Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: http://www.tandfonline.com/loi/uasa20

The Role of CPS Nonresponse in the Measurement of Poverty

Charles Hokayem, Christopher Bollinger & James P. Ziliak

To cite this article: Charles Hokayem, Christopher Bollinger & James P. Ziliak (2015) The Role of CPS Nonresponse in the Measurement of Poverty, Journal of the American Statistical Association, 110:511, 935-945, DOI: 10.1080/01621459.2015.1029576

To link to this article: http://dx.doi.org/10.1080/01621459.2015.1029576

Accepted author version posted online: 16 Apr 2015.
Published online: 07 Nov 2015.

The Role of CPS Nonresponse in the Measurement of Poverty

Charles HOKAYEM, Christopher BOLLINGER, and James P. ZILIAK

The Current Population Survey Annual Social and Economic Supplement (CPS ASEC) serves as the data source for official income, poverty, and inequality statistics in the United States. There is a concern that the rise in nonresponse to earnings questions could deteriorate data quality and distort estimates of these important metrics. We use a dataset of internal ASEC records matched to Social Security Detailed Earnings Records (DER) to study the impact of earnings nonresponse on estimates of poverty from 1997–2008. Our analysis does not treat the administrative data as the "truth"; instead, we rely on information from both administrative and survey data. We compare a "full response" poverty rate that assumes all ASEC respondents provided earnings data to the official poverty rate to gauge the nonresponse bias. On average, we find the nonresponse bias is about 1.0 percentage point.

KEY WORDS: Administrative data; Hot deck; Imputation; Nonrandom selection

1. INTRODUCTION

The Current Population Survey Annual Social and Economic Supplement (CPS ASEC) serves as the official source of income and poverty statistics for the United States. ASEC respondents may be reluctant to answer income questions, or indeed any questions, out of concern for response confidentiality, or they may just have insufficient knowledge of the answers (Groves 2001; Korinek, Mistiaen, and Ravallion 2007). As seen in Figure 1, the nonresponse rate for ASEC earnings among workers (both item nonresponse and supplement nonresponse) has risen dramatically since the early 1990s. The earnings imputation rate has reached 20%, and nonresponse of the entire ASEC supplement adds an additional 10 percentage points to make total nonresponse about 30% in a typical year over the past decade. Rates of item nonresponse for other earnings (e.g., self-employment) trended upward in the 1990s, but they only contribute 1–2 percentage points per year, implying that 95% of earnings nonresponse is due to wage and salary workers. Earnings accounts for over 80% of total income in national income accounts; thus, failure to accurately measure it may significantly bias estimates of the income distribution.

This article assesses whether and to what extent there is bias in official poverty rates caused by earnings nonresponse. The poverty rate, which has been measured consistently since the late 1960s, is not only the key statistical barometer of the well being of low-income families in the United States, but also is used in allocating billions of dollars annually in intergovernmental transfers for nearly 40 federal programs (Citro and Michael 1995; Ziliak 2006; Gabe 2007; Meyer and Sullivan 2012; Short 2013). Thus, knowledge of potential bias from earnings nonresponse is important as it could have substantive budgetary implications. We note that bias could also arise from nonresponse to the initial CPS interview (unit nonresponse) or in other income sources used in constructing the poverty rate such as government transfer programs or private nonlabor income (e.g., retirement, rent/interest/dividends). Initial interview noncontact/refusal rates are in the range 8%–9% (Dixon 2012), while rates of nonresponse in the ASEC among other income sources used in constructing poverty estimates generally range from 0.5% to 4% depending on source, and thus are much less common than earnings imputation.

The current approach of the U.S. Census Bureau is to retain earnings nonrespondents in the sample and to assign them earnings via a matched "donor" with similar demographic characteristics using a sequential "hot deck" procedure (Little and Rubin 2002; Turek et al. 2009). The advantage of this approach is that with weights the sample retains population representativeness, and there may be efficiency gains from retaining the whole sample (Andridge and Little 2010). However, the hot deck procedure may bias estimates of population statistics if the missing at random (MAR) assumption does not hold (Bollinger and Hirsch 2013). Hirsch and Shumacher (2004) and Bollinger and Hirsch (2006) studied the hot deck procedure in both the ASEC and the CPS Outgoing Rotation Group, and showed the hot deck procedure causes earnings regression parameters to be biased even when the MAR assumption holds. In the event that

Charles Hokayem is Visiting Assistant Professor of Economics, Centre College, Danville, KY 40422 (E-mail: *charles.hokayem@centre.edu*). Christopher Bollinger is Professor, Department of Economics, University of Kentucky, Lexington, KY 40506-0034 (E-mail: *crboll@uky.edu*). James P. Ziliak is Chair, Department of Economics, and Director, Center for Poverty Research, University of Kentucky, Lexington, KY 40506-0047 (E-mail: *jziliak@uky.edu*).
We thank Martha Stinson, Graton Gathright, and Gary Benedetto for their help in understanding the Detailed Earnings Record files. The matched ASEC-DER file was made possible by funding to Census from the Office of the Assistant Secretary for Planning and Evaluation in the U.S. Department of Health and Human Services under the direction of Joan Turek at ASPE and Chuck Nelson and Ed Welniak at Census. We thank the editor, two anonymous referees, Ed Welniak, Trudi Renwick, Chuck Nelson, Adam Bee, and session participants at the 2012 Joint Statistical Meetings, the 2013 Society of Labor Economists Meetings, 2013 Institute for Research on Poverty Summer Workshop, and University of Kentucky for helpful comments. We also thank the Summer at Census program for support. Much of this article was written while Hokayem was employed by the U.S. Census Bureau. The views expressed in this research, including those related to statistical, methodological, technical, or operational issues, are solely those of the authors and do not necessarily reflect the official positions or policies of the Census Bureau. The authors accept responsibility for all errors.

Color versions of one or more of the figures in the article can be found online at *www.tandfonline.com/r/jasa*.

Journal of the American Statistical Association
September 2015, Vol. 110, No. 511, Applications and Case Studies
DOI: 10.1080/01621459.2015.1029576

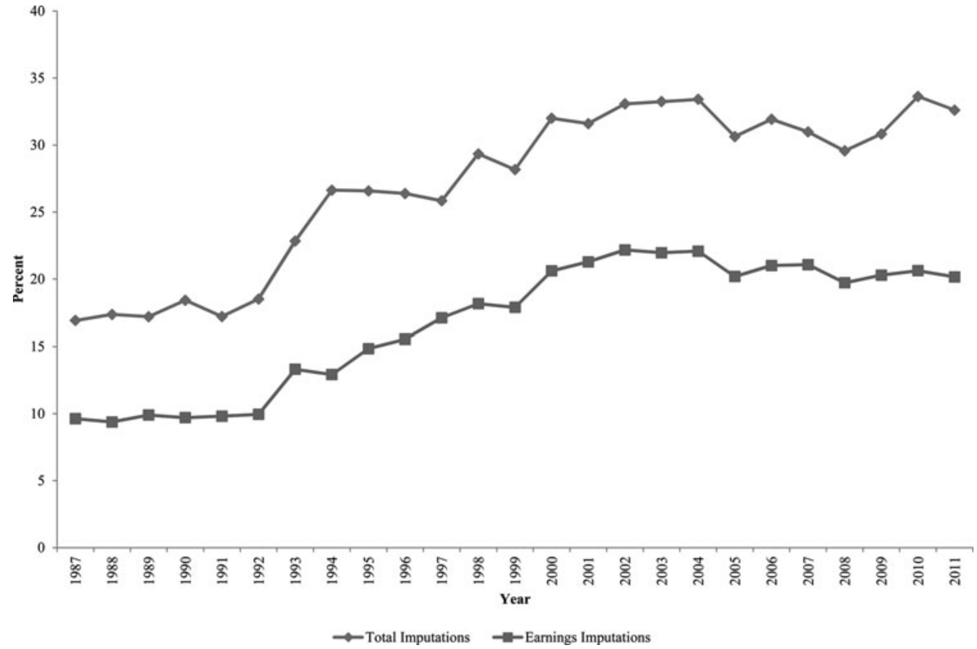

Figure 1. Trends in earnings and total (item + supplement) imputations in the ASEC among workers. Source: Authors' calculations; U.S. Census Bureau, Current Population Survey, 1988–2012 Annual Social and Economic Supplement.

the researcher is using the exact variable definitions in their regression models as employed by Census for the hot deck, there is no bias under MAR. However, Hirsch and Schumacher (2004) and Bollinger and Hirsch (2006) pointed out that this is rare because of the coarseness of categories used in the hot deck. Given the bias in regression parameters, there is a possibility the hot deck procedure could bias estimates of statistics derived from income such as poverty rates.

We propose a new approach to address the effect of earnings nonresponse on the measurement of poverty. Similar to the hot deck approach, we seek the missing counterfactual owing to nonresponse: what would the poverty rate be if all nonrespondents reported their earnings? To estimate what we call the "full-response poverty rate," we assemble a proprietary dataset of internal ASEC records matched to Social Security Detailed Earnings Records (DER) that covers survey years 1998–2009 and allows for the systematic study of long-term trends in income imputation and poverty rates. The DER file contains earnings from all jobs reported on a worker's W-2 forms as well as income from self-employment subject to Federal Insurance Contributions Act (FICA) and/or Self-Employment Contributions Act (SECA) taxation. The DER data are central to our analysis; however, our procedure does not treat the administrative data as correct or the "truth," and instead uses ASEC earnings as the baseline. While some research on wages have treated administrative records like the DER as correct (see Bound and Krueger 1991; Bollinger 1998), these analyses typically attempted to remove individuals whose characteristics (industry and occupation) were likely to indicate substantial under-the-table earnings. This approach will not work in estimation of poverty rates, since all individuals and families must be included. Moreover, recent research (Roemer 2002; Abowd and Stinson 2013) has suggested that this is not necessarily an appropriate approach, and, in fact, the alleged "over-reporting" of CPS earnings among low-income persons may reflect actual earnings not in the DER such as unreported and/or uncovered earnings (both legal and illegal).

There is the possibility that earnings reports in the ASEC are also underreports if the respondent suffers from recall bias or also feels compelled to shelter earnings from the Census field representative, perhaps over confidentiality concerns. We are not aware of any evidence on whether income sheltering is more prevalent in the DER or ASEC, and if they are the same then we still would expect earnings reports in the ASEC to exceed the DER at the low-end of the distribution because some sources are not required to be reported to tax authorities. Thus, our approach is to control for differences between the ASEC and DER in coverage of different income categories, some of which do not require reporting to the tax authorities in the DER, and some of which are required to be reported, but are not, holding constant any other forms of measurement error in either survey.

Three major issues arise in establishing the full-response poverty series: the DER data are at the individual level, whereas the Census poverty rate is a family-level construct; the DER data differ in content from the ASEC data; and not all nonrespondents are matched to the DER records. For the first issue, we match DER and ASEC data at the individual level and then use family identifiers to aggregate DER earnings to the family level. For the second two issues, we address them most simply by comparing ASEC poverty rates to DER poverty rates for those who both report earnings to the ASEC and are matched to the DER. This provides a simple correction to the DER poverty rates for those who fail to report earnings in the ASEC. Similarly, we can compare ASEC poverty rates for those who are matched to the DER and those who are not. Doing so provides a correction to account for nonrespondents who are not matched. Together, the corrections provide an estimate of the poverty rate that would emerge in the absence of earnings nonresponse.

We compare our approach to several alternatives, drawing out important differences in the various assumptions underlying the reasons for the missing data and the attendant implications for the measurement of poverty. For example, Nicholas and Wiseman (2009, 2010) and Turek et al. (2012), each examined the effect of nonresponse on poverty in select years using what we call a "plug-in" approach. That is, they replace ASEC earnings with DER earnings, or the maximum of ASEC and DER earnings, to construct an alternative poverty series, finding little effect of earnings nonresponse on poverty. In addition to our new method of accounting for nonresponse, our study differs from Nicholas and Wiseman (2009, 2010), and Turek et al. (2012), in several ways. First, we examine a longer time series. Second, we examine the entire poverty universe and not just workers. Third, consistent with the Census construction of poverty as a family concept, we derive measures of nonresponse at the family level. Fourth, we distinguish the contribution of matched versus unmatched, respondent versus nonrespondent, and working versus nonworking families to the poverty rate.

Our results suggest that the assumption of MAR, even conditional on known characteristics, is not valid in earnings data in the ASEC. Hence, any correction that assumes missing completely at random or MAR, such as the hot deck procedure, is likely to be biased. We show that the ASEC underestimates the rate of poverty by an average of about 1.0 percentage point, or roughly 3 million persons.

2. POVERTY, NONRESPONSE, AND LINKED SURVEY AND ADMINISTRATIVE DATA

The official poverty rate is based both on the actual earnings of those persons who respond to the ASEC earnings questions and the imputed earnings of those persons who do not respond to the ASEC earnings questions. The poverty rate can be written as a weighted average of these two groups:

$$P^C = P_{\mathrm{R}}^{\mathrm{ASEC}} * \Pr\{\mathrm{R}\} + P_{\mathrm{NR}}^{\mathrm{ASEC}} * \Pr\{\mathrm{NR}\}, \qquad (1)$$

where P^C is the official Census poverty rate, $P_{\mathrm{R}}^{\mathrm{ASEC}}$ is the poverty rate among respondents (R) using ASEC earnings, $\Pr\{\mathrm{R}\}$ is the probability of earnings response, $P_{\mathrm{NR}}^{\mathrm{ASEC}}$ is the poverty rate among nonrespondents (NR) using ASEC earnings, and $\Pr\{\mathrm{NR}\}$ is the probability of nonresponse. The ASEC data provide consistent estimates of three of the terms on the right-hand side: $P_{\mathrm{R}}^{\mathrm{ASEC}}$, $\Pr\{\mathrm{R}\}$, and $\Pr\{\mathrm{NR}\}$. The term $P_{\mathrm{NR}}^{\mathrm{ASEC}}$ is not measured in the ASEC data for earnings nonrespondents and thus must be estimated if these observations are to be retained.

The Census Bureau implements a hot deck procedure to replace the missing earnings to derive an estimate of the poverty rate (Welniak 1990). Specifically, Census uses a sequential match procedure by dividing individuals with missing data into one of 12 allocation groups based on values of socioeconomic variables such as age, gender, race, marital status, and employment status, and then an observation in each allocation group is matched to another observation with complete data (called the donor). If no match is found based on the full set of match variables, then a match variable is dropped and variable definitions are collapsed to be less restrictive until a match is made. A more streamlined procedure based on eight allocation groups is used for whole supplement nonresponse. The sequential hot deck provides a consistent estimate of earnings under the MAR assumption.

A transparent alternative to the hot deck is the naïve bounds cases motivated by the work of Cochran (1977) and Manski (1989). Because the poverty rate falls between 0 and 1, we can place bounds on the official series by making the polar assumptions in Equation (1) that the poverty rate among nonrespondents is 0 or 1. As depicted in Figure 2, making these substitutions results in a lower bound of poverty falling below the official rate by about 3 percentage points on average in our sample, while the upper bound is three times the official rate in a typical year. Although our estimates are for person-level poverty rates, a recent paper by Manski (2014) found similarly wide bounds for family poverty rates in the United States. However, the naïve bounds are extreme because they assume we know nothing about the poverty status of nonrespondents. A significant innovation of our article is that we know a lot about nonreponders as we have administrative tax data on their earnings from the DER.

Specifically, we link CPS ASEC survey data to the Social Security DER data to assess how earnings nonresponse affects the official poverty rate in the ASEC by providing estimates of $P_{\mathrm{NR}}^{\mathrm{ASEC}}$ under plausibly weaker assumptions than used in the hot deck. The sample for the analysis consists of the entire Census poverty universe: that is, all noninstitutionalized families and unrelated individuals aged 15 and older from the ASEC for survey years 1998–2009 (reporting income for 1997–2008). The ASEC is then matched to the DER file, which is an extract of Social Security Administration's Master Earning File (MEF) and includes data on total earnings such as wages and salaries and income from self-employment subject to Federal Insurance Contributions Act (FICA) and/or Self-Employment Contributions Act (SECA) taxation. These earnings are not top-coded either at the FICA contribution limits or by Census as is done for ASEC earnings, though self-employment earnings are only reported if they are nonnegative. There is a separate DER record for each job held by an individual, and thus we aggregate them into a single, annual record. Nonworkers and those who do not pay into Social Security are not in the DER.

Like the match to the DER, imputations of earnings occur at the individual level as well. For our purposes we classify a worker as having imputed earnings if either wages and salary from longest job is imputed, wages and salary from other jobs is imputed, self-employment earnings is imputed, or the whole ASEC supplement is imputed. However, since the official poverty rate in the United States is a family concept, Census sums individual income across all persons in the family to create family income that is compared to the official poverty threshold. Thus, to be consistent with the family definition of poverty, we aggregate individual income nonresponse and match status to create family-level variables. That is, a family is considered imputed if *any* member in the family has imputed earnings, or has the entire supplement imputed. A family is considered matched to the DER data if *all* earners in the family are matched to a DER record. An implication is that it is possible for families to contain no workers, especially among retirees and the disabled, and thus by construction no match with the DER is possible for the family.

Figure 3 depicts trends in the family level ASEC-DER match rate conditional on earners in the family (recall that by

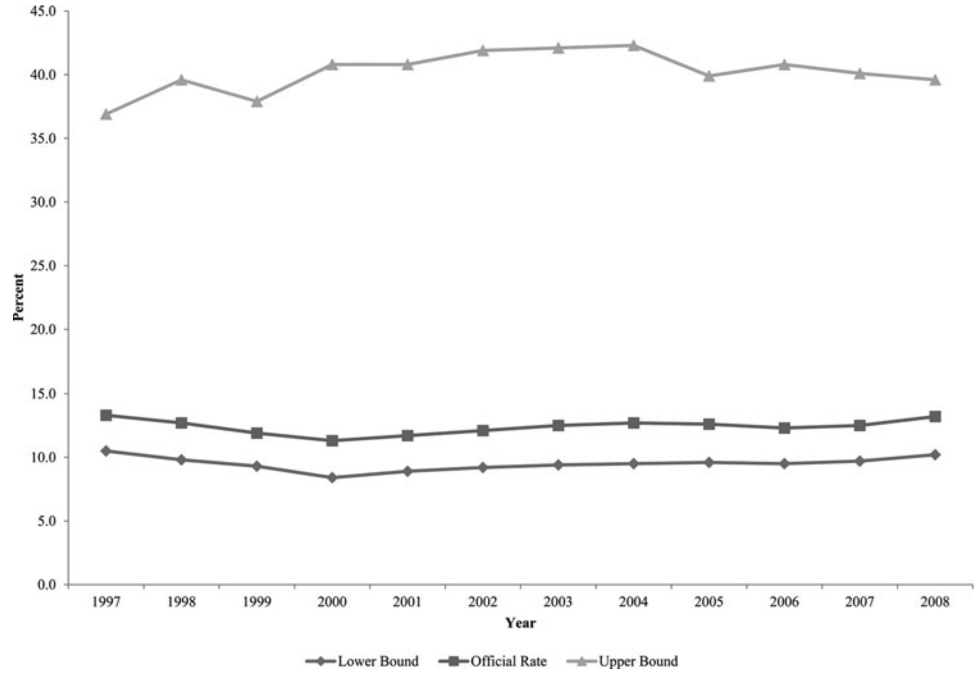

Figure 2. Naïve bounds on the official poverty rate. Source: Authors' calculations; U.S. Census Bureau, Current Population Survey, 1998–2009 Annual Social and Economic Supplement.

construction a family cannot be matched if there are no earners). In 1997, just over 60% of earner families in the ASEC were matched to the DER. We note that this family-level DER match rate of earners is about 10–12 percentage points lower than individual-level match rates. This occurs because about 10% of families have more earners than DER matches, and thus we classify the whole family as nonmatched. However, the match rate rose to 74% starting in 2005 and held steady thereafter. The increase in the rate of ASEC-DER matches most likely occurred because Census changed the consent process for linking to SSA data from "opt-in" to "opt-out," that is, starting in 2005 sample members were automatically enrolled in the link process and had to request that they be removed. Importantly, there is a 20–25 percentage point difference in DER match rates depending on whether the family is a respondent or nonrespondent, highlighting the importance of distinguishing match/nonmatch by respondent status in our full-response poverty rate described below.

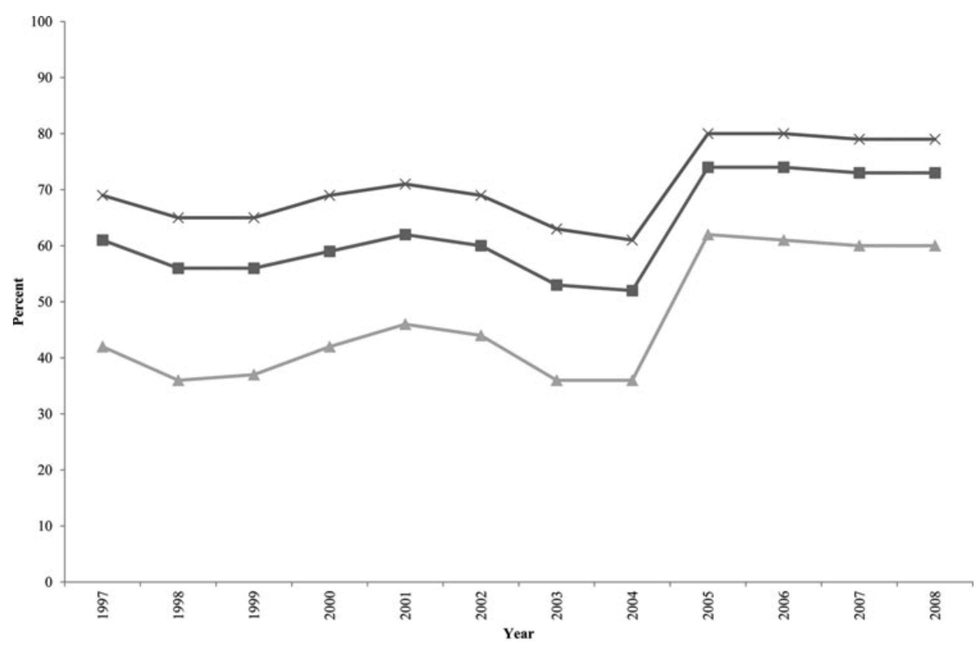

Figure 3. Family level ASEC-DER match rate (earners). Source: Authors' calculations; U.S. Census Bureau, Current Population Survey, 1998–2009 Annual Social and Economic Supplement, and Social Security Administration, Detailed Earnings Record, 1997–2008.

Table 1. Summary statistics (head of family)

	Respondent, DER match		Respondent, DER nonmatch		Nonrespondent, DER match		Nonrespondent, DER nonmatch	
Characteristic	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Age	43.07	0.03	57.11	0.05	45.57	0.06	48.64	0.06
Gender								
Male (%)	55.70	0.12	49.13	0.12	55.17	0.22	54.99	0.18
Female (%)	44.30	0.12	50.87	0.12	44.83	0.22	45.01	0.18
Race								
White (%)	84.83	0.08	83.21	0.09	80.52	0.17	79.81	0.15
Black (%)	10.65	0.07	12.61	0.08	14.42	0.15	14.43	0.13
Other race (%)	4.52	0.05	4.18	0.05	5.06	0.09	5.76	0.08
Marital status								
Married (%)	54.74	0.19	46.43	0.20	58.51	0.35	59.80	0.29
Widowed (%)	3.82	0.07	21.15	0.16	5.07	0.16	8.33	0.16
Separated or divorced (%)	19.31	0.15	16.79	0.15	17.53	0.27	14.89	0.21
Single, never-married (%)	22.13	0.16	15.63	0.14	18.89	0.27	16.97	0.22
Educational attainment								
Less than high school (%)	9.13	0.07	23.28	0.10	11.65	0.14	15.57	0.13
High school completed (%)	28.00	0.10	32.18	0.11	31.46	0.20	32.67	0.17
More than high school (%)	62.87	0.11	44.54	0.12	56.89	0.21	51.77	0.18
Employment status								
Employed (%)	83.14	0.14	34.62	0.18	80.91	0.27	67.82	0.26
Unemployed (%)	3.78	0.07	2.65	0.06	3.22	0.12	2.68	0.09
Not in labor force								
Retired (%)	4.84	0.08	43.66	0.19	6.89	0.18	15.77	0.21
Disabled (%)	1.88	0.05	10.19	0.11	2.41	0.10	4.28	0.11
Other reason (%)	5.73	0.08	8.68	0.11	6.11	0.16	9.13	0.16
Family size	2.51	0.00	2.13	0.00	2.69	0.01	2.69	0.01
Number of related children under 18	0.77	0.00	0.48	0.00	0.74	0.00	0.67	0.00
Official poverty status (%)	6.50	0.12	21.19	0.20	7.06	0.23	11.19	0.23
Family type								
Married couple (%)	53.47	0.11	44.49	0.11	57.07	0.20	57.93	0.16
Female householder, no spouse present (%)	26.74	0.10	36.87	0.11	25.41	0.17	25.34	0.14
Male householder, no spouse present (%)	19.79	0.09	18.64	0.09	17.52	0.15	16.74	0.12
ASEC family earnings ($)	58,417	154	18,164	128	62,220	314	53,614	271
DER family earnings ($)	55,978	321	N/A	N/A	61,974	1521	N/A	N/A
ASEC family income ($)	65,504	162	39,421	130	70,124	328	63,711	281
DER family income ($)	63,065	327	N/A	N/A	69,877	1527	N/A	N/A

NOTE: SE, standard errors (estimated using generalized function parameters).
Source: Authors' calculations; U.S. Census Bureau, Current Population Survey, 1998–2009 Annual Social and Economic Supplement. For information on sampling and nonsampling error, see http://www.census.gov/prod/techdoc/cps/cpsmar13.pdf. Social Security Administration, Detailed Earnings Record, 1997–2008.

Table 1 presents detailed summary statistics of the sample family head based on match and respondent status. Across most demographic characteristics, the differences between matched and nonmatched families are much more pronounced among respondents than nonrespondents. For example, matched respondents are 14 years younger on average than nonmatched respondents, reflecting the fact that the latter group is much more likely to be retired or disabled, while there is only a 3-year age gap between matched and nonmatched nonrespondents. In both cases, the differences are statistically significant at the 5% level. Likewise matched respondents are statistically much less likely to be a high-school dropout or to be living in poverty than nonmatched respondents. These gaps are relatively small among nonrespondents. As a consequence, ASEC earnings and family income are substantially and statistically higher for matched than nonmatched respondents. Interestingly, even though the difference in earnings and income among nonrepondents is comparatively small across match status, the level is higher than among respondents, suggesting that high-income persons are less likely to respond to the ASEC.

3. A "FULL-RESPONSE" MEASURE OF POVERTY

With the linked survey and administrative data we seek to identify the missing counterfactual of what the poverty rate would be had all respondents provided their earnings, and as such to establish what we term as the "full-response" poverty rate, denoted as P^{Full}. This is complicated by the fact that earnings in the DER differ from those in the ASEC and that not all respondents in the ASEC have a linked record in the DER.

Specifically, ASEC earnings reports can differ from DER reports both because not all jobs are covered by Social Security (e.g., railroad workers, teachers in certain states), and thus are not required to be recorded in the DER, and that "under-the-table" earnings could be reported to the ASEC that are not reported to the IRS. There is evidence that under-the-table earnings can comprise a significant portion of the family budget among certain segments of the low-income population (Edin and Lein 1997; Venkatesh 2006). Likewise, ASEC sample members will not be matched to the DER if (a) they did not give consent to be linked to the DER, (b) they had no earnings covered by the DER (legal or illegal), or (c) they did not work for pay. The difference between ASEC and DER reports implies that we need to make an adjustment for measurement differences across the two series, while not being matched to the DER implies we need to make an adjustment for sample composition differences across the series.

Formally, we expand our decomposition of the poverty rate in Equation (1) from two groups to four groups defined by respondent/nonrespondent status and DER match/nonmatch status as

$$P^{C} = P^{\mathrm{ASEC}}_{\mathrm{R,M}} * \Pr\{\mathrm{R}\,\&\,\mathrm{M}\} + P^{\mathrm{ASEC}}_{\mathrm{R,NM}} * \Pr\{\mathrm{R}\,\&\,\mathrm{NM}\} + P^{\mathrm{ASEC}}_{\mathrm{NR,M}} * \Pr\{\mathrm{NR}\,\&\,\mathrm{M}\} + P^{\mathrm{ASEC}}_{\mathrm{NR,NM}} * \Pr\{\mathrm{NR}\,\&\,\mathrm{NM}\}, \quad (2)$$

where the subscript M refers to an ASEC sample member matched to the DER and NM is not matched to the DER. For example, $P^{\mathrm{ASEC}}_{\mathrm{R,M}}$ is the poverty rate of respondents matched to the DER using ASEC earnings, and $\Pr\{\mathrm{R\&M}\}$ is the probability of responding to the ASEC *and* matched to the DER. We observe $P^{\mathrm{ASEC}}_{\mathrm{R,M}}$ and $P^{\mathrm{ASEC}}_{\mathrm{R,NM}}$ in the ASEC regardless of match status, and hereafter we collapse the first two terms in Equation (2) as $P^{\mathrm{ASEC}}_{\mathrm{R}} * \Pr\{\mathrm{R}\}$, which is simply the first term in Equation (1). However, we do not observe $P^{\mathrm{ASEC}}_{\mathrm{NR,M}}$ or $P^{\mathrm{ASEC}}_{\mathrm{NR,NM}}$, and thus use the DER earnings to provide an alternative measure of the earnings for these two unobservable poverty rates.

In Equation (2), we replace $P^{\mathrm{ASEC}}_{\mathrm{NR,M}}$ with $P^{\mathrm{DER}}_{\mathrm{NR,M}}$, which is the poverty rate of matched nonrespondents using the DER as the measure of earnings. To account for measurement differences between the DER and the ASEC, we add a correction for measurement error among matched respondents: $(P^{\mathrm{ASEC}}_{\mathrm{R,M}} - P^{\mathrm{DER}}_{\mathrm{R,M}})$. Putting this together gives an estimator for the term $P^{\mathrm{ASEC}}_{\mathrm{NR,M}}$:

$$\hat{P}^{\mathrm{ASEC}}_{\mathrm{NR,M}} = P^{\mathrm{DER}}_{\mathrm{NR,M}} + (P^{\mathrm{ASEC}}_{\mathrm{R,M}} - P^{\mathrm{DER}}_{\mathrm{R,M}}). \quad (3)$$

We would like to make a similar substitution for $P^{\mathrm{ASEC}}_{\mathrm{NR,NM}}$ with $P^{\mathrm{DER}}_{\mathrm{NR,NM}}$ in the final term of Equation (2), but $P^{\mathrm{DER}}_{\mathrm{NR,NM}}$ will never be observed. If we assume that nonmatched nonrespondents are similar to matched nonrespondents, we could use the estimator in Equation (3): $P^{\mathrm{DER}}_{\mathrm{NR,M}} + (P^{\mathrm{ASEC}}_{\mathrm{R,M}} - P^{\mathrm{DER}}_{\mathrm{R,M}})$. However, as shown in Table 1, workers who are not matched to the DER differ from those who are matched to the DER in both demographic characteristics and in earnings levels. To correct for these differences, we compare the poverty rates based on ASEC family earnings of nonmatched respondents to matched respondents $(P^{\mathrm{ASEC}}_{\mathrm{R,NM}} - P^{\mathrm{ASEC}}_{\mathrm{R,M}})$. Substituting these expressions into the term for $P^{\mathrm{ASEC}}_{\mathrm{NR,NM}}$ gives our estimator,

$$\hat{P}^{\mathrm{ASEC}}_{\mathrm{NR,NM}} = P^{\mathrm{DER}}_{\mathrm{NR,M}} + \left(P^{\mathrm{ASEC}}_{\mathrm{R,M}} - P^{\mathrm{DER}}_{\mathrm{R,M}}\right) + \left(P^{\mathrm{ASEC}}_{\mathrm{R,NM}} - P^{\mathrm{ASEC}}_{\mathrm{R,M}}\right). \quad (4)$$

Our approach here allows for nonresponse to be related not only to demographic characteristics (as does the hot deck procedure), but also to unobservable characteristics and the income level or poverty rate itself. Our approach also allows matching or failure to match to be related to demographic characteristics as well as unobservable characteristics. As such, it also allows the DER and ASEC measures to differ and corrects for those differences. However, we assume that there is no interaction between these three mechanisms. That is, we are assuming that measurement difference nonresponse and nonmatch are additive once we condition on poverty status. In Equations (3) and (4), the term $(P^{\mathrm{ASEC}}_{\mathrm{R,M}} - P^{\mathrm{DER}}_{\mathrm{R,M}})$ implies we are assuming that measurement difference between the DER and the ASEC does not differ between respondents and nonrespondents. That is, if nonrespondents were to respond, the differences between their DER record and their ASEC response would be similar to the differences between current respondents DER and ASEC reports. In Equation (4), the term $(P^{\mathrm{ASEC}}_{\mathrm{R,NM}} - P^{\mathrm{ASEC}}_{\mathrm{R,M}})$ implies we are assuming that the differences in poverty rates between the matched and nonmatched populations are the same in both the DER and the ASEC. The first set of assumptions, which allow Equation (3) to provide an estimate of the term $P^{\mathrm{ASEC}}_{\mathrm{NR,M}}$, is weaker than the MAR assumption used in the hot deck procedure. Indeed, if the MAR assumption holds, the results in Equation (3) should be equivalent (up to sampling error) to using the hot deck procedure. The second set of assumptions regarding the match is not required for the hot deck procedure since the hot deck does not involve matching to the DER. However, if MAR holds, and nonmatched at random were also to hold, then again our procedure should be similar to the hot deck. Our procedure allows both of these assumptions to fail, but does not allow nonmatch, nonresponse, and measurement error processes to covary, conditional on poverty status.

Substituting (3) and (4) into (2) gives the full-response poverty rate as

$$P^{\mathrm{Full}} = P^{\mathrm{ASEC}}_{\mathrm{R}} * \Pr\{\mathrm{R}\} + \left(P^{\mathrm{DER}}_{\mathrm{NR,M}} + \left(P^{\mathrm{ASEC}}_{\mathrm{R,M}} - P^{\mathrm{DER}}_{\mathrm{R,M}}\right)\right) * \Pr\{\mathrm{NR\&M}\} + \left(P^{\mathrm{DER}}_{\mathrm{NR,M}}\left(P^{\mathrm{ASEC}}_{\mathrm{R,M}} - P^{\mathrm{DER}}_{\mathrm{R,M}}\right) + (P^{\mathrm{ASEC}}_{\mathrm{R,NM}} - P^{\mathrm{ASEC}}_{\mathrm{R,M}})\right) * \Pr\{\mathrm{NR\&NM}\}. \quad (5)$$

It is worth emphasizing that the expression in (5) consists solely of *observed* data—both survey and administrative—and thus serves as our estimate of the benchmark poverty rate in the United States. Recall, though, that a family is considered matched to the DER data if *all* earners in the family are matched to a DER record. An implication is that it is possible for families to contain no workers, especially among retirees and the disabled, and thus by construction no match with the DER is possible for the family. Consequently, we have to modify our full poverty rate in Equation (5) to be conditional on earner status in the family, that is,

$$P^{\mathrm{Full}} = \left(P^{\mathrm{Full}}|\mathrm{earner} \geq 1\right) \Pr\{\mathrm{earner} \geq 1\} + \left(P^{\mathrm{Full}}|\mathrm{earner} = 0\right) \Pr\{\mathrm{earner} = 0\}. \quad (6)$$

4. RESULTS

Table 2 presents our benchmark full-response poverty estimates from Equation (6), along with the standard error computed

Table 2. Estimates of full-response poverty rate

Year	P^{Full} (%)	SE	P^{C} (%)	SE
1997	14.0	(0.216)	13.3	(0.211)***
1998	13.7	(0.213)	12.7	(0.206)***
1999	12.5	(0.203)	11.9	(0.199)***
2000	11.9	(0.198)	11.3	(0.193)***
2001	12.5	(0.143)	11.7	(0.139)***
2002	13.3	(0.146)	12.1	(0.140)***
2003	13.3	(0.146)	12.5	(0.142)***
2004	13.8	(0.147)	12.7	(0.142)***
2005	13.4	(0.145)	12.6	(0.141)***
2006	13.3	(0.143)	12.3	(0.139)***
2007	13.9	(0.145)	12.5	(0.139)***
2008	14.2	(0.146)	13.2	(0.142)***

NOTE: P^{Full} is our full response poverty measure from Equation (6) in the text; P^{C} is the official Census poverty measure; standard errors (SE) in parentheses are estimated using generalized function parameters; significance reflects statistical test for comparison to P^{Full}; $^{***}p < 0.01$.

Source: Authors' calculations; U.S. Census Bureau, Current Population Survey, 1998–2009 Annual Social and Economic Supplement. Social Security Administration, Detailed Earnings Record, 1997–2008.

using generalized variance parameters, and compares it to the official Census poverty rate in each year. The table makes clear that the official poverty rate is statistically significantly lower in each year, averaging about 1.0 percentage point lower than the full-response benchmark, and this gap seems to have widened over time. Given that the U.S. population averaged about 286 million people over our sample period, our results suggest that the official rate is undercounting the number of poor persons by roughly 2–3 million per year compared to a rate in which all sample members respond to the earnings questions. In our earlier working paper (Hokayem, Bollinger, and Ziliak 2014), we explore whether the difference between the full-response poverty rate and the official rate is driven by certain demographic groups. There we show that on average families with children have full poverty rates 1.4 percentage points higher than the official rate, families headed by a female have full rates 1 percentage point higher, and families headed by a nonwhite or nonblack (other race) have full rates 1.5 percentage points higher on average. There is no difference among the elderly as most are not in the labor force and thus do not contribute to earnings nonresponse (Hokayem, Bollinger, and Ziliak 2014).

4.1 Components of Full-Response Poverty

In Tables 3(a) and 3(b), we explore in finer detail the components of the full-response poverty rate that might shed light on why the official poverty rate is systematically lower. In Table 3(a), we present the components of Equation (6) for families with at least one earner. The numbers in columns (1), (4), and (6) sum up to the number in column (7), subject to rounding error. Of particular note is column (3) where we compare the poverty rates of matched respondents using ASEC earnings versus DER earnings. The difference is negative and statistically significant in each year, which means that ASEC earnings are higher and poverty rates lower than in the DER, suggesting that ASEC earnings captures income sources not reported to the DER either because they are not taxable or they are "under the table." On the other hand, in column (5) we report the difference in poverty rates of nonmatched respondents and matched respondents in the ASEC. This difference is statistically significant and positive, suggesting that nonmatched respondent families are systematically poorer than matched families. This correction grows over time, especially after 2004 when the Census changed from the "opt-in" to the "opt-out" consent of being linked. In Table 3(b), we present the same calculations for nonearner families. Note that most of the terms are zero since by construction nonearner families are not matched to the DER. Also notable is the fact that the poverty rates of nonearner families are more than three times higher and statistically different than earner families. The full-response poverty rate reported in Table 2 is much closer to the earner rates because the probability of a family containing at least one worker averages over 85% in each year so that the earner sample receives nearly six times more weight in the full poverty calculation.

4.2 Alternative Approaches

Against the benchmark in (6) we compare two alternative point-estimators of poverty rates using the DER, and a third that refines the naïve bounds presented in Figure 2. The first DER alternative we call the "plug-in" estimate of poverty. Specifically, we replace ASEC earnings with DER earnings only for those nonrespondents with a DER match (bold term in Equation (7)), and use reported ASEC earnings for respondents and (hot deck) ASEC earnings for persons without a DER match:

$$P_{\text{NR}}^{\text{Plug-in}} = P_{\text{R,M}}^{\text{ASEC}} * \Pr\{\text{R} \,\&\, \text{M}\} + P_{\text{R,NM}}^{\text{ASEC}} * \Pr\{\text{R} \,\&\, \text{NM}\} + \boldsymbol{P_{\text{NR,M}}^{\text{DER}} * \Pr\{\text{NR} \,\&\, \text{M}\}} + P_{\text{NR,NM}}^{\text{ASEC}} * \Pr\{\text{NR} \,\&\, \text{NM}\}. \tag{7}$$

The logic here is that DER earnings for the actual worker dominate imputed earnings from an unrelated person, especially if the MAR assumption is violated (either because the imputation algorithm uses too sparse a set of demographics, or there is selection on unobservables).

In Table 4, we reproduce the full-response poverty rate in column (1) and in column (2) we report estimates of the plug-in poverty rate using the DER for matched nonrespondents. We see that in each year the plug-in series is statistically much lower than the full-response poverty rate (but higher than the official Censusrate). In our working paper, we also considered the case where we replace ASEC earnings with the DER for both matched respondents and nonrespondents (Hokayem, Bollinger, and Ziliak 2014). This approach implicitly assumes that survey reports in the ASEC are mismeasured, and the DER records provide a superior measure of earnings. This may not be true, however, both because some earnings reported in CPS are not taxable, and some earnings may be reported to the Census but not to the IRS, especially self-employment earnings and "under-the-table" earnings (Bound and Krueger 1991; Bollinger 1998; Roemer 2002). Hokayem, Bollinger, and Ziliak (2014) reported that this approach yields poverty rates that are generally higher than the full-response poverty rate. We also examined the approach used by Turek et al. (2012) and Nicholas and Wiseman (2009, 2010) whereby the maximum of ASEC and DER earnings is used, which makes the strong assumption that measurement error in the ASEC is always negative (not simply an underreport of true earnings on average, but never an

Table 3a. Components of full-response poverty conditional on at least one earner (P^{Full}|Earner> = 1)

Year	(1) $P_{\text{R}}^{\text{ASEC}} * \Pr\{\text{R}\}$	(2) $P_{\text{NR,M}}^{\text{DER}}$	(3) $(P_{\text{R,M}}^{\text{ASEC}} - P_{\text{R,M}}^{\text{DER}})$	(4) ((2)+(3))* $\Pr\{\text{NR\&M}\}$	(5) $(P_{\text{R,NM}}^{\text{ASEC}} - P_{\text{R,M}}^{\text{ASEC}})$	(6) ((2)+(3)+(5))* $\Pr\{\text{NR\&DM}\}$	(7) (P^{Full}|Earner> = 1)
1997	6.8	12.0	−1.4	1.2	2.9	2.3	10.4
1998	6.4	11.6	−1.4	1.2	2.7	2.7	10.3
1999	6.2	10.2	−1.4	1.0	2.9	2.4	9.5
2000	5.2	10.0	−1.4	1.3	2.5	2.4	8.8
2001	5.5	10.3	−1.8	1.3	3.1	2.3	9.1
2002	5.5	11.2	−1.9	1.4	3.9	2.8	9.7
2003	5.5	11.5	−2.1	1.2	1.6	2.6	9.3
2004	5.5	12.9	−1.8	1.4	1.3	3.0	9.9
2005	5.5	10.6	−2.0	1.7	7.2	2.1	9.4
2006	5.6	10.1	−1.8	1.7	6.7	2.1	9.4
2007	5.8	10.9	−1.8	1.8	8.0	2.5	10.0
2008	6.2	10.2	−1.7	1.6	8.1	2.3	10.2

NOTE: See the text and Equations (5) and (6) for description of the poverty rates; columns (1), (4), and (6) sum to (P^{Full}|Earner> = 1) in column (7) subject to rounding error.
Source: Authors' calculations; U.S. Census Bureau, Current Population Survey, 1998–2009 Annual Social and Economic Supplement. Social Security Administration, Detailed Earnings Record, 1997–2008.

overreport). This is particularly strong when considering the nonrespondents. In these cases, the DER is used only when it exceeds the hot deck imputation. Since the hot deck is a random match, we expect it to contain differences that are both positive and negative. It also makes the somewhat weaker assumption that earnings in the DER are always an understatement of true earnings. Since the procedure also uses the hot deck earnings for individuals who are not matched to the DER, this assumes that nonmatched, nonrespondents are missing and unmatched at random. This approach, by construction, will necessarily result in a lower poverty rate than that achieved by the hot deck procedure. Taken together, these results suggest that the MAR assumption does not appear to hold: nonrespondents are more likely to be in poverty than their matched counterparts.

The second alternative approach with the DER, which we call the "probability" approach, is to use the DER and other sample information to predict the poverty status for those households who are nonrespondents. Formally, we replace the terms $P_{\text{NR,M}}^{\text{ASEC}}$ and $P_{\text{NR,NM}}^{\text{ASEC}}$ from Equation (2) with aggregates from the prediction of a model of poverty status estimated on the respondent sample. Modeling $P_{\text{NR,M}}^{\text{ASEC}}$, the poverty status of nonrespondents who are matched, is relatively straightforward. We use a simple probit model, where the dependent variable is the poverty indicator from the ASEC, and explanatory variables include the DER earnings and ASEC demographic characteristics of the household. We estimate the model on the subsample of matched respondents. We then use the model parameters to predict the probability of a household being in poverty for the matched nonrespondent sample. Since we condition on the DER earnings, this should capture any selection on earnings between respondents and nonrespondents.

Modeling $P_{\text{NR,NM}}^{\text{ASEC}}$, the poverty rate for nonmatched nonrespondents, is significantly more difficult since we have no measure of earnings. We attempt to model the process using the

Table 3b. Components of full-response poverty conditional on no earners (P^{Full}|Earner = 0)

Year	(1) $P_{\text{R}}^{\text{ASEC}} * \Pr\{\text{R}\}$	(2) $P_{\text{NR,M}}^{\text{DER}}$	(3) $(P_{\text{R,M}}^{\text{ASEC}} - P_{\text{R,M}}^{\text{DER}})$	(4) ((2)+(3))* $\Pr\{\text{NR\&M}\}$	(5) $(P_{\text{R,NM}}^{\text{ASEC}} - P_{\text{R,M}}^{\text{ASEC}})$	(6) ((2)+(3)+(5))* $\Pr\{\text{NR\&DM}\}$	(7) (P^{Full}|Earner = 0)
1997	32.2	0.0	0.0	0.0	35.8	3.6	35.8
1998	30.8	0.0	0.0	0.0	34.6	3.8	34.6
1999	29.2	0.0	0.0	0.0	32.3	3.1	32.3
2000	29.2	0.0	0.0	0.0	32.3	3.1	32.3
2001	30.1	0.0	0.0	0.0	33.9	3.8	33.9
2002	31.3	0.0	0.0	0.0	35.1	3.8	35.1
2003	32.1	0.0	0.0	0.0	36.2	4.1	36.2
2004	32.2	0.0	0.0	0.0	36.1	3.9	36.1
2005	33.3	0.0	0.0	0.0	36.9	3.6	36.9
2006	32.7	0.0	0.0	0.0	36.5	3.8	36.5
2007	33.5	0.0	0.0	0.0	36.8	3.3	36.8
2008	33.4	0.0	0.0	0.0	37.0	3.6	37.0

NOTE: See the text and Equations (5) and (6) for description of the poverty rates; columns (1), (4), and (6) sum to (P^{Full}|Earner = 0) in column (7) subject to rounding error.
Source: Authors' calculations; U.S. Census Bureau, Current Population Survey, 1998–2009 Annual Social and Economic Supplement. Social Security Administration, Detailed Earnings Record, 1997–2008.

Table 4. Alternative approaches to estimating poverty rates (%) using ASEC and/or DER data

	(1)		(2)		(3)		(4)	
Year	P^{Full}	SE	$P_{\text{NR}}{}^{\text{Plug-in}}$	SE	P^{Prob}	SE	P^{IPW}	SE
1997	14.0	(0.216)	13.5	(0.213)***	13.6	(0.213)***	13.5	(0.222)***
1998	13.7	(0.213)	12.9	(0.208)***	13.1	(0.209)***	13.2	(0.219)***
1999	12.5	(0.203)	12.1	(0.200)***	12.2	(0.201)***	12.3	(0.211)**
2000	11.9	(0.198)	11.6	(0.196)***	11.6	(0.195)***	11.6	(0.206)***
2001	12.5	(0.143)	12.2	(0.142)***	12.1	(0.141)***	12.3	(0.150)***
2002	13.3	(0.146)	12.6	(0.143)***	12.5	(0.142)***	12.7	(0.151)***
2003	13.3	(0.146)	12.8	(0.143)***	12.7	(0.143)***	13.0	(0.152)***
2004	13.8	(0.147)	13.2	(0.144)***	12.9	(0.143)***	13.2	(0.152)***
2005	13.4	(0.145)	13.2	(0.144)***	13.6	(0.145)***	13.1	(0.151)***
2006	13.3	(0.143)	13.0	(0.142)***	13.4	(0.144)	13.0	(0.149)***
2007	13.9	(0.145)	13.2	(0.142)***	13.5	(0.144)***	13.2	(0.150)***
2008	14.2	(0.146)	13.7	(0.144)***	14.2	(0.146)	13.8	(0.151)***

NOTE: See the text for description of the poverty rates; standard errors (SE) in parentheses are estimated using generalized function parameters; significance reflects statistical test for comparison to P^{Full}; $^{***}p < 0.01$, $^{**}p < 0.05$.
Source: Authors' calculations; U.S. Census Bureau, Current Population Survey, 1998–2009 Annual Social and Economic Supplement. Social Security Administration, Detailed Earnings Record, 1997–2008.

sample of nonmatched respondents; that is, a sample where the poverty status is observed. We use a probit model, similar to the structure above, but control only on the ASEC demographic characteristics. To account for the potential sample selection into ASEC response, we use a standard Heckman selection correction. Following Bollinger and Hirsch (2013), we use ASEC month in sample as our exclusion restriction. The strong assumptions necessary for the Heckman approach to be valid, however, is the most tenuous aspect of this approach. We then use the model to predict (including the selection term) the poverty status for the nonmatched nonrespondents.

The results of this approach are found in column (3) of Table 4, labeled P^{Prob}. The probability approach yields estimates that are slightly lower than the P^{Full} approach in column (1). The standard errors and tests reveal that in all cases except 2006 and 2008, the probability approach is lower than the P^{Full} estimates. However, these estimates are slightly larger than the plug-in estimates in general (although not statistically significant). The major challenge with both approaches is that we have very little information about nonmatched nonrespondents, and thus the key difference between the "full" approach and the probability approach resides in this term.

In our full-response poverty measure, we address this by including corrections both for measurement differences in earnings between the DER and ASEC data and for sample composition differences between those who are matched to the DER

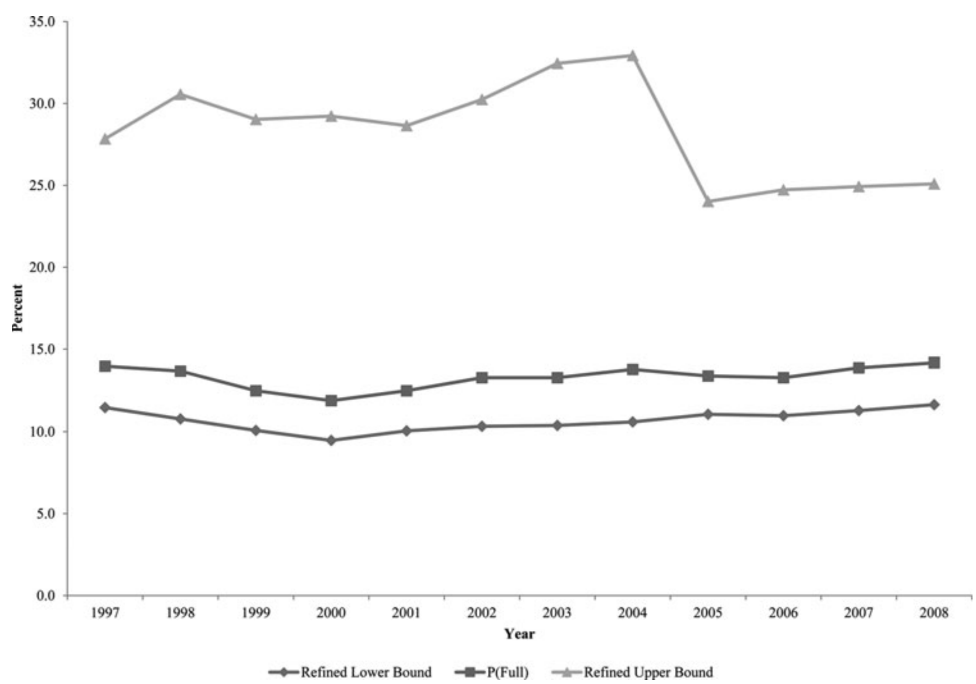

Figure 4. Refined bounds on the full-response poverty rate. Source: Authors' calculations; U.S. Census Bureau, Current Population Survey, 1998–2009 Annual Social and Economic Supplement, and Social Security Administration, Detailed Earnings Record, 1997–2008.

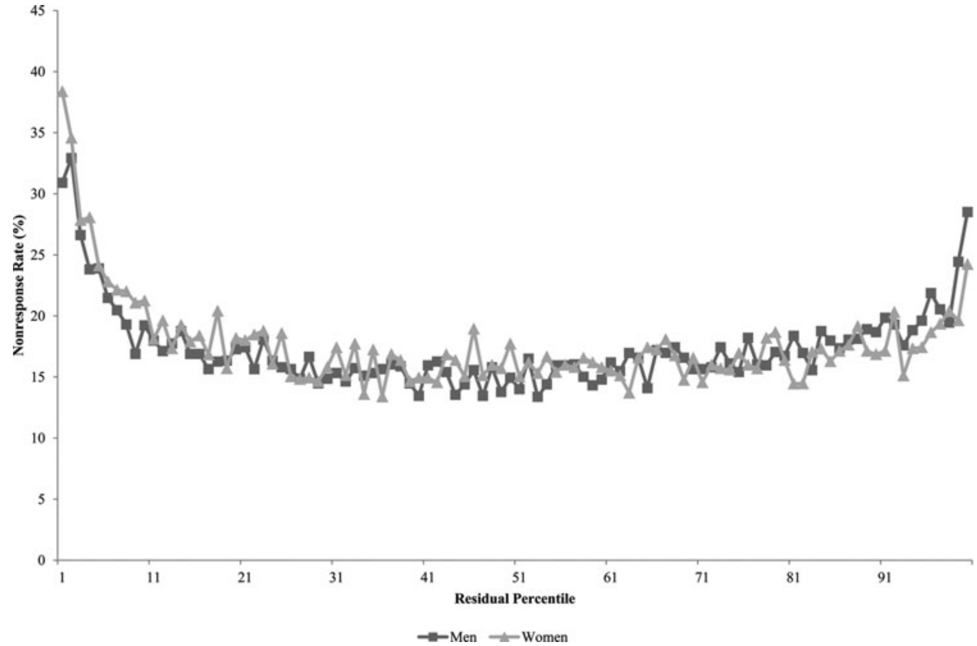

Figure 5. Nonresponse rate across residual log wage distribution of workers in the DER. Source: Authors' calculations; Detailed Earnings Records, Social Security Administration, 2005–2009.

and those not matched. However, we acknowledge that our estimates rest on these assumptions, and that an alternative approach would instead be to refine the naïve bounds by using information from the DER when a match to the ASEC is possible, and remaining more modest with our ability to predict poverty for nonmatch nonrespondents. That is, in lieu of the estimator for $\hat{P}^{\mathrm{ASEC}}_{\mathrm{NR,NM}}$ in Equation (4), which is the fourth term in Equation (2), we simply consider the two alternatives that all nonmatched nonresponders are poor against the alternative that none are poor. We depict this scenario against the full-response poverty rate in Figure 4. On average, the refined lower bound in Figure 4 is 1 percentage point higher than the naïve version in Figure 2, and the refined upper bound is 12 percentage points lower than the naïve bound. There is a notable convergence of the bounds starting in 2005, which is strictly an artifact of the improved match rates between the ASEC and DER.

We last consider a popular approach to missing data problems that does not rely on proprietary DER data—inverse probability weighting (IPW; Rosenbaum 1987; Bang and Robins 2005; Wooldridge 2007, 2010). A potential pitfall of the Census hot deck procedure under MAR is the finite set of covariates that are used to find a matched donor. IPW has the advantage over the hot deck approach in that the set of covariates can be expanded while not increasing the computational burden. As such, IPW is closely related to the propensity score method in the treatment effects literature (Rosenbaum and Rubin 1983). It assumes that conditional on a set of demographic factors, z_i, the resulting prediction of response probability is everywhere random. With these assumptions, we can obtain a consistent estimate of the population weighted poverty rate (Wooldridge 2010, pp. 822–823)

$$P^{\mathrm{IPW}} = \sum_{i=1}^{n} \left(R_i P_i w_i / \Pr\{z_i\}) / \sum_{i=1}^{n} (R_i w_i / \Pr\{z_i\} \right), \quad (8)$$

where R_i takes a value of 1 if the person is an ASEC respondent (0 otherwise), P_i is the poverty status of the individual, and w_i is the (adjusted) inverse probability of sample inclusion for the individual; that is, Equation (8) is the poverty rate of respondents weighted by the inverse probability of response.

In column (4) of Table 4, we present IPW estimates of the poverty rate. To implement the IPW approach we need to fit a flexible model of the probability of response, Pr{R}, which can include higher-order powers and interactions of the z_i. As reported in Hokayem, Bollinger, and Ziliak (2014), we examined combinations of three different sets of demographic characteristics (e.g., age, race, gender, education, marital status), three different models for the probability of nonresponse (OLS, probit, logit), and each yielded similar outcomes. That is, as seen in column (4) the IPW poverty rate is statistically and qualitatively systematically lower than the benchmark full-response rate by about 0.5 percentage point.

To better understand why both the unweighted ASEC respondent sample and the IPW ASEC respondent sample understate the poverty rate, in Figure 5 we present the residuals from a log DER wage equation by gender using the covariates in the inverse probability weighting models and in the hot deck procedure. Note that this wage equation uses DER earnings rather than CPS earnings, and is for workers only. Figure 5 shows a pronounced U-shape, with high nonresponse in both tails of the residual distribution. Even after conditioning on the covariates typically used in the hot deck procedure and in the IPW approach, we still see a double selection found in the extreme tails of the residual distribution. This suggests a violation of the MAR assumption necessary for both the hot deck procedure and the IPW approach. This also suggests there are unobservables causing nonresponse among individuals with low earnings than what observables would predict. The hot deck procedure and the weighting adjustments in the IPW approach do not account for these unobservables.

5. CONCLUSION

This article uses a unique dataset of administrative earnings data matched to internal ASEC to study the effects of earnings imputation on poverty measurement. Our analysis estimates the bias caused by earnings nonresponse. We compare a "full-response" poverty rate that assumes all ASEC respondents provided earnings data to the official poverty rate to gauge this bias. On average, we find the nonresponse bias to be about 1.0 percentage point. This bias seems more pronounced among more economically disadvantaged groups such as single female-headed families and those families headed by a nonwhite.

Our study is somewhat unique in that we take the earnings reported in the ASEC as the baseline compared to administrative reports in the DER. This stems from the fact that not all earnings are subject to Social Security taxation and thus not reported in the DER, and "under-the-table" earnings may show up in the ASEC but are not reported to tax authorities. This seems borne out in our sample in that poverty rates across the 12 years among matched respondents averages a statistically significant 1.7 percentage points lower using ASEC earnings than DER earnings, suggesting that simply replacing ASEC earnings with DER earnings is not the best solution to earnings nonresponse.

However, even though ASEC earnings may have some advantages compared to DER earnings, simply dropping nonrespondents is not ideal either. Our estimates suggest that dropping nonrespondents results in a poverty rate systematically higher than our full-response poverty rate. The bias caused by dropping nonrespondents stems from the loss of earners, and particularly high earners, who overall are more likely to be nonrespondents. Moreover, Little and Rubin (2002) made a compelling case against such practice because of the potential loss of efficiency and representativeness. To address the latter concern, we constructed an inverse probability weighted poverty series and found that this series results in too low of a poverty rate relative to our benchmark (typically in the middle of the range between the official poverty rate and the full-response rate). Most importantly, however, the accuracy of official poverty estimates in the United States would benefit greatly from reduced nonresponse of earnings and other income sources.

[Received April 2014. Revised January 2015.]

REFERENCES

Abowd, J., and Stinson, M. (2013), "Estimating Measurement Error in Annual Job Earnings: A Comparison of Survey and Administrative Data," *Review of Economics and Statistics*, 95, 1451–1467. [936]

Andridge, R., and Little, R. (2010), "A Review of Hot-Deck Imputation for Survey Nonresponse," *International Statistical Review*, 78, 40–64. [935]

Bang, H., and Robins, J. M. (2005), "Doubly Robust Estimation in Missing Data and Causal Inference Models," *Biometrics*, 61, 962–973. [944]

Bollinger, C. (1998), "Measurement Error in the CPS: A Nonparametric Look," *Journal of Labor Economics*, 16, 576–594. [936,941]

Bollinger, C., and Hirsch, B. (2006), "Match Bias From Earnings Imputation in the Current Population Survey: The Case of Imperfect Matching," *Journal of Labor Economics*, 24, 483–519. [935]

——— (2013), "Is Earnings Nonresponse Ignorable?" *Review of Economics and Statistics*, 95, 407–416. [935,943]

Bound, J., and Krueger, A. B. (1991), "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics* 9, 1–24. [936,941]

Citro, C., and Michael, R. (1995), *Measuring Poverty: A New Approach*, Washington, D.C.: National Academies Press. [935]

Cochran, W. G. (1977), *Sampling Techniques,* Wiley Series in Probability and Statistics, New York: Wiley. [937]

Dixon, J. (2012), "Using Contact History Information to Adjust for Nonresponse in the Current Population Survey," in *JSM Proceedings: Section on Government Statistics*, Alexandria, VA: American Statistical Association, pp. 1977–1982. [935]

Edin, K., and Lein, L. (1997), *Making Ends Meet: How Single Mothers Survive Welfare and Low-Wage Work*. New York: Russell Sage. [940]

Gabe, T. (2007), "Federal Programs That Use the "Official" Poverty Definition for Determining Eligibility Or for Allocating Funds," Memo From the Congressional Research Service to the U.S. House Ways and Means Subcommittee on Income Security and Family Support (October 30). [935]

Groves, R. (2001), *Survey Nonresponse*, New York: Wiley. [935]

Hirsch, B., and Schumacher, E. (2004), "Match Bias in Wage Gap Estimates Due to Earnings Imputation," *Journal of Labor Economics*, 22, 689–722. [935]

Hokayem, C., Bollinger, C., and Ziliak, J. P. (2014), *The Role of CPS Nonresponse on the Level and Trend in Poverty*, Discussion Paper 2014-05, Lexington, KY: University of Kentucky Center for Poverty Research. [941,944]

Korinek, A., Mistiaen, J., and Ravallion, M. (2007), "An Econometric Method of Correcting for Unit Nonresponse in Surveys," *Journal of Econometrics*, 136, 213–235. [935]

Little, R., and Rubin, D. (2002), *Statistical Analysis With Missing Data*, New York: Wiley. [935,945]

Manski, C. (1989), "Anatomy of the Selection Problem," *Journal of Human Resources*, 24, 343–360. [937]

——— (2014), *Credible Interval Estimates for Official Statistics With Survey Nonresponse*, Evanston, IL: Department of Economics, Northwestern University. [937]

Meyer, B., and Sullivan, J. (2012), "Identifying the Disadvantaged: Official Poverty, Consumption Poverty, and the New Supplemental Poverty Measure," *Journal of Economic Perspectives*, 26, 1–27. [935]

Nicholas, J., and Wiseman, M. (2009), "Elderly Poverty and Supplemental Security Income" *Social Security Bulletin*, 69, 45–73. [937,941]

——— (2010), "Elderly Poverty and Supplemental Security Income: 2002–2005," *Social Security Bulletin*, 70, 1–29. [937,941]

Roemer, M. (2002), *Using Administrative Earnings Records to Assess Wage Data Quality in the March Current Population Survey and the Survey of Income and Program Participation*, Longitudinal Employer-Household Dynamics Program, Technical Paper No. TP-2002-22, Washington, DC: U.S. Census Bureau. [936,941]

Rosenbaum, P. (1987), "Model-Based Direct Adjustment," *Journal of the American Statistical Association*, 82, 387–394. [944]

Rosenbaum, P., and Rubin, D. (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70, 41–55. [944]

Short, K. (2013), *The Research Supplemental Poverty Measure: 2012*, Current Population Reports, Washington, DC: U.S. Census Bureau. [935]

Turek, J., Scheuren, F., Snclair-James, B., Ghose, B., and Desale, S. (2009), "The Effects of Imputation of CPS Income and Poverty Series: 1981–2007," in *JSM Proceedings: Section on Survey Research Methods*, Alexandria, VA: American Statistical Association, pp. 1892–1900. [935]

Turek, J., Swenson, K., Ghose, B., Scheuren, F., and Lee, D. (2012), "How Good are ASEC Earnings Data? A Comparison to SSA Detailed Earnings Records" in *Proceedings of the 2012 Federal Committee on Statistical Methodology Research Conference*. [937,941]

Venkatesh, S. (2006), *Off the Books: The Underground Economy of the Urban Poor*, Cambridge, MA: Harvard University Press. [940]

Welniak, E. J. (1990), *Effects of the March Current Population Survey's New Processing System on Estimates of Income and Poverty*, Washington, DC: U.S. Census Bureau. [937]

Wooldridge, J. M. (2007), "Inverse Probability Weighted M-Estimation for General Missing Data Problems," *Journal of Econometrics*, 141, 1281–1301. [944]

——— (2010), *Econometric Analysis of Cross Section and Panel Data* (2nd ed.), Cambridge, MA: MIT Press. [944]

Ziliak, J. P. (2006), "Understanding Poverty Rates and Gaps: Concepts, Trends, and Challenges," *Foundations and Trends in Microeconomics*, 1, 127–199. [935]