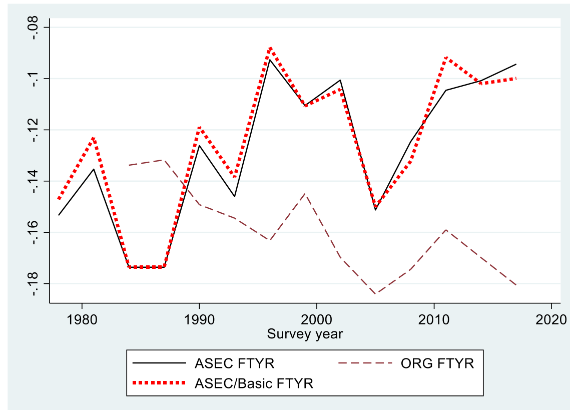
Note: The error rate is defined as the fraction of persons who have different classifications of their FTYR status in the ASEC and Basic data. The ASEC classification of FTYR status requires a person to work 52 weeks in the previous year with a usual workweek of at least 35 hours. The Basic classification requires the person to have worked all four monthly interviews and to have been classified by the Census Bureau as being in "full-time status" in each of the four months.

Figure 6. Wage Differentials for Alternative Definitions of the FTYR Workforce (Adjusted for Covariates)

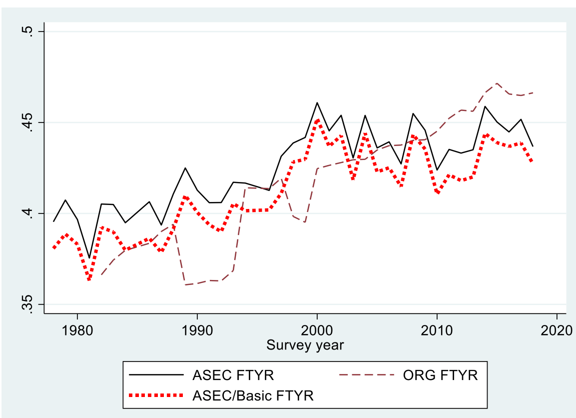
a. log (Female wage/Male wage)



b. log Male (Black wage/White wage)



c. Standard Deviation log (Weekly Wage)



Note: The ASEC sample consists of workers who reported working 52 weeks and a usual workweek of at least 35 hours of work in the previous calendar year. The Basic sample consists of persons who worked in all four monthly interviews and reported full-time status for weekly hours in each of the interviews. The "ASEC/Basic FTYR" sample excludes ASEC FTYR respondents who did not work 35+ hours in each Basic CPS appearance.