Preliminary work. Do not cite without permission.

The Implications of Misreporting for Longitudinal Studies of SNAP

Brian Stacy, Economic Research Service, USDA

Erik Scherpf, Economic Research Service, USDA

December 20, 2018

Abstract

Researchers studying a variety of important economics, nutrition, and health topics use survey data containing information on SNAP participation. In order to study the dynamics of SNAP participation or recognizing possible selection bias in cross-sectional estimators, many researchers use longitudinal estimators to estimate the causal effects of SNAP. However, misreporting of SNAP participation is common in survey datasets, and bias from misreporting can be larger for longitudinal estimators. In an analysis of data combining newly compiled administrative datasets on SNAP participation from nine states and covering the years 2005-2015 with individual records from the CPS ASEC survey, we confirm findings in previous studies of substantial misreporting and find evidence that the misreporting is not done at random. Additionally, we examine bias caused by misreporting in a longitudinal estimators and find severe bias, much greater in magnitude than bias caused by misreporting in cross-sectional estimators. We find that a longitudinal conditional distribution estimator may be an attractive solution for researchers using public use survey datasets.

Keywords: Food Assistance Programs, Survey Datasets, Misreporting, Panel Data Econometrics

The findings and conclusions in this preliminary publication have not been formally disseminated by the USDA and should not be construed to represent any agency determination or policy. This research was supported in part by the intramural research program of the USDA, Economic Research Service.

The question of whether the Supplemental Nutrition Assistance Program (SNAP), formerly known as food stamps, improves food security and other outcomes, and whether it has other undesirable effects, such as a negative effect on labor supply or obesity, continues to be a topic of fierce debate. SNAP is the largest near cash, low-income assistance program in the United States, and provided benefits to 44.2 million individuals in fiscal year 2017 according to Oliveira (2018). Any serious study of the topic recognizes that a simple comparison of outcomes between SNAP participants and non-SNAP participants would yield biased estimates of the effects of interest, because SNAP participation is not randomly assigned. A common approach to dealing with this selection bias is to use longitudinal methods, such as first differencing, where for instance outcomes are examined before and after an individual participates in SNAP.[1] These methods can help account for selection on the basis of time-constant individual factors that may relate both to SNAP participation and the outcomes of interest. Aside from issues related to selection bias, many researchers are directly interested in using longitudinal estimators of SNAP to examine the dynamics of SNAP participation and how it relates to policies or individual characteristics.

Additionally, several recent studies have found substantial misreporting of SNAP participation in commonly used data sets, including the Current Population Survey, the National Longitudinal Survey of Youth, the Survey of Income and Program Participation, and the American Community Survey.[2] Moreover, several studies have pointed out that bias induced by misreporting can potentially be larger for longitudinal estimators than cross sectional estimators.[3] By using a longitudinal estimator, for instance a first difference OLS estimator or a conditional logit estimator, researchers may be helping to address one bias, selection bias, while exacerbating another, bias from using a misreported measure of SNAP participation.

In this paper, we examine the extent of bias in longitudinal estimators caused by misreported SNAP participation and offer recommendations for researchers. Our paper is the

[1] A few examples include Heflin & Ziliak (2008) who use a first difference estimator and data from the Panel Study of Income Dynamics, Nord & Golla (2009) and Wilde & Nord (2005) who apply fixed effects methods using a longitudinally linked version of the December CPS, Baum (2011) uses an individual fixed effects estimator using data from the NLSY, 1979.

[2] See Meyer & Mittag (2015), Meyer, Mittag, & Goerge (2018), Mittag (2018), Courtemanche, Denteh, & Tchernis (2018), Kreider, Pepper, Gundersen, & Jolliffe (2012), Bollinger & David (1997), and Almada, McCarthy, & Tchernis (2016).

[3] See Bound & Krueger (1991), Kim & Solon (2005), or Freeman (1984).

first in which we are aware to directly examine errors in *changes* in reported SNAP participation over time, and the impacts of this type of reporting error on longitudinal estimators of the effects of SNAP. Using a novel linkage of SNAP administrative data from nine states between 2005 and 2015 combined with survey data from the Current Population Survey, Annual Social and Economic Supplement (CPS ASEC), we compare estimates making use of the survey based measure of SNAP participation to estimates based on an administrative data verified measure of SNAP participation. Because of the structure of the CPS survey, where survey respondents are interviewed each month for four consecutive months, are out of the survey for eight months, and then interviewed again for another four months. This survey design allows data users to match household interviews in consecutive years, and we are able to construct a short (two-period) panel dataset.

Our paper has three main findings. First, we replicate earlier work and find substantial misreporting error in an indicator for SNAP participation based on survey data. However, the degree of error for the first difference of SNAP participation is much more severe. Of those individuals identified as having switched onto SNAP in the survey, meaning they reported being off SNAP in period 1 and on SNAP in period 2, roughly 86% of these switches were false according to the administrative data. Of those identified as switching off SNAP based on the survey reports, roughly 87% of these switches were false according to the administrative data, meaning the vast majority of survey based reports of changes in SNAP participation are noise rather than signal. These results imply that a first-differenced OLS estimator using a survey based measure of SNAP participation could be attenuated by nearly a factor of 10, because of bias from misreporting. Second, as an application, we examine estimates of the effect of SNAP participation on labor supply using our CPS ASEC data, and find mild/modest differences between cross-sectional estimates using a survey report and an administrative data verified report of SNAP participation, and large differences when using longitudinal estimators. Third, we examine potential solutions to reduce bias in longitudinal estimators. Specifically, we find that conditional distribution methods, similar to those in Mittag (2018), but modified for longitudinal data analysis, perform well in reducing bias from misreporting.

Our work is most closely linked to Kreider et al. (2012) and Almada et al. (2016), who examine the impacts of misreporting on estimates of SNAP. Kreider et al. (2012) use a non-parametric bounding approach combined with information on the size of the SNAP caseload to place bounds on the effects of SNAP on child health outcomes. Almada et al. (2016) examine the impacts of misreporting in an individual fixed effects IV regression of a measure of obesity on SNAP participation using National Longitudinal Survey of Youth - 1979 data. The authors use a parametric approach of modeling misreporting, as well as a non-parametric bounding approach, to examine the impacts of misreporting. Instead of parametric and non-parametric modeling, our paper will be able to observe misreporting directly using linked SNAP administrative records on SNAP receipt to survey reports from the Current Population Survey, Annual Social and Economic Supplement (CPS ASEC) for nine states from 2004-2015. Our paper is also related to Meyer & Mittag (2017), who use linked survey data from the ACS and CPS to SNAP administrative records. These authors specifically examine bias from misreporting of SNAP on the left-hand side of a linear probability and a probit model. Our paper differs by examining bias from misreporting of SNAP when it is included as a right hand side variable and will also examine bias in longitudinal estimators.

In the remainder of the paper, we will discuss econometric issues associated with misreporting in cross-sectional and first-differences OLS estimators. We discuss our novel data linkage between the CPS ASEC survey and monthly SNAP administrative data on benefit issuance. We show estimates of commonly used cross-sectional and longitudinal estimators using both survey reports of SNAP and our SNAP administrative data verified measure. We conclude by offering a solution for researchers, which is to use conditional distribution methods to adjust for misreporting.

Econometric Issues

We will begin by comparing the econometric issues associated with misreporting of an independent variable for a cross-sectional OLS estimator and an OLS estimator based on first-differencing the outcome and the independent variables. To highlight the primary issues, consider the simple model:

$$Y_{it} = \alpha + \beta S_{it} + u_i + \epsilon_{it} \tag{1}$$

Where Y_{it} is our outcome of interest for an individual, a measure of food insecurity, or obesity, or a labor supply outcome as the case may be, and let S_{it} be the "true", binary measure of participation in SNAP. Let u_i be a time constant, individual heterogeneity term, or fixed effect, that may be correlated with S_{it}, and ϵ_{it} is an idiosyncratic error term. Also, let

$$R_{it} = S_{it} + v_{it} \tag{2}$$

be a measure of SNAP participation that is potentially misreported. In this case the possible errors are: $v_{it} = -1$ which is an under-report where the individual is truly on SNAP ($S_{it} = 1$) but the individual is reported to not be on SNAP ($R_{it} = 0$) or $v_{it} = 1$ which is an over-report where individual is truly not on SNAP ($S_{it} = 0$), but the individual is reported to be on SNAP ($R_{it} = 1$). $v_{it} = 0$ would indicate that the survey report is accurate. Because of the binary nature of the SNAP participation measure, the misreporting error, v_{it}, is mechanically correlated with the "true" measure,S_{it}, and as a result typical attenuation bias results based on classical measurement error assumptions do not hold.

However, it is possible to characterize bias caused by misreporting of a binary independent variable. As shown in Freeman (1984) and Aigner (1973), in the cross-sectional case the simple OLS estimator of Y_{it} on R_{it}, which would need to assume that $u_i \; and \; \epsilon_{it}$ are uncorrelated with R_{it} has the following properties:

$$E(Y_{it} \mid R_{it} = 1) = \alpha + \beta P(S_{it} = 1 \mid R_{it} = 1) \tag{3}$$

$$E(Y_{it} \mid R_{it} = 0) = \alpha + \beta P(S_{it} = 1 \mid R_{it} = 0) \tag{4}$$

Which, because the cross-sectional simple OLS estimator is the difference of the two, implies:

$$plim \; \hat{\beta}_{OLS} = \beta[P(S_{it} = 1 \mid R_{it} = 1) - P(S_{it} = 1 \mid R_{it} = 0)] \tag{5}$$

Which after some simple rearrangement is

$$plim \; \hat{\beta}_{OLS} = \beta[1 - (P(S_{it} = 0 \mid R_{it} = 1) + P(S_{it} = 1 \mid R_{it} = 0))] \tag{6}$$

This implies the bias caused by misreporting in the simple OLS estimator is proportional to one minus the probability of an under-report, $P(S_{it} = 0 \mid R_{it} = 1)$, plus the probability of an over-report, $P(S_{it} = 1 \mid R_{it} = 0)$. This also implies that the cross-sectional estimator will suffer an attenuation bias as a result of misreporting. As also shown in Freeman (1984), when covariates are added to the regression the magnitude of the bias generally increases.

However, researcher may be concerned that S_{it} is correlated with u_i and therefore apply longitudinal methods to account for this. The first differenced model then is

$$\Delta Y_{it} = \beta \Delta S_{it} + \Delta \epsilon_{\text{it}} \tag{7}$$

And

$$\Delta R_{it} = \Delta S_{it} + \Delta v_{it} \tag{8}$$

Under some strong assumptions about the independence of the misreporting error, v_{it}, over time, it is again possible to characterize bias caused by misreporting in the first-differenced estimator.[4] However, there is reason to be skeptical that v_{it} is really uncorrelated over time. For instance, it is plausible that one reason an individual would misreport SNAP participation is the stigma associated with participation, which may lead an individual to continuously misreport over time. Fortunately, Bound & Krueger (1991) and Bound, Brown, Duncan, & Rodgers (1994) offer another way to characterize the bias from misreporting, which does not place assumptions on the autocorrelation of the misreporting.

The simple OLS estimator using the reported measure of SNAP participation, R_{it}, is

$$\hat{\beta}_{OLS} = \frac{Cov(\mathrm{R_{it}}, Y_{it})}{Var(R_{it})} = \underbrace{\beta \frac{Cov(\mathrm{R_{it}}, S_{it})}{Var(R_{it})}}_{Misreport\ Bias} + \underbrace{\frac{Cov(\mathrm{R_{it}}, u_i)}{Var(R_{it})} + \frac{Cov(\mathrm{R_{it}}, \epsilon_{it})}{Var(R_{it})}}_{Selection/Omitted\ Variable\ Bias} \tag{9}$$

In this case, assuming no correlation between R_{it} and u_i *and* ϵ_{it},

$$plim\ \hat{\beta}_{OLS} = \beta \lambda_{OLS} \tag{10}$$

Where

$$\lambda_{OLS} = \frac{Cov(\mathrm{R_{it}}, S_{it})}{Var(R_{it})} \tag{10}$$

Similarly,

$$plim\ \hat{\beta}_{FD} = \beta \lambda_{FD} \tag{11}$$

and

---

[4] As shown in Freeman (1984), the bias in the first-differenced estimator, assuming independence of the misreporting error over time, can be described as:

$$plim\ \hat{\beta}_{FD} = \beta[1 - (P(S_{it} = 0 \mid R_{it} = 1) + P(S_{it} = 1 \mid R_{it} = 0))] \frac{\sigma^2_{\Delta_S}}{\sigma^2_{\Delta_R}}$$

Where $\boldsymbol{\sigma^2_{\Delta_S}}$ is the variance of true changes in SNAP participation and $\boldsymbol{\sigma^2_{\Delta_R}}$ is the variance of reported changes. This implies that the bias in the first difference OLS estimator will exceed the bias in the cross-sectional estimator as long as $\boldsymbol{\sigma^2_{\Delta_S} < \sigma^2_{\Delta_R}}$.

$$\lambda_{FD} = \frac{Cov(\Delta \mathrm{R}_{\mathrm{it}}, \Delta S_{it})}{Var(\Delta R_{it})} \tag{13}$$

Which is simply the regression coefficient of a regression of an individual's change in true SNAP participation status, ΔS_{it}, on the survey based measure of change in SNAP participation, ΔR_{it}. The intuition, as described in Bound & Krueger (1991), is that the bias term, λ, is the degree that a change in reported SNAP participation translates into an actual change in SNAP participation. We will directly estimate λ_{OLS} and λ_{FD} in order to characterize the bias caused by misreporting of SNAP participation using our linked CPS ASEC and SNAP administrative data, as we will describe below.

Data

We make use of a novel linkage between SNAP administrative records from nine states on SNAP receipt for individuals to a restricted use version of the Current Population Survey Annual Social and Economic Supplement (CPS ASEC) files from 2005 to 2015. Our dataset combining SNAP administrative records to individual survey reports from the CPS ASEC allow us to verify SNAP receipt in the CPS ASEC using our admin records. A number of studies (Meyer & Mittag (2015), Meyer, Mok, & Sullivan (2015), Kreider et al. (2012)) have chronicled misreporting of public assistance program participation in survey datasets, including the CPS ASEC. Meyer et al. (2018) find that around 50 percent of true SNAP participants do not report SNAP receipt in the CPS ASEC, and the authors find that misreports of SNAP participation are non-random, varying with household characteristics, which may introduce complicated biases in estimation of the impacts of SNAP. By linking our SNAP administrative records to the CPS, we are able to utilize a more accurate measure of SNAP participation, that coming from administrative data. Our CPS sample was limited to individuals who were between the ages of 18 and 64.

We additionally use longitudinally linked March CPS data from 2005 to 2015 to answer our research questions. The CPS samples roughly 50,000 households every month. Sampled households are interviewed each month for four consecutive months, are out of the survey for eight months, and then interviewed again for another four months. This survey design allows data users to match household interviews in consecutive years; for instance, matching March of previous year to March of the current year. Taking advantage of this design, we construct a set of a short (two-period) panel dataset, in which we observe individuals in consecutive surveys following the approach laid out in Madrian & Lefgren (1999).

The March CPS was chosen for its Annual Social and Economic Supplemental (ASEC), which contains information on income, workforce participation, hours worked, poverty, program participation, and benefits received during the previous calendar year. The supplement asks whether anyone in the household received food stamps during the previous twelve months. It also asks each household member whether he or she participated in the labor force during the previous year and about typical hours worked per week and weeks worked.

Our study also makes use of SNAP administrative records from a diverse set of nine anonymous states, spanning the years 2005 to 2015. Unfortunately, not all states have data available for all years. The states, with the span of years available for each state in parenthesis, include: State 1 (2005-2015), State 2 (2005-2015), State 3 (2006-2015), State 4 (2007-2015), State 5 (2008-2015), State 6 (2009-2015), State 7 (2009-2015), State 8 (2009-2013), and State 9 (2010-2015). From 2010 to 2013, when all nine states had data available, these states represented approximately 32% of the total individual SNAP caseload. In 2005, this still represented 5% of the total caseload. The data include all monthly SNAP payments to individuals in each of the nine states. Using this data, we are able to form a measure of SNAP participation that corresponds to the CPS ASEC survey question, by creating an indicator of whether an individual was enrolled in SNAP according to the administrative records during the 12 month reference period in the CPS ASEC. Because the ASEC is fielded annually in March, this means creating a measure based on whether or not an individual can be found as receiving SNAP in the administrative records from March of the previous year to the March in which they are surveyed.

We linked the SNAP administrative records and individuals records from the CPS ASEC using the Person Identification Validation System (PVS) of the U.S. Census Bureau. The PVS system uses information available in the SNAP administrative records, including name, address, and birth date to match to a database of social security numbers, which are then anonymized (known as a PIK), and can be linked to the confidential version of the CPS. PIKs are available for around 99 percent of administrative records. Using the approach in Meyer et al. (2018) and

Mittag (2018), we adjust the CPS survey weights to reflect that our data excludes some observations with missing PIKs.[5]

Misreporting of SNAP Participation & Changes in SNAP Participation in the CPS ASEC

In the remainder of this article, we classify the measure of SNAP participation based on the SNAP administrative records as the "true" value of participation, and we view a disagreement between the survey based measure and the administrative measure as a misreport. We acknowledge that administrative data themselves can be imperfect, and that different administrative data sources can disagree, as reported in Courtemanche et al. (2018). However, because States are held accountable by the USDA Food and Nutrition Service for accuracy in benefit issuance through the SNAP Quality Control system, in which States provide a sample of SNAP cases from their administrative data to FNS for review of errors, we will maintain that our monthly SNAP administrative records on benefit receipt and payment is the more accurate source.

Consistent with Meyer et al. (2018), our linked CPS ASEC and administrative files reveal substantial under-reporting of SNAP in the CPS ASEC survey. Meyer et al. (2018) find a false negative rate of 48.98% in the CPS ASEC in their linkage of SNAP administrative data from Illinois and Maryland and the CPS ASEC, and a false positive rate of 0.84% for households. Our unit of analysis is the individual rather than the household, but we find similar levels of under-reporting 43.9%. This means that among individuals we can identify as receiving SNAP in the administrative records, 43.9% do not report SNAP receipt in the ASEC survey. We find slightly higher levels of over-reporting, 2.9%, in our linked data, but we attribute this higher over-reporting rate to differences in the unit, individual versus household. The CPS ASEC measure of SNAP participation is asked at the household level, and so part of the reason for over-reporting in individuals may be that an individual is part of the survey household, but not actually part of the SNAP case unit. In the appendix, we report summary statistics by reported SNAP status and the admin verified SNAP status for individuals in our nine states with administrative data linkages by the individual's work status, hours worked, and demographics. Consistent with

[5] We use a logit model of a binary indicator for receiving a PIK on individual characteristics in the CPS including age, education, gender, race, the county unemployment rate, state fixed effects, and year fixed effects. We then generate predicted probabilities of receiving a PIK from this logit model, and multiply the CPS weights by the inverse of the probability of receiving a PIK. This approach is based on that in Wooldridge (2007).

Meyer et al. (2018), we see that misreporting appears to be related to many individual characteristics. We find that those reporting SNAP in the CPS ASEC have slightly lower probabilities of employment (46.4%) than individuals identified as receiving SNAP according to the administrative records (47.8%). Meyer et al. (2018) find that higher income is positively related to misreporting, and our result is consistent with this. Other notable differences between our measures of SNAP participation are that the sample receiving SNAP according to the administrative reports are more likely to be female (59.3% reported SNAP versus 62.5% for the admin measure) and less likely to have less than a HS degree (27.2% reported SNAP versus 25.3% for the admin measure.

In Table 1 below, we examine reporting of SNAP participation longitudinally using our two period panels in the CPS ASEC. We report a frequency table for those identified as changing in SNAP participation from one year to the next based on survey reports and those verified to have changed SNAP participation according to the administrative records. Each cell in the table show the percentage of all individuals linked longitudinally in the CPS ASEC classified based on the CPS ASEC reports and the SNAP admin data reports. Because longitudinal estimators of the effects of SNAP participation on outcomes rely on changes in SNAP participation for identification, understanding the extent that reports of changes in SNAP participation in the CPS ASEC are accurate is important for assessing potential bias in these estimators. Because of U.S. Census Bureau disclosure rules, we cannot report exact observations counts for these cells, but we can say that overall, we have approximately 36,500 total observations.

Table 1 suggests errors in changes in SNAP participation due to misreporting is a severe issue in survey data. We discuss misreporting of SNAP participation in terms of four groups: those who switch off SNAP, those who are on SNAP in both periods, those who are off SNAP in both periods, and those who switch onto SNAP. Row 1 of Table 1 suggest that overall those identified as switching off SNAP based on survey reports, meaning they reported to be on SNAP in period 1 but not period 2, make up 2.7 percent of individuals. However, among this group only 14.4% of them (0.39/2.7) can be verified to have actually switched off SNAP in the administrative records. This suggests 85.6% of the reports of individuals switching off SNAP are false reports, higher than the 43.9% under-reporting SNAP participation in the cross-section.

We see that the majority of those identified as switching off SNAP were actually either on the program in both periods (1.03% out of 2.7%) or on the program in neither period (1.2% out of 2.7%). In the second row, we examine those who remain on SNAP for both period. For this group, we see that roughly 27.5% of those who report being on SNAP for both period are false reports. Most of these false reports are individuals who were actually on SNAP in neither period (1.19% out of 5.53%). Although over-reporting SNAP participation is rare, because the group of individuals who are off SNAP is so large, a sizable proportion of the individuals reporting to be on SNAP in both periods are actually those who are off SNAP in both periods. In row 3, the percentage of individuals misreporting being off SNAP in both periods is 4.7%. This group has the smallest proportion of errors. When examining individuals reported to have switched on to SNAP in Row 4, we find similar rates of error for switches onto SNAP as for those who switch off SNAP. We find that 2.78% of the sample report switching onto SNAP according to the CPS ASEC survey, but only 13.3% of them (0.37/2.78) are actual switches according to the administrative records, suggesting 86.7% are false reports. Again, most of these false switches were individuals who were either on SNAP in both periods or off SNAP in both periods.

Table 2 reports summary statistics of individuals identified as changing in SNAP participation based on the survey reports and the administrative data reports. We see some evidence of a relationship between individual characteristics and the misreporting. Column 1 shows summary statistics for individuals identified as switching off SNAP in the survey and Column 2 shows the statistic for individuals switching off according to the administrative data. We see that individuals switching off SNAP according to the CPS data are more likely to be employed (59.4% versus 52.8%), older (39.9 years old versus 36.2 years old), and Whiter (60.5% versus 57.6%) compared to sample according to the administrative data. In Columns 7 and 8, we see the comparison for those reporting to switch on to SNAP. Those reporting they switched on to SNAP are less likely to work (53.6% versus 59.5%), older (39 versus 37.3) and Whiter (57.5 versus 60.4) than the sample based on the administrative data. These summary statistics suggest that errors in changes in SNAP participation may be non-random as well, which could introduce complicated biases in longitudinal estimators.

Table 1. Frequency Table for Changes in SNAP Participation Based on Survey Reports and Administrative Data

		Admin Verified SNAP Measure				
		Switch off SNAP	On SNAP-No Change	Not On SNAP-No Change	Switch on SNAP	Total
Survey SNAP Measure	Switch off SNAP	0.39%	1.03%	1.2%	0.07%	2.7%
	On SNAP - No Change	0.18%	4.01%	1.19%	0.15%	5.53%
	Not On SNAP - No Change	0.82%	2.45%	84.88%	0.84%	88.99%
	Switch on SNAP	0.06%	1.11%	1.24%	0.37%	2.78%
	Total	1.45%	8.6%	88.51%	1.43%	100%

Notes: Summary statistics derived from longitudinally linked CPS ASEC microdata and SNAP administrative records from 2005-2015.

Table 2. Summary Statistics of Key Variables Based on Data Source (Survey or Administrative Data) – Changes in SNAP

	Switch off SNAP-Survey	Switch off SNAP-Admin	On SNAP-No Change-Survey	On SNAP-No Change-Admin	Not On SNAP-No Change-Survey	Not On SNAP-No Change-Admin	Switch on SNAP-Survey	Switch on SNAP-Admin
Employed	59.4%	52.8%	46.5%	46.0%	76.5%	76.4%	53.6%	59.5%
Usual Weekly Hours Worked	37.8	37.2	35.1	34.5	39.6	39.6	35.3	36.9
Age	39.9	36.2	37.9	38.0	42.7	42.7	39.0	37.3
Female	56.7%	56.1%	62.2%	63.3%	49.2%	49.2%	57.1%	56.2%
Less HS	19.6%	21.7%	25.4%	26.8%	6.9%	7.1%	25.5%	17.7%
High School Diploma	42.0%	37.8%	41.7%	41.4%	27.8%	28.0%	37.4%	39.2%
Some College	21.1%	23.6%	19.9%	19.6%	19.6%	19.6%	22.8%	26.5%
Associates Degree	6.5%	7.3%	6.9%	6.8%	10.0%	9.9%	5.8%	6.6%
Bachelor Degree	8.3%	8.4%	4.9%	4.3%	22.9%	22.8%	6.9%	8.2%
White	60.5%	57.6%	56.0%	55.1%	78.5%	78.3%	57.5%	60.4%
Black	27.6%	27.5%	32.9%	33.2%	11.1%	11.2%	27.7%	30.0%
Asian	3.1%	2.8%	1.6%	1.8%	5.2%	5.3%	3.8%	2.4%
Unemployment Rate	8.7%	8.6%	8.1%	8.2%	7.7%	7.7%	8.6%	8.7%
Observations	1,000	550	2,700	3,100	31,500	32,000	1,000	500

Notes: Summary statistics derived from longitudinally linked CPS ASEC microdata and SNAP administrative records from 2005-2015. Observation counts rounded in accordance with U.S. Census Bureau disclosure rules.

Table 3 shows estimates of the degree of bias one can expect caused by misreporting, λ_{OLS} and λ_{FD}, which is based on SNAP administrative data and survey reports. As a reminder, λ is the regression coefficient from a regression of an individual's "true" SNAP participation status or change in SNAP participation based on the administrative data on the reported measure of participation or change in SNAP participation based on the survey data. The intuition, as described in Bound & Krueger (1991), is that the bias term, λ, is the degree with which reported SNAP participation translates into actual SNAP participation. A $\lambda = 1$, would imply no bias caused by misreporting of the OLS estimator, and a λ close to zero indicates severe bias of the effect of interest toward zero. Importantly, λ does not depend on the particular outcome studied.

Our results suggest moderate bias caused by misreporting for the cross-sectional estimator and severe bias for the first-differenced estimator. Column 1 of Table 3 shows the estimate of λ_{OLS}, the degree that the effect of interest is biased toward zero, to be 0.634. As an example, if the true effect of SNAP on an outcome, say food security, were 0.10, then the OLS estimator would suggest an effect around 0.0634 because of misreporting of SNAP in the survey data, assuming there was no other source of bias. For comparison to another context, misreporting in earnings produced a λ of around 0.95-1.0 in the cross-sectional OLS estimator according to Bound & Krueger (1991), suggesting that misreporting of SNAP participation may be a worse problem. Column 2 reports the estimate of λ_{FD}, the degree of bias for the first-difference estimator, to be 0.116, suggesting severe bias. This suggest, as in the example above, the first difference OLS estimator would estimate the effect of SNAP to be 0.0116, rather than the true effect of 0.10, because of misreporting. For context, the degree of bias for the first-difference estimator using misreported earnings was around 0.7-0.85 according to Bound & Krueger (1991).

Table 3. Regression of Administrative Data Based SNAP Measure on Reported SNAP Measure in CPS ASEC.

VARIABLES	(1) Admin Verified SNAP – Cross Section	(2) Admin Verfied SNAP – First Difference
Reported SNAP –Cross Section	0.634***	
	(0.006)	
Reported SNAP - FD		0.116***
		(0.009)
Observations	73,000	36,500
R-squared	0.337	0.026

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Empirical Application: Effects of SNAP Participation on Labor Supply

In our empirical application, we model the impact of SNAP participation on the labor supply of participants. A large previous literature examines the impacts of low-income assistance programs on work effort. Classical labor economics theory suggests that participation in a means tested assistance program may discourage work effort. First because of an income effect, where the additional resources provided may induce individuals to consume more leisure, because leisure is a normal good. Second because of a substitution effect, where because benefits fade with income, the return to work is reduced, and consequently the opportunity cost of leisure is decreased, leading to an increase in leisure. SNAP is structured so that, after deductions, $1 of wage income leads to a loss of $0.30 in SNAP benefits. A wide body of empirical evidence does provide some evidence of a work disincentive for SNAP. For example, Hoynes & Schanzenbach (2012) examine the effects of enrollment in the food stamps program during the initial program rollout in the late 1960s and early 1970s. The authors find a large negative effect on labor supply, particularly for female heads of household. Fraker & Moffitt

(1988), Keane & Moffitt (1998), and Hagstrom (1996), however, find smaller negative impacts on labor supply.

Assume a workers labor supply is Y_i, which is subject to covariates X_i, and that labor supply differs depending on whether or not they choose to enroll in SNAP, S_i. In our case, labor supply, Y_i, will mean either whether an individual has worked over the past year or the individuals usual hours worked per year. A general model can be written as

$$Y_{it} = \Gamma(\alpha_0 + \tau S_{it} + X_{it}\beta + \mu_i + \epsilon_{it}) \tag{14}$$

Where Γ() is an unspecified functional form, or link function, for the labor supply model. In our application, we will examine three cases: a linear model using OLS, a logistic model, and a linear model in which we take the log of our hours worked outcome. X_{it} is a vector of covariates, μ_i is a time constant, individual specific heterogeneity term or "individual fixed effect", and ϵ_{it} is an idiosyncratic error term. Regressions with covariates control for age, gender, race, an indicator for completing high school, some college, college, or an advanced degree, and the county level unemployment rate.[6] A word of caution in interpreting our findings. The main point of this analysis is to study how misreporting affects cross-sectional and longitudinal estimates, so we keep our specification relatively simple. A more careful analysis may be needed to produce unbiased estimates of the effects of SNAP on labor supply.

Table 4 reports estimates of the effects of SNAP on labor supply outcomes using either the reported SNAP measure or the "true" SNAP measure based on the administrative data. The table reports estimates for the coefficient on SNAP participation from 32 separate regressions, which differ in the outcome, covariates, model, or SNAP measure used. Full results are available in appendix tables A2 to A9, including coefficients for the other covariates in the model. Column 1 and 2 show estimates of the impact of SNAP on a binary indicator for whether the individual worked during the 12 month ASEC reference period. Column 1 includes no covariates in the model, while Column 2 does include the covariates described above. Column 3 and 4 show estimates of the impact of SNAP on hours worked per week, and column 4 includes additional covariates in the model. In each panel of the table, the top row reports the estimate based on the survey reports of SNAP participation, while the lower row shows the estimate based

[6] In the case of our estimators taking first differences of our outcomes and covariates, the gender and race indicators drop from the model.

on the administrative data based measure of SNAP. The top panel shows the cross-sectional OLS linear regression results. The second panel shows the first-differenced OLS estimator. Panel 3 shows marginal effect estimates from a cross-sectional logit model. The fourth panel shows marginal effect estimates from a fixed effect logit estimator, which makes use of the longitudinal nature of the data to account for time constant, individual level heterogeneity, u_i. The fifth and sixth panels show estimates of a cross-section and first-difference OLS estimator that takes the log of hours worked, instead of the levels as in Panels 1 and 2.

Two findings are notable. First, the differences are much larger in magnitude for the longitudinal estimators than the cross-sectional estimators. Second, the bias is more complicated than a simple attenuation bias that might be expected based on equation (6). This more complicated bias may be caused by the fact that the misreporting errors are correlated with our work outcomes, which are evident in Table 2 and appendix Table A1. For instance, in Panel 1 and column 2, the estimated effect of SNAP on a binary indicator for work when covariates are included using the reported SNAP measure is -0.222 and the estimated effect is -0.221 using the admin data based SNAP measure, a change of just 0.001 percentage points. However, in the first-difference model in panel 2 and column 2, the estimated effect changes from -0.039 using the self-reports to 0.048 using the administrative data. This is an 8.7 percentage point change and the sign switched from negative to positive. The cross-sectional logit regressions show point estimates that are fairly similar to the cross-sectional OLS estimator and do not greatly differ depending on whether one uses the reported SNAP or admin SNAP measure. The fixed effect logit estimates in column 1 again show a large effect of using the administrative data measure of SNAP. The estimated coefficients shift 19.8 percentage points and again change signs. In column 2 of this panel, there is little difference but both estimators show an imprecisely estimated effect near zero. The cross-sectional OLS estimator using log hours worked as the dependent variable in panel 5 again shows relatively small differences between the self-reported SNAP measure and the admin based measure. The difference is slightly more noticeable in column 4. There again are relatively large differences in the first-difference estimates in Panel 6, which use log hours worked as the outcome. In this case, the estimates change from roughly -0.05 to a roughly null effect when using the administrative data based measure of SNAP.

Given that the results in Table 3 suggest the cross-sectional estimator should be biased toward zero by approximately 0.6, it is worthwhile to discuss the reason the cross-sectional estimators do not show more bias. Additionally, it may be surprising to see the estimates from the longitudinal estimator change signs, when the expected bias was a severe attenuation toward zero. Equation (9) above shows that the bias formula for misreporting assumes no correlation between the reported measure of SNAP participation, R_{it}, and the error components, $u_i \; or \; \epsilon_{it}$. However, our summary statistics, Tables A1 and Table 2, provide reason to be skeptical of this. We see notable differences in probability of work for individuals depending on which measure is used. If misreporting error is correlated with labor supply, then the bias may be more complicated than that suggested by equation (9). For different outcomes, with non-random misreporting error, the misreporting bias may be greater or less than that suggested by our estimates of λ.

Table 4. Regression Estimates of Effects of SNAP on Work and Hours Worked using Reported SNAP Measure from CPS ASEC versus Admin Verified SNAP Measure. Results from 32 Separate Regressions.

Estimated Effects	(1) Worked Past Year	(2) Worked Past Year	(3) Hours Worked	(4) Hours Worked
Cross Sectional w/ OLS:				
Reported SNAP	-0.291***	-0.222***	-13.82***	-9.912***
	(0.007)	(0.007)	(0.257)	(0.267)
Admin Verified SNAP	-0.286***	-0.221***	-13.60***	-9.556***
	(0.006)	(0.006)	(0.238)	(0.251)
First Differences w/ OLS:				
Reported SNAP	-0.040***	-0.039***	-2.373***	-2.330***
	(0.010)	(0.010)	(0.413)	(0.419)
Admin Verified SNAP	0.050***	0.048***	1.769***	1.710***
	(0.015)	(0.015)	(0.607)	(0.612)
Cross Sectional w/ Logit:				
Reported SNAP	-0.242***	-0.177***		
	(0.005)	(0.005)		
Admin Verified SNAP	-0.238***	-0.177***		
	(0.004)	(0.005)		
Panel w/ Fixed Effect Logit:				
Reported SNAP	-0.092***	-0.000		
	(0.024)	(0.119)		
Admin Verified SNAP	0.106***	0.000		
	(0.032)	(0.117)		
Cross Sectional OLS w/ Logs of Outcome:				
Reported SNAP			-0.167***	-0.115***
			(0.009)	(0.009)
Admin Verified SNAP			-0.162***	-0.097***
			(0.008)	(0.008)
First Differences OLS w/ Logs of Outcome:				
Reported SNAP			-0.055***	-0.053***
			(0.013)	(0.013)
Admin Verified SNAP			-0.010	-0.006
			(0.019)	(0.020)
Covariates	NO	YES	NO	YES

Robust Standard Errors in Parenthesis. Covariates include controls for age, education, race, marital status, and the county unemployment rate.

*** p<0.01, ** p<0.05, * p<0.1

Solutions for Researchers

Estimates in Table 4 suggest that estimators of the effects of SNAP, particularly longitudinal estimators, can be severely biased by misreporting of SNAP participation in survey datasets. While administrative data can be an attractive solution to this problem, accessing these databases can be difficult for many researchers due to the required legal agreements and disclosure review processes inherent with administrative data. Moreover, administrative data on their own have limitations, particularly a sparse set of covariates compared to survey data, and linking administrative data to survey datasets is difficult, because this requires access to confidential information on survey respondents in order to match to individuals. In the following section, we examine a promising solution for misreporting of SNAP participation in survey data, that does not requires access to linked administrative data.

Mittag (2018) proposes the use of a two-step method to accounting for misreporting, the conditional distribution method. As discussed in Mittag (2018), this approach can be seen as a multiple imputation method (Rubin (1996); Rubin (2004)). In the first step, researchers with access to linked administrative and survey datasets fits a model of the conditional distribution of the "true" SNAP measure, and reports estimates from the model in a public forum. In the second step, another set of researchers without access to the confidential linkages use the estimates from the model of the conditional distribution of the "true" SNAP measure to create a simulated dataset containing an imputed SNAP measure and uses these imputed values in their regression models. The key to this approach is that the imputed SNAP measures, based on the estimated conditional density, have approximately the same covariance structure as the "true" SNAP measure based on the administrative data. In Mittag (2018), the author shows the method works well using American Community Survey data in New York state, in producing approximately unbiased cross-sectional regression estimates.

We extend the Mittag (2018) approach to the longitudinal case, where instead of modeling a cross-sectional, binary "true" SNAP participation measure, we model the changes in "true" SNAP participation over time. We model changes in "true" SNAP participation using a Multinomial Logit regression, which takes on three outcomes, $\Delta S_{it} = -1$, an individual cycles off SNAP, $\Delta S_{it} = 0$, no change in SNAP participation, and $\Delta S_{it} = 1$, the individual cycles onto SNAP. We include the change in the reported SNAP measure as a covariate in the model, as well as the changes in hours worked, work status, age, education, marital status, and the unemployment rate. Estimates from our Multinomial Logit regression model are available in Table A10. Note that these covariates are the covariates that can be found in the CPS ASEC files alone. Following Mittag (2018), we then produce fitted probabilities of each of our three outcomes using the CPS ASEC data, and based on these probabilities generate 100 simulated datasets containing an imputed change in SNAP participation for each individual based on the covariates.[7] The researchers than stack the 100 simulated datasets and estimate their regression models using standard regression software commands. This approach is based on two assumptions. First, that the conditional density be identical for the linked data and public use data. In our application, this is trivially met, because we use the linked data for estimation of the conditional density and for the regression application. Second, the conditional distribution model must be properly specified. While a parametric model specified by researchers will almost never be exactly correct, a reasonable model may still perform well in practice at reducing bias.

Results from applying the conditional distribution method are available in Table 5 and show that the conditional distribution method performs well in recovering estimates based on the administrative data. The table shows estimates from twelve separate first-difference OLS regressions. The top row shows estimates of the effect of SNAP on whether or not an individual works, and the bottom row shows estimates for hours worked. In columns (1) – (3) no covariates are included, and in columns (4) – (6) we include age, education, marital status, and the unemployment rate in the regression models. Column (1) and (4) show the OLS estimate based on the survey based measure of SNAP participation. Column (2) and (5) show the estimate based on administrative data, and column (3) and (6) show the estimate based on the imputed values from the conditional distribution method. Estimates based on the conditional distribution method are very similar to those based on the administrative data. In most cases, differences can only be found in the third decimal place.

[7] Imputed values are formed using the inverse transformation method. This involves first drawing random values from a uniform distribution. Using fitted values from our Multinomial Logit model, we then form fitted values for the conditional density function, and then draw random values based on the random values from the uniform distribution. Stata code is attached in the appendix.

Table 5. Regression Estimates of First Differences in Work and Hours Worked on First Differences in SNAP on using OLS with Reported SNAP Measure, OLS with Admin Verified Measure, and the Mittag (2018) Conditional Distribution Method

Estimated Effects	(1) Worked Past Year	(2) Worked Past Year	(3) Hours Worked	(4) Hours Worked
First Differences w/ OLS:				
Reported SNAP	-0.040***	-0.039***	-2.373***	-2.330***
	(0.010)	(0.010)	(0.413)	(0.419)
Admin Verified SNAP	0.050***	0.048***	1.769***	1.710***
	(0.015)	(0.015)	(0.607)	(0.612)
Conditional Distribution Method	0.054 ***	0.047 ***	1.947***	1.698***
	(0.006)	(0.003)	(0.288)	(0.126)
Covariates	NO	YES	NO	YES
Observations	36,000	36,000	36,000	36,000

Bootstrap standard errors in parenthesis from 100 bootstrap replications. In each bootstrap replication, both the multinomial logit estimation and regression estimation steps are performed. Covariates include controls for age, education, race, marital status, and the county unemployment rate.

*** p<0.01, ** p<0.05, * p<0.1

Conclusions

In this paper we examine the impacts of misreporting of SNAP participation in survey datasets on longitudinal estimators of the causal impacts of SNAP. Confirming a strand of recent literature, we find strong evidence of misreporting in SNAP survey datasets and produce evidence suggesting the misreporting is non-random. Additionally, we produce evidence that errors in changes of SNAP participation over time based on survey reports is a potentially even more serious issue than in the cross-section. For instance, we find that 85.6% of the individuals identified as switching off SNAP based on survey reports are false reports, higher than the 43.9% under-reporting SNAP participation in the cross-section. We estimate λ, a measure of misreporting bias used in previous literature that does not depend on a specific outcome variable, and we find significantly more severe bias in longitudinal estimators compared to cross-sectional estimators. In an application estimating the effect of SNAP participation on labor supply, we find little bias in cross-sectional estimator, and severe bias in longitudinal estimators.

The good news is that conditional distribution methods, which can be used by researchers with access to publicly available survey data, seem to work well in reducing bias caused by misreporting of SNAP. Researchers with access to linked administrative and survey datasets may want to publish conditional density model estimates, as we have done in this article, so that other researchers can use these models to improve their own research. Future work could examine misreporting in other commonly used datasets, including the Survey of Income and Program Participation (SIPP) and the December CPS food security supplement.

Bibliography

Aigner, D. J. (1973). Regression with a binary independent variable subject to errors of observation. *Journal of Econometrics, 1*(1), 49-59.

Almada, L., McCarthy, I., & Tchernis, R. (2016). What can we learn about the effects of food stamps on obesity in the presence of misreporting? *American Journal of Agricultural Economics, 98*(4), 997-1017.

Baum, C. L. (2011). The effects of food stamps on obesity. *Southern Economic Journal, 77*(3), 623-651.

Bollinger, C. R., & David, M. H. (1997). Modeling discrete choice with response error: Food Stamp participation. *Journal of the American Statistical Association, 92*(439), 827-835.

Bound, J., Brown, C., Duncan, G. J., & Rodgers, W. L. (1994). Evidence on the validity of cross-sectional and longitudinal labor market data. *Journal of labor Economics, 12*(3), 345-368.

Bound, J., & 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), 1-24.

Courtemanche, C. J., Denteh, A., & Tchernis, R. (2018). Estimating the Associations between SNAP and Food Insecurity, Obesity, and Food Purchases with Imperfect Administrative Measures of Participation: National Bureau of Economic Research.

Fraker, T., & Moffitt, R. (1988). The effect of food stamps on labor supply: A bivariate selection model. *Journal of Public Economics, 35*(1), 25-56.

Freeman, R. B. (1984). Longitudinal analyses of the effects of trade unions. *Journal of labor Economics, 2*(1), 1-26.

Hagstrom, P. A. (1996). The food stamp participation and labor supply of married couples: An empirical analysis of joint decisions. *Journal of Human Resources*, 383-403.

Heflin, C. M., & Ziliak, J. P. (2008). Food Insufficiency, Food Stamp Participation, and Mental Health*. *Social Science Quarterly, 89*(3), 706-727. doi: doi:10.1111/j.1540-6237.2008.00556.x

Hoynes, H. W., & Schanzenbach, D. W. (2012). Work incentives and the Food Stamp Program. *Journal of Public Economics, 96*(1–2), 151-162. doi: http://dx.doi.org/10.1016/j.jpubeco.2011.08.006

Keane, M., & Moffitt, R. (1998). A structural model of multiple welfare program participation and labor supply. *International economic review*, 553-589.

Kim, B., & Solon, G. (2005). Implications of mean-reverting measurement error for longitudinal studies of wages and employment. *Review of Economics and statistics, 87*(1), 193-196.

Kreider, B., Pepper, J. V., Gundersen, C., & Jolliffe, D. (2012). Identifying the effects of SNAP (food stamps) on child health outcomes when participation is endogenous and misreported. *Journal of the American Statistical Association, 107*(499), 958-975.

Madrian, B. C., & Lefgren, L. (1999). A note on longitudinally matching Current Population Survey (CPS) respondents: National Bureau of Economic Research Cambridge, Mass., USA.

Meyer, B. D., & Mittag, N. (2015). Using linked survey and administrative data to better measure income: Implications for poverty, program effectiveness and holes in the safety net: National Bureau of Economic Research.

Meyer, B. D., & Mittag, N. (2017). Misclassification in binary choice models. *Journal of Econometrics*, *200*(2), 295-311.

Meyer, B. D., Mittag, N., & Goerge, R. M. (2018). Errors in survey reporting and imputation and their effects on estimates of food stamp program participation: National Bureau of Economic Research.

Meyer, B. D., Mok, W. K., & Sullivan, J. X. (2015). Household surveys in crisis. *Journal of Economic perspectives, 29*(4), 199-226.

Mittag, N. (2018). Correcting for Misreporting of Government Benefits. *American Economic Journal: Economic Policy, Forthcoming*.

Nord, M., & Golla, A. (2009). Does SNAP decrease food insecurity? Untangling the self-selection effect. US Dept. of Agriculture. *Economic Research Service ERR*(85).

Oliveira, V. (2018). The Food Assistance Landscape: FY 2017 Annual Report *EIB-190*. U.S. Department of Agriculture, Economics Research Service.

Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association, 91*(434), 473-489.

Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys* (Vol. 81): John Wiley & Sons.

Wilde, P., & Nord, M. (2005). The effect of food stamps on food security: a panel data approach. *Review of Agricultural Economics, 27*(3), 425-432.

Wooldridge, J. M. (2007). Inverse probability weighted estimation for general missing data problems. *Journal of Econometrics, 141*(2), 1281-1301.

Appendix of Tables

Table A1. Summary Statistics of Key Variables By Survey Reported SNAP Use and Admin Verified SNAP Use

	All	Not on SNAP - Reported	Not On SNAP - Admin	On SNAP Reported	On SNAP - Admin
Employed	72.7	75.4	75.9	46.4	47.8
Usual Weekly Hours Worked	39.2	39.5	39.6	34.7	34.8
Age	42.3	42.6	42.8	39.3	38.0
Female	50.8	49.9	49.3	59.3	62.5
Less HS	9.4	7.6	7.4	27.2	25.3
High School Diploma	29.9	28.9	28.4	39.3	41.6
Some College	19.5	19.5	19.5	19.9	20.0
Associate Degree	9.7	9.9	10	7.0	6.9
Bachelor Degree	20.3	21.9	22.3	5.2	4.9
Advanced Degree	11.2	12.2	12.4	1.4	1.2
White	75.4	77.2	77.9	58.3	56.0
Black	14.0	12.4	11.6	29.4	32.6
Asian	4.8	5.1	5.2	2.0	1.9
Unemployment Rate	7.8	7.8	7.8	8.3	8.2
Observations	73,000	67,000	66,000	6,000	7,300

Notes: Summary statistics derived from linked CPS ASEC microdata and SNAP administrative records from 2005-2015. Observation counts rounded in accordance with U.S. Census Bureau disclosure rules.

Table A2. OLS Regression of Work Status on SNAP participation Measure (either Self Reported of Verified by Admin Data) and Other Covariates using CPS ASEC Data

VARIABLES	(1) Self Reported SNAP: worked	(2) Self Reported SNAP: worked	(3) Admin Verified SNAP: worked	(4) Admin Verified SNAP: worked
SNAP Measure	-0.291***	-0.222***	-0.286***	-0.221***
	(0.007)	(0.007)	(0.006)	(0.006)
Age: 20-30 years		0.224***		0.223***
		(0.011)		(0.011)
Age: 30-40 years		0.275***		0.272***
		(0.011)		(0.011)
Age: 40-50 years		0.274***		0.269***
		(0.011)		(0.011)
Age: 50-60 years		0.229***		0.220***
		(0.011)		(0.011)
Age: 60+ years		0.054***		0.045***
		(0.012)		(0.012)
Female		-0.090***		-0.086***
		(0.003)		(0.003)
High School Diploma		0.129***		0.130***
		(0.007)		(0.007)
Some College		0.148***		0.145***
		(0.007)		(0.007)
Associates Degree		0.205***		0.202***
		(0.008)		(0.008)
Bachelor's Degree		0.210***		0.205***
		(0.007)		(0.007)
Advanced Degree		0.252***		0.247***
		(0.008)		(0.008)
White		0.010		0.010
		(0.015)		(0.015)
Black		-0.022		-0.012
		(0.015)		(0.016)
American Indian		-0.041		-0.051
		(0.032)		(0.032)
Asian		-0.064***		-0.065***
		(0.017)		(0.017)
married		0.009***		0.008**
		(0.004)		(0.004)
Unemployment Rate		0.000		0.000
		(0.001)		(0.001)
Constant	0.754***	0.404***	0.759***	0.412***
	(0.002)	(0.019)	(0.002)	(0.019)
Observations	73,000	72,000	73,000	72,000
R-squared	0.032	0.099	0.037	0.101

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A3. OLS Regression of First Differences of Work Status on First Differences of SNAP participation Measure (either Self Reported of Verified by Admin Data) and Other Covariates using CPS ASEC Data

VARIABLES	(1) Self Reported SNAP: Δ worked	(2) Self Reported SNAP: Δ worked	(3) Admin Verified SNAP: Δ worked	(4) Admin Verified SNAP: Δ worked
Δ SNAP Measure	-0.040***	-0.039***	0.050***	0.048***
	(0.010)	(0.010)	(0.015)	(0.015)
Δ Age: 20-30 years		0.083***		0.082***
		(0.023)		(0.023)
Δ Age: 30-40 years		0.140***		0.138***
		(0.030)		(0.030)
Δ Age: 40-50 years		0.220***		0.217***
		(0.034)		(0.034)
Δ Age: 50-60 years		0.327***		0.323***
		(0.036)		(0.036)
Δ Age: 60+ years		0.383***		0.379***
		(0.038)		(0.038)
Δ High School Diploma		0.079***		0.079***
		(0.017)		(0.017)
Δ Some College		0.105***		0.105***
		(0.020)		(0.020)
Δ Associates Degree		0.130***		0.131***
		(0.023)		(0.023)
Δ Bachelor's Degree		0.179***		0.180***
		(0.024)		(0.024)
Δ Advanced Degree		0.175***		0.176***
		(0.030)		(0.030)
Δ married		0.029*		0.031*
		(0.017)		(0.017)
Δ Unemployment Rate		0.026***		0.026***
		(0.002)		(0.002)
Constant	-0.003	-0.003	-0.003	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)
Observations	36,500	35,500	36,500	35,500
R-squared	0.001	0.015	0.000	0.015

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A4. Logistic Regression of Work Status on SNAP participation Measure (either Self Reported of Verified by Admin Data) and Other Covariates using CPS ASEC Data

VARIABLES	(1) Self Reported SNAP: worked	(2) Self Reported SNAP: worked	(3) Admin Verified SNAP: worked	(4) Admin Verified SNAP: worked
SNAP Measure	-0.242***	-0.177***	-0.238***	-0.177***
	(0.005)	(0.005)	(0.004)	(0.005)
Age: 20-30 years		0.171***		0.173***
		(0.009)		(0.009)
Age: 30-40 years		0.225***		0.224***
		(0.009)		(0.009)
Age: 40-50 years		0.223***		0.220***
		(0.009)		(0.009)
Age: 50-60 years		0.176***		0.170***
		(0.009)		(0.009)
Age: 60+ years		0.027***		0.021**
		(0.009)		(0.009)
Female		-0.089***		-0.085***
		(0.003)		(0.003)
High School Diploma		0.104***		0.105***
		(0.006)		(0.006)
Some College		0.121***		0.118***
		(0.006)		(0.006)
Associates Degree		0.179***		0.176***
		(0.007)		(0.007)
Bachelor’s Degree		0.186***		0.181***
		(0.006)		(0.006)
Advanced Degree		0.238***		0.234***
		(0.007)		(0.007)
White		0.010		0.011
		(0.014)		(0.014)
Black		-0.019		-0.011
		(0.015)		(0.015)
American Indian		-0.039		-0.048*
		(0.028)		(0.028)
Asian		-0.062***		-0.063***
		(0.016)		(0.016)
married		0.010***		0.008**
		(0.004)		(0.004)
Unemployment Rate		-0.000		0.000
		(0.001)		(0.001)
Observations	73,000	72,500	73,000	72,500

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A5. Fixed Effect Logit Regression of Work Status on SNAP participation Measure (either Self Reported of Verified by Admin Data) and Other Covariates using CPS ASEC Data

VARIABLES	(1) Self Reported SNAP: worked	(2) Self Reported SNAP: worked	(3) Admin Verified SNAP: worked	(4) Admin Verified SNAP: worked
SNAP Measure	-0.092***	-0.000	0.106***	0.000
	(0.024)	(0.119)	(0.032)	(0.117)
Age: 20-30 years		0.000		0.000
		(0.159)		(0.154)
Age: 30-40 years		0.001		0.001
		(0.317)		(0.299)
Age: 40-50 years		0.002		0.001
		(0.630)		(0.603)
Age: 50-60 years		0.003		0.003
		(1.072)		(1.035)
Age: 60+ years		0.003		0.003
		(1.292)		(1.251)
High School Diploma		0.000		0.000
		(0.191)		(0.188)
Some College		0.001		0.001
		(0.247)		(0.241)
Associates Degree		0.001		0.001
		(0.335)		(0.333)
Bachelor's Degree		0.001		0.001
		(0.481)		(0.474)
Advanced Degree		0.001		0.001
		(0.487)		(0.485)
married		0.000		0.000
		(0.118)		(0.125)
Unemployment Rate		0.000		0.000
		(0.058)		(0.057)
Observations	12,000	12,000	12,000	12,000

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A6. OLS Regression of Hours Worked on SNAP participation Measure (either Self Reported of Verified by Admin Data) and Other Covariates using CPS ASEC Data

VARIABLES	(1) Self Reported SNAP: Hours Worked	(2) Self Reported SNAP: Hours Worked	(3) Admin Verified SNAP: Hours Worked	(4) Admin Verified SNAP: Hours Worked
SNAP Measure	-13.82***	-9.912***	-13.60***	-9.556***
	(0.257)	(0.267)	(0.238)	(0.251)
Age: 20-30 years		12.60***		12.56***
		(0.366)		(0.367)
Age: 30-40 years		17.16***		16.97***
		(0.371)		(0.371)
Age: 40-50 years		17.52***		17.29***
		(0.366)		(0.366)
Age: 50-60 years		15.64***		15.27***
		(0.366)		(0.366)
Age: 60+ years		7.503***		7.115***
		(0.407)		(0.406)
Female		-7.253***		-7.095***
		(0.137)		(0.137)
High School Diploma		5.146***		5.205***
		(0.286)		(0.283)
Some College		5.395***		5.309***
		(0.299)		(0.297)
Associates Degree		8.174***		8.083***
		(0.335)		(0.333)
Bachelor’s Degree		9.595***		9.433***
		(0.303)		(0.301)
Advanced Degree		12.36***		12.21***
		(0.326)		(0.324)
White		0.554		0.586
		(0.639)		(0.643)
Black		-0.553		-0.135
		(0.662)		(0.666)
American Indian		-2.162		-2.596*
		(1.360)		(1.365)
Asian		-2.445***		-2.471***
		(0.716)		(0.721)
married		0.611***		0.571***
		(0.156)		(0.156)
Unemployment Rate		-0.030		-0.025
		(0.032)		(0.032)
Constant	30.13***	11.75***	30.35***	12.01***
	(0.075)	(0.758)	(0.076)	(0.761)
Observations	73,000	72,000	73,000	72,000
R-squared	0.037	0.147	0.042	0.148

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A7. OLS Regression of First Differences of Hours Worked on First Differences of SNAP participation Measure (either Self Reported of Verified by Admin Data) and Other Covariates using CPS ASEC Data
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	(1) Self Reported SNAP: Δ Hours Worked	(2) Self Reported SNAP: Δ Hours Worked	(3) Admin Verified SNAP: Δ Hours Worked	(4) Admin Verified SNAP: Δ Hours Worked
Δ SNAP Measure	-2.373***	-2.330***	1.769***	1.710***
	(0.413)	(0.419)	(0.607)	(0.612)
Δ Age: 20-30 years		3.663***		3.597***
		(0.731)		(0.732)
Δ Age: 30-40 years		5.730***		5.656***
		(1.103)		(1.104)
Δ Age: 40-50 years		8.679***		8.539***
		(1.275)		(1.276)
Δ Age: 50-60 years		12.513***		12.343***
		(1.407)		(1.408)
Δ Age: 60+ years		14.095***		13.913***
		(1.527)		(1.528)
Δ High School Diploma		1.969***		1.978***
		(0.637)		(0.635)
Δ Some College		2.580***		2.585***
		(0.719)		(0.717)
Δ Associates Degree		3.930***		3.991***
		(0.892)		(0.891)
Δ Bachelor's Degree		6.169***		6.199***
		(0.945)		(0.944)
Δ Advanced Degree		5.417***		5.472***
		(1.272)		(1.271)
Δ married		1.346*		1.419*
		(0.744)		(0.745)
Δ Unemployment Rate		1.031***		1.024***
		(0.079)		(0.079)
Constant	-0.120	-0.096	-0.121	-0.098
	(0.091)	(0.091)	(0.091)	(0.091)
Observations	36,500	35,500	36,500	35,500
R-squared	0.001	0.013	0.000	0.012

Table A8. OLS Regression of Log Hours Worked on SNAP participation Measure (either Self Reported of Verified by Admin Data) and Other Covariates using CPS ASEC Data

VARIABLES	(1) Self Reported SNAP: Log Hours Worked	(2) Self Reported SNAP: Log Hours Worked	(3) Admin Verified SNAP: Log Hours Worked	(4) Admin Verified SNAP: Log Hours Worked
SNAP Measure	-0.167***	-0.115***	-0.162***	-0.097***
	(0.009)	(0.009)	(0.008)	(0.008)
Age: 20-30 years		0.456***		0.456***
		(0.020)		(0.020)
Age: 30-40 years		0.577***		0.575***
		(0.019)		(0.019)
Age: 40-50 years		0.592***		0.590***
		(0.019)		(0.019)
Age: 50-60 years		0.587***		0.584***
		(0.019)		(0.019)
Age: 60+ years		0.514***		0.511***
		(0.020)		(0.020)
Female		-0.153***		-0.151***
		(0.003)		(0.003)
High School Diploma		0.041***		0.044***
		(0.008)		(0.008)
Some College		0.008		0.011
		(0.009)		(0.009)
Associates Degree		0.046***		0.049***
		(0.009)		(0.009)
Bachelor’s Degree		0.080***		0.083***
		(0.008)		(0.008)
Advanced Degree		0.110***		0.113***
		(0.009)		(0.009)
White		-0.001		-0.002
		(0.015)		(0.015)
Black		0.019		0.021
		(0.016)		(0.016)
American Indian		-0.033		-0.037
		(0.040)		(0.040)
Asian		-0.013		-0.014
		(0.017)		(0.017)
married		0.001		0.002
		(0.003)		(0.003)
Unemployment Rate		-0.001*		-0.001*
		(0.001)		(0.001)
Constant	3.627***	3.106***	3.629***	3.105***
	(0.002)	(0.026)	(0.002)	(0.026)
Observations	54,000	53,000	54,000	53,000
R-squared	0.009	0.114	0.011	0.114

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A9. OLS Regression of First Differences in Log Hours Worked on First Differences in SNAP participation Measure (either Self Reported of Verified by Admin Data) and Other First Differenced Covariates using CPS ASEC Data

VARIABLES	(1) Self Reported SNAP: Δ Log Hours Worked	(2) Self Reported SNAP: Δ Log Hours Worked	(3) Admin Verified SNAP: Δ Log Hours Worked	(4) Admin Verified SNAP: Δ Log Hours Worked
Δ SNAP Measure	-0.055***	-0.053***	-0.010	-0.006
	(0.013)	(0.013)	(0.019)	(0.020)
Δ Age: 20-30 years		0.113**		0.113**
		(0.045)		(0.045)
Δ Age: 30-40 years		0.088*		0.089*
		(0.049)		(0.049)
Δ Age: 40-50 years		0.075		0.075
		(0.050)		(0.050)
Δ Age: 50-60 years		0.053		0.054
		(0.052)		(0.052)
Δ Age: 60+ years		0.008		0.008
		(0.055)		(0.055)
Δ High School Diploma		-0.011		-0.010
		(0.026)		(0.026)
Δ Some College		-0.022		-0.021
		(0.028)		(0.027)
Δ Associates Degree		-0.008		-0.007
		(0.029)		(0.029)
Δ Bachelor's Degree		0.030		0.031
		(0.030)		(0.030)
Δ Advanced Degree		0.015		0.017
		(0.034)		(0.034)
Δ married		0.002		0.003
		(0.014)		(0.014)
Δ Unemployment Rate		-0.001		-0.001
		(0.002)		(0.002)
Constant	-0.001	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Observations	24,000	23,000	24,000	23,000
R-squared	0.001	0.004	0.000	0.003

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A10. Multinomial Logit Regressions of Admin Verified Changes in SNAP Receipt on Covariates using CPS ASEC Data. For Purposes of Mittag (2018) Conditional Distribution Method.

VARIABLES	(1) Switch Off SNAP	(2) No Change	(3) Switch on SNAP
Δ SNAP Measure	-2.540***		2.366***
	(0.110)		(0.113)
Δ Usual Hours Worked	-0.001		0.005
	(0.005)		(0.005)
Δ Worked	-0.391**		0.056
	(0.193)		(0.197)
Δ Age: 20-30 years	-0.214		0.594*
	(0.343)		(0.311)
Δ Age: 30-40 years	-0.798*		0.576
	(0.470)		(0.460)
Δ Age: 40-50 years	-1.444***		0.743
	(0.556)		(0.563)
Δ Age: 50-60 years	-1.902***		0.379
	(0.615)		(0.629)
Δ Age: 60+ years	-1.941***		0.466
	(0.684)		(0.698)
Δ High School Diploma	-0.599**		0.370
	(0.245)		(0.278)
Δ Some College	-0.138		0.244
	(0.302)		(0.320)
Δ Associates Degree	-0.346		-0.168
	(0.383)		(0.385)
Δ Bachelor's Degree	-0.393		0.320
	(0.412)		(0.426)
Δ Advanced Degree	-0.663		-0.300
	(0.594)		(0.575)
Δ Married	0.151		-0.681**
	(0.344)		(0.306)
Δ Unemployment Rate	-0.050		0.053*
	(0.031)		(0.031)
Constant	-4.484***		-4.450***
	(0.052)		(0.051)
Observations	36,000	36,000	36,000

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Stata Code Used for Simulated Draws from the Conditional Distribution Method

```
*Estimate Multinomial Logit Model (See Appendix Table A10 for
estimates)
mlogit snap_admin_diff snap_reported_diff usual_hours_worked_diff
worked_diff $covariates_diff, base(0)

predict snap_admin_diff_xb1, xb outcome(-1)
predict snap_admin_diff_xb2, xb outcome(1)

*Generate 100 Simulated Datasets Containing Imputed SNAP Measure
expand 100

*Use Multinomial Logit coefficients to generate simulated observations
gen denom=1+exp(snap_admin_diff_xb1)+exp(snap_admin_diff_xb2)
gen p1=exp(snap_admin_diff_xb1)/denom
gen p2=1/denom
gen p3=exp(snap_admin_diff_xb2)/denom
gen u=runiform()
drop if p1==.
gen snap_admin_diff_sim=-1 if u<=p1
gen p12=p1+p2
replace snap_admin_diff_sim=0 if snap_admin_diff_sim==. & u<=p12
replace snap_admin_diff_sim=1 if snap_admin_diff_sim==.
```