Fixing Errors in a SNAP: Addressing SNAP Under-reporting to Evaluate Poverty*

Jonathan Rothbaum
U.S. Census Bureau†

Liana Fox
U.S. Census Bureau‡

Kathryn Shantz
Urban Institute§

October 18, 2021

Abstract

The Supplemental Nutrition Assistance Program (SNAP) is a major piece of the social safety net. At peak participation in 2013, SNAP annual benefits exceeded $76 billion, with 47.6 million monthly participants. However, SNAP is vastly under-reported in surveys. Because SNAP is administered at the state level, nationwide administrative data is not available. We address this issue by using administrative SNAP data from eight states to impute "true" SNAP participation for the rest of the country. We validate our approach by implementing a leave-one-out imputation and evaluate how SNAP affects resources and poverty and the relationship between SNAP receipt and earnings.

Keywords: Poverty, Supplemental Nutrition Assistance Program
JEL Codes: I3

---

*This report is released to inform interested parties of ongoing research and to encourage discussion. Any views expressed on statistical, methodological, technical, or operational issues are those of the author and not necessarily those of the U.S. Census Bureau. The data in this paper has been cleared by the Census Bureau's Disclosure Review Board release authorization numbers CBDRB-FY19-199, CBDRB-FY19-566, CBDRB-FY20-134, CBDRB-FY20-191, and CBDRB-FY20-254.

For more information on confidentiality protection, sampling error, nonsampling error, and definitions, see www2.census.gov/programs-surveys/cps/techdocs/cpsmar14.pdf. All comparisons in this paper are significant at the 0.1 level unless otherwise noted in the text.

We would like to thank Kalee Burns and Mel Stephens for helpful comments and the participants of the Census Research Seminar, the 2021 USDA Expert Panel on SNAP Administrative Data, the 2021 AEA session on Data Quality and Measurement of Earnings, Poverty, and Inequality, the 2020 Southern Economics Association session on the Measurement of U.S. Income and Earnings Distributions, the 2019 CNSTAT Data Linkage Day, and the 2019 PAA session on Measuring Poverty.

†4600 Silver Hill Road, Washington, DC 20233 (email: jonathan.l.rothbaum@census.gov)

‡4600 Silver Hill Road, Washington, DC 20233 (email: liana.e.fox@census.gov)

§500 L'Enfant Plaza SW Washington, DC 20024 (email: kshantz@urban.org)

1 Introduction

A large body of research has documented substantial measurement error in survey reports of income and program participation. For earnings, that literature goes back decades.[1] Although earnings has been the most studied, measurement error and misreporting can be more severe in other income types and for program participation. For example, Bee and Mitchell (2017) show that 46 percent of those with retirement income in administrative data fail to report it in the Current Population Survey Annual Social and Economic Supplement (CPS ASEC). Likewise, Fox et al. (2017), Shantz and Fox (2018), Meyer and Mittag (2018), Meyer et al. (2018), among others, have shown high rates of under-reporting of the receipt of means-tested program benefits, including in the CPS ASEC, the Survey of Income and Program Participation (SIPP), and the American Community Survey (ACS).

Much of this work requires direct access to the survey data linked to administrative records. However, not all researchers and policymakers will have access to this linked data. Furthermore, for some programs, such as the Supplemental Nutrition Assistance Program (SNAP), administrative data is generally only available for some states in some years.[2] Several attempts have been made to correct survey data for under-reporting. The Transfer Income Model (TRIM3) developed by the Urban Institute under contract with the Department of Health and Human Services (HHS) (Zedlewski and Giannarelli 2015) is one approach. For SNAP misreporting, TRIM3 uses aggregate statistics on state program spending and quality control data on the relationship between reported income and SNAP receipt and the amount received to distribute the missing SNAP benefits to CPS ASEC households.[3] The Congressional Budget Office (CBO) has developed another approach (Habib 2018). The CBO model assigns an expected probability to each household that did not report benefit receipt using a

[1]See Kilss and Scheuren (1978), Duncan and Hill (1985), Bound and Krueger (1991), Bound et al. (2001), Kapteyn and Ypma (2007), Abowd and Stinson (2013), among many many others.

[2]The SNAP program is administered at the state level. Access to program data requires an agreement with each individual state.

[3]SNAP quality control data is available at https://www.fns.usda.gov/resource/snap-quality-control-data, accessed on 3/9/2021.

probit regression on survey reports of receipt. The model then imputes program receipt and benefits to households with the highest expected probability of receipt until they match program aggregates. Mittag (2019) uses linked survey and administrative data from New York to produce conditional relationships between true SNAP receipt and survey characteristics. These relationships can be used to adjust individual observations for survey under-reporting.

Without access to the linked administrative data, the TRIM3 and CBO models require assumptions about the relationship between observable characteristics and misreporting that cannot be directly verified. By releasing conditional relationships, Mittag (2019) does not suffer from this shortcoming. His approach also has the advantage that it entails only the release of conditional relationships that can be combined with the public-use data to impute SNAP receipt and benefit amounts to correct for survey misreporting. However, the Mittag approach requires researchers have the knowledge and expertise to apply the correction. In addition, his approach cannot condition on linked administrative data (such as W-2 earnings) to better model and predict SNAP receipt and correct for benefit misreporting. This is important because any imputation model is mis-specified, as it is an approximation (hopefully close) of the true underlying process of misreporting. Having better predictors of misreporting (including administrative data) in the model can result in better models and help address issues with dimensionality.[4]

We take an approach that is similar to Mittag (2019), but with additional information on survey respondents (from other administrative data) and for additional states (8 states rather than 1). Similar to Mittag, we estimate models of the conditional relationships between 1) SNAP receipt and 2) benefit amounts and observable information in the survey, along with linked administrative income data. To estimate those models, we use a resampling and feature-selection process that allows us to include high-dimensional imputation models while also accounting for uncertainty in the optimal set of covariates and their conditional relationships. We use those conditional relationships to impute "true" SNAP receipt and

[4] There is a limit to how many covariates can be included in an imputation model given concerns about overfitting.

benefit amounts for households in locations where SNAP administrative data is not available. We validate our approach by implementing a leave-one-out imputation. We leave each state with administrative data out of the imputation model separately and compare our imputed SNAP benefits in those states to the administrative data. We create five separate implicates, or independent imputations, to properly estimate uncertainty in the resulting estimates. We then show how estimates of the Supplemental Poverty Measure (SPM), a measure of poverty that includes resources from SNAP and some other in-kind benefits, are affected by correcting for SNAP under-reporting. To evaluate how the imputed data might affect the kinds of analyses of interest to researchers, we also compare regression results using the imputed values to regressions using the administrative values.

To address the fact that researchers do not have access to the corrected data, we hope to release public-use extract files with misreporting-corrected SNAP receipt and benefit amounts. This file would be comprised of 1) imputed "true" receipt and benefits amounts for households in states without available SNAP administrative data and 2) synthetic SNAP receipt and benefit amounts for households in states where administrative data is available, to protect respondent confidentiality.[5]

Furthermore, we plan to integrate this approach to address SNAP under-reporting into forthcoming experimental income and resource statistic estimates as part of the National Experimental Wellbeing Statistics (NEWS) project.[6]

2 Data

We use the 2014 CPS ASEC linked to a variety of administrative records. This linkage relies on Protected Identification Keys (PIK), assigned by the Census Bureau. PIKs are assigned by a probabilistic matching algorithm that compares characteristics of records in

[5] There are several examples of similar public-use data microdata releases. Benedetto et al. (2013) released synthetic data, including synthetic administrative data, as part of the SIPP Synthetic Beta (SSB) project, and Rothbaum and Bee (2020) released public-use weights modeled using administrative data.

[6] For background on work at the U.S. Census Bureau to integrate administrative and survey data into improved income and resource estimates, see Bee and Rothbaum (2019).

census and survey data to characteristics of records in a reference file constructed from the Social Security Administration (SSA) Numerical Identification System (or Numident) as well as other federal administrative data. The Numident is a record of every Social Security Number (SSN) ever issued by the SSA. Once assigned, the PIK uniquely identifies a particular person and is consistent for that person over time (PIKs correspond one-to-one with a specific SSN), allowing us to link individuals across other de-identified data sources using only the PIK.[7] All estimates in this paper, except for Table 1, are based on the sample of SPM Resource Units whose household head was successfully assigned a PIK.[8]

We link the CPS ASEC to IRS data on 1040 tax filings and 1099-R forms on distributions and withdrawals from defined-benefit and defined-contribution retirement plans. We also use Detailed Earnings Record (DER) data from SSA on job-level earnings reported on W-2 forms for each linked CPS ASEC respondent.

This gives us information on administrative earnings, interest income, dividends, rental income, retirement income, survivor and disability benefits, and taxable adjusted gross income.[9]

We also use information from the U.S. Department of Agriculture (USDA) Food and Nutrition Service (FNS) on SNAP participation and benefit payments.[10] This data includes aggregates for state-level monthly beneficiary counts and aggregate benefit amounts. We calculate average monthly SNAP recipients as the average number of SNAP households in the FNS data divided by the number of households in the CPS ASEC.

For household-level SNAP receipt and benefit amounts, we link to administrative program data from eight states: Arizona, Idaho, Maryland, Michigan, New York, North Dakota,

[7] The linkage rate in the 2014 CPS ASEC is 87 percent.

[8] See Fox (2018) for more information about the Supplemental Poverty Measure and SPM Resource Units.

[9] We list each income type with the administrative source in parenthesis: earnings (1040, DER), social security benefits (1040), interest (1040, both taxable and non-taxable), dividends (1040), rental income (gross rental income on the 1040), retirement income (1099-R), survivor benefits (1099-R), disability benefits (1099-R), and taxable adjusted gross income (1040). Note that any information from 1040s filings is only available for tax filers.

[10] Available at https://www.fns.usda.gov/pd/supplemental-nutrition-assistance-program-snap, accessed on 3/24/2020.

Tennessee and Virginia, shown in Figure 1. Because we are focusing on SNAP receipt and benefits in this paper, we call these our states with administrative data, even though the other administrative data is available for all linked respondents, regardless of their state of residence. We only observe survey reports of SNAP receipt and benefit amounts in the other 42 states and the District of Columbia.

Table 1 compares summary statistics of various demographic and socioeconomic characteristics of individuals in states with and without SNAP administrative data. The first two rows show the average monthly rate of SNAP receipt and, conditional on receipt, the annualized average amount received in the USDA FNS data. The eight states with administrative SNAP data had higher SNAP participation rates (29.4 percent) than in the remaining 42 states and DC (26.3 percent). Conditional on SNAP receipt, states with administrative data had lower annualized average monthly benefits ($3,192) than the other states ($3,305).

The SPM rate for both groups of states was 15.5 percent and the difference was not statistically significant. Reported SNAP receipt was lower in both groups of states than in the FNS data, but the states with administrative data did have higher survey reports of SNAP participation than other states (14.2 vs. 12.5 percent).

There are other differences in the demographic and socioeconomic characteristics of individuals in states with and without administrative SNAP data. Individuals in states with administrative records are more likely to be working age (18-64), Black, Asian, foreign born, renters, to have health insurance, to have a bachelor's degree, to live inside an MSA and principal city and not to work. They are less likely to be young (under 18), White, Hispanic, native born, own their home with no mortgage, to live outside an MSA, and to work less than full-time and year-round.

We use mobility curves, from Foster and Rothbaum (2014), to show how SNAP affects SPM resources in Figure 2. The curve evaluates the share of the population that would be lifted out of poverty due to SNAP benefits. However, rather than focusing on a single poverty line for each SPM unit, the mobility curve traces out the share that would be lifted out of poverty for all possible poverty lines. For example, if \$27,000 were the poverty line, administrative reports of SNAP would move 1.7 percent of individuals out of poverty. At the same resource level, the survey reported SNAP moves 1.1 percent of individuals out of poverty. Survey reports of SNAP understate the impact of SNAP across much of the resource distribution, and especially from about \$20,000 to \$75,000.[11] Mobility curves provide an efficient visual summary of how SNAP impacts utility or welfare.[12]

3 Methods

3.1 Missing Data and Imputation

Suppose O is a collection of observable variables with no missing values with $O = (O_1, {}_2, \ldots, O_R)$ and $Y_1, Y_2, \ldots Y_p$ are variables with missing values, with $Y = (Y_1, Y_2, \ldots, Y_p)$. In our case, Y contains administrative SNAP information. O can contain information from the survey and other administrative data. Let $f(Y|O; \ )$ be the conditional joint density of administrative SNAP information, with $\ = ({}_1, {}_2, \ldots, {}_p)$ and where ${}_j$ is a vector of parameters in the conditional distribution for Y_j such as regression coefficients and dispersion parameters. Let A be an indicator for whether a household resides in a state with administrative SNAP data.

Administrative SNAP participation and benefits received is information that is observed for some households ($A = 1$) but not observed for others ($A = 0$). We can treat this as a

[11]The curve is evaluated at \$1,500 intervals from \$0 to \$150,000. At 86 percent of the points between \$21,000 and \$75,000, the impact of SNAP is greater in the administrative data than in the survey reports.

[12]Much like stochastic dominance for comparisons of distributions, mobility curves are based on a notion of mobility dominance for comparisons of pairs of distributions, with a pair being an initial and final distribution. In our case, the pair of distributions: 1) initial = resources, excluding SNAP and 2) final = resources, with SNAP. Foster and Rothbaum define first-order mobility dominance as the case when under any monotonically increasing utility function, mobility (changes in income/resources) has increased utility more under one pair of distributions than another. Similarly, second-order mobility dominance is present when for any monotonically increasing and concave utility function, mobility increases utility under one pair of distributions than another. First-order mobility dominance is present when one curve is greater than or equal to another at all points. Second-order mobility dominance is present when the integral of one curve is greater than another at all points. We do not formally test for mobility dominance, but the results shown in Figure 2 are consistent with administrative SNAP data first-order mobility dominating survey reports. Stated plainly, this would imply that for any increasing utility function, the true impact of SNAP benefits on utility (taking the administrative records as true) is greater than the impact of SNAP as reported on the survey.

problem of missing information, much like item nonresponse in surveys, that can be handled with imputation. Under any imputation model, one must impose assumptions on f and to assign plausible values to Y where data are missing.

From Rubin (1976), data is Missing at Random (MAR) if missingness can be accounted for by observable characteristics. In our notation, SNAP administrative data is MAR if $f(Y|O,A=1) = f(Y|O,A=0) = f(Y|O)$.

Another way to view imputation is through the lens of a researcher or data user. Consider a statistic Q, which could be a distributional statistic - such as a mean or median, a regression coefficient, or any other statistic or parameter of interest to the researcher. An imputation model is congenial or proper and results in unbiased estimates of Q if $E(\hat{Q}|O,A=1) = E(\hat{Q}|O,A=0) = E(\hat{Q}|O) = Q$ and has valid confidence intervals for $\hat{Q}$ (Meng, 1994; Rubin, 1996).

An important aspect of proper imputation is that misclassification (imputed value $\neq$ true value) does not necessarily bias estimates of $\hat{Q}$, contrary to the case of many standard measurement error and misclassification models in economics. To give a simple example for the purposes of intuition, suppose we are interested in estimating the correlation between a given variable O_r and one of the imputed variables Y_j, where $Y_j=\{0,1\}$. Now suppose there is a set of individuals with the same value for $O_r = c_r$, who have $Y_j = 1$ with some probability $E(Y_j|O_r = c_r) = p$. An imputation model using only O_r could correctly assign the probability of $Y_j = 1$ for this group of individuals. However, if that was all of the information used in the model, there would be no way to improve the prediction for this group. If we randomly assigned $Y_j = 1$ to these individuals with probability p, we would expect a misclassification rate of $2p(1-p)$. However, these individuals would not bias the expected correlation between Y_j and O_r in the imputed data. The key insight is that the misclassification is not random overall (although it may be random for this group). We should expect that the correct proportion of observably similar individuals will be misclassified in both directions such that they do not bias estimates of the correlation between O_r and Y_j.

This is only true when the imputation model is proper for the analysis being conducted. There are many examples in the literature where this condition fails for a given statistic or set of statistics. An example is match bias in the CPS. Bollinger and Hirsch (2006) show that because the imputation model in the CPS does not include union status, estimates of the relationship between union status and earnings are attenuated in the imputed data. Even in this case, the issue is not that their earnings are misclassified (as very rarely will imputed earnings match the true value for a given individual), but that they are drawn from the wrong distribution – one that does not condition on union status.

However, uncongeniality for one statistic does not indicate bias for other related statistics. For example match bias on union status does not necessarily mean that the CPS imputation model will cause bias for statistics of the unconditional earnings distribution. It is impossible for congeniality to hold for all possible statistics Q, unless the model perfectly predicts the missing values, i.e. there is no misclassification.[13] However, we could assess the quality of an imputation model by comparing a set of the resulting $\hat{Q}$ estimates against known Q values. We take this approach, using a variety of statistics to evaluate our modeling strategy, which we discuss in more detail in Section 4.1.

This approach differs from others that use unlinked auxiliary data to correct for under-reported benefits (such as in Habib 2018). For example, using a regression based on the survey requires an assumption about the relationship of true receipt and observable information to reported receipt. For example, if households are assigned receipt randomly based on their predicted probability of receipt until the survey receipt rate matches external aggregates, then given reported receipt Y_R, true receipt Y_T, and under-reporting rate , the assumption is that $P(Y_T jO) = \frac{1}{1-} P(Y_R jO)$. We show that this is not the case. For example, when we regress having earnings on SNAP receipt , we do not find the same relationship when we use the survey and administrative data (see Table 5).

[13] In this sense, misclassification can be important. If the imputed value equals true value for all cases, the data is not truly "imputed." However, in practice, imputations are unlikely to have extremely low misclassification rates, and we must evaluate the potential bias of each $\hat{Q}$ with the available information.

Other approaches condition on individual characteristics in unlinked microdata. Given a set of characteristics that are observable in two data sets O_C, let O_C^1 and O_C^2 be the distribution of those characteristics in the two data sets. One can model Y as a function of the shared variables under the assumption that $P(Y_T|O_C^1) = P(Y_T|O_C^2)$. Given Y_R, this involves imputing receipt to households until the $P(Y_T|O_C^1) = P(Y_T|O_C^2)$ condition holds. TRIM3 (Zedlewski and Giannarelli, 2015), for example, uses SNAP quality control data to model the relationship between income and SNAP receipt. However, if income reporting to the survey differs from income reporting as part of SNAP administration, $P(Y_T|O_C^1)$ may not equal $P(Y_T|O_C^2)$. This can be due to misreporting (either on the survey or in the administrative data) or temporal misalignment (SNAP qualification is not based on calendar-year income, whereas survey income in the CPS ASEC is), among other possible reasons. Fox (2018) find that the TRIM3 model overstates the impact of SNAP on poverty, possibly as a result of the fact that $P(Y_T|O_C^1) \neq P(Y_T|O_C^2)$.[14]

3.2 Imputation Model

We impute values for three variables in Y: 1) SNAP receipt, and conditional on receipt, 2) months of SNAP receipt, and 3) annual SNAP benefit amount. We impute each Y_j at the household level. We use predictive mean matching. For each variable, we predict $\hat{Y}_j$ for all households, whether $A = 1$ or $A = 0$. For each household i where $A = 0$, we find the 10 closest households k where $A = 1$, where distance is defined as $|\hat{Y}_{j,i} - \hat{Y}_{j,k}|$. We randomly select a household m from this set of 10 and assign $Y_{j,i} = Y_{j,m}$.

For the model variables (O), we include many variables from survey responses. These include household level information and data from the household head and spouse. The variables include: SNAP receipt and amounts, household income, receipt of each income type on the CPS ASEC, disability status, hours and weeks worked, occupation, industry, health

[14] Although the high rate of income imputation in the CPS ASEC and the limited number of covariates used for imputation could also result in $P(Y_T|O_C^1) \neq P(Y_T|O_C^2)$ due to non-congeniality of the CPS ASEC income imputation process for this purpose.

insurance information, means-tested program participation, age, education, race, marital status, family and household composition, information on item nonresponse, union status, citizenship status, etc. We also include in O information on each income type on 1040s and income reported on W-2s (from the DER) and from 1099-Rs. As a proxy for state-level variation in SNAP administration, O also includes USDA summary information on monthly SNAP participation and benefit payments at the state level.[15] We also include many two- and three-way interactions, as well as handful of four-way interactions.[16]

However, as a practical matter, there are too many potential variables in O to be used in our model. We reduce the set of variables to be used to impute each Y_j in two stages. In the first stage, we take all of the possible variables in O using a stepwise selection OLS regression model to prune the list to $\hat{O}_j$ that predict Y_j. In this first stage, the selection criteria are relatively permissive, and thus the set of variables in $\hat{O}_j$ is relatively large (hundreds of variables and interactions). In terms of the general notation $f(Y|O;\ )$, this process places constraints on .[17]

The next step is to estimate the values in $\hat{\ }$. As $\hat{\ }$ is a set of unknown parameters, we also must incorporate the uncertainty in $\hat{\ }$ into the imputation process. We do this as follows. In each implicate c, we start by taking a Bayesian Bootstrap of the CPS ASEC sample, we then do a second-stage variable selection process to further reduce the number of variables in $\hat{O}_j$ to $\hat{O}_{j;c}$.[18] From the OLS regression of Y_j on $\hat{O}_{j;c}$, we estimate $\hat{\ }_{j;c}$. Doing this on a Bayesian

[15] We hope to include information on state- and county-level SNAP rules in future work.

[16] For example, there may be a non-linear relationship between error in reported SNAP receipt and income. That relationship could also vary by race, age, education, etc. Capturing that variation in our imputation model likely requires higher level interactions.

[17] This is primarily done for practical speed considerations. Reducing the number of candidate variables upfront considerably speeds up the process of imputation for each variable in each implicate. Taking into account the leave-one-out models discussed in Section 4.1, we run nine separate imputation models. The nine models are each set up the same, but differ only in the sample used to predict "true" SNAP benefits. Each model has five independent implicates. Therefore, each variable Y_j is imputed 45 times. This first-stage selection means that instead of taking days or weeks to run, the full set of models can be run in hours.

[18] The Bayesian Bootstrap (Rubin, 1981) is the Bayesian analogue of the bootstrap. Each observation is drawn (with replacement) with an expected probability of $1/n$, but with variability. The probabilities of being drawn are defined by taking $n\ \ 1$ draws from the uniform distribution (0,1), ordering draws from lowest to highest, where $u = u_0, u_1, u_2, ..., u_n$ given $u_0 = 0$ and $u_n = 1$. The probability of being drawn for each observation i is based on the gaps between each adjacent value in u, so that for observation i the probability of being drawn is $g_i = u_i \ \ u_{i-1}$. As noted in Benedetto et al. (2013), using the Bayesian Boostrap adds

Bootstrap sample enables us to account for the uncertainty present in each step of this process, including which variables are used as model predictors ($\hat{O}_{j;c}$) and to draw from the distribution of parameters values $\hat{\ }_{j;c}$. This resampling approach to estimating uncertainty in regression-based imputation has been taken in other data products and research, including SIPP topic flag imputation (Benedetto et al., 2016), the SIPP Gold Standard and SIPP Synthetic Beta (Benedetto et al., 2013), and imputation research on missing income in the CPS ASEC (Hokayem et al., 2020).

With the estimates of $\hat{O}_{j;c}$ and $\hat{\ }_{j;c}$, we can estimate $\hat{Y}_{j;c}$ and take a random draw for each household in i where $A = 0$ from the ten nearest households k where $A = 1$.

We repeat this process five times, to create the five independent implicates. Therefore, we have five separate data sets with imputed SNAP administrative data for all households where $A = 0$. For any statistic or parameter estimate, we can account for the uncertainty in the imputation process (Rubin, 1976). This approach involves calculating the total variance by combining the within implicate variation (for example, the standard error of an estimate in one implicate) with the between implicate variation (the variance of the estimates for that parameter across the five implicates).

4 Results

4.1 Leave-One-Out Imputation and Diagnostics

In this section we evaluate the validity of our assumptions and the quality of the imputation model, to the extent possible. Our imputation model is based on the assumption that there are not unobservable differences between households in states with and without administrative data that would bias our imputes. This is the basis of the MAR assumption discussed above, that $f(Y jO; A = 1) = f(Y jO; A = 0) = f(Y jO)$.

additional variability to the imputation process to account for the fact that the sample distribution may not be the same as the population distribution. Without the use of the Bayesian Bootstrap, the confidence intervals would not be proper.

By definition, we cannot directly test this assumption as we do not observe the unobservable information in the states without administrative SNAP data. However, we can indirectly test whether this assumption is plausible. Given SNAP administrative data for state s, we can leave state s out of the imputation model and impute SNAP data for households in state s based on the data from the other states where $A = 1$. We can test whether $f(Y|O, A = 1, state \neq s) = f(Y|O, state = s)$, which is similar to the untestable assumption of $f(Y|O, A = 1) = f(Y|O, A = 0)$.

We call this our leave-one-out (LOO) imputation. We run the imputation model leaving out each state s with have SNAP administrative data, one at a time. For example, we can treat the administrative SNAP data from New York as unobserved and impute SNAP benefits in New York from the other seven states with administrative SNAP data. We then compare the estimates using the imputation model to estimates using the administrative data. In doing so, we test whether $E(\hat{Q}|O, A = 1, state \neq s) = E(\hat{Q}|O, state = s)$.

To increase the statistical power of the comparisons we make, we pool the LOO imputations and compare the administrative data to the LOO results. In each of the five pooled implicates, we have imputed administrative SNAP data for each household that was estimated based on the administrative data from the households in the seven other states.[19]

Because of the nature and severity of misreporting in the data, even basic statistics about SNAP cannot be estimated accurately in the survey data. For example, estimates of the income distribution of SNAP recipients, the anti-poverty impact of SNAP, or the relationship between SNAP and work cannot be estimated accurately when 40 percent or more of SNAP recipients do not report their benefits. Therefore, we validate our model on these kinds of relatively simple summary statistics and conditional relationships.

We revisit the SNAP mobility curve. In Figure 2, we showed that survey reports of SNAP benefits understates their impact on SPM resources, especially from \$20,000 to \$75,000. In

[19]Some of the comparisons discussed below are also at the state level. All of the unpooled results that correspond to those discussed in this section are available upon request, that is for each state with administrative data.

Figure 3, we add the pooled LOO results to the curve. The curves estimated with the administrative data and the LOO imputes are very close to each other at nearly every point.[20]

Next we evaluate how the pooled LOO imputations compares to the administrative data using the SPM in Table 2. In column (1), we show the SPM rates estimated using survey responses for the states with administrative data. Overall, the SPM rate in this sample is 14.5 percent. Using the administrative data, the SPM estimate is 14.1 percent, a difference of 0.4 percent. With the pooled LOO imputations, we estimate an SPM rate of 14.0 percent. For all but one of the 47 subgroups (unrelated individuals), the pooled LOO estimate is not statistically different from the estimate using administrative data. However, for the majority of subgroups, LOO and administrative estimates lower SPM poverty than the survey-only estimate.

We also evaluate estimates of SNAP receipt and benefits, comparing the administrative data to the LOO imputations in Table 3. The "Adrec States" row compares the administrative data to the pooled LOO estimates.[21] The administrative to LOO comparisons are shown for each individual state as well. We do not see significant differences in estimates of SNAP receipt when comparing the administrative data to the pooled LOO estimates or in comparisons for each state individually. However, conditional on SNAP receipt, we do see some significant differences in estimates of the distribution of benefits received in the pooled sample and individually for New York and Virginia.

We also test whether $E(\hat{Q}|O, A = 1, state \neq s) \neq E(\hat{Q}|O, state = s)$ for two sets of relatively simple regressions using SNAP data. In the first, we regress administrative SNAP participation on survey reports of SNAP and various demographic and socioeconomic characteristics, shown in Table 4. Columns (1) and (2) show the estimates for the states with administrative data, with column (1) using the administrative data and (2) using the pooled LOO imputations. Column (3) shows the difference. For none of the coefficients

[20] They are statistically different at only two of the 101 points evaluated.
[21] Adrec indicates administrative record.

is there a statistically significant difference in the models. In each case, survey reports of SNAP receipt are a strong predictor of administrative receipt, but other characteristics, such as race, gender, nonresponse to SNAP receipt and income also predict the presence of SNAP benefits in the administrative data.

In the second set of regression results, we regress the presence of earnings on SNAP receipt and a small set of demographic and socioeconomic characteristics, shown in Table 5. As above, columns (1) and (2) show the estimates for the states with administrative data, with column (1) using the administrative data and (2) using the pooled LOO imputations. Column (6) shows the difference. Again there are no coefficients with a significant difference. The coefficient on SNAP receipt is -0.18 using both the administrative data and from the pooled LOO imputations.

We do not claim that there would be no statistically significant difference between any given statistic or parameter $E(\hat{Q}jO)$, as estimated in our imputation model, and the true underlying Q. We found several cases in Table 3 for benefit amounts where the estimates do differ. However, the results, especially from Figure 3, give us confidence in our modeling strategy. Next, we turn to estimates in the full sample of CPS ASEC respondent, rather than in the eight states with administrative data that we have focused on to this point.

4.2 National Estimates of the Impact of SNAP

Figure 4 shows the mobility curve estimates of the impact of survey-reported SNAP on SPM resources in the states with administrative records and in all states, in the dashed blue and solid red lines respectively. The impact of SNAP is statistically significantly greater in states with administrative data at 10 percent of all points analyzed, including several points below $5,000, between $40,000 and $50,000, and around $80,000 to $90,000. This is consistent with greater reported SNAP receipt in states with administrative data, shown in Table 1.

For the states without administrative records, we use the values from our imputation model to estimate the impact of SNAP benefits. We show summary statistics for the imputed

SNAP receipt and benefits in Table 3 under "All Imputed." We compare the imputed estimates from states without administrative data to the estimates using the administrative data for the states where it is available. The estimates for SNAP receipt and average SNAP benefits are not statistically different.

For the national estimates that correct for SNAP under-reporting, we use the administrative data in states where they are available and the imputed values in the remaining states and DC. The impact of SNAP at each level of SPM resources is shown in Figure 4. The two lines are statistically different at 15 percent of the points analyzed, with differences around $25,000 and $35,000 of SPM resources that could affect poverty estimates.

Table 6 shows the national SPM estimates using the combined administrative and imputed data. We estimate that SNAP under-reporting causes the SPM to understate poverty by 0.2 percent overall. We also compare the impact of administrative SNAP data in the national estimates to their effect in the states with administrative SNAP data with a difference-in-difference comparison. For each estimate, we first calculate the difference in SPM rates using administrative vs. survey data (for states with administrative data, Table 2, Column (4) and at the national level, Table 6, Column (3)). Then we compare this difference in the national estimates to the difference in the estimate for states with administrative data. The results are shown in Table 6, Column (5). Overall, and for most of the subgroups analyzed, we do not find a statistically significant difference in the estimates.

4.3 SNAP Imputations and Conditional Relationships

Next, we compare how conditional relationships differ between the imputed and observed administrative data. First we regress administrative SNAP receipt on survey reported receipt and various other household characteristics, shown in Table 4. We find few statistically significant differences between our estimate for all states and for states with administrative data.[22]

[22]The comparisons are shown in column (5). The differences we do find are in the coefficients for Pacific Islander and household income squared (after the inverse hyperbolic sine transformation).

Next, we examine a relationship of considerable interest to policy-makers and researchers - the relationship between having earnings and SNAP receipt. We do so with a linear probability OLS regression, with the results shown in Table 5. We evaluate this relationship for five different data source-sample combinations (in the numbered columns): (1) SNAP administrative data in states with SNAP administrative data, (2) leave-one-out imputation in states with SNAP administrative data, (3) imputed SNAP receipt in states without SNAP administrative data, (4) survey-reported SNAP receipt in states with SNAP administrative data, and (5) survey-reported SNAP receipt in states without SNAP administrative data. The regressions sample includes working aged (18-64) householders with reported (not imputed) earnings and SNAP receipt.

Using the administrative data, we find that SNAP recipients are 18 percentage points less likely to have earnings, conditional on the other model variables. However, using survey reports, SNAP receipt is associated with a 23 percentage point lower probability of having earnings. Survey estimates of the relationship between SNAP and earnings receipt overestimate the true impact by 5 percentage points (which is about 30 percent of the true estimate of 18 percentage points). Those with earnings are more likely to misreport SNAP receipt as non-receipt.[23]

In validating our imputation model, we do not find significant differences in regression coefficients comparing the estimates from the leave-one-out SNAP imputations to the administrative data. However, there are differences between the regression coefficients using the imputed data in states without administrative data compared to states with administrative data, shown in Column (7). In states without administrative data, Blacks and householders with an Associate's degree are more likely not to work (conditional on SNAP receipt) than in states with administrative data. However, these differences do not appear to be related to SNAP reporting, as shown in Column (10).

[23]This is consistent with prior work. Fox et al. (2017) found that households with a lower share of workers and a lower share of full-time, year-round workers were more accurate in their survey reports of SNAP receipt.

These results suggest that the imputation model is capturing some of the kinds of conditional correlations that are of interest to researchers and policy-makers. This has the potential to help researchers better understand relationships between outcomes like earnings and SNAP receipt, which could be biased in the survey data.

5 Conclusion

In this paper, we propose using methods developed for handling missing data to correct for SNAP under-reporting in the absence of complete administrative data coverage and given limited access to the administrative data that is available. We use administrative data available in a subset of states to correct for under-reporting nationally. Because the assumptions that underpin our imputation approach cannot be directly tested, we implement a set of leave-one-out estimates to indirectly test the MAR and congeniality assumptions that are required for an imputation model to result in unbiased estimates of statistics of interest to researchers and policy-makers.

We test these assumptions using a measure of distributional change (the mobility curve), a measure of poverty (the SPM), and with measures of conditional correlations (regressions with SNAP receipt as the dependent and independent variables). For nearly all of the statistics we estimate, we do not find statistically significant differences between the leave-one-out imputation estimates and the estimates using the administrative data. This supports the MAR and congeniality assumptions that underpin our imputation approach. We then impute missing administrative SNAP data for states where that data is not available. Specifically, we show that this imputation approach can be used to accurately evaluate the impact of SNAP on the Supplemental Poverty Measure, the relationship between SNAP and other resources available for consumption, and basic conditional relationships (such as between SNAP receipt and having earnings).

This approach offers a possible solution to addressing survey misreporting when administrative data is not available for all individuals. This approach also yields microdata that can be used by other researchers directly in their analyses. We believe this can improve the quality of those analyses by reducing bias due to survey misreporting as well as provide feedback to us in our imputation models to increase the number of statistics ($\hat{Q}$) for which the model is congenial and unbiased.

In this paper, the availability of administrative benefits was based on geography, but the same methodology could be used to impute administrative income or benefits over time as well, with available administrative data in some time periods, but not others. We believe this has the potential to improve estimates of income and resources when administrative data is not available. This can help us correct for survey misreporting in two situations: 1) timely estimates can be released before the administrative data is available, and 2) historic estimates can be released for years when some or all of the administrative data is unavailable.

References

Abowd, John M and Martha H Stinson, "Estimating measurement error in annual job earnings: A comparison of survey and administrative data," *Review of Economics and Statistics*, 2013, *95* (5), 1451–1467.

Bee, Adam and Jonathan Rothbaum, "The Administrative Income Statistics (AIS) Project: Research on the Use of Administrative Records to Improve Income and Resource Estimates," *U.S. Census Bureau SEHSD Working Paper #2019-36*, 2019.

_ **and Joshua Mitchell**, "Do Older Americans Have More Income Than We Think?," *U.S. Census Bureau SEHSD Working Paper #2017-39*, 2017.

Benedetto, Gary, Joanna Motro, and Martha Stinson, "Introducing Parametric Models and Administrative Records into 2014 SIPP Imputations," in "Proceedings of the 2015 Federal committee on Statistical Methodology Research conference" 2016.

_ **, Martha Stinson, and John M Abowd**, "The creation and use of the SIPP Synthetic Beta," *U.S. Census Bureau Working Paper*, 2013.

Bollinger, Christopher R and Barry T Hirsch, "Match bias from earnings imputation in the Current Population Survey: The case of imperfect matching," *Journal of Labor Economics*, 2006, *24* (3), 483–519.

Bound, John and Alan B Krueger, "The extent of measurement error in longitudinal earnings data: Do two wrongs make a right?," *Journal of Labor Economics*, 1991, *9* (1), 1–24.

_ **, Charles Brown, and Nancy Mathiowetz**, "Measurement error in survey data," in "Handbook of econometrics," Vol. 5 2001, pp. 3705–3843.

Duncan, Greg J and Daniel H Hill, "An investigation of the extent and consequences of measurement error in labor-economic survey data," *Journal of Labor Economics*, 1985, *3* (4), 508–532.

Foster, James and Jonathan Rothbaum, "The mobility curve: Measuring the impact of mobility on welfare," *Working Paper*, 2014.

Fox, Liana, "The supplemental poverty measure: 2018," *Current population reports*, 2018.

Fox, Liana E, Misty L Heggeness, and Kathryn Stevens, "Precision in measurement: Using SNAP administrative records to evaluate poverty measurement," *U.S. Census Bureau SEHSD Working Paper #2017-49*, 2017.

Habib, Bilal, "How CBO Adjusts for Survey Underreporting of Transfer Income in Its Distributional Analyses," *Congressional Budget O ce Working Paper 2018-07*, 2018.

Hokayem, Charles, Trivellore Raghunathan, and Jonathan Rothbaum, "Match Bias or Nonignorable Nonresponse? Improved Imputation and Administrative Data in the CPS ASEC," *Journal of Survey Statistics and Methodology*, 2020.

Kapteyn, Arie and Jelmer Y Ypma, "Measurement error and misclassification: A comparison of survey and administrative data," *Journal of Labor Economics*, 2007, *25* (3), 513–551.

Kilss, Beth and Frederick J Scheuren, "The 1973 CPS-IRS-SSA exact match study," *Social Security Bulletin*, 1978, *41*, 14.

Meng, Xiao-Li, "Multiple-imputation inferences with uncongenial sources of input," *Statistical Science*, 1994, pp. 538–558.

Meyer, Bruce D and Nikolas Mittag, "Misreporting of Government Transfers: How Important Are Survey Design and Geography?," 2018.

_ , _ , and Robert M Goerge, "Errors in survey reporting and imputation and their effects on estimates of food stamp program participation," *National Bureau of Economic Research Working Paper #25143*, 2018.

Mittag, Nikolas, "Correcting for misreporting of government benefits," *American Economic Journal: Economic Policy*, 2019, *11* (2), 142–64.

Rothbaum, Jonathan and Adam Bee, "Coronavirus infects surveys, too: Nonresponse bias during the pandemic in the CPS ASEC," *U.S. Census Bureau SEHSD Working Paper #2020-10*, 2020.

Rubin, Donald B, "Inference and missing data," *Biometrika*, 1976, *63* (3), 581–592.

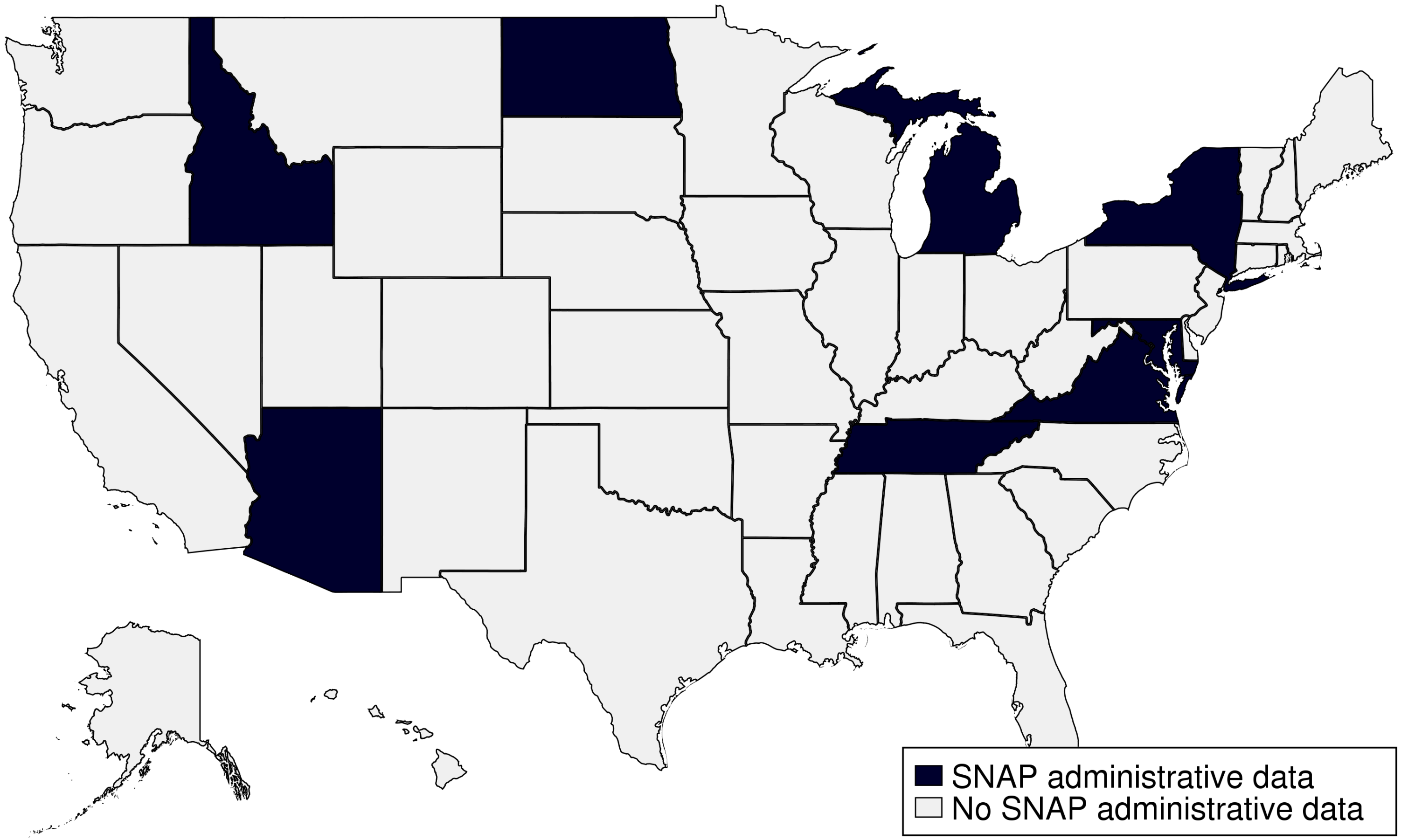
_ , "The bayesian bootstrap," *The annals of statistics*, 1981, pp. 130–134.

_ , "Multiple imputation after 18+ years," *Journal of the American statistical Association*, 1996, *91* (434), 473–489.

Shantz, Kathryn and Liana E Fox, "Precision in Measurement: Using State-Level Supplemental Nutrition Assistance Program and Temporary Assistance for Needy Families Administrative Records and the Transfer Income Model (TRIM3) to Evaluate Poverty Measurement," *U.S. Census Bureau SEHSD Working Paper #2018-30*, 2018.

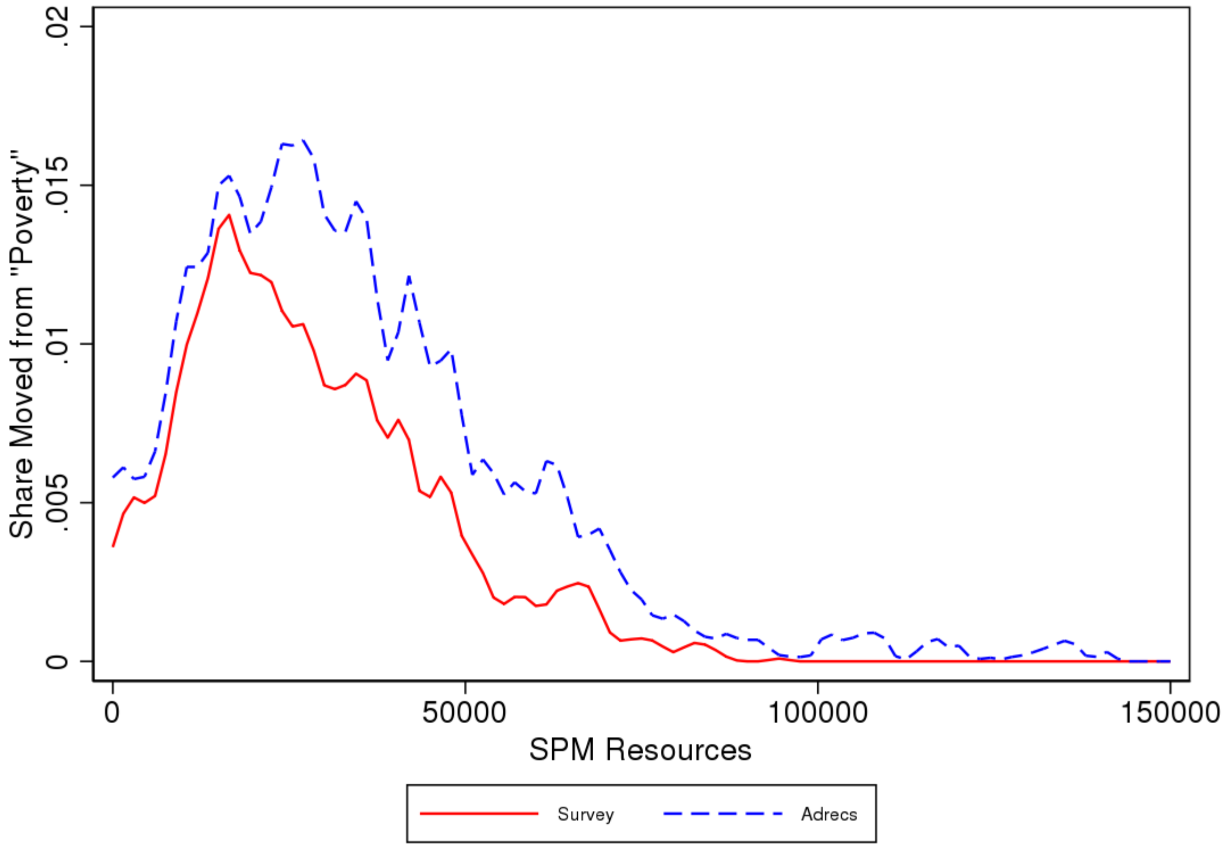
Zedlewski, Sheila and Linda Giannarelli, "TRIM: A tool for social policy analysis," *Urban Institute Research Report*, 2015.

Figure 1: States with Household-level SNAP Administrative Data



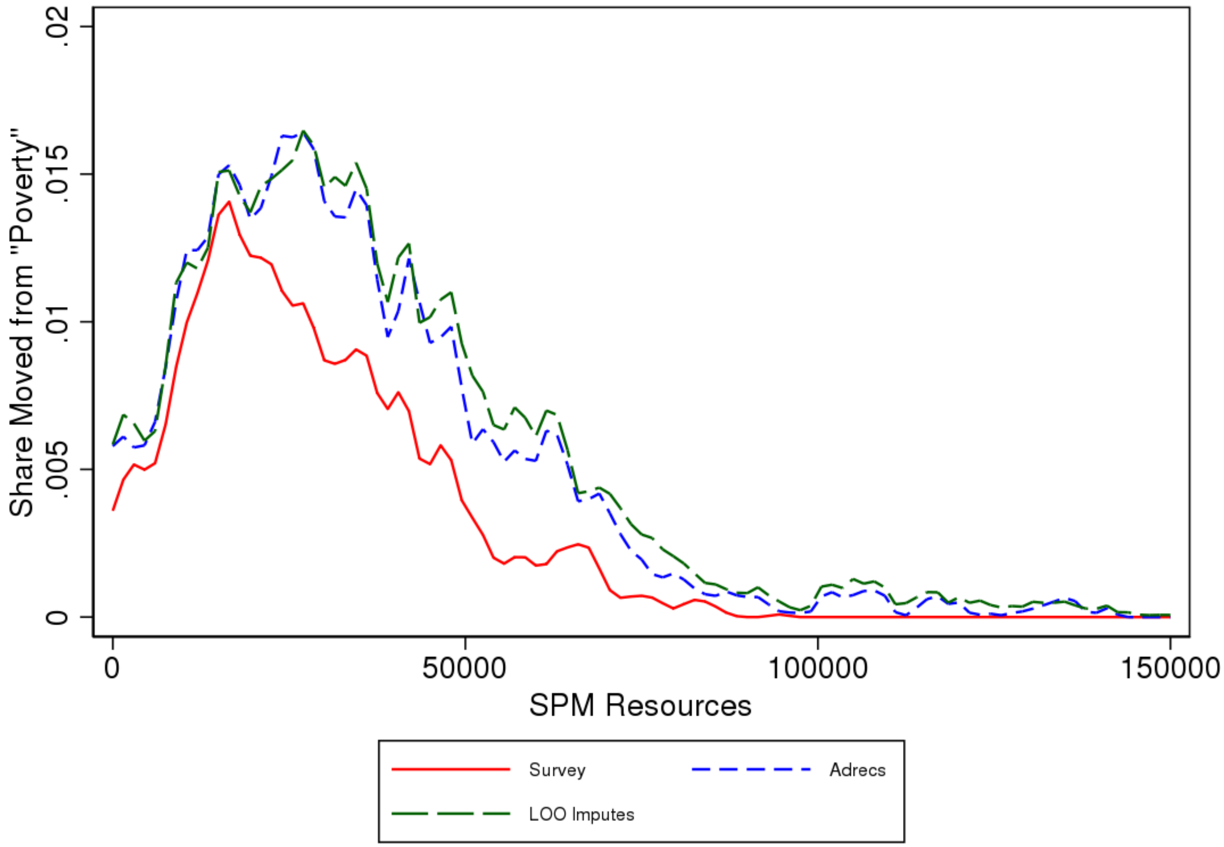
Source: The administrative SNAP data for 2013 comes from Arizona, Idaho, Maryland, Michigan, New York, North Dakota, Tennessee and Virginia.

Figure 2: SNAP Impact on Poverty For All Possible Poverty Lines:
Comparing Survey Reports to Administrative Records



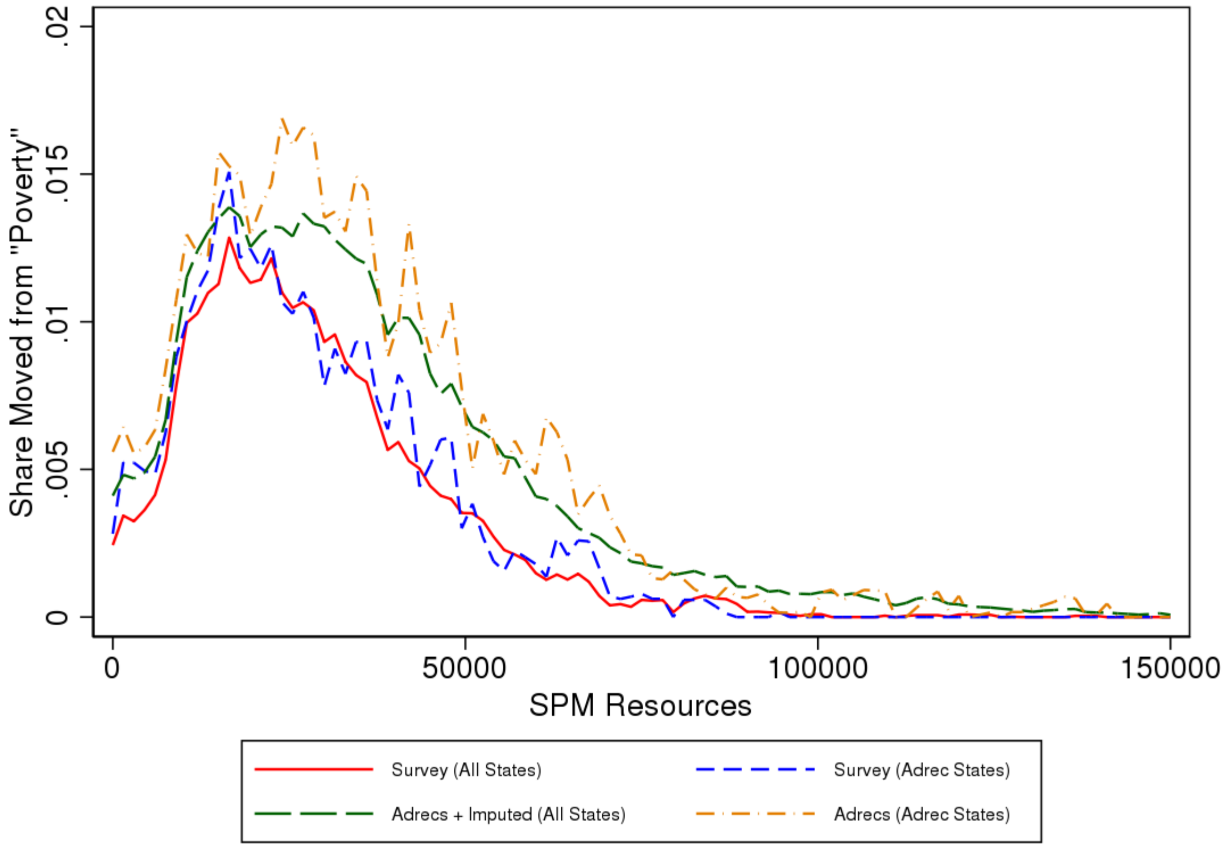
Notes: This figure shows the Mobility Curve (Foster and Rothbaum, 2014), which traces out the share of the population that SNAP moves out of "poverty" measured at all poverty lines from \$0 to \$150,000 in \$1,500 intervals in the pooled sample of the eight states whose administrative SNAP data we use in this paper. For example, if \$27,000 were the poverty line, administrative reports (Adrecs) of SNAP would move 1.7 percent of individuals out of poverty. At the same income level, the survey reports move 1.1 percent of individuals out of poverty.
Source: 2014 CPS ASEC Traditional File linked to state SNAP administrative records for eight states: Arizona, Idaho, Maryland, Michigan, New York, North Dakota, Tennessee and Virginia.

Figure 3: SNAP Impact on Poverty For All Possible Poverty Lines:
Comparing Leave-One-Out Imputations to Administrative Records



Notes: This figure shows the Mobility Curve (Foster and Rothbaum, 2014), which traces out the share of the population that SNAP moves out of "poverty" measured at all poverty lines from \$0 to \$150,000 in \$1,500 intervals in the pooled sample of the eight states whose administrative SNAP data we use in this paper. For example, if \$27,000 were the poverty line, administrative reports (Adrecs) of SNAP would move 1.7 percent of individuals out of poverty). At the same income level, the survey reports move 1.1 percent of individuals out of poverty and the leave-one-out imputed SNAP benefits move 1.7 percent out of poverty. The leave-one-out and administrative estimates are statistically different at only two percent of the points tested. Both are statistically different from the survey estimates at about 50 percent of the points tested.
Source: 2014 CPS ASEC Traditional File linked to state SNAP administrative records for eight states: Arizona, Idaho, Maryland, Michigan, New York, North Dakota, Tennessee and Virginia.

Figure 4: SNAP Impact on Poverty For All Possible Poverty Lines:
Comparing Survey, Administrative, and Imputed Estimates



Notes: This figure shows the Mobility Curve (Foster and Rothbaum, 2014), which traces out the share of the population that SNAP moves out of "poverty" measured at all poverty lines from $0 to $150,000 in $1,500 intervals in the pooled sample of the eight states whose administrative SNAP data we use in this paper. For example, for the eight states with administrative data, if $27,000 were the poverty line, administrative reports (Adrecs) of SNAP would move 1.7 percent of individuals out of poverty (the highest point estimate of the administrative record (Adrecs) line above). At the same income level, the survey reports move 1.1 percent of individuals out of poverty - both in states with administrative records and all states (the survey estimates for the two sets of states are not statistically different).
Source: 2014 CPS ASEC Traditional File linked to state SNAP administrative records for eight states: Arizona, Idaho, Maryland, Michigan, New York, North Dakota, Tennessee and Virginia.

Table 1: Summary Statistics - States with and without SNAP Administrative Data

	States with Admin Records		Other States		Difference	
	Estimate	SE	Estimate	SE	Estimate	SE
	(1)	(2)	(3)	(4)	(5)	(6)
USDA State-Level Aggregates						
Average Monthly SNAP Recipiency Rate	29.4	0.03	26.3	0.02	3.10***	0.03
Average Monthly Receipt Amount*12	3192	5	3305	8	-113***	9
SPM Rates	15.5	0.5	15.5	0.2	0.01	0.55
SNAP Recipiency Rate	14.2	0.7	12.5	0.2	-1.72**	0.72
Cond. Mean HH Annual SNAP Value	3,508	112	3,565	53	58	126
Male	48.9	0.1	49.0	Z	0.12	0.17
Female	51.1	0.1	51.0	Z	-0.12	0.17
Under 18 years	22.9	0.1	23.8	Z	0.90***	0.14
18 to 64 years	62.6	0.2	62.1	0.1	-0.51*	0.30
65 years and older	14.5	0.2	14.1	0.1	-0.39	0.31
Married couple unit	59.6	0.7	60.7	0.3	1.14	0.74
Cohabiting partner unit	7.8	0.4	8.1	0.2	0.27	0.39
Female reference person unit	13.5	0.5	13.3	0.2	-0.26	0.55
Male reference person unit	4.9	0.3	4.4	0.1	-0.43	0.30
Unrelated individuals	14.2	0.4	13.5	0.1	-0.72*	0.40
White	73.5	0.5	78.6	0.1	5.09***	0.64
White, not Hispanic	62.7	0.8	62.3	0.2	-0.37	0.97
Black	16.2	0.1	12.2	Z	-3.94***	0.11
Asian	6.3	0.4	5.3	0.1	-1.02**	0.46
Hispanic (any race)	13.9	0.7	18.1	0.2	4.26***	0.85
Native born	85.4	0.4	87.3	0.2	1.82***	0.47
Foreign born	14.6	0.4	12.7	0.2	-1.82***	0.47
Naturalized citizen	7.2	0.3	5.9	0.1	-1.36***	0.33
Not a citizen	7.4	0.3	6.9	0.1	-0.47	0.32
Total, aged 25 and older	67.7	0.3	66.6	0.1	-1.11***	0.34
No high school diploma	7.9	0.3	7.8	0.1	-0.06	0.32
High school, no college	20.0	0.3	19.8	0.2	-0.13	0.36
Some college, no degree	17.3	0.3	17.9	0.2	0.56	0.38
Bachelor's degree or higher	22.5	0.4	21.1	0.2	-1.47***	0.46
Owner	66.0	0.6	68.0	0.3	2.02***	0.71
Owner/mortgage	44.1	0.7	43.3	0.3	-0.85	0.77
Owner/no mortgage/rentfree	21.9	0.5	24.8	0.3	2.87***	0.52
Renter	34.0	0.6	32.0	0.3	-2.02***	0.71
Inside MSAs	87.6	1.6	84.4	0.6	-3.25*	1.70
Inside principal cities	36.9	1.2	31.7	0.5	-5.19***	1.26
Outside principal cities	50.7	1.1	52.7	0.5	1.94*	1.16
Outside MSAs	12.4	1.6	15.7	0.6	3.25*	1.70
Northeast	33.1	0.1	14.2	Z	-18.97***	0.14
Midwest	17.9	0.1	22.1	Z	4.20***	0.13
South	34.9	0.1	37.9	0.1	2.99***	0.14
West	14.0	0.1	25.8	Z	11.78***	0.13
With private insurance	65.7	0.6	63.8	0.3	-1.85***	0.64
With public, no private insurance	22.7	0.6	22.4	0.2	-0.30	0.56
Not insured	11.6	0.3	13.8	0.1	2.15***	0.38
Total, 18 to 64 years	62.6	0.2	62.1	0.1	-0.51*	0.30
All workers	46.4	0.3	46.7	0.2	0.28	0.39
Worked full-time, year-round	32.5	0.4	32.1	0.2	-0.41	0.47
Less than full-time, year-round	13.9	0.3	14.6	0.1	0.69**	0.32
Did not work at least 1 week	16.1	0.3	15.4	0.2	-0.79**	0.34
Total, 18 to 64 years	62.6	0.2	62.1	0.1	-0.51*	0.30
With a disability	4.7	0.2	4.8	0.1	0.09	0.20
With no disability	57.5	0.3	56.9	0.1	-0.57	0.35

Notes: This table compares summary statistics for CPS ASEC respondents in states with and without SNAP administrative records. The first two rows show data from the USDA on state-level monthly aggregates for SNAP recipiency rates and average benefits received in 2013. All other rows are calculated from the survey responses to the CPS ASEC only, without using administrative data. In the difference estimates, ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively. Z indicates an estimate rounds to zero.
Source: 2014 CPS ASEC Traditional File linked to state SNAP administrative records for eight states: Arizona, Idaho, Maryland, Michigan, New York, North Dakota, Tennessee and Virginia.

Table 2: Comparison of SPM using Survey, Administrative, and Imputed SNAP Data - States with Administrative Data Only

	SPM Estimate			Differences Between Estimates					
	Survey Only		Leave-One-Out (LOO)	Adrecs - Survey		LOO - Survey		LOO - Adrecs	
	(Official) (1)	SNAP Adrecs (2)	Imputations (3)	Estimate (4)	SE (5)	Estimate (6)	SE (7)	Estimate (8)	SE (9)
All People	14.49	14.09	13.95	-0.4062***	(0.1391)	-0.5395***	(0.2002)	-0.1333	(0.1806)
Male	13.67	13.20	13.17	-0.4766***	(0.1487)	-0.5033**	(0.2361)	-0.02673	(0.2176)
Female	15.28	14.94	14.70	-0.3389**	(0.1592)	-0.5741***	(0.1996)	-0.2352	(0.1731)
Under 18 years	15.35	14.73	14.70	-0.6170**	(0.2665)	-0.6526**	(0.3098)	-0.03553	(0.2790)
18 to 64 years	14.48	14.10	14.00	-0.3707***	(0.1418)	-0.4780**	(0.2173)	-0.1072	(0.2052)
65 years and older	13.18	12.96	12.56	-0.2202	(0.1495)	-0.6232	(0.3820)	-0.4031	(0.3761)
Married couple unit	8.61	8.20	8.22	-0.4082***	(0.1420)	-0.3913**	(0.1769)	0.01687	(0.1645)
Cohabiting partner unit	14.17	13.52	13.81	-0.6469	(0.5796)	-0.3598	(0.9464)	0.2871	(0.9192)
Female reference person unit	30.01	29.01	29.04	-1.000	(0.6433)	-0.9767	(0.8087)	0.02359	(0.6722)
Male reference person unit	21.34	21.66	20.85	0.3226	(0.7936)	-0.4907	(1.004)	-0.8133	(1.148)
Unrelated individuals	23.31	23.38	22.41	0.07397	(0.2357)	-0.8977**	(0.4170)	-0.9717**	(0.3891)
White	12.17	11.72	11.60	-0.4477***	(0.1574)	-0.5741**	(0.2456)	-0.1264	(0.2140)
White, not Hispanic	9.99	9.63	9.50	-0.3568***	(0.1384)	-0.4887**	(0.1960)	-0.1319	(0.1741)
Black	23.66	23.35	23.17	-0.3189	(0.5270)	-0.4948	(0.7270)	-0.1759	(0.7409)
Asian	15.78	15.78	15.32	Z	(0.3451)	-0.4679	(0.4760)	-0.4679	(0.4760)
Hispanic (any race)	24.89	24.09	24.29	-0.7990*	(0.4557)	-0.5927	(0.8630)	0.2064	(0.8198)
Native born	13.37	13.00	12.90	-0.3778**	(0.1527)	-0.4770**	(0.2036)	-0.09922	(0.1825)
Foreign born	21.65	21.07	20.71	-0.5880***	(0.2037)	-0.9396**	(0.4628)	-0.3516	(0.4343)
Naturalized citizen	17.66	17.34	17.05	-0.3199	(0.2452)	-0.6105*	(0.3510)	-0.2906	(0.3471)
Not a citizen	25.98	25.11	24.69	-0.8790***	(0.3271)	-1.297	(0.8326)	-0.4178	(0.7647)
Total, aged 25 and older	13.34	13.04	12.84	-0.3051**	(0.1216)	-0.5027***	(0.1842)	-0.1976	(0.1729)
No high school diploma	33.01	31.21	30.89	-1.807***	(0.6806)	-2.125**	(0.9990)	-0.3179	(0.9089)
High school, no college	16.00	15.79	15.54	-0.2121	(0.2341)	-0.4614*	(0.2770)	-0.2493	(0.2961)
Some college, no degree	11.31	11.27	11.05	-0.03749	(0.1645)	-0.2581	(0.3023)	-0.2206	(0.2778)
Bachelor's degree or higher	6.30	6.19	6.09	-0.1095**	(0.04638)	-0.2061**	(0.09825)	-0.09662	(0.08783)
Owner	9.29	9.17	9.04	-0.1182	(0.1009)	-0.2548	(0.1981)	-0.1366	(0.1880)
Owner/mortgage	8.11	8.02	8.05	-0.09565	(0.09877)	-0.06450	(0.1914)	0.03115	(0.1746)
Owner/no mortgage/rentfree	11.70	11.54	11.06	-0.1642	(0.2304)	-0.6436	(0.4061)	-0.4795	(0.3897)
Renter	25.36	24.36	24.23	-1.008***	(0.3892)	-1.135***	(0.4341)	-0.1264	(0.3738)
Inside MSAs	14.59	14.27	14.11	-0.3223**	(0.1517)	-0.4828**	(0.2226)	-0.1605	(0.2099)
Inside principal cities	19.71	19.28	19.14	-0.4241*	(0.2456)	-0.5643	(0.3721)	-0.1401	(0.3584)
Outside principal cities	11.13	10.88	10.70	-0.2536	(0.1902)	-0.4278	(0.2687)	-0.1742	(0.2494)
Outside MSAs	13.83	12.84	12.90	-0.9849**	(0.4700)	-0.9307	(0.6804)	0.05420	(0.5205)
Northeast	14.34	13.92	13.71	-0.4147	(0.2743)	-0.6287*	(0.3280)	-0.2140	(0.2992)
Midwest	13.66	13.18	13.32	-0.4814*	(0.2503)	-0.3414	(0.5734)	0.1400	(0.5217)
South	13.68	13.33	13.01	-0.3481	(0.2538)	-0.6689**	(0.2685)	-0.3208	(0.2327)
West	17.98	17.54	17.71	-0.4343	(0.3417)	-0.2696	(0.5836)	0.1646	(0.4936)
With private insurance	7.04	6.83	6.84	-0.2112***	(0.07666)	-0.2071*	(0.1170)	0.004072	(0.1025)
With public, no private insurance	29.52	28.77	28.40	-0.7433	(0.5144)	-1.118*	(0.6500)	-0.3745	(0.5907)
Not insured	29.66	28.74	28.25	-0.9193**	(0.4243)	-1.406**	(0.7143)	-0.4871	(0.6618)
Total 18 to 64 years	14.48	14.10	14.00	-0.3707***	(0.1418)	-0.4780**	(0.2173)	-0.1072	(0.2052)
All workers	8.81	8.52	8.45	-0.2880***	(0.08744)	-0.3566**	(0.1445)	-0.06859	(0.1353)
Worked full-time, year-round	5.02	4.87	4.78	-0.1564***	(0.05475)	-0.2389**	(0.1073)	-0.08252	(0.09937)
Less than full-time, year-round	17.54	16.95	16.91	-0.5917**	(0.2451)	-0.6281*	(0.3360)	-0.03643	(0.3151)
Did not work at least 1 week	31.76	31.13	30.91	-0.6229	(0.4379)	-0.8479	(0.5546)	-0.2250	(0.5632)
Total 18 to 64 years	14.48	14.10	14.00	-0.3707***	(0.1418)	-0.4780**	(0.2173)	-0.1072	(0.2052)
With a disability	25.80	25.39	24.51	-0.4102	(0.8085)	-1.285	(0.9660)	-0.8749	(0.9680)
With no disability	13.60	13.23	13.19	-0.3697***	(0.1324)	-0.4139*	(0.2184)	-0.04416	(0.2086)

Notes: This table compares the estimated SPM rates using survey SNAP response (1), state administrative data, or Adrec states, (2) for the eight states studied in this paper, and imputed SNAP benefits from the pooled leave-one-out models (3) for the same eight states. Columns (4) and (6) shows the difference between the survey estimates and administrative records estimates and leave-one-out imputed estimates, respectively. Column (8) shows the differences in estimates using the administrative data compared to the pooled leave-one-out imputations. SPM estimated with administrative data and with the leave-one-out imputations is lower for most groups than using the survey data, due to under-reporting of SNAP benefits. However, the leave-one-out estimates are not statistically different from the administrative estimates for any group except unrelated individuals. In the difference estimates, ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively. Z indicates an estimate rounds to zero.
Source: 2014 CPS ASEC Traditional File linked to state SNAP administrative records for eight states: Arizona, Idaho, Maryland, Michigan, New York, North Dakota, Tennessee and Virginia.

Table 3: Comparison of SNAP Administrative Data to Imputes

				Annual SNAP Amount																	
							Percentile														
	SNAP Receipt			Average			10th			25th			Median			75th			90th		
	Adrec (1)	Imputed (2)	Difference (3)	Adrec (4)	Imputed (5)	Difference (6)	Adrec (7)	Imputed (8)	Difference (9)	Adrec (10)	Imputed (11)	Difference (12)	Adrec (13)	Imputed (14)	Difference (15)	Adrec (16)	Imputed (17)	Difference (18)	Adrec (19)	Imputed (20)	Difference (21)
Adrec States	0.200	0.201	0.002	2,613	2,388	-225**	396	315	-82*	992	833	-160*	2,060	1,870	-189**	3,693	3,200	-493*	5,821	5,479	-342
	(0.008)	(0.010)	(0.009)	(64)	(104)	(111)	(33)	(35)	(47)	(54)	(70)	(84)	(48)	(80)	(87)	(127)	(243)	(257)	(185)	(449)	(458)
Each Individually																					
Arizona	0.217	0.223	0.006	2,688	2,323	-365	285	347	62	884	890	5	1,842	1,902	60	3,950	3,173	-777	6,134	4,982	-1,152
	(0.056)	(0.048)	(0.028)	(194)	(256)	(294)	(47)	(89)	(102)	(186)	(167)	(242)	(127)	(222)	(250)	(219)	(578)	(593)	(719)	(853)	(951)
Idaho	0.166	0.159	-0.007	2,361	1,991	-371	270	288	18	657	699	42	1,510	1,544	35	3,462	2,526	-936	5,742	4,184	-1,558
	(0.020)	(0.025)	(0.024)	(321)	(320)	(434)	(70)	(114)	(125)	(115)	(204)	(217)	(204)	(264)	(318)	(766)	(617)	(903)	(943)	(887)	(1,269)
Maryland	0.141	0.142	0.001	2,630	2,670	40	281	430	149	771	1,098	328	1,934	2,069	135	3,902	3,659	-243	6,286	5,805	-480
	(0.011)	(0.022)	(0.022)	(198)	(355)	(376)	(84)	(176)	(192)	(163)	(255)	(280)	(184)	(248)	(290)	(300)	(731)	(754)	(298)	(840)	(876)
Michigan	0.230	0.218	-0.012	2,760	2,562	-198	503	345	-157	1,053	999	-53	2,060	1,974	-86	3,794	3,428	-367	6,193	5,828	-365
	(0.015)	(0.027)	(0.027)	(151)	(396)	(409)	(78)	(123)	(142)	(129)	(382)	(386)	(102)	(336)	(344)	(398)	(731)	(804)	(223)	(1,048)	(1,050)
New York	0.216	0.216	-0.001	2,588	2,281	-307**	560	305	-255***	1,219	783	-436***	2,192	1,750	-442***	3,511	3,024	-487	4,939	5,104	165
	(0.010)	(0.015)	(0.015)	(109)	(121)	(142)	(74)	(42)	(82)	(122)	(84)	(139)	(47)	(113)	(116)	(251)	(342)	(385)	(356)	(535)	(610)
North Dakota	0.130	0.142	0.013	2,384	2,657	273	472	387	-85	1,028	908	-120	1,999	1,952	-47	3,157	3,602	445	4,729	6,322	1,593
	(0.019)	(0.021)	(0.017)	(252)	(473)	(433)	(144)	(185)	(220)	(293)	(394)	(460)	(196)	(432)	(429)	(317)	(870)	(841)	(634)	(1,067)	(989)
Tennessee	0.253	0.274	0.021	2,532	2,607	75	264	379	115	745	969	224	1,786	2,076	290	3,868	3,519	-350	6,158	5,943	-215
	(0.021)	(0.046)	(0.045)	(208)	(252)	(265)	(55)	(121)	(126)	(139)	(203)	(217)	(176)	(141)	(178)	(356)	(461)	(488)	(421)	(825)	(888)
Virginia	0.118	0.122	0.004	2,453	1,989	-465	414	218	-195	1,021	558	-463*	2,027	1,456	-571**	3,281	2,521	-760*	5,006	4,777	-230
	(0.011)	(0.017)	(0.015)	(194)	(309)	(322)	(127)	(61)	(138)	(210)	(156)	(249)	(152)	(257)	(279)	(315)	(369)	(440)	(911)	(1,260)	(1,388)
All Imputed	0.192			2,676			341			900			2,059			3,792			6,077		
	(0.004)			(83)			(23)			(52)			(48)			(174)			(122)		
All States (Adrec + Imputed)	0.194			2,663			351			917			2,059			3,771			6,035		
	(0.004)			(68)			(20)			(43)			(39)			(141)			(104)		

Notes: This table compares various estimates of SNAP receipt and benefits from administrative SNAP data and the imputation model. "Adrec States" indicate estimates from the pooled sample of states with administrative SNAP data available. For the pooled adrec states and each individually, column (2) shows the results from the leave-one-out (LOO) imputation. In the LOO model, the state's administrative SNAP records are ignored and SNAP receipt and benefit amounts are imputed using data from the remaining seven states. This allows us to test the quality of the imputation by comparing the model results to the SNAP administrative data for each state. The "All Imputed" row shows the imputation estimates for the 42 states and DC where administrative records are not available. The estimates in "All States" combine the administrative data for the eight states where it is available with the imputed data for the remaining states and DC. In the difference estimates, ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

Source: 2014 CPS ASEC Traditional File linked to state SNAP administrative records for eight states: Arizona, Idaho, Maryland, Michigan, New York, North Dakota, Tennessee and Virginia.

Table 4: Predicting "True" SNAP Receipt from Survey Responses

	Pooled Adrec States			All States	
	Adrecs	Leave-One-Out	Difference	Adrec + Imputed	Difference from Adrec States
	(1)	(2)	(3)	(4)	(5)
Survey SNAP Recipient	0.729***	0.752***	0.023	0.738***	0.009
	(0.032)	(0.041)	(0.053)	(0.023)	(0.040)
Imputed Survey SNAP Receipt	0.195***	0.243***	0.048	0.211***	0.017
	(0.022)	(0.026)	(0.031)	(0.035)	(0.041)
Female	-0.019***	-0.014	0.005	-0.017**	0.002
	(0.007)	(0.009)	(0.011)	(0.008)	(0.011)
Race					
Black	0.079***	0.055***	-0.025	0.087***	0.008
	(0.014)	(0.013)	(0.018)	(0.015)	(0.020)
Native American	0.013	0.016	0.003	0.016	0.003
	(0.032)	(0.039)	(0.041)	(0.024)	(0.040)
Asian	0.013	0.022	0.009	0.037	0.024
	(0.016)	(0.019)	(0.024)	(0.037)	(0.040)
Pacific Islander	-0.128***	-0.008	0.120	0.046	0.174***
	(0.038)	(0.188)	(0.204)	(0.045)	(0.058)
Hispanic	0.074***	0.064***	-0.010	0.080***	0.006
	(0.013)	(0.020)	(0.023)	(0.008)	(0.015)
Survey SNAP Amount	0.007	0.006	-0.001	0.011	0.003
	(0.014)	(0.017)	(0.022)	(0.011)	(0.018)
Survey SNAP Amount^2	-0.00062	-0.00064	-0.00001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Age	-0.001	0.001	0.002	-0.001	0.000
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
Age^2	-0.000017	-0.000029	-0.000012	-0.000015	0.000002
	(0.000015)	(0.000022)	(0.000025)	(0.000017)	(0.000023)
Household Income	0.019*	0.022	0.003	-0.0004	-0.0197
	(0.010)	(0.022)	(0.026)	(0.009)	(0.013)
Household Income^2	-0.00306***	-0.00303***	0.00003	-0.0020***	0.0011*
	(0.0005)	(0.0010)	(0.0012)	(0.0004)	(0.0006)
Household Income $\neq$ 0	0.090	-0.002	-0.092	0.114	0.024
	(0.083)	(0.142)	(0.167)	(0.097)	(0.127)
Constant	0.280***	0.275***	-0.005	0.323***	0.043
	(0.062)	(0.089)	(0.104)	(0.060)	(0.085)
R2	0.48	0.48		0.43	
N	7,400	7,400		39,000	

Notes: This table compares regressions results using SNAP administrative data and imputed SNAP benefits. In each column, we regress a dummy for "true" SNAP receipt on administrative SNAP reports (Adrecs) and various individual characteristics, for household heads. In column (1), we show the results using the SNAP administrative data in the eight available states. In column (2), we show the results in the same states using the pooled leave-one-out imputations. In column (3), we show the difference between (2) and (1). In column (4), we show the regression results using the available SNAP administrative data as well as the imputed data for the remaining states and DC. Column (5) shows the difference between (4) and (1). Household income and survey SNAP amount transformed using the inverse hyperbolic sine function for the regressions. ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.
Source: 2014 CPS ASEC Traditional File linked to state SNAP administrative records for eight states: Arizona, Idaho, Maryland, Michigan, New York, North Dakota, Tennessee and Virginia.

Table 5: Association between Earnings and SNAP Receipt

	SNAP Receipt Source					Comparisons (Columns Compared in Brackets)				Difference in Differences
	Administrative			Survey		Administrative		Survey - Administrative		Survey - Adrecs
	Adrec States		Non-Adrec States	Adrec States	Non-Adrec States	LOO - Adrecs	Imputed - Adrecs	Adrec States	Non-Adrec States	Non-Adrec - Adrec States
	Adrecs	Leave-One-Out	Imputed			[(2)-(1)]	[(3)-(1)]	[(4)-(1)]	[(5)-(3)]	[(9)-(8)]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
SNAP Receipt	-0.175***	-0.178***	-0.172***	-0.227***	-0.223***	-0.00309	0.00302	-0.0521***	-0.0512***	0.000857
	(0.0185)	(0.0184)	(0.0141)	(0.0225)	(0.00984)	(0.0175)	(0.0234)	(0.0148)	(0.0146)	(0.0291)
Female	0.0284***	0.0269***	0.0436***	0.0256***	0.0413***	-0.00150	0.0151	-0.00285	-0.00227	0.000583
	(0.00951)	(0.00971)	(0.00447)	(0.00972)	(0.00440)	(0.00263)	(0.0104)	(0.00195)	(0.00146)	(0.00306)
Urban	0.00465	0.00169	0.00714	0.00373	0.00819	-0.00297	0.00249	-0.000921	0.00105	0.00197
	(0.0110)	(0.0117)	(0.00665)	(0.0115)	(0.00650)	(0.00331)	(0.0135)	(0.00227)	(0.00121)	(0.00481)
Race/Ethnicity										
Black	-0.00384	-0.00541	-0.0486***	-0.00519	-0.0522***	-0.00157	-0.0447**	-0.00135	-0.00359	-0.00224
	(0.0149)	(0.0146)	(0.00927)	(0.0147)	(0.00769)	(0.00435)	(0.0185)	(0.00364)	(0.00525)	(0.0110)
Native American	-0.0280	-0.0275	-0.0468**	-0.0174	-0.0432**	0.000478	-0.0188	0.0106	0.00357	-0.00701
	(0.0336)	(0.0352)	(0.0191)	(0.0348)	(0.0190)	(0.00819)	(0.0385)	(0.00679)	(0.00524)	(0.0177)
Asian	-0.0153	-0.00975	-0.00147	-0.0200	-0.0123	0.00553	0.0138	-0.00472*	-0.0108*	-0.00612
	(0.0192)	(0.0189)	(0.0110)	(0.0192)	(0.00953)	(0.00461)	(0.0226)	(0.00276)	(0.00612)	(0.0170)
Pacific Islander	0.0682*	0.0967	0.0153	0.0789**	0.0122	0.0286	-0.0529	0.0107*	-0.00303	-0.0137
	(0.0410)	(0.0601)	(0.0238)	(0.0375)	(0.0222)	(0.0406)	(0.0465)	(0.00631)	(0.00980)	(0.0323)
Hispanic	0.00884	0.00787	0.0316***	0.00304	0.0231***	-0.000971	0.0228	-0.00580	-0.00849***	-0.00269
	(0.0199)	(0.0191)	(0.00648)	(0.0199)	(0.00620)	(0.00550)	(0.0216)	(0.00400)	(0.00168)	(0.00690)
Education										
High School	0.118***	0.119***	0.101***	0.109***	0.101***	0.000688	-0.0177	-0.00878*	0.000360	0.00914
	(0.0257)	(0.0257)	(0.0112)	(0.0255)	(0.0108)	(0.00698)	(0.0289)	(0.00502)	(0.00368)	(0.0195)
Some College	-0.0206	-0.0215	-0.00163	-0.0215	-0.00323	-0.000909	0.0189	-0.000938	-0.00159	-0.000655
	(0.0182)	(0.0184)	(0.00735)	(0.0183)	(0.00694)	(0.00446)	(0.0193)	(0.00306)	(0.00284)	(0.00595)
Associates	0.0630***	0.0645***	0.0224***	0.0609***	0.0229***	0.00151	-0.0406**	-0.00208	0.000499	0.00258
	(0.0177)	(0.0184)	(0.00763)	(0.0186)	(0.00753)	(0.00602)	(0.0188)	(0.00463)	(0.00217)	(0.00730)
Bachelors	-0.00478	-0.00358	0.0133**	-0.000165	0.0148**	0.00119	0.0181	0.00461	0.00149	-0.00312
	(0.0138)	(0.0143)	(0.00658)	(0.0138)	(0.00650)	(0.00498)	(0.0147)	(0.00364)	(0.00180)	(0.00743)
Masters	0.00919	0.0109	0.0123**	0.0122	0.0143**	0.00166	0.00311	0.00303*	0.00200	-0.00103
	(0.0124)	(0.0127)	(0.00602)	(0.0125)	(0.00598)	(0.00308)	(0.0138)	(0.00174)	(0.00137)	(0.00338)
Age	0.0306***	0.0312***	0.0244***	0.0306***	0.0249***	0.000565	-0.00615	0.0000306	0.000408	0.000377
	(0.00393)	(0.00398)	(0.00168)	(0.00395)	(0.00168)	(0.000918)	(0.00440)	(0.000634)	(0.000384)	(0.00117)
Age^2	-0.000399***	-0.000405***	-0.000333***	-0.000400***	-0.000337***	-0.00000599	0.0000667	-0.000000315	-0.00000433	-0.00000402
	(0.0000449)	(0.0000454)	(0.0000193)	(0.0000451)	(0.0000194)	(0.0000103)	(0.0000508)	(0.00000691)	(0.00000399)	(0.0000124)
Constant	0.258***	0.248***	0.405***	0.263***	0.390***	-0.0103	0.146	0.00434	-0.0149	-0.0192
	(0.0914)	(0.0910)	(0.0373)	(0.0905)	(0.0364)	(0.0234)	(0.0999)	(0.0163)	(0.0135)	(0.0478)
R^2	0.14	0.14	0.14	0.15	0.14					
N	5,200	5,200	27,500	5,200	27,500					

Notes: This table compares regression results using three sources of information on SNAP receipt: 1) administrative data (Adrecs), 2) imputed data, and 3) survey responses. In each column, we regress a dummy for having earnings on SNAP receipt and various individual characteristics, for working-age (18-64 years old) household heads with survey-reported SNAP and earnings receipt (not imputed). In column (1), we show the results using the SNAP administrative data in the eight available states. In column (2), we show the results in the same states using the pooled leave-one-out imputations. In column (3), we show the results using the imputed data in the other 42 states and DC. Columns (4) and (5) show the same results using survey-reported SNAP receipt in the adrec and non-adrec states, respectively. Columns (6)-(10) show comparisons of the results shown in (1)-(5). ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.
Source: 2014 CPS ASEC Traditional File linked to state SNAP administrative records for eight states: Arizona, Idaho, Maryland, Michigan, New York, North Dakota, Tennessee and Virginia.

Table 6: Comparison of SPM using Survey and Combined Administrative and Imputed SNAP Data

	SPM Estimate		Differences Between Estimates			
	Survey Only	SNAP Adrecs	Adrecs with Imputes - Survey		Difference in Difference Comparison Adrecs - Survey Estimates for All States - Adrec States	
	(Official) (1)	with Imputes (2)	Estimate (3)	SE (4)	Estimate (5)	SE (6)
All People	14.58	14.36	-0.22***	(0.08)	0.18	(0.13)
Male	13.93	13.65	-0.27***	(0.08)	0.20	(0.15)
Female	15.21	15.04	-0.18*	(0.09)	0.16	(0.15)
Under 18 years	15.40	15.38	-0.02	(0.17)	0.60**	(0.26)
18 to 64 years	14.35	14.08	-0.28***	(0.07)	0.09	(0.13)
65 years and older	14.24	13.89	-0.34***	(0.10)	-0.12	(0.16)
Married couple unit	8.88	8.71	-0.17***	(0.06)	0.24*	(0.13)
Cohabiting partner unit	16.15	16.17	0.02	(0.35)	0.66	(0.59)
Female reference person unit	29.08	28.94	-0.14	(0.46)	0.86	(0.67)
Male reference person unit	20.17	19.87	-0.30	(0.38)	-0.62	(0.71)
Unrelated individuals	23.77	23.09	-0.69***	(0.17)	-0.76***	(0.26)
White	12.75	12.55	-0.20**	(0.08)	0.24	(0.15)
White, not Hispanic	10.23	10.11	-0.12	(0.09)	0.24*	(0.14)
Black	24.14	23.56	-0.58	(0.43)	-0.26	(0.57)
Asian	15.38	14.97	-0.40	(0.40)	-0.40	(0.40)
Hispanic (any race)	23.84	23.43	-0.41	(0.34)	0.39	(0.52)
Native born	13.58	13.40	-0.19**	(0.08)	0.19	(0.14)
Foreign born	21.84	21.34	-0.49***	(0.19)	0.09	(0.24)
Naturalized citizen	17.13	16.66	-0.47	(0.29)	-0.15	(0.34)
Not a citizen	26.56	26.03	-0.52	(0.38)	0.36	(0.45)
Total, aged 25 and older	13.32	13.04	-0.28***	(0.06)	0.02	(0.11)
No high school diploma	29.86	28.36	-1.50***	(0.37)	0.31	(0.65)
High school, no college	16.02	15.83	-0.20*	(0.10)	0.02	(0.21)
Some college, no degree	11.76	11.69	-0.08	(0.13)	-0.04	(0.19)
Bachelor's degree or higher	6.43	6.31	-0.12**	(0.05)	-0.01	(0.06)
Owner	9.66	9.57	-0.09	(0.07)	0.03	(0.11)
Owner/mortgage	8.03	7.95	-0.08	(0.06)	0.01	(0.11)
Owner/no mortgage/rentfree	12.60	12.49	-0.10	(0.17)	0.06	(0.26)
Renter	25.65	25.12	-0.52**	(0.21)	0.48	(0.36)
Inside MSAs	14.91	14.68	-0.22**	(0.09)	0.10	(0.15)
Inside principal cities	18.68	18.33	-0.35*	(0.20)	0.08	(0.27)
Outside principal cities	12.63	12.48	-0.15*	(0.09)	0.10	(0.18)
Outside MSAs	12.83	12.61	-0.22	(0.26)	0.76	(0.47)
Northeast	13.00	13.02	0.02	(0.21)	0.43*	(0.25)
Midwest	11.99	11.62	-0.36**	(0.14)	0.12	(0.26)
South	15.04	14.79	-0.25	(0.21)	0.10	(0.29)
West	17.56	17.33	-0.23	(0.19)	0.20	(0.36)
With private insurance	7.70	7.55	-0.14***	(0.05)	0.07	(0.08)
With public, no private insurance	27.44	27.29	-0.15	(0.26)	0.59	(0.48)
Not insured	27.75	26.98	-0.78***	(0.22)	0.14	(0.42)
Total 18 to 64 years	14.35	14.08	-0.28***	(0.07)	0.09	(0.13)
All workers	9.19	8.92	-0.27***	(0.06)	0.02	(0.09)
Worked full-time, year-round	4.95	4.77	-0.18***	(0.05)	-0.02	(0.06)
Less than full-time, year-round	18.51	18.05	-0.47***	(0.14)	0.13	(0.24)
Did not work at least 1 week	30.70	30.39	-0.31*	(0.18)	0.32	(0.40)
Total 18 to 64 years	14.35	14.08	-0.28***	(0.07)	0.09	(0.13)
With a disability	26.40	25.57	-0.83**	(0.34)	-0.42	(0.73)
With no disability	13.37	13.14	-0.23***	(0.07)	0.14	(0.13)

Notes: This table compares the estimated SPM rates using survey SNAP response (1), state administrative data (Adrecs) for the available eight states and imputed SNAP benefits in the remaining states (2). Column (3) shows the difference between the estimates using the administrative data-based estimates in (2) and the survey-based estimates in (1), with the standard errors shown in (4). Column (5) shows the difference in difference comparison between the estimates in (3) and the same estimates in Table 2 using the sample of households in states with administrative data, with standard errors in (6). In the estimates in (3) and (5), ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.
Source: 2014 CPS ASEC Traditional File linked to state SNAP administrative records for eight states: Arizona, Idaho, Maryland, Michigan, New York, North Dakota, Tennessee and Virginia.