



10-2019

Trouble in the Tails? What We Know about Earnings Nonresponse 30 Years after Lillard, Smith, and Welch

Christopher R. Bollinger
University of Kentucky, crboll@email.uky.edu

Barry T. Hirsch
Georgia State University

Charles M. Hokayem
U.S. Census Bureau

James P. Ziliak
University of Kentucky, jziliak@email.uky.edu

Follow this and additional works at: https://uknowledge.uky.edu/economics_facpub



Part of the [Economics Commons](#), and the [Political Science Commons](#)

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Repository Citation

Bollinger, Christopher R.; Hirsch, Barry T.; Hokayem, Charles M.; and Ziliak, James P., "Trouble in the Tails? What We Know about Earnings Nonresponse 30 Years after Lillard, Smith, and Welch" (2019). *Economics Faculty Publications*. 7.

https://uknowledge.uky.edu/economics_facpub/7

This Article is brought to you for free and open access by the Economics at UKnowledge. It has been accepted for inclusion in Economics Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

Trouble in the Tails? What We Know about Earnings Nonresponse 30 Years after Lillard, Smith, and Welch

Notes/Citation Information

Published in *Journal of Political Economy*, v. 127, no. 5.

© 2019 by The University of Chicago. All rights reserved.

The copyright holder has granted the permission to post the article here.

Digital Object Identifier (DOI)

<https://doi.org/10.1086/701807>

Trouble in the Tails? What We Know about Earnings Nonresponse 30 Years after Lillard, Smith, and Welch

Christopher R. Bollinger

University of Kentucky

Barry T. Hirsch

Georgia State University and Institute of Labor Economics (IZA)

Charles M. Hokayem

United States Census Bureau

James P. Ziliak

University of Kentucky

Earnings nonresponse in household surveys is widespread, yet there is limited knowledge of how nonresponse biases earnings measures. We examine the consequences of nonresponse on earnings gaps and inequality using Current Population Survey individual records linked to administrative earnings data. The common assumption that earnings are missing at random is rejected. Nonresponse across the earnings distribution is U-shaped, highest in the left and right tails. Inequality measures differ between household and administrative data due in part to nonresponse. Nonresponse biases earnings differentials by race, gender, and education, particularly in the tails. Flexible copula-based models can account for nonrandom nonresponse.

We thank Adam Bee, Dan Black, Charlie Brown, James Heckman, Bruce Meyer, Chuck Nelson, Trudi Renwick, James Spletzer, Ed Welniak, and five anonymous reviewers for helpful comments, plus participants at presentations at the United States Census Bureau, Soci-

Electronically published July 17, 2019

[*Journal of Political Economy*, 2019, vol. 127, no. 5]

© 2019 by The University of Chicago. All rights reserved. 0022-3808/2019/12705-0004\$10.00

I. Introduction

Thirty-plus years ago, Lillard, Smith, and Welch (1986) brought to the forefront the issue of earnings nonresponse in the Current Population Survey (CPS), providing a sharp critique of Census Bureau (hereafter, Census) imputation procedures. Since that time much has changed, some for the better and some not. Census responded to the critique of Lillard et al. and substantially improved the quality of its imputation procedures. For the Annual Social and Economic Supplement (ASEC; known historically as the March supplement), Census uses a sequential hot-deck procedure to address item nonresponse for missing earnings data by assigning individuals with missing earnings values that come from individuals (“donors”) with similar characteristics.¹ Less well known is that in addition to item nonresponse, there exists ASEC supplement nonresponse. This occurs when households participating and responding in the monthly CPS refuse to participate in the ASEC supplement. Census also uses a hot-deck procedure for whole-supplement nonresponse. Offsetting the progress in data processing, however, were sharply rising rates of earnings nonresponse. As depicted in figure 1, there was a substantial increase in the 1990s and then again after 2011, such that by 2015 total (item and whole) nonresponse in the ASEC reached 43 percent.² The item nonresponse rate of 25 percent is more than double that at the time of the critique of Lillard et al. Additionally, the CPS monthly outgoing rotation group files have earnings-item nonresponse rates currently above 35 percent, while the much larger American Community Survey (ACS) has earnings nonresponse rates of about 20 percent, suggesting that nonresponse is pervasive across the most important federal household surveys.

Unfortunately, we know surprisingly little about patterns of earnings nonresponse, or its potential consequences for important labor-market issues such as earnings gaps by gender and race, or inequality. Lillard et al.

ety of Labor Economists, Joint Statistical Meetings, American Economic Association Meetings, Brigham Young University, Emory University, Federal Reserve Bank of Cleveland, Federal Reserve Board, Institute for Fiscal Studies, University of Essex, University of New South Wales, and University of Sydney. The analysis provided in this paper has been conducted at the United States Census Bureau, the Atlanta Research Data Center, and the Kentucky Research Data Center. All results shown in the paper have received clearance from the Census Bureau. The opinions and conclusions are solely those of the authors. Programs are provided as supplementary material online at the journal’s website.

¹ Welniak (1990) documents changes over time in Census hot-deck methods for the CPS ASEC.

² For a careful analysis, see Bee, Gathright, and Meyer (2015). An additional form of nonresponse is so-called unit nonresponse, which occurs when there is a noninterview or refusal to participate even in the monthly CPS survey. These rates for the basic CPS were between 8 and 9 percent during our sample period (Dixon 2012). Also, as a point of comparison, nonresponse rates for typical labor supply variables (weeks worked or hours per week) were in the 3 to 5 percent range over the past two decades.

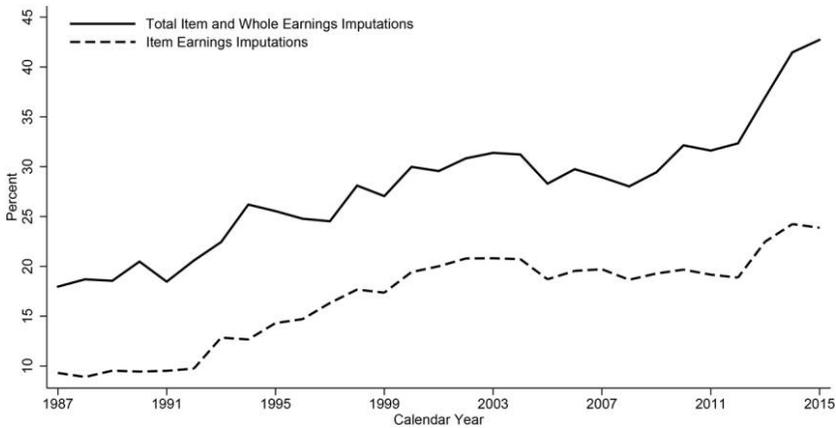


FIG. 1.—Trends in item and total earnings imputations in the ASEC. This figure displays trends in item earnings and total (item + whole supplement) imputations in the ASEC among workers. The imputation rate is weighted using the ASEC supplement weight. Sources: Authors' calculations and the US Census Bureau Current Population Survey, 1988–2016 Annual Social and Economic Supplement.

(1986, 492) suggested that ASEC nonresponse is likely to be highest in the tails of the distribution, but provided limited evidence since they could not observe earnings for nonrespondents. Lillard et al. (1986, 493, table 2) place white men in eight earnings intervals based on a combination of reported and predicted earnings. They find a U-shaped nonresponse pattern with respect to earnings, with the highest rates in the top three earnings categories.

Whether and to what extent earnings nonresponse is of economic consequence depends on the questions being addressed and the reasons for nonresponse. Prior research has shown that use of imputed earnings can seriously bias average wage gap estimates studied widely by social science researchers (Hirsch and Schumacher 2004; Bollinger and Hirsch 2006; Heckman and LaFontaine 2006), even if the earnings data are missing completely at random.³ If earnings are missing completely at random,

³ Hirsch and Schumacher (2004) labeled this “match bias,” which occurs most notably in earnings equations in which the dependent variable includes nonrespondents whose earnings are assigned (imputed) based on a set of match attributes. If one includes attributes on the right-hand side that are not matched in the earnings imputation process, coefficients on these attributes are attenuated. Hirsch and Schumacher used union status as an example; nonrespondent union members are mostly assigned nonunion earnings, while nonunion workers are sometimes assigned the earnings of a union member. Heckman and LaFontaine (2006) focus on bias to GED estimated returns; nonrespondent GED holders are rarely assigned the earnings of a GED donor. Bollinger and Hirsch (2006) expand the analysis to forms of imperfect matching (e.g., education or age) and derive precise measures of match bias.

then nonresponse is not dependent on earnings, even absent covariates; if earnings are missing at random (MAR), then nonresponse is not dependent on earnings after conditioning on covariates; and if earnings are not missing at random (NMAR), then nonresponse is dependent on the value of missing earnings even after conditioning on covariates (Rubin 1976; Little and Rubin 2002). It is this latter case that is referred to as having nonresponse bias or nonignorable nonresponse. Both Census imputation procedures and common inverse-probability weighting methods to deal with nonresponse assume that nonresponse is ignorable; that is, those not reporting earnings have earnings similar to respondents with equivalent measured attributes. If the MAR assumption is violated, measures of earnings gaps and distributions will be biased.

Given the high earnings nonresponse rates in Census household surveys, coupled with a paucity of evidence on nonresponse patterns and its consequences, we address three important and closely related questions. We do so using restricted-access ASEC household files linked to administrative tax data from the Social Security Administration Detailed Earnings Record (DER) for March 2006–11 (corresponding to calendar years 2005–10). First, how does earnings nonresponse vary across the earnings distribution? Access to the DER is uniquely advantageous to address this question as it affords the opportunity to fill in missing earnings for nonrespondents, and to compare survey responses to administrative tax records for respondents. We align each worker's ASEC earnings response status against the worker's corresponding earnings level from the DER. We examine this relationship for men and women separately, as well as for full-time, full-year workers and those whose ASEC reports are provided by a proxy (i.e., another household member). Nearly half of ASEC earnings reports are from proxies. The extent to which proxy reporting affects nonresponse patterns and earnings accuracy is not well understood.

The second question we address is whether nonresponse is ignorable. That is, do respondents and nonrespondents have equivalent conditional earnings distributions, and if so, can the earnings of survey respondents accurately describe the missing counterfactual distribution of a combined respondent and nonrespondent sample as if the nonrespondents had responded? This question directly addresses the efficacy of the MAR assumption used in Census imputation procedures and, more broadly, in many related missing-data procedures. MAR relies on the assumption that the joint distribution of earnings and response status, conditional on covariates, can be expressed as the product of the conditional marginal distribution of response status and the conditional marginal distribution of earnings. This leads to our two complementary tests of MAR made possible with access to administrative data, one that examines whether the decision to respond to the ASEC earnings question is independent of

earnings, and a second that examines whether the distribution of earnings is independent of response status. Furthermore, we also estimate parametric and nonparametric earnings regressions using the DER, and then test differences in the residuals from those regressions based on response status. This provides estimates of summary statistics for the conditional distribution of earnings for both respondents and nonrespondents. Absent the link to the DER these tests are not possible because of missing earnings of ASEC nonrespondents.

The third question is whether earnings nonresponse affects standard estimates of earnings gaps (by gender, race, and ethnicity), returns to schooling, and earnings inequality and volatility. To address this question, we estimate saturated quantile earnings models to test how nonresponse affects outcomes in the tails of the distribution, alongside models of central tendency. In addition, we also present estimates of how nonresponse impacts standard measures of inequality such as the Gini and 90-10 (and 90-50 and 50-10) ratios and top income shares. For the volatility analysis, we exploit the longitudinal dimension of the ASEC whereby it is possible to match up to half of the sample from one year to the next to examine both the dynamics of nonresponse and implications for summary measures of volatility. Answers to the inequality and volatility questions have taken on increasing importance in recent years with the expansion of distributional research, whether using standard summary measures of unconditional or conditional inequality (e.g., Piketty and Saez 2003; Lemieux 2006; Autor, Katz, and Kearney 2008; Burkhauser et al. 2012) or fully specified quantile regression models of earnings (Buchinsky 1994; Kline and Santos 2013; Arellano and Bonhomme 2017). Under MAR, unconditioned measures of inequality may differ between the full sample with imputations and a sample omitting imputed earners. The full sample is likely to provide a consistent estimate of unconditional inequality if the covariates used in the imputation procedure provide an unbiased measure of earnings and maintain variance. Using only respondents provides more accurate earnings responses, but risks bias (absent reweighting) to the extent that nonresponse rates differ across the earnings distribution, as we subsequently show. The full sample with imputations results in biased estimates of conditional inequality, however, because the relationship between inequality and the multivariate correlations with respect to demographic, geographical, and job attributes not used (or not used fully) in the imputation process will be biased (Hirsch and Schumacher 2004; Bollinger and Hirsch 2006). Data from the DER are particularly helpful here both because we can fill in the missing ASEC earnings with the DER and because, unlike the public-release and internal versions of ASEC, earnings from the DER are not top coded, which improves our estimates of the importance of nonresponse in the right tail of the earnings distribution.

We are not the first to examine nonresponse using a validation study or to find deviations from MAR, but prior studies are generally old, use small samples, and/or examine restricted populations (e.g., married white males). Most similar to our initial analysis is a paper by Greenlees, Reece, and Zieschang (1982), who examined the 1973 ASEC and compared wage-and-salary earnings the previous year with 1972 linked income tax records of full-time, full-year male heads of households in the private non-agricultural sector whose spouse did not work. They found evidence that selection into response declined weakly with respect to earnings. No distributional analysis was provided. David et al. (1986) conducted a related validation study using the 1981 ASEC linked to 1980 IRS reports, also finding evidence of negative selection into response. More recently, Kline and Santos (2013) examined whether returns to schooling and other earnings-equation parameters are sensitive to departures from the MAR assumption, using the exact match of the 1973 ASEC linked to IRS earnings data. They provided evidence that missing data probabilities among men are U-shaped, with very low and high wage men least likely to report. Hokayem, Bollinger, and Ziliak (2015) used the linked ASEC-DER data to examine how treatment of nonrespondents affects family poverty rate estimates and noted the U-shaped nonresponse pattern. Although informative and suggestive, it is not known whether results from studies that examined response bias using older data can be generalized outside their time period and narrow demographic samples. In short, there is little validation evidence using recent data to examine the extent and consequences of CPS nonresponse bias across the earnings distribution. Given the increasing rates of nonresponse over time, it is important to know whether nonresponse is ignorable and, if not, the size and patterns of bias.

In general, we find that nonresponse is not ignorable—earnings are not missing at random, even conditional on a rich set of covariates—and as we allude to in the title, the highest rates of nonresponse are in the tails of the earnings distribution. While on average, male (female) nonrespondents have slightly higher (lower) earnings than respondents, nonresponse is not simply an up or down shift in the distribution. Individuals with earnings that differ substantially from the average (either the gross or conditional mean) are the most likely not to report earnings. This U-shaped pattern is in evidence across gender, race, ethnicity, employment status (hourly and full time, full year), month-in-sample, proxy earnings status, and panel status (year 1 or year 2).

Our finding of NMAR suggests that reliance on respondent samples (even if reweighted) also may provide biased estimates of population earnings. While we find the impact of nonresponse bias on *averages* is small, the bias on conditional quantile estimates of gender, race, and education earnings gaps associated with very high or low earnings is upward of 20 percent and statistically significant. Moreover, between one-third and one-half

the difference in inequality measures between the CPS and administrative data is accounted for by nonresponse in the CPS. We conclude by demonstrating that public users of the ASEC can approximate the population distribution of earnings using a copula-based selection model recently proposed by Arellano and Bonhomme (2017).

II. Earnings Nonresponse and Response Bias

Official government statistics, as well as most research analyzing earnings (and income) differences, include both respondents and nonrespondents, replacing the missing earnings with an imputed value. Researchers typically assume (usually implicitly) that nonresponse does not produce systematic biases in the measurement of earnings. The aim of our paper is to determine whether this assumption is justified.

Formally the ignorability of missing earnings underlying the MAR assumption is a statement about the joint distribution of earnings (Y) and response status (R), conditional on covariates (X):

$$f(Y, R|X) = f(Y|X) * f(R|X), \quad (1)$$

which means that once we condition on known covariates, earnings and response status are independent. Because the Bayes theorem permits us to relate the joint distribution of (Y, R) to conditional distributions regardless of whether MAR holds, we can write the joint distribution as

$$f(Y, R|X) = f(Y|R, X) * f(R|X) = f(R|Y, X) * f(Y|X). \quad (2)$$

The implication of MAR is then readily seen by equating equations (1) and (2):

$$f(Y|R, X) = f(Y|X) \quad \text{and} \quad f(R|Y, X) = f(R|X).$$

If either of these conditions fails, then the MAR assumption fails. When the only data available are survey reports, one method for testing for the presence of nonresponse bias across the joint distribution in equation (2) is to treat response as a form of sample selection and to estimate a flexible quantile model via copula methods (Joe 2014; Arellano and Bonhomme 2017). Bollinger and Hirsch (2013) adopted a restrictive version of this approach by estimating the conditional mean of earnings controlling for selection via a standard two-step Heckman (1979) method. Below we demonstrate the efficacy of the copula-based method to recover the unbiased distribution of earnings using respondents only, but the main analysis in this study takes advantage of our access to linked administrative earnings for both respondents and nonrespondents, permitting direct tests of nonresponse bias via validation methods.

Specifically, because the MAR assumption is conditioned on covariates, it is sufficient to test MAR by focusing on the conditional distributions on the right-hand side of equation (2). Simply put, does response status depend on earnings or does the distribution of earnings depend on response status? For the former we estimate models of the form

$$\Pr(R_i = 1 | Y_i^{\text{DER}}, X_i) = F(\alpha + \gamma Y_i^{\text{DER}} + X_i \beta), \quad (3)$$

where Y_i^{DER} is administrative earnings reports from the DER described in the next section. Because of the very large number of covariates we restrict these tests to parametric estimators (probit and linear probability) so that a test that $f(R|Y, X) = f(R|X)$ amounts to a test of $\gamma = 0$. We consider specifications that control for Y_i^{DER} in both logarithmic form and flexible percentiles, as well as models that relax separability between Y_i^{DER} and X_i . Greenlees et al. (1982) and David et al. (1986) implemented tests along the lines of equation (3) as it provides the simplest and most straightforward way to answer the question of independence. Since R is a binary variable, its entire distribution is summarized by $\Pr(R = 1 | Y, X)$. If earnings have any predictive power, then earnings and response are not independent and the MAR assumption fails.

For the test of conditional independence of earnings from response we estimate both parametric and nonparametric models of the form

$$Y_i^{\text{DER}} = \delta + \theta R_i + X_i \pi + v_i. \quad (4)$$

Summary measures of $f(Y|R, X)$ are the key for understanding sample selection when Y is the dependent variable in a regression. Unlike $f(R|Y, X)$, the conditional distribution of earnings given response may be summarized by a variety of parameters (e.g., mean, median, quantiles, variance, and skewness) that describe different features of the distribution, such as center, spread, or symmetry. The classic paper by Heckman (1979) and later papers (for a survey, see Vella [1998]) suggest that a key parameter is $E[Y|R = 1, X]$, in which case the test that $f(Y|R, X) = f(Y|X)$ amounts to a test of $\theta = 0$. When the regression of interest is a quantile regression such as the median or other percentiles, it is less clear what the most important parameters will be. For the nonparametric models we estimate kernel density functions separately for respondents and nonrespondents and conduct Kolmogorov-Smirnov tests of the null that $f(Y|R = 1, X) = f(Y|R = 0, X)$. Rejecting the null of equality is a sufficient condition to reject the hypothesis that $f(Y|R, X) = f(Y|X)$.

III. Data: The ASEC-DER Link Files

The data used in our analysis are restricted-access CPS ASEC person records linked to the Social Security Administration Detailed Earnings Record

(DER) for survey years 2006–11 (reporting earnings for calendar years 2005–10).⁴ The ASEC is a survey of roughly 60,000 households (plus an additional 30,000 households as part of the Children’s Health Insurance Program) conducted in March of each year. It serves as the source of official federal statistics on income, poverty, inequality, and health insurance coverage, and has been the workhorse data set for earnings-inequality research in the United States. The primary difference between the internal ASEC we use and the version available publicly is that the internal file has higher top-code values on income components (Larrimore et al. 2008). We link the internal ASEC to the DER file, which is an extract of the Master Earnings File and includes data on total earnings as reported on a worker’s W-2 form, wages and salaries and income from self-employment subject to Federal Insurance Contributions Act and/or Self-Employment Contributions Act taxation, as well as deferred wage (tax) contributions to 401(k), 403(b), 408(k), 457(b), and 501(c) retirement and trust plans, all of which we include in our earnings measure. Only positive self-employment earnings are reported in the DER because individuals do not make self-employment tax contributions if they have self-employment losses (Nicholas and Wiseman 2009). In addition, some parts of gross compensation do not appear in the DER file such as pretax health insurance premiums and education benefits (Abowd and Stinson 2013), nor do “off the books” earnings appear in the DER, though they could be reported in the ASEC.⁵ Unlike the internal ASEC earnings records, DER earnings are not top coded.⁶ This is important given substantial concerns regarding nonresponse and response bias in the right tail of the distribution.

The principal sample used in our analysis includes civilian wage-and-salary workers ages 18 to 65 who have reported or imputed positive earnings in the prior year. We exclude workers who are full-time students, as well as a small number of workers identified in the ASEC and linked to the DER who show zero DER earnings but positive deferred compensation. We also exclude individuals with whole imputations of the ASEC, that is, those for whom all ASEC supplement data are imputed. We provide

⁴ The linked ASEC-DER files were obtained as part of an internal-to-Census project and analyzed in a secure facility at the US Census Bureau in Suitland, MD. Researchers outside the Census Bureau interested in accessing such data must have their project approved by the bureau and the Social Security Administration for analysis conducted in a secure Federal Statistical Research Data Center. For more information, see <https://www.census.gov/fsrdc>.

⁵ Whether survey reports of earnings differ from tax reports is an important, open issue. Recent evidence in Hurst, Li, and Pugsley (2014) suggests that among the self-employed, survey and tax reports do not differ substantively, but whether this holds for the general labor force is not established and should be the subject of future research.

⁶ Confidentiality protections preclude us from disclosing individual earnings values such as the maximum earnings values in the DER. The two components of our internal ASEC total earnings variable—earnings on the primary job and all other earnings—are each capped at \$1.1 million.

a separate analysis of this subsample in the appendix, which is available as an online supplement. The full sample, including those with no ASEC-DER link, consists of 508,288 individuals (270,409 men and 237,879 women).

Since a worker can appear multiple times per year in the DER file if they have multiple jobs, we collapse the DER file into one earnings observation per worker per year by aggregating total earnings (box 1 of the W-2, labeled “wages, tips, other compensation”), total self-employment earnings, and total deferred contributions across all employers. In this way, DER earnings are most compatible with ASEC earnings from all wage-and-salary jobs (WSAL-VAL) plus nonnegative self-employment earnings. We classify a worker as having imputed earnings if wage-and-salary income from the longest job (I-ERNVAL), from other jobs (I-WSVAL), or from self-employment earnings is imputed. For much of our analysis, we focus on annual earnings because this measure is available in both the ASEC and DER, but we also examine earnings among full-time, full-year workers, as well as average hourly earnings found by dividing annual ASEC or DER earnings by annual hours worked. Annual hours worked is constructed by multiplying weeks worked (WKSWORK) by the usual hours worked per week (HRSWK); these ASEC labor-market measures are available for earnings nonrespondents as well as respondents.

Table 1 provides summary statistics for our full sample in column 1, weighted by the ASEC person supplement weight. The average worker is 41 years old, slightly more likely to be a male, and has an average of nearly 14 years of education. The majority are married with spouse present (58 percent), native born (84 percent), and work full time, full year (71 percent). Nonresponse to either the wage-and-salary questions or the self-employment earnings question totals 23 percent of the sample. Nonresponse is concentrated on the wage-and-salary questions (22 percent) largely because relatively few individuals are self-employed. ASEC interviews identify for each household a single respondent who provides information about other members of the household; hence, 48 percent of the earnings responses are proxy responses, an issue we return to in a subsequent section. Inflation-adjusted ASEC total earnings are \$45,897, while average real DER earnings are a higher \$48,478.

Table 1 also presents descriptive statistics for the sample broken down by ASEC response status and by DER link status. In general, nonrespondents are not markedly different from the full sample (or the respondent sample). They are slightly less likely to be a Hispanic or a female, more likely to be never married, and more likely to be full-time, full-year workers. In both the ASEC and the DER measures, nonrespondents have slightly higher annual earnings. It is unsurprising that the ASEC difference is small since the imputed earnings derive from the earnings of the respondents. On average, 86 percent of the ASEC sample is successfully linked to the DER, though appendix figure 1 demonstrates that the linkage rate is considerably lower at low earnings, rising from about 72 percent

TABLE 1
 SAMPLE AVERAGES BY RESPONSE STATUS AND BY LINKAGE STATUS

	FULL SAMPLE (1)	RESPONSE STATUS		LINKAGE STATUS	
		Respondent (2)	Non- respondent (3)	Linked (4)	Nonlinked (5)
Age	41.4	41.3	41.8	41.7	39.6
Race/ethnicity:					
White, non-Hispanic	68.7	69.3	66.4	71.8	49.0
Black, non-Hispanic	10.6	9.9	13.0	10.6	10.9
Asian, non-Hispanic	4.5	4.3	5.2	4.2	6.2
Other race, non-Hispanic	1.9	1.9	1.7	1.9	1.7
Hispanic	14.4	14.5	13.7	11.5	32.2
Gender:					
Female	46.8	47.5	44.7	47.9	39.9
Education (years)	13.7	13.7	13.6	13.9	12.6
Marital status:					
Married, spouse present	57.6	58.5	54.3	58.9	49.4
Married, spouse absent	16.9	16.9	16.9	16.6	18.6
Single, never married	25.6	24.6	28.9	24.5	32.0
Nativity:					
Native	84.2	84.5	83.4	87.6	63.6
Foreign born, US citizen	6.5	6.3	7.2	6.3	7.7
Foreign born, not a US citizen	9.2	9.2	9.4	6.1	28.8
Employment:					
Full time, full year	71.2	70.6	73.4	72.3	64.3
Work hours (per week)	39.9	40.0	39.8	40.2	38.5
Nonresponse:					
Nonresponse rate (wages and salaries or self-employment income)	22.6	0	100	20.1	38.0
Nonresponse rate (wages and salaries)	22.4	0	99.1	19.9	37.9
Linkage rate	86.2	88.9	76.7	100	0
Proxy	48.0	45.1	58.1	47.2	53.1
ASEC total earnings (2010 dollars)	45,897	45,838	46,099	47,665	34,884
DER total earnings (2010 dollars)	48,478	47,895	50,796	48,478	NA
DER average hourly total earnings (2010 dollars)	25.61	25.45	28.70	26.10	NA
Observations	508,288	399,823	108,465	440,227	68,061

SOURCES.—US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

NOTE.—This table shows sample descriptive statistics for the full sample and broken down by ASEC response status and ASEC-DER linkage status. Full ASEC averages include imputed nonrespondent earnings. Each average is weighted by the ASEC supplement weight. For information on confidentiality protection, sampling error, nonsampling error, and definitions, see <https://www2.census.gov/programs-surveys/cps/techdocs/cpsmar17.pdf>.

to 92 percent across the ASEC earnings distribution. In table 1, the non-linked sample shows more striking differences from the full sample than the nonrespondent sample. Individuals for whom a link was not found are 2 years younger, 8 percentage points more likely to be male, and have

1.3 fewer years of education. Most notably they are more than twice as likely to be Hispanic, and over three times more likely to be foreign born and not a citizen. Nonlinked workers are almost twice as likely to be an earnings nonrespondent, and they report ASEC earnings nearly \$13,000 lower than those reported by linked workers.

In appendix tables 1 and 2 we document that link failure between ASEC and DER is concentrated among noncitizen immigrants. Because the opt-out rate to agree to link the ASEC and DER is a trivial 0.5 percent, most link failures are due to lack of personally identifiable information used in constructing a linkage indicator. To address this, we estimate a saturated probit model of the probability of an ASEC-DER link as a function of a full array of demographic characteristics, including nativity, Hispanic ethnicity, and their interaction (see appendix table 3). As described in the appendix, we then use the fitted values to construct inverse-probability weights (IPWs) to rebalance the ASEC-DER linked sample for the missing nonlink sample (i.e., the ratio of the ASEC weight to the fitted probability of a link). Because most of the linkage failures are not due to an opt-out choice by the respondent and instead are accounted for by observed demographics, we believe any potential bias from selection on unobservables, which would not be corrected by IPWs, is minimal.

IV. Is Response a Function of Earnings, and Are Earnings a Function of Response?

We begin our analysis by examining the conditional distribution of response given earnings, where in table 2 we present estimates of equation (3) using the linked ASEC-DER sample and both unweighted and IPW linear-probability models.⁷ In this first test, we begin by conditioning only on the logarithm of DER earnings. Columns 1 and 2 do not control for any confounders, while in columns 3 and 4 we control for a rich set of covariates in X_i , including a quartic in potential experience, race, marital status, citizenship, education, metropolitan area size, occupation, industry, and year. Column 4 also interacts the covariates with DER earnings, relaxing separability. We recognize that this is a relatively simple model of the joint distribution, so subsequent analysis moves from use of a single linear log earnings term to categorical measures for earnings percentiles that allow for different responses throughout the distribution. This allows for a less parametric relationship between nonresponse and earnings.

The results in table 2 suggest an average tendency of positive rather than negative selection into response. That said, the coefficients for both men and women are close to zero (with or without controls). The effect of DER earnings for men with controls is a precisely estimated -0.014

⁷ Probit models yield observationally equivalent marginal effects to the linear-probability models presented.

TABLE 2
ASEC MEAN NONRESPONSE WITH RESPECT TO DER EARNINGS
FOR MEN AND WOMEN, 2006–11

	OLS			
	Unweighted (1)	Weighted (2)	Weighted, with X's (3)	Weighted, with X's Interacted with DER (4)
Men				
ln Earnings ^{DER}	-.017*** (.001)	-.018*** (.001)	-.014*** (.001)	-.012*** (.001)
Constant	.374*** (.009)	.398*** (.011)	.436*** (.017)	.401*** (.054)
Observations	224,852	224,852	224,852	224,852
R ²	.002	.002	.018	.019
Women				
ln Earnings ^{DER}	-.009*** (.001)	-.010*** (.001)	-.008*** (.001)	-.008*** (.001)
Constant	.276*** (.008)	.292*** (.010)	.349*** (.017)	.412 (.051)
Observations	214,869	214,869	214,869	214,869
R ²	.001	.001	.015	.016

SOURCES.—US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

NOTE.—This table shows OLS estimation of eq. (3) described in Sec. IV of the text. Cols. 1 and 2 include a single control, ln Earnings^{DER}. Cols. 3 and 4 include additional controls for potential experience, race, marital status, citizenship, education, metropolitan area size, occupation, industry, and year. Col. 4 interacts these controls with ln Earnings^{DER}. Robust standard errors in parentheses. Weighted estimates are weighted using inverse-probability weights for ASEC-DER linkage.

*** Significance $p < .01$.

(a 10 percent increase in earnings decreases the probability of nonresponse by just over a tenth of a percentage point). The effect for women is roughly half that size (-0.008). Although these results provide what we believe are accurate measures of central tendency for these broad samples of men and women, our results for men appear to be just the opposite of that found by Greenlees et al. (1982), who found negative selection into response. Their small sample of married white men with nonworking spouses in 1972, however, is not representative of today's workforce. In order to compare our estimates with those of Greenlees et al., in results not shown we create a similar sample restricted to married white male citizens with spouse present. Unlike Greenlees et al., we include those with working spouses since married women's labor force participation is now closer to the norm rather than the exception. In contrast to the negative coefficients on log earnings for all men, using the restrictive married white male sample flips the signs and produces positive coefficients, meaning negative selection into response. The latter results are

qualitatively consistent with Greenlees et al., as well as previous studies finding negative selection into response, though again we emphasize that their sample is not representative of the modern labor force.

A. Nonresponse across the DER Distribution

Rather than focusing on central tendency, it is more informative to examine how nonresponse varies across the distribution. Grouping observations by DER earnings centile for the linked sample and estimating nonresponse rates for each centile (by gender) produces nonresponse rates that vary across the distribution nonparametrically. Panel A of figure 2 plots these results for the entire sample (smoothed using 3 percentile point moving averages). We note that the highest nonresponse rates are for men through the lowest 30 centiles (as high as 28 percent) and for both men and women at the highest 5 centiles, reaching 25 percent. Throughout the middle of the distribution, the graph is relatively flat. This is suggestive of our main result—“trouble is in the tails”—which is underscored in more dramatic fashion in panels B and C of figure 2. In panel B we focus on earnings among full-time, full-year workers, and in panel C we adjust for hours of work regardless of work status and depict nonresponse rates across the distribution of average real hourly earnings. Here the “trouble in the tails” is most evident: nonresponse rates rise dramatically in the left and right tails. Although similar to panel A, in panel C both men and women in the highest centiles have nonresponse rates reaching 30 percent. Through the middle of the distribution, however, the nonresponse rates are remarkably flat. The linear models reported in table 4 will necessarily fit this part of the distribution, thus explaining the apparent absence of substantive nonresponse bias when focusing on central tendency. The less pronounced trouble in the lower tails in panel A, which includes part-time and part-year workers, is largely explained by the fact that low earnings are caused not only by low pay but also by few weeks worked and low weekly hours. Both panels B and C adjust for annual hours and thus reveal a more striking pattern of U-shaped nonresponse across the distribution.⁸ In short, nonresponse in the left tail is associated primarily with a low wage, not low earnings resulting from low hours worked. This pattern is widespread. Appendix figures 2–4 show that U-shaped patterns hold across race, ethnicity, interview month, and proxy report status.

⁸ Reported hours are concentrated at 2080 (full-time, full-year). While nonresponse is somewhat higher for workers at 2080 hours (3 percentage points), there is no other obvious pattern across hours worked. Mean annual hours worked systematically increase across the DER earnings distribution, as expected. That said, for those with low DER earnings, mean hours worked are substantial, about 1000 hours for men and 650 for women in the lowest 3 earnings percentiles. This suggests that hours worked are not driving the U-shape.

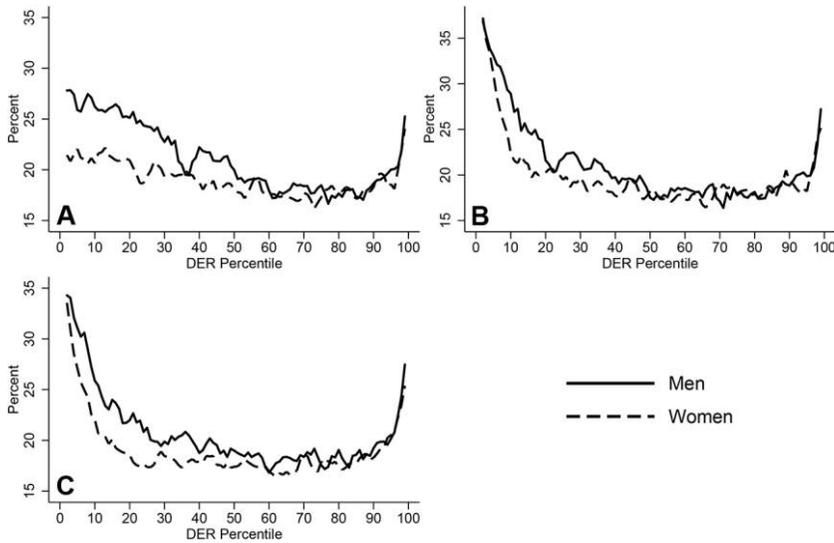


FIG. 2.—Nonresponse rates by gender for joint DER earnings distribution. *A*, Earnings: all workers. *B*, Earnings: full-time, full-year workers. *C*, Average hourly earnings: all workers. Each panel shows the nonresponse rate for a 3 point moving percentile average across a common DER earnings distribution for men and women. The nonresponse rate is weighted using inverse-probability weights for ASEC-DER linkage. Sources: US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

The nonresponse rates in figure 2 do not control for other factors, many of which are known to be associated with earnings and nonresponse. To address this, we modify the nonresponse equation specification seen previously in table 2 by grouping the bottom 90 percent of earners into earnings deciles, while breaking up the top decile into finer percentile increments. Table 3 presents results both with and without human capital, demographic, and location controls, separately for men and women. In all cases, we include a full set of decile/percentile dummy variables, rather than including an intercept. Hence, each coefficient provides an estimate of the nonresponse rate at the given DER earnings level. Readily evident from the coefficients is that nonresponse rates are not constant across the distribution, with the highest earnings deciles producing the highest nonresponse. The U-shapes are highly similar with and without controls.⁹ Among men, the lowest decile has a 14 percent nonresponse rate, while the typical range through the rest of the distribution is roughly

⁹ Note that our unconditioned figures showing nonresponse rates across the earnings distribution also have been constructed conditioned on a detailed set of covariates. While conditioning affects the level of nonresponse, curvature of the conditioned and unconditioned nonresponse figures is indistinguishable to the eye.

TABLE 3
ASEC NONRESPONSE ACROSS THE DER EARNINGS DISTRIBUTION
FOR MEN AND WOMEN, 2006–11

DER EARNINGS DECILES AND PERCENTILES	OLS			
	Men		Women	
	Earnings Decile Dummies (1)	Earnings Decile Dummies and X 's (2)	Earnings Decile Dummies (3)	Earnings Decile Dummies and X 's (4)
Decile 10	.267*** (.004)	.140*** (.009)	.213*** (.003)	.120*** (.009)
Decile 20	.253*** (.004)	.130*** (.009)	.214*** (.004)	.119*** (.009)
Decile 30	.217*** (.003)	.098*** (.009)	.202*** (.003)	.109*** (.009)
Decile 40	.210*** (.003)	.095*** (.009)	.198*** (.003)	.107*** (.009)
Decile 50	.189*** (.003)	.074*** (.009)	.194*** (.003)	.103*** (.009)
Decile 60	.179*** (.003)	.064*** (.009)	.184*** (.003)	.095*** (.009)
Decile 70	.178*** (.003)	.065*** (.009)	.180*** (.003)	.094*** (.009)
Decile 80	.176*** (.003)	.064*** (.009)	.174*** (.003)	.089*** (.009)
Decile 90	.185*** (.003)	.073*** (.009)	.176*** (.003)	.091*** (.009)
Percentiles 91–95	.200*** (.004)	.088*** (.010)	.176*** (.004)	.090*** (.010)
Percentile 96	.199*** (.010)	.086*** (.013)	.197*** (.010)	.106*** (.013)
Percentile 97	.227*** (.011)	.111*** (.014)	.193*** (.010)	.102*** (.013)
Percentile 98	.252*** (.011)	.133*** (.014)	.183*** (.010)	.091*** (.013)
Percentile 99	.271*** (.011)	.149*** (.014)	.189*** (.010)	.097*** (.013)
Percentile 100	.320*** (.011)	.192*** (.015)	.243*** (.011)	.146*** (.014)
Observations	224,852	224,852	214,869	214,869
R^2	.212	.225	.192	.204

SOURCES.—US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

NOTE.—This table shows OLS estimation of eq. (3), which includes DER earnings decile dummy variables described in Sec. IV.A of the text. Cols. 1 and 3 include only decile dummy variables, while cols. 2 and 4 add controls described in table 2. Robust standard errors in parentheses. Estimates are weighted using inverse-probability weights for ASEC-DER linkage.

*** Significance $p < .01$.

half that at 6.5 percent to 7.5 percent. For men in the highest 3 percentiles, the nonresponse rate again rises over 14 percent with the top 1 percent having a 19 percent nonresponse rate. For women, the results with controls are less pronounced, but again we see the U-shape. At the lowest decile, the nonresponse rate is 12 percent, while through the middle

of the distribution it falls to around 9 percent, and in the highest percentile, it rises to 14 percent. While we do not reject the null hypothesis that these rates are equal through the middle of the decile range (40th through 70th deciles), we do reject the null that all deciles are equal.

Our final evidence in this section is to show nonresponse rates for men and women with respect to percentiles across the *predicted* earnings distribution, seen in figure 3. We do this to test whether or not the U-shape is largely a result of observable covariates. The linked ASEC-DER sample is used to estimate conditional mean earnings equations along the lines of equation (4) using the same rich set of demographic controls, as well as controls for both full-time/part-time and full-year/part-year status. The predicted DER earnings for each worker, which can be thought of as an “attribute index,” are then used similarly to the actual DER wage in figure 2. Workers are grouped by (3 point moving average) centile, and the resulting nonresponse rate is plotted, along with a smoothed quadratic trend function. Panel A of figure 3 makes it clear that nonresponse is somewhat higher in the tails of the attribute distribution of men compared to that in women in panel B. For the most part, though, nonresponse for

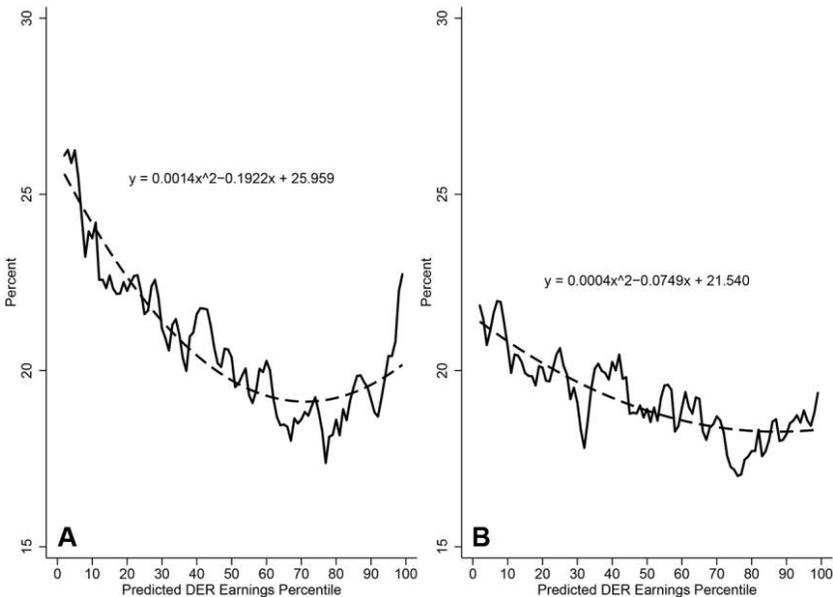


FIG. 3.—Nonresponse rate by predicted DER earnings. *A*, Men. *B*, Women. Each panel shows the nonresponse rate for a 3 point moving average across the predicted DER earnings distribution, along with a smoothed polynomial. The nonresponse rate is weighted using inverse-probability weights for ASEC-DER linkage. Predicted DER earnings come from an OLS estimation of equation (4) described in Section IV.A in the text. Sources: US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

men and women demonstrates less of a U-shape across the attribute distribution than it does across the earnings distribution. The U-shaped non-response (i.e., trouble in the tails) is not driven primarily by observable earnings attributes; rather, it results from the *realization* of either very low or very high earnings.

B. DER Earnings Residuals across the Distribution

We next examine the distribution of earnings conditional on response and earnings covariates, $f(Y|R, X)$, again using the linked ASEC-DER data with inverse-probability weights. We estimate earnings regressions specified in equation (4) using $\ln \text{Earnings}_i^{\text{DER}}$, and in figure 4 provide kernel density estimates of residuals for respondents and nonrespondents.

The left panel of figure 4 presents the administrative earnings distributions by ASEC response status among men, while the right panel does so for women. In both panels, peaks of the respondent distribution are higher than peaks of the nonrespondent distribution. Similarly, the tails of the nonrespondent distribution are generally longer, indicating a higher variance for nonrespondents. Appendix table 4 supports this, demonstrating

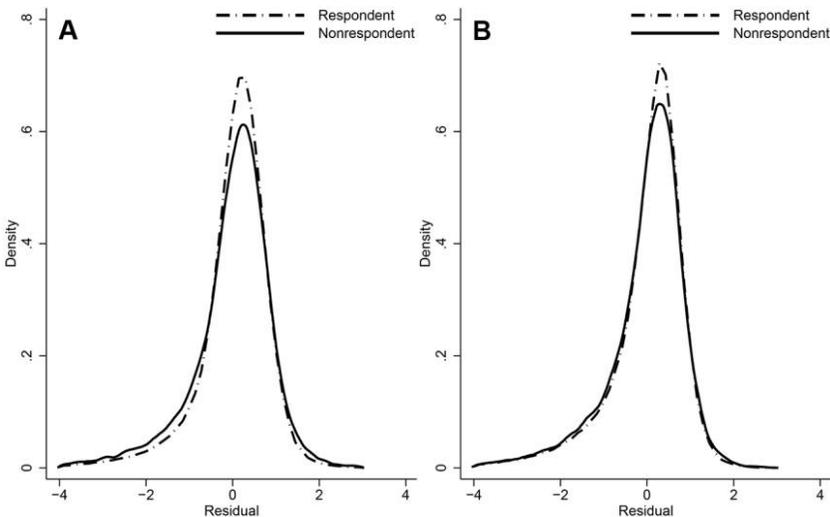


FIG. 4.—Residuals of log earnings regressions by ASEC response status and gender. *A*, Male kernel density of residuals; kernel = Epanechnikov, bandwidth = 0.0475. *B*, Female kernel density of residuals; kernel = Epanechnikov, bandwidth = 0.0502. Each panel shows the kernel density estimate of residuals for respondent and nonrespondent distributions. Residuals come from an OLS estimation of equation (4) described in Section IV.B in the text and appendix A.5. The OLS estimation uses inverse-probability weights for ASEC-DER linkage. Sources: US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

that the variance for male (female) nonrespondents is 1.37 (1.12) times the variance of male (female) respondents. Testing differences between these variances using either the standard F -test or Levine's test rejects the null hypothesis of equivalence at conventional levels. Tests for differences in means reject the null hypothesis as well. A simple test of the difference in the medians fails to reject for men, but does reject for women. Examining the percentiles shows that the major differences occur in the tails, as seen in figure 4. We conclude that there is strong evidence of differences between these distributions, with the most substantive differences in the variances and percentiles. Furthermore, Kolmogorov-Smirnov tests reject the null (p -value < 0.00) that $f(Y|R = 1, X) = f(Y|R = 0, X)$, which is a sufficient condition to reject the hypothesis that $f(Y|R, X) = f(Y|X)$.

C. Proxy Respondents and Measurement Error

Census interviewers designate a single person to be the respondent for all household members in a bid to lower the time and money costs of conducting household surveys. Although a single person is recorded as providing answers to survey questions, the designee may rely on input from other household members in providing requested information. In the ASEC sample used in our analysis, 54 percent of men have their earnings reported by a proxy, while 42 percent of women rely on proxy reports.¹⁰ As seen in appendix figure 4, earnings nonresponse is substantially higher among individuals with proxy earnings responses than among self-respondents. For our combined sample of women and men, earnings nonresponse rates are 24.2 percent for proxy respondents versus 16.4 percent for self-respondents. The gap in nonresponse rates between proxies and self-respondents is about 2 percentage points greater among men than among women; this gap varies little across the earnings distribution.

There exists rather limited information on the reliability of proxy earnings responses (Mellow and Sider 1983; Lee and Lee 2012; Reynolds and Wenger 2012). Using the linked ASEC-DER sample, we can observe whether administrative earnings in the DER, where there are no proxies, vary with respect to proxy use in the ASEC. That is, we estimate two equations, each separately for men and women: (1) an ASEC wage equation with spouse and nonspouse proxy variables and (2) a DER wage equation with ASEC spouse and nonspouse proxy variables. Each wage regression also controls for a saturated set of confounders. The proxy variables in the DER equation act as "phantom" dummies; if ASEC proxy coefficients only measured

¹⁰ We designate a response as a proxy response when an individual's line number differs from the line number of the household respondent. This method is not 100 percent accurate. Census identifies a respondent at the end of an interview. If there has been a change in the respondent after the survey collects earnings information, this method need not identify correctly the household member providing the earnings information.

reporting differences between proxies and self-respondents, the DER proxy coefficients should be zero. Proxy coefficients in ASEC wage equations reflect the combined effects of proxy misreporting and worker heterogeneity. Inclusion of phantom ASEC proxy variables in DER administrative earnings regressions thus provides estimates of worker earnings heterogeneity correlated with proxy status (conditional on measured attributes). Thus, in order to estimate proxy misreporting error, we simply subtract the DER phantom proxy coefficients from the corresponding ASEC proxy coefficients. Note that we exclude imputed earners since we cannot know whether the donor's earnings used in the ASEC imputation were self-reported or from a proxy.

These results are summarized in columns 5 and 6 of table 4, using both a single proxy variable and distinguishing between spouse and non-spouse proxies. In general, the DER and ASEC proxy coefficients differ substantively, particularly so in the male regressions. In models with a single proxy variable (i.e., proxy use versus self-response), we find that proxies understate both men's annual earnings and men's hourly earnings by 0.062 log points. Underreporting of men's earnings are moderately larger when there are nonspouse rather than spousal proxies. For women, underreporting by proxies is a comparatively small 0.014 log points. Underreporting by nonspouse proxies is about a third larger than by spouse proxies. One clear result from this analysis is that inclusion of dummies for spouse and nonspouse proxy reports captures substantive unobserved heterogeneity, as seen by the DER coefficients. Both women and men with earnings reported by spousal proxies have higher administrative (DER) earnings (note that we control for marital status in all earnings equations).

The substantive underreporting of men's wages and earnings by proxies, coupled with minimal underreporting of women's earnings, has obvious implications for measurement of the gender gap, which is frequently measured using the CPS.¹¹ From above, the difference-in-difference in the male-female earnings gap from proxy reports is 0.048 log points ($0.062 - 0.014$). Were all earnings reported by proxies, these results would imply that the gender gap is understated by the full 0.048 log points. Based on sample averages of proxy use among men of 54.4 percent and 41.5 percent among women, a back-of-the-envelope calculation implies that gender-asymmetric underreporting of earnings by proxies understates the gender wage gap by about 0.028 log points ($0.544 \times 0.0621 - 0.415 \times 0.0145 = 0.0278$), or about 14 percent of the regression-adjusted ASEC average wage gap of .20 (the adjusted DER gender gap is .19). We return to the gender gap in a later section, focusing on the gap across the distribution using quantile models.

¹¹ Blau and Kahn (2017) provide a comprehensive survey of the gender wage gap, with a focus on CPS estimates.

TABLE 4
 PROXY MISREPORTING OF MALE AND FEMALE ANNUAL AND HOURLY EARNINGS
 BASED ON CPS-ASEC AND DER DIFFERENCES IN PROXY COEFFICIENTS

	VARIABLE					
	CPS-ASEC		DER		CPS Proxy Misreport	
	Men (1)	Women (2)	Men (3)	Women (4)	Men (5)	Women (6)
Annual earnings:						
Earnings equations with proxy coefficients:						
Proxy	-.0638	.0502	-.0017	.0647	-.0621	-.0145
Earnings equations with spouse and nonspouse proxy coefficients:						
Spouse proxy	.0233	.1210	.0807	.1330	-.0574	-.0120
Nonspouse proxy	-.2010	-.0576	-.1320	-.0392	-.0690	-.0184
Hourly earnings:						
Wage equations with proxy coefficients:						
Proxy	-.0501	.0031	.0117	.0172	-.0618	-.0141
Wage equations with spouse and nonspouse proxy coefficients:						
Spouse proxy	-.0091	.0360	.0483	.0476	-.0574	-.0116
Nonspouse proxy	-.1150	-.0473	-.0459	-.0289	-.0691	-.0184

SOURCES.—US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

NOTE.—This table shows the OLS estimation of earnings and wage regressions that include controls for proxy, spouse proxy, and nonspouse proxy. Cols. 1 and 2 use ASEC earnings, while cols. 3 and 4 use DER earnings. CPS proxy misreporting estimates (cols. 5 and 6) are calculated as the difference between the ASEC and DER proxy coefficients (col. 1 minus col. 3 for men and col. 2 minus col. 4 for women). See Sec. IV.C in the text for further explanation. The CPS-ASEC equations exclude imputed earners since we cannot know whether the donor's earnings were self-reported or from a proxy. The DER equations include the same sample. All columns include additional controls described in table 2. Estimates are weighted using inverse-probability weights for ASEC-DER linkage.

We conclude this section with a brief discussion of differences in the reported earnings in the ASEC and in the DER. Empirical investigation of measurement error in earnings in the CPS and other surveys has a long history (Herriot and Spiers 1975; Halsey 1978; Alvey and Cobleigh 1980; Mellow and Sider 1983; Duncan and Hill 1985; Poterba and Summers 1986; Rogers and Herzog 1987; Mathiowetz and Duncan 1988; Marquis and Moore 1990; Bound and Krueger 1991; Bound et al. 1994; Bollinger 1998; Bound, Brown, and Mathiowetz 2001; Roemer 2002). The goal of this exercise is to examine the relationship between the survey respondents for those with linked surveys using nonparametric kernel regression. We use ordinary least squares (OLS) to estimate models of both ASEC and DER earnings on the same covariates as previously, and the residuals from each model are then used for the nonparametric regression of ASEC on DER. Appendix figure 5, using a log-earnings scale, shows that the “common man” hypothesis (i.e., mean reverting measurement error) found in the validation literature is supported: individuals with low earnings tend to overreport their earnings, while individuals with high earnings tend to underreport. Since this analysis was conducted on residuals, these are

not associated with demographic characteristics such as education or race. This evidence provides some interesting qualifications on our main finding that nonresponse is concentrated in the tails of the distribution. Here we see that for respondents, measurement error is also concentrated in the tails of the distribution. Previous authors (Bollinger and David 2001; Kapteyn and Ypma 2007) have found similar overlaps in the population of “noncooperative” survey respondents. This suggests, perhaps, that the Census imputation procedure may reflect the response that typical nonrespondents *would make*, were they to participate, measurement error and all. It does, however, highlight that individuals in the extreme parts of the earnings distribution (both unconditional and conditional) are not responding to the survey in ways we might hope. Our prior results show that many simply do not respond, while appendix figure 5 shows that those who do respond are not appropriately revealing their earnings. This evidence adds support to the idea that survey response and nonresponse are correlated with the level of income, even controlling for demographic factors.

D. Earnings Nonresponse over Time and Earnings Growth

One advantage of the rotation group structure of the ASEC is the overlapping nature of the sample, allowing up to 50 percent of sample individuals to be followed across adjacent years. There is a small literature examining either measurement error or nonresponse in panel settings (Bound and Krueger 1991; Fitzgerald, Gottschalk, and Moffitt 1998; Bound et al. 2001; Bollinger and David 2005). We briefly examine the rates of nonresponse for the 2 year panels covered by our data, the relationship between earnings and nonresponse, and the impact of nonresponse on simple measures of earnings growth. Several authors (Peracchi and Welch 1995; Cameron and Tracy 1998; Hardy and Ziliak 2014) have pointed out that the subsample of individuals who can be followed across adjacent years in the ASEC are not fully representative because the sample frame is the household address and not the person, and thus movers are not followed. Nonetheless, the longitudinal sample is widely used and thus it is important to assess nonresponse, and indeed, as appendix table 6 demonstrates, there are few observable differences between the panel and cross-sectional samples. We find that the linkage rate for panel individuals rises to 88.3 percent (compared to 87.4 percent for the full ASEC sample). The earnings nonresponse rate is 16.9 percent in year 1 of the panel and 18.2 percent in year 2, as compared to 22.6 percent in the full cross-section sample.

Column 1 of table 5 presents the (unweighted) response status in the first year, cross-tabulated with the response status in the second year. Overall, 72.8 percent of the sample responds in both years and 7.9 percent

TABLE 5
JOINT DISTRIBUTION OF RESPONSE AND LOG EARNINGS GROWTH
BY RESPONSE STATUS

	RATE (1)	LOG EARNINGS GROWTH	
		ASEC (2)	DER (3)
Full sample		.006	.013
Nonrespondent in both years	7.9	.037	.020
Respondent in both years	72.8	.0001	.011
Respondent only in year 1	10.3	-.008	-.017
Respondent only in year 2	9.0	.037	.055

SOURCES.—US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

NOTE.—This table shows nonresponse and response rates (col. 1) and log earnings growth for ASEC earnings (col. 2) and DER earnings (col. 3) by response status for the 2 year ASEC panel. Col. 1 is unweighted, while cols. 2 and 3 are weighted by inverse-probability weights for ASEC-DER linkage.

does not respond in either year. The joint-year response rate is of course lower than the single-year response rates (83.1 percent in the first and 81.8 percent in the second year, as reported in appendix table 6). Many individuals change their response status and such changes are approximately symmetric. We find that 10 percent of individuals respond in the first year but become nonrespondents in the second year; 9 percent of individuals do not respond in year 1 but then do so in year 2. Figure 5 displays panel nonresponse rates plotted against the DER earnings centile for the first year in the panel. Panel A combines full-time, full-year workers with part-time and part-year workers, but unlike the earlier figures, here we combine the male and female samples. The year 1 and year 2 rates are (unsurprisingly) very comparable in shape to our prior results seen in figure 2. The third line tracks the percentage of those who failed to respond in both years. Although multiyear nonresponse is obviously lower than annual nonresponse, we again find that such nonresponse is U-shaped with respect to the level of earnings. Panel B presents the same breakdown for the full-time, full-year sample, while panel C shows the full sample with respect to hourly wage centiles. As in comparable panels in figure 2, we find more pronounced U-shaped patterns in panels B and C.

Using IPW weights to account for individuals not linked to the DER, columns 2 and 3 of table 5 present average earnings growth between the first and second year of the panel. We focus primarily on column 3, examining the growth of inflation-adjusted earnings in the DER. Overall, the average earnings growth was 0.013 log points. Most notable is the striking pattern between those who respond only in one year: low (negative) DER earnings growth for those who respond only in year 1 and high (positive) DER earnings growth for those who respond in year 2. This

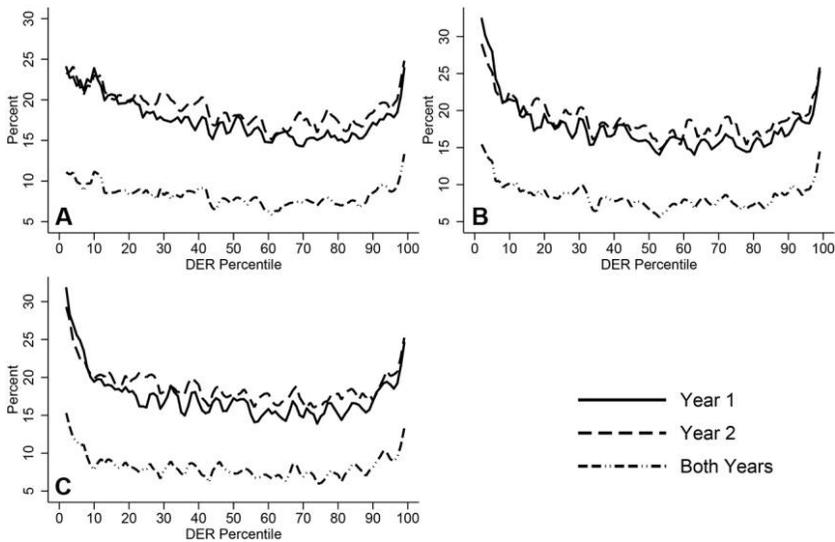


FIG. 5.—Nonresponse rates by panel status for year 1 joint DER earnings distribution. *A*, Earnings: all workers. *B*, Earnings: full-time, full-year workers. *C*, Average hourly earnings: all workers. Each panel shows the nonresponse rate for a 3 point moving average for the year 1 DER earnings distribution in the 2 year ASEC panel. The nonresponse rate is shown for year 1, year 2, and both years of the panel. The nonresponse rate is weighted using inverse-probability weights for ASEC-DER linkage. Sources: US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

pattern suggests strong selection into response based on changes in earnings, and is consistent with the U-shaped pattern found in the cross-sectional analysis as well; those who have very low or very high earnings may fail to respond if that is an unusual or new situation. Earnings growth for nonrespondents in either year is higher in absolute value than for those who respond in both years. This provides further evidence that nonresponse in the CPS should be treated as NMAR.

Here, unlike the previous analyses, the ASEC earnings growth includes the imputations for nonrespondents. We include the ASEC growth rates in column 2 for comparison and evaluation of the imputation process. Comparison of growth rates between the ASEC and DER confound both measurement error and imputations in the two categories in which response switches. For those who respond in both periods, measurement differences lead to the ASEC having strikingly lower estimates of earnings growth. In the case of nonresponse in both periods, the ASEC imputation procedure appears to impute higher earnings growth than observed. While one can take a variety of perspectives on whether administrative earnings are the correct measure, the marked difference in relative growth suggests that the imputations are extremely poor in capturing earnings dynamics.

V. How Troubling Is Trouble in the Tails? The Consequences of Nonresponse

The linked ASEC-DER data permit us to examine directly whether relying solely on respondents' earnings may produce in some circumstances results similar to what would be produced using complete (but unobtainable) data. Because the DER sample includes administrative earnings for nonrespondents as well as respondents, we can compare estimates from respondent-only samples with those from complete samples, something not possible with publicly available data. Here we focus on three main types of estimation that should provide researchers with guidelines for judging the importance of nonresponse in their research (above and beyond that demonstrated in the prior section on proxy responses and longitudinal earnings growth). In Section V.A we examine the implications for linear models of earnings fitted with least-squares estimators. We find a modest impact from using a respondent-only sample, as the symmetric nonresponse in the tails has little impact on estimation of the means. In Section V.B, we consider the impact on coefficient estimates from quantile regressions. Here we find estimates in the lower and upper quantiles from respondent-only samples to be problematic, as compared to use of a full sample from the ASEC-DER link. Our concerns regarding use of a respondent-only ASEC sample are reinforced in Section V.C, where we examine earnings inequality. This conclusion is not surprising given that measures of inequality are sensitive to earnings in the tails.

A. Mean Earnings Estimates

Using the ASEC-DER sample and the IPW weighting to account for representativeness, we estimate least-squares log annual DER earnings equations by gender, separately for the linked respondents, linked nonrespondents, and all linked workers samples, again controlling for the same set of covariates used previously in the analysis. In table 6 we provide the predicted DER earnings for men and women using means from the full sample multiplied by coefficient estimates from (1) regressions using the full sample in column 1, (2) regressions on the subsample of respondents in column 2, and (3) regressions on the subsample of nonrespondents in column 3. We use as our benchmark the predicted DER earnings based on coefficients from the full sample in column 1.

Focusing first on men, the use of full-sample coefficients with the full-sample worker attributes (X 's) results in a predicted mean log earnings of 10.488. This is close to that obtained using respondent-only β 's, which leads to a predicted mean log earnings of 10.502, or 0.014 (1 percent) higher than obtained with the full sample. The equivalent values for women are 10.053 using full sample β 's and 10.061 using respondent β 's, a 0.008 difference. However, selection on observables is readily evident comparing

TABLE 6
 PREDICTED LOG DER EARNINGS WITH FULL SAMPLE, ASEC RESPONDENTS,
 AND ASEC NONRESPONDENTS, 2006–11

	VARIABLE		
	β 's from $\ln \text{Earnings}^{\text{DER}}$		
	All Workers (1)	Respondents (2)	Nonrespondents (3)
Men:			
Prediction with full sample X 's	10.488	10.502	10.437
Observations	224,852	180,564	44,288
R^2 of earnings equation	.321	.327	.306
Women:			
Prediction with full sample X 's	10.053	10.061	10.018
Observations	214,869	175,253	39,616
R^2 of earnings equation	.268	.274	.253

SOURCES.—US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

NOTE.—This table shows predicted mean $\ln \text{Earnings}^{\text{DER}}$ from earnings regressions for all linked workers (col. 1), linked respondents (col. 2), and linked nonrespondents (col. 3). Predicted $\ln \text{Earnings}^{\text{DER}}$ are based on sample means from the full ASEC sample. All columns include additional controls described in table 2. Regression estimates use inverse-probability weights for ASEC-DER linkage.

columns 2 and 3 using respondent (R) and nonrespondent (NR) β 's, respectively. The R – NR predicted earnings difference is $10.502 - 10.437 = 0.065$ for men and $10.061 - 10.018 = 0.043$ for women. These differences are substantive. Because the nonrespondent shares of the total samples are relatively small (roughly 20 percent), the respondent-only sample provides coefficient estimates close to what would be produced using the full sample, the latter not being an option with public-use data. In short, users of public data can avoid substantial bias by removing imputed earnings. One can rebalance the respondent sample using inverse-probability weights, adjusting the ASEC supplement weight with model-based estimates of the probability of response. The analysis comparing male and female earnings is particularly interesting because gender is the one worker attribute always matched correctly in Census imputations (Bollinger and Hirsch 2006). That is, there exists no “match bias” (i.e., wage gap attenuation) resulting from assignment of imputed earnings from a different-sex donor.

B. *Earnings Gaps across the Distribution*

We next examine the implications of nonresponse across the distribution of earnings for a host of widely studied outcomes such as earnings gaps across gender, race, and education. Figures 6–8 depict estimates of coefficients from quantile regressions of log annual earnings on the same set of covariates used in our earlier conditional analyses at the 5th, 10th,

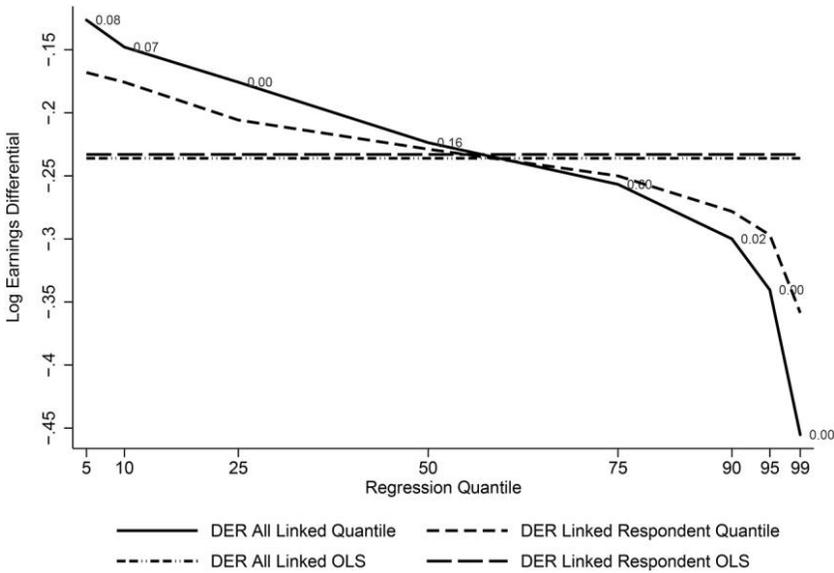


FIG. 6.—Female-male earnings gap by quantile for full-time, full-year workers. This figure plots the female coefficient from saturated quantile and OLS regressions using DER earnings for two samples: (1) linked ASEC respondents and nonrespondents (“DER All Linked Quantile” and “DER All Linked OLS”) and (2) linked respondents only (“DER Linked Respondent Quantile” and “DER Linked Respondent OLS”). OLS and quantile estimates are weighted using inverse-probability weights for ASEC-DER linkage. Numbers in the figure show p -values of the difference in the quantile estimates between the “DER All Linked” and “DER Linked Respondent” samples at each quantile. Sources: US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

25th, 50th, 75th, 90th, 95th, and 99th quantiles. Each figure contains estimates from two samples—one using DER earnings on both linked ASEC respondents and nonrespondents (“All Linked”), and the other using linked respondents only (“Linked Respondent”). We focus here on the full-time, full-year subsample, in part because earnings distributions including part-time and part-year workers confound the level of earnings with hours worked and thus are difficult to interpret. While wages are often used in applications, concern arises there too with differences in wage distributions between full-time/full-year workers and those who work less, as well as potential measurement error in annual hours worked. It should be noted that quantile estimates are measuring differences in the conditional distribution and hence do not match the unconditional quantiles.

Figure 6 presents the estimated coefficients on the female indicator variable from pooled earnings quantile regressions, along with the p -value of the difference in coefficient estimates from the two samples. The OLS coefficients are presented as horizontal lines for comparison. In general,

produces biased estimates that understate the racial gap among men. The largest impact in panel A is at the high end of the distribution, where the bias is 0.038 log points, or nearly 20 percent relative to the combined respondent-nonrespondent sample. As with the male-female differential, this is likely driven by missing high-earning men. Although black men are less likely to report than white men in general, it appears that conditional on other factors, nonrespondents are disproportionately white men at the highest earnings. Here, along with the consistent underestimation of the differential in the respondent-only sample, the OLS estimates display modest underestimation as well. In panel B, the black-white differential for women displays a slightly different pattern. While at the higher quantiles, the respondent-only sample continues to slightly understate the gap, we note that at the lower quantiles the bias is reversed, with the respondent-only subsample slightly overstating the gap by about 0.012 log points, or about 10 percent of the combined sample gap. Panels C and D depict the respective gaps for men and women between Hispanics and whites. As we saw for the female black-white differential, the respondent-only sample understates the differential at the highest quantiles but overstates it at the lower quantiles for both men and women. For Hispanic men, the bias in the differential is most pronounced at the highest quantiles (0.03 log points at the 99th percentile, or 20 percent of the combined respondent/nonrespondent gap), while for women the bias is largest at the lowest quantiles.

Finally, figure 8 examines the earnings differential between those whose highest degree is high school (excluding GEDs) compared to high school dropouts (panel A) and college graduates (with that being the highest degree) compared to high school graduates (panel B). High school returns are systematically understated using the respondent sample, particularly so in the bottom half of the distribution, but with minimal difference at the top of the distribution. The same qualitative pattern is seen for estimates of the return to college, but with a modest downward bias throughout the entire distribution (being largest at the 90th and 95th percentiles). In both schooling return cases, the respondent sample understates the return at the means (OLS).

C. Earnings Inequality

There is limited evidence regarding how earnings nonresponse affects the measurement of inequality; a priori it is not readily apparent how it should do so. One needs to identify who fails to respond, how nonresponse differs with respect to true and typically unobserved earnings (conditional on covariates), how any such nonresponse bias might differ across the earnings distribution, and how one can best treat top-coded earnings. Census uses different top-code values depending on earnings source, and these

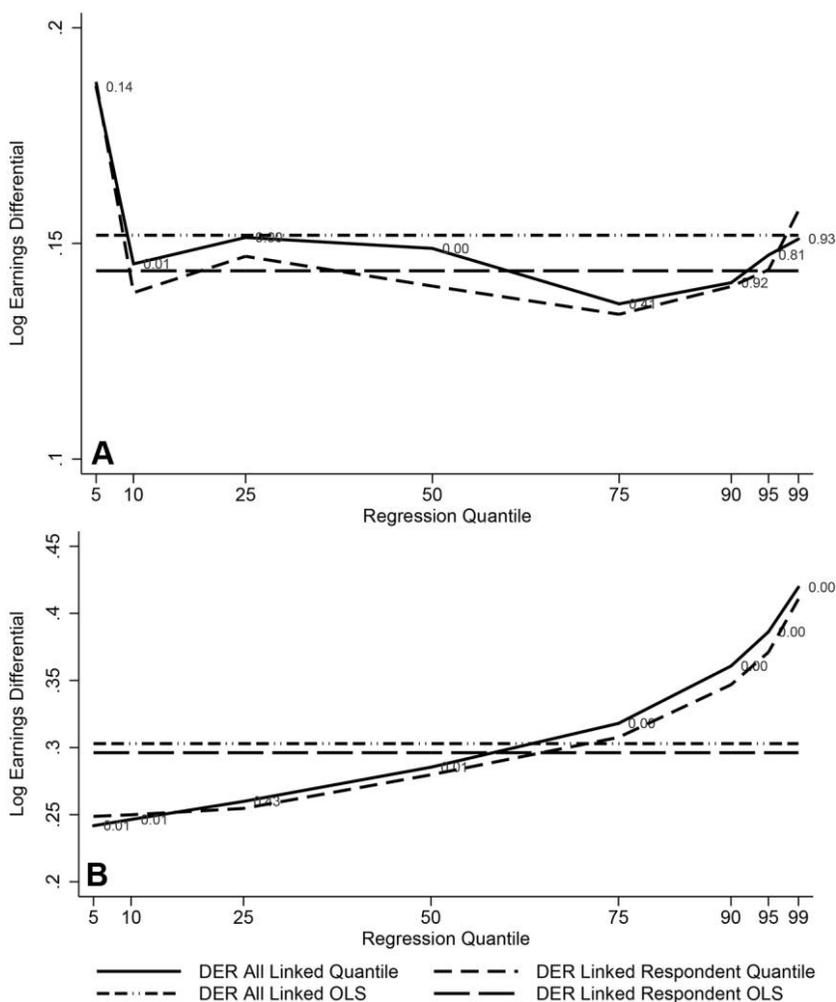


FIG. 8.—Earnings gaps by education and quantile: full-time, full-year workers. *A*, High school return. *B*, College return. Panel *A* plots the coefficient comparing high school to high school dropouts from saturated quantile and OLS regressions using DER earnings for the two samples described in figure 6. Similarly, panel *B* plots the coefficient comparing college graduates to high school graduates from saturated quantile and OLS regressions using DER earnings for the two samples described in figure 6. OLS and quantile estimates are weighted using inverse-probability weights for ASEC-DER linkage. Numbers in the figure show *p*-values of the difference in the quantile estimates between the “DER All Linked” and “DER Linked Respondent” samples at each quantile. Sources: US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

values differ between internal and public-release versions of the ASEC. A key advantage of the DER data is that earnings are not top coded, thus permitting a direct comparison of estimates of upper-tail inequality from tax records to top-coded survey responses. Some inequality studies have excluded imputed earners (Lemieux 2006; Autor et al. 2008), while others have not (Burkhauser et al. 2012). This is the first such direct comparison from linked individual survey and tax data on how nonresponse and top coding affects earnings-inequality estimates. We estimate several leading measures of inequality emphasized in the recent literature—including the Gini coefficient (fig. 9) and 90-10 ratio (fig. 10), along with 90-50, 50-10, and top 1 percent share in appendix figures 7–9. For brevity, we restrict our discussion here to the Gini coefficient results as similar patterns are obtained for the 90-10 ratio.

In panel A of figure 9 we show the earnings Gini for the full sample of workers. Shown with the dash-dotted line is the full ASEC sample, with the long-dashed line is the ASEC for respondents only, with the solid line is the DER for all linked workers (and ASEC for nonlinked), and with the short-dashed line is the DER for linked respondents. Comparing the full ASEC with imputations versus ASEC respondents only, one sees that the respondent-only sample shows too low a level of inequality owing to the omission of nonrespondents disproportionately represented in the far left and right tails. Hence, inclusion of imputations is appropriate for measuring unconditioned inequality, despite the severe biases that can arise from inclusion of imputations in conditioned analyses (Bollinger and Hirsch 2006). As with the ASEC, removing nonrespondents from the DER reduces the Gini measure. The larger impact in the DER reflects the fact that the imputations in the ASEC do not capture the NMAR aspect of nonresponse. As compared to the two DER measures, the ASEC measures show a substantially lower level of inequality and somewhat different trends. Earnings inequality in the ASEC is roughly flat over the full sample period, and everywhere below the DER. Using DER earnings, we find a higher level of inequality (about 10 percent higher) and a modest upward trend after 2007. Panel A establishes that NMAR nonresponse has an impact on measures of inequality. Removing those missing values results in a downward bias in estimating inequality. Although inclusion of the imputations fails to account for NMAR bias, it does correct for MAR bias with respect to those attributes matched in the Census imputations.

In panels B and C of figure 9, we explore whether the gap between ASEC and DER earnings inequality is due to nonresponse (panel B) or due to differences in measurement of earnings, including top coding (panel C). Panel B shows three series—the ASEC inclusive of nonrespondents, the DER for linked respondents and nonrespondents (and ASEC for nonlinked), and a hybrid DER measure that uses DER earnings for linked nonrespondents and ASEC earnings for respondents (and ASEC

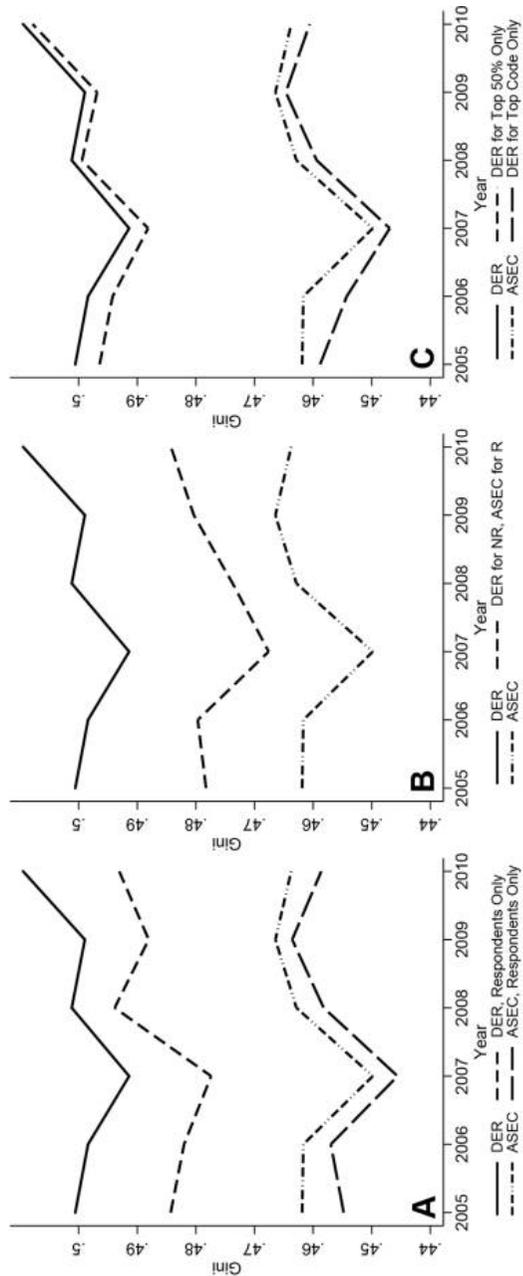


FIG. 9.—Trends in Gini earnings inequality. A, All workers and respondents only. B, Role of nonresponse. C, Role of measurement. Panel A shows the earnings Gini for the following series: (1) full ASEC sample (“ASEC”); (2) ASEC earnings for only respondents (“ASEC, Respondents Only”); (3) DER earnings for all linked workers and ASEC earnings for nonlinked (“DER”); and (4) DER earnings for only linked respondents (“DER, Respondents Only”). Panel B includes a series that uses DER earnings for nonrespondents and ASEC earnings for respondents (“DER for NR, ASEC for R”). Panel C includes a series that uses DER earnings only for top-coded ASEC earnings (“DER for Top Code Only”) and a series that uses DER earnings for workers in the top half of the ASEC earnings distribution (“DER for Top 50% Only”). The Gini is weighted using the ASEC supplement weight. Sources: US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

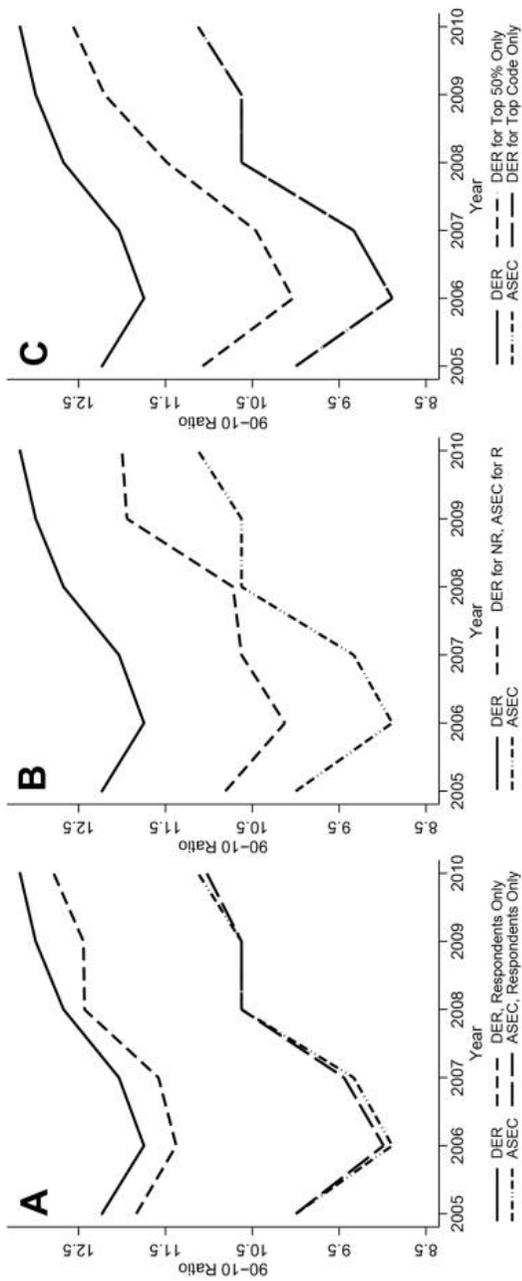


FIG. 10.—Trends in 90-10 earnings inequality. A, All workers and respondents only. B, Role of nonresponse. C, Role of measurement. This figure shows the 90-10 percentile ratio for various series. See note to figure 9 for series descriptions. Sources: US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

for the nonlinked). In all cases the sample size is held constant by using the ASEC for the nonlinked, whether a respondent or nonrespondent. Here we see that the hybrid measure produces a Gini level roughly one-third to halfway between the pure ASEC and DER measures. Comparing the hybrid measures to the pure ASEC measures supports the conclusion that nonrandom, nonresponse bias (NMAR) causes an understatement in the level and trend in earnings inequality based solely on ASEC.

Panel C presents the original ASEC and DER series, along with two additional series. In one series the DER is used only for top-coded ASEC values with a DER link, and in the other series we replace the ASEC with the DER for workers in the top half but not bottom half of the ASEC earnings distribution, regardless of imputation or top-code status. The former case is of interest because the full ASEC and DER groups include a convolution of nonrespondents and top-coded workers, and thus it is less obvious what direct role the top code in the internal ASEC plays vis-à-vis administrative tax data. The latter case is of interest because the DER does not capture off-the-book earnings, and thus the higher level of inequality observed in the DER might be an artifact of underreported earnings in the lower half of the distribution. The results in panel C demonstrate that top-coded earnings alone in the internal ASEC are not the primary cause of the gap in inequality estimates from tax data in the DER versus ASEC survey data. The DER-only series shows substantially higher and (to a lesser extent) rising inequality as compared to ASEC earnings with DER replacing ASEC top codes. In addition, the majority of the gap between the DER and ASEC earnings inequality arises from earnings in the upper half of the ASEC distribution, and not from off-the-books underreporting in the lower half. This conclusion is based on the minimal differences between the DER-only series and the hybrid ASEC-DER series with DER earnings replacing the ASEC in the top half of the ASEC distribution. ASEC measures of inequality tend to understate inequality because the Census hot deck (owing to nonresponse bias) imputes earnings for nonrespondents that are too high in the left tail and too low in the right tail.

VI. Recommendations for Users of Public ASEC

Our results indicate that nonresponse bias causes both earnings gaps and inequality measures estimated with ASEC earnings responses to be understated. Because of nonresponse, the observed data include too few low earners and too few very high earners. The Census hot-deck procedure, based on the MAR assumption, fails to correct this problem because MAR conditions are not met. The general CPS user community does not have access to either the internal ASEC used in this paper or the DER. The advantage of the former comes primarily from data with higher top-code values compared to the public ASEC. Since 1996 Census has attempted

to address this discrepancy, while still maintaining confidentiality, by releasing “proxy values” for those individuals with earnings in between the public and internal top codes. During survey years 1996–2010 the proxy came in the form of cell means, while from 2011 onward via rank swapping. The latter approach is preferred because it preserves the distribution of earnings above the top code. Recently Census released rank-swap values for all the top-code income components (not just earnings) back to 1975, and we recommend that public researchers using the ASEC prior to the 2011 survey year adopt these top codes.¹²

Hirsch and Schumacher (2004) and Bollinger and Hirsch (2006) recommended dropping the imputed nonrespondents in the ASEC because of the attenuation bias that imputations impart on regression coefficients, and then reweighting the sample with inverse-probability weights to retain population representativeness. Their recommendation to drop imputations was based on analyses focusing on models of central tendency (OLS, median regression). Overall, our results here strongly suggest that MAR is violated across the distribution, and thus dropping nonrespondents and reweighting will not correct for nonrandom nonresponse. In practice, we demonstrate that the economic bias from nonresponse may be small in well-specified linear models of wages and earnings. As a general recommendation for distributional research, however, public ASEC users are advised to implement a flexible selection model that corrects for nonrandom nonresponse. In this section, we demonstrate the utility of one such approach in an application to earnings inequality.

Specifically, we implement a procedure recently proposed in Arellano and Bonhomme (2017) whereby one first estimates quantile regressions corrected for nonresponse using copula methods, and then uses predictions from those regressions to create simulated earnings data. The key assumptions to identification are an exclusion restriction, continuity of the residuals of the two equations (main and selection), continuity of the dependent variable in the main outcome (earnings), and the propensity of being observed being conditionally nonzero (e.g., there is not some group that can be correctly predicted to never be observed). For our application, we estimate conditional quantile models of earnings that include controls for a quartic in age, nine education categories, race, gender, immigration status, region and metro status, and industry, and correct for nonrandom selection using the Frank copula as it allows for nonresponse being concentrated in a tail.¹³ To identify the selection model, we use the month-in-sample in which the respondent is observed in the ASEC as an

¹² See <https://www2.census.gov/programs-surveys/demo/datasets/income-poverty/time-series/data-extracts/asec-incometopcodes-swappingmethod-corrected-110514.zip>.

¹³ Our programs, which are based on those provided by Arellano and Bonhomme (2017), are available as supplementary material online at the journal's website.

exclusion restriction. All else being equal, we expect nonresponse to be lower in month 1 or 5 of the rotation cycle as interviews are done in person, while the other six months are conducted over the phone. Appendix figure 3 seems to confirm this as it shows that rates of nonresponse in months 1 and 5 lie everywhere below months 2–4 and 6–8. Thus month-in-sample should be correlated with nonresponse, and at the same time, we do not expect month-in-sample to be related to true individual earnings (Bollinger and Hirsch 2013).¹⁴ We have not observed any variable or combination of variables that perfectly predicts nonresponse or response. Thus, we are confident that the key assumptions for identifying the Arellano and Bonhomme model are met.

After estimating the quantiles, we randomly generate an integer q between 1 and 99 for each individual in the full sample. Following the conditional quantile decomposition method of Machado and Mata (2005), we use the quantile coefficients associated with the draw of q for each individual to produce a prediction of the q th quantile of the earnings distribution. This provides a simulated distribution that can then be used to estimate a variety of statistics, including measures of income inequality. Because the nonresponse throughout the distribution is addressed differentially at each quantile, the Arellano-Bonhomme approach will provide a simulated distribution that has higher dispersion compared to the more restrictive approach in Buchinsky (1998).

In evaluating the efficacy of this approach, there are two possible benchmarks against which to compare our estimates, the latter of which is based solely on survey responses from the ASEC. The first is the administrative records of the DER. The DER provides a source of information on income that is official and is a natural comparison. As noted in Section IV.C, however, there are differences in how individuals report ASEC earnings relative to their DER earnings. While one perspective is that the DER earnings are “correct,” there is the potential that ASEC earnings contain earnings that are not reported to the government (at the low end) or that DER earnings contain other errors (see Kapteyn and Ypma 2007). The fundamental question being addressed in this paper is how to account for nonresponse. Hence the ideal “benchmark” would be the ASEC in which everyone answered the earnings question. The closest approximation to that would be to use the ASEC earnings, but replace linked nonrespondents with their DER. This is the benchmark we adopt.

We estimate the quantile selection model for each year, and in table 7 we present the 6 year average Gini coefficient, and 90-10, 90-50, and

¹⁴ Krueger, Mas, and Niu (2017) and Hirsch and Winters (2016) find substantial differences across the CPS month-in-sample reports of unemployment and multiple job holding, respectively. We find no such pattern of rotation group bias with respect to earnings.

TABLE 7
 PERFORMANCE OF SELECTION CORRECTION METHODS
 FOR NONRESPONSE IN THE ASEC ON INEQUALITY

SAMPLE	INEQUALITY MEASURES			
	Gini	90-10	90-50	50-10
ASEC	.461	10.099	2.607	3.870
ASEC, only respondents with IPW	.464	10.227	2.641	3.869
ASEC, only respondents with copula	.482	10.521	2.676	3.929
ASEC for respondents, DER for nonrespondents (benchmark)	.477	11.038	2.683	4.112

SOURCES.—US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

NOTE.—This table shows inequality measures for the following series: the full ASEC sample, ASEC earnings only for respondents weighted by inverse-probability weights for nonresponse, ASEC earnings only for respondents using the copula selection model, and DER earnings for nonrespondents and ASEC earnings for respondents. See Sec. VI in the text for further details about each series. Unless otherwise noted, all series are estimated using the ASEC supplement weight.

50-10 ratios. We present the estimates for the full ASEC including both respondents and imputed nonrespondents, the ASEC for respondents only but using inverse-probability weights to adjust for nonresponse, the ASEC for respondents only using the copula selection model, and the benchmark of ASEC for respondents and DER for nonrespondents. Table 7 demonstrates that while the ASEC with IPWs brings the inequality estimates closer to the benchmark compared to the full ASEC, the IPW approach falls short compared to the quantile copula selection model that captures nonrandom selection into response in the tails. Figure 11 presents the annual estimates of the inequality measures, where we see that for several measures (i.e., Gini and 90-50 ratio) our method sometimes exceeds inequality from the benchmark, and in some years falls below, so that on average it aligns closely with the benchmark. The 90-10 and 50-10 ratios using the copula method lie below the benchmark in each year, suggesting that some of the measurement differences between the DER and ASEC at low earnings persist. The copula method still performs better, relative to the benchmark, compared to either ASEC or the ASEC with IPWs.

VII. Conclusion

This paper set out to examine the progress in earnings measurement in the CPS in the more than three decades since the important critique of Lillard et al. (1986). In our analysis we address three questions relying on a unique restricted-access data set that links ASEC household files to administrative earnings tax records. First, how do nonresponse and patterns

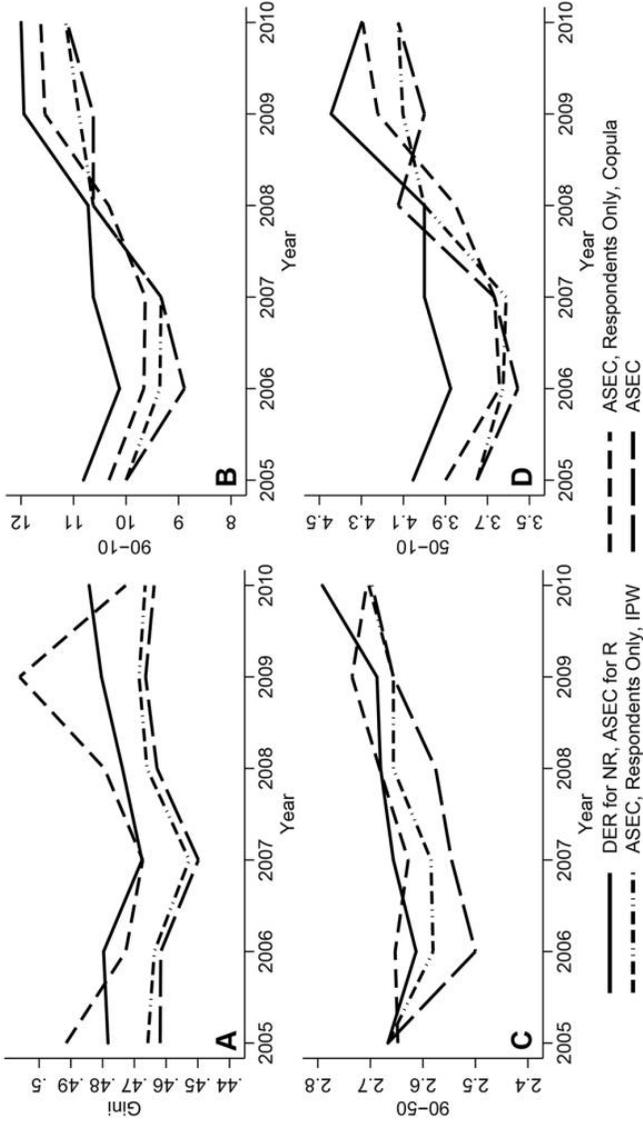


FIG. 11.—Performance of inequality measures in ASEC. *A*, Gini coefficient. *B*, 90-10 ratio. *C*, 90-50 ratio. *D*, 50-10 ratio. Each panel shows inequality measures for the following series: (1) full ASEC sample (“ASEC”); (2) DER earnings for nonrespondents and ASEC earnings for respondents (“DER for NR, ASEC for R”); (3) ASEC earnings for only respondents weighted by inverse-probability weights for nonresponse (“ASEC, Respondents Only, IPW”); and (4) ASEC earnings for only respondents using the copula selection model (“ASEC, Respondents Only, Copula”). See Section VI in the text for further details about each series. Unless otherwise noted, all series are estimated using the ASEC supplement weight. Sources: US Census Bureau Current Population Survey, 2006–11 Annual Social and Economic Supplement; Social Security Administration Detailed Earnings Record, 2005–10.

of nonresponse bias vary across the earnings distribution, and are these patterns similar for women and men (and other groups)? Although levels of nonresponse differ based on gender, race, and ethnicity, U-shaped patterns of nonresponse across the earnings distribution are highly similar across groups. Likewise, we see substantial differences in the level of nonresponse based on the CPS survey month-in-sample and for proxy versus self-respondents, yet we see highly similar U-shaped patterns of nonresponse with respect to earnings for each of these groups. With or without conditioning on covariates, we find a U-shaped nonresponse pattern, with left-tail “strugglers” and right-tail “stars” being least likely to report earnings. Women and men have similar U-shaped nonresponse patterns across the distribution, although men have a higher level of nonresponse.

Second, is nonresponse ignorable? The short answer is no. As stated above, nonresponse is not independent of realized earnings, with or without control for covariates. Relatedly, earnings differ with respect to response status, conditional on covariates. Our third question asks whether there are economic implications of nonrandom nonresponse on estimates of earnings gaps and inequality. We do find small biases at the means for some earnings gaps (e.g., schooling returns and racial/ethnic wage gaps). Gender gaps are slightly understated throughout much of the distribution, but substantively understated in both the left and right tails. Because those with unusually low and high earnings, conditional on measured attributes, are disproportionately missing from the sample, wage equation coefficient estimates on attributes associated with very low (high) earnings are understated in absolute value. Race, gender, and returns to schooling gaps in the tails can be off by as much as 20 percent due to nonresponse. Particularly pronounced are estimates of upper-tail inequality, where nonresponse accounts for one-third to one-half of the 30 percent gap between survey and tax record estimates. Moreover, our evidence from matched panels shows that earnings-growth estimates in the ASEC are substantially understated from imputations. There is trouble in the tails.

The analysis in this paper has implications for researchers using the CPS, as well as similar household data sets such as the American Community Survey (ACS). As emphasized in prior work, even if nonresponse were completely missing at random, severe “match bias” can arise in the estimation of earnings-equation coefficients if researchers include nonrespondents whose earnings are imputed by Census. The simplest and most widely used solution in this case is to throw out imputed earnings. The respondent-only sample can be reweighted by the inverse probability of response, although in practice this typically makes little difference. This easy fix, however, does not provide consistent estimates when there is nonrandom nonresponse. This is particularly true for research focusing on the upper and lower tails of the earnings distribution. Solving the problem of

survey nonresponse is much more difficult absent access to linked administrative data. Progress on this front can continue with additional efforts to link household surveys, tax records, and federal and state-level administrative data on transfers, as recommended recently by the bipartisan Commission on Evidence-Based Policymaking (2017). In the interim, we demonstrated that a flexible copula-based model to correct for nonrandom selection into response offers promise for researchers conducting distributional analysis using the ASEC.

References

- Abowd, John M., and Martha H. Stinson. 2013. "Estimating Measurement Error in Annual Job Earnings: A Comparison of Survey and Administrative Data." *Rev. Econ. and Statis.* 95 (December): 1451–67.
- Alvey, Wendy, and Cynthia Cobleigh. 1980. "Exploration of Differences between Linked Social Security and CPS Earnings Data for 1972." In *Studies from Interagency Data Linkages*, report no. 11, 11–18. Washington, DC: U.S. Dept. Health, Educ., and Welfare.
- Arellano, Manuel, and Stéphane Bonhomme. 2017. "Quantile Selection Models with an Application to Understanding Changes in Wage Inequality." *Econometrica* 85 (January): 1–28.
- Autor, D., L. Katz, and M. Kearney. 2008. "Trends in U.S. Wage Inequality: Revisiting the Revisionists." *Rev. Econ. and Statis.* 90 (May): 300–323.
- Bee, C. Adam, Graton M. R. Gathright, and Bruce D. Meyer. 2015. "Bias from Unit Non-Response in the Measurement of Income in Household Surveys." https://harris.uchicago.edu/files/jsm2015_bgm_unit_non-response_in_cps.pdf.
- Blau, Francine, and Lawrence Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Sources." *J. Econ. Literature* 55 (September): 789–865.
- Bollinger, Christopher R. 1998. "Measurement Error in the Current Population Survey: A Nonparametric Look." *J. Labor Econ.* 16 (July): 576–94.
- Bollinger, Christopher R., and Martin H. David. 2001. "Estimation with Response Error and Nonresponse: Food-Stamp Participation in the SIPP." *J. Bus. and Econ. Statis.* 19 (April): 129–141.
- . 2005. "I Didn't Tell and I Won't Tell: Dynamic Response Error in the SIPP." *J. Appl. Econometrics* 20 (May/June): 563–69.
- Bollinger, Christopher R., and Barry T. Hirsch. 2006. "Match Bias from Earnings Imputation in the Current Population Survey: The Case of Imperfect Matching." *J. Labor Econ.* 24 (July): 483–519.
- . 2013. "Is Earnings Nonresponse Ignorable?" *Rev. Econ. and Statis.* 95 (May): 407–16.
- Bound, John, Charles Brown, Greg J. Duncan, and Willard L. Rodgers. 1994. "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data." *J. Labor Econ.* 12 (July): 345–68.
- Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. "Measurement Error in Survey Data." In *Handbook of Econometrics*, vol. 5, edited by E. E. Leamer and J. J. Heckman, 3705–843. Amsterdam: Elsevier.
- Bound, John, and Alan B. Krueger. 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *J. Labor Econ.* 9 (January): 1–24.

- Buchinsky, Moshe. 1994. "Changes in the U.S. Wage Structure 1963–1987: An Application of Quantile Regression." *Econometrica* 62 (March): 405–58.
- . 1998. "The Dynamics of Changes in the Female Wage Distribution in the USA: A Quantile Regression Approach." *J. Appl. Econometrics* 13 (January): 1–30.
- Burkhauser, Richard V., Shuaizhang Feng, Stephen Jenkins, and Jeff Larrimore. 2012. "Recent Trends in Top Income Shares in the USA: Reconciling Estimates from March CPS and IRS Tax Return Data." *Rev. Econ. and Statis.* 94 (May): 371–88.
- Cameron, Stephen, and Joseph Tracy. 1998. "Earnings Variability in the United States: An Examination Using Matched-CPS Data." Working paper, Fed. Reserve Bank New York.
- Commission on Evidence-Based Policymaking. 2017. "The Promise of Evidence-Based Policymaking." <https://www.cep.gov/cep-final-report.html>.
- David, Martin, Roderick J. A. Little, Michael E. Samuhel, and Robert K. Triest. 1986. "Alternative Methods for CPS Income Imputation." *J. American Statis. Assoc.* 81 (March): 29–41.
- Dixon, John. 2012. "Using Contact History Information to Adjust for Nonresponse in the Current Population Survey." In *2012 JSM Proceedings: Papers Presented at the Joint Statistical Meetings, 1977–1982*. Alexandria, VA: American Statis. Assoc.
- Duncan, Greg J., and Daniel H. Hill. 1985. "An Investigation of the Extent and Consequences of Measurement Error in Labor-Economic Survey Data." *J. Labor Econ.* 3 (October): 508–32.
- Fitzgerald, John, Peter Gottschalk, and Robert Moffitt. 1998. "The Michigan Panel Study of Income Dynamics." *J. Human Resources* 33 (Spring): 251–99.
- Greenlees, John, William Reece, and Kimberly Zieschang. 1982. "Imputation of Missing Values when the Probability of Response Depends on the Variable Being Imputed." *J. American Statis. Assoc.* 77 (June): 251–61.
- Halsey, H. 1978. "Validating Income Data: Lessons from the Seattle and Denver Income Maintenance Experiment." In *Proceedings of the Survey of Income and Program Participation Workshop, Survey Research Issues in Income Measurement: Field Techniques, Questionnaire Design, and Income Validation*. Washington, DC: U.S. Dept. Health, Educ., and Welfare.
- Hardy, Bradley, and James P. Ziliak. 2014. "Decomposing Trends in Income Volatility: The 'Wild Ride' at the Top and the Bottom." *Econ. Inquiry* 52 (January): 459–76.
- Heckman, James J. 1979. "Sample Selection Bias as Specification Error." *Econometrica* 47 (January): 153–61.
- Heckman, James J., and Paul A. LaFontaine. 2006. "Bias-Corrected Estimates of GED Returns." *J. Labor Econ.* 24 (July): 661–700.
- Herriot, R. A., and E. F. Spiers. 1975. "Measuring the Impact on Income Statistics of Reporting Differences between the Current Population Survey and Administrative Sources." *Proceedings, American Statistical Association Social Statistics Section*, 147–58. Alexandria, VA: American Statis. Assoc.
- Hirsch, Barry T., and Edward J. Schumacher. 2004. "Match Bias in Wage Gap Estimates due to Earnings Imputation." *J. Labor Econ.* 22 (July): 689–722.
- Hirsch, Barry T., and John V. Winters. 2016. "Rotation Group Bias in Measures of Multiple Job Holding." *Econ. Letters* 147 (October): 160–63.
- Hokayem, Charles, Christopher Bollinger, and James P. Ziliak. 2015. "The Role of CPS Nonresponse in the Measurement of Poverty." *J. American Statis. Assoc.* 110 (September): 935–45.

- Hurst, Erik, Geng Li, and Ben Pugsley. 2014. "Are Household Surveys Like Tax Forms? Evidence from Income Underreporting of the Self-Employed." *Rev. Econ. and Statis.* 96 (March): 19–33.
- Joe, Harry. 2014. *Dependence Modeling with Copulas*. Boca Raton, FL: CRC Press.
- Kapteyn, Arie, and Jelmer Y. Ypma. 2007. "Measurement Error and Misclassification: A Comparison of Survey and Administrative Data." *J. Labor Econ.* 25 (July): 513–50.
- Kline, Patrick, and Andres Santos. 2013. "Sensitivity to Missing Data Assumptions: Theory and an Evaluation of the U.S. Wage Structure." *Quantitative Econ.* 4 (July): 231–67.
- Krueger, Alan, Alexandre Mas, and Xiaotong Niu. 2017. "The Evolution of Rotation Group Bias: Will the Real Unemployment Rate Please Stand Up?" *Rev. Econ. and Statis.* 99 (May): 258–64.
- Larrimore, Jeff, Richard V. Burkhauser, Shuaizhang Feng, and Laura Zayatz. 2008. "Consistent Cell Means for Topcoded Incomes in the Public Use March CPS (1976–2007)." *J. Econ. and Social Measurement* 33:89–128.
- Lee, Jungmin, and Sokbae Lee. 2012. "Does It Matter Who Responded to the Survey? Trends in the U.S. Gender Earnings Gap Revisited." *Indus. and Labor Relations Rev.* 65 (January): 148–60.
- Lemieux, Thomas. 2006. "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill." *A.E.R.* 96 (June): 461–98.
- Lillard, Lee, James P. Smith, and Finis Welch. 1986. "What Do We Really Know about Wages? The Importance of Nonreporting and Census Imputation." *J.P.E.* 94 (June): 489–506.
- Little, Roderick J. A., and Donald B. Rubin. 2002. *Statistical Analysis with Missing Data*. 2nd ed. Hoboken, NJ: Wiley-Interscience.
- Machado, José A. F., and José Mata. 2005. "Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression." *J. Appl. Econometrics* 20 (May/June): 445–65.
- Marquis, Kent H., and Jeffrey C. Moore. 1990. "Measurement Errors in SIPP Program Reports." In *Proceedings of the 1990 Annual Research Conference*, 721–45. Washington, DC: U.S. Bur. Census.
- Mathiowetz, Nancy A., and Greg J. Duncan. 1988. "Out of Work, Out of Mind: Response Error in Retrospective Reports of Unemployment." *J. Bus. and Econ. Statis.* 6 (April): 221–29.
- Mellow, Wesley, and Hal Sider. 1983. "Accuracy of Response in Labor Market Surveys: Evidence and Implications." *J. Labor Econ.* 1 (October): 331–44.
- Nicholas, Joyce, and Michael Wiseman. 2009. "Elderly Poverty and Supplemental Security Income." *Social Security Bull.* 69 (May): 45–73.
- Peracchi, Franco, and Finis Welch. 1995. "How Representative are Matched Cross-Sections? Evidence from the Current Population Survey." *J. Econometrics* 68 (July): 153–79.
- Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States." *Q.J.E.* 118 (February): 1–39.
- Poterba, James M., and Lawrence H. Summers. 1986. "Reporting Errors and Labor Market Dynamics." *Econometrica* 6 (November): 221–29.
- Reynolds, Jeremy, and Jeffrey B. Wenger. 2012. "He Said, She Said: The Gender Wage Gap According to Self and Proxy Reports in the Current Population Survey." *Social Sci. Res.* 41 (March): 392–411.
- Roemer, Mark. 2002. "Using Administrative Earnings Records to Assess Wage Data Quality in the Current Population Survey and the Survey of Income

- and Program Participation.” Longitudinal Employer-Household Dynamics Program Technical Paper No. TP-2002-22, U.S. Census Bureau.
- Rogers, Willard L., and A. Regula Herzog. 1987. “Covariances of Measurement Errors in Survey Responses.” *J. Official Statis.* 3 (October): 403–18.
- Rubin, Donald B. 1976. “Inference and Missing Data (with Discussion).” *Biometrika* 63 (3): 581–92.
- Vella, Francis. 1998. “Estimating Models with Sample Selection Bias: A Survey.” *J. Human Resources* 33 (Winter): 127–69.
- Welniak, Edward J. 1990. “Effects of the March Current Population Survey’s New Processing System on Estimates of Income and Poverty.” In *Proceedings of the American Statistical Association*. Alexandria, VA: American Statis. Assoc.