The Effect of SNAP and School Food Programs on Food Security, Diet Quality, and Food Spending: Sensitivity to Program Reporting Error *

Kyung Min Kang †
Department of Economics, Johns Hopkins University

Robert A Moffitt ‡
Department of Economics, Johns Hopkins University and NBER

November 29, 2017

Abstract: There is an extensive research literature on the effects of the Supplemental Nutrition Assistance Program (SNAP) on food-related outcomes which has shown somewhat mixed results but generally favorable effects. However, most of the research has used data sets whose information on SNAP participation is gathered from responses on household surveys, and such responses are subject to reporting error. This study uses the FoodAPS data set to examine the effect of reporting error, for that data set contains information on SNAP participation gathered from government administrative records. Our analysis shows that the degree of reporting error is extremely small and has very little effect on the estimated impact of participation in the SNAP program on food security, diet quality, and food spending. A supplemental analysis of the effect of school food programs is also provided.

Paper prepared for presentation at the NBER Conference on Using FoodAPS for Research in Diet, Health, Nutrition, and Food Security, Washington DC., December 7-8, 2017.

*This research was supported by USDA grant no. 59-5000-5-0115 to the National Bureau of Economic Research, entitled "Using FoodAPS for Research in Diet, Health, Nutrition, and Food Security." The views expressed are those of the authors and not necessarily those of the Economic Research Service, Food and Nutrition Service, or the U.S. Department of Agriculture.We thank Marianne Bitler and the participants of a pre-conference in February, 2017 for valuable comments.

†kkang15@jhu.edu.
‡moffitt@jhu.edu.

A reasonably large literature has developed on the effect of the Supplemental Nutrition Assistance Program (SNAP) on a variety of food outcomes, including food consumption and spending, diet quality, and food security. The findings from this literature have been reviewed by Bitler (2016); Gregory et al. (2015); Hoynes et al. (2015); Hoynes and Schanzenbach (2016). The reviews show often positive effects of SNAP participation on these food outcomes, although in many cases the estimates from different studies form an uncomfortably wide range and include studies showing no statisticially significant effects. A number of reasons that the results are not stronger have been suggested. Perhaps the most important is that SNAP participation is self-selected and those with worse food outcomes may be more likely to participate in the program in the first place, biasing the effects of SNAP participation in a negative direction. Almost all of the existing studies make at least some attempt to address this selection bias issue, most often by some type of instrumental variable method

But another hypothesis is that there is measurement error in the variables measuring SNAP participation. In most household surveys, the number of families reporting that they receive SNAP benefits, when aggregated up to national or state totals, falls below the level reported in the administrative records of the SNAP program. This implies that at least some respondents are not reporting receipt of benefits, which implies measurement error in the measure of SNAP participation. The textbook case of classical measurement error in an independent variable is well-known to bias the OLS coefficient on the variable in question toward zero and, therefore, this could be an alternative, or additional, explanation for why the SNAP effects found in the literature have not been stronger.

This paper uses recently collected data from the National Household Food Acquisition and Purchase Survey (FoodAPS) to examine the effect of SNAP reporting error on food security, diet quality, and food spending. The FoodAPS is a household survey which collected information on 4,826 families on these food related variables as well as asking conventional survey questions on participation in the SNAP program. However, administrative records on SNAP participation were also collected for a large fraction of the families in the survey, allowing an examination of the degree of reporting error in the data. Because the three types of food-related outcomes were also collected, a determination can be made of how misreporting of SNAP receipt affects estimates of the impact of the prograam on those outcomes.

The paper proceeds as follows. First, a simple theoretical analysis of the impact of measurement error in a binary program participation variable on OLS estimates of the impact of the program on outcomes is provided. While the theoretical impact differs in exact form from the textbook model because the participation variable is binary and not continuous, the same bias toward zero occurs when measurement error is classical. A second analysis of the case where the misreporting is endogenous and correlated with the outcomes of interest is presented, which shows that the bias can be moved upward or downward, and therefore the bias-toward-zero result may not hold. A third case occurs when true program participation is endogenous and correlated with the unobservables in the outcome equations and when instrumental variable (IV) methods are used to correct the problem. It is shown that when the measurement error is classical, the IV estimator is biased in the opposite direction from the OLS estimator, namely, away from zero. A final case is analyzed where true program participation is endogenous but the measurement error is correlated with the unobservables in the outcome equation. In this case, the bias from the IV estimator is unchanged as long as the instruments used are independent of the measurement error.

Following the theoretical analysis, the sample to be used in the analysis is described, both the FoodAPS survey variables and the administrative data on SNAP participation. The data show that measurement error in the data occurs in both directions: some survey respondents do not report SNAP participation even though they appear in the administrative data ("false negatives") but some families report participation in the survey but do not appear in the administrative data ("false positives"). However, the rates are very low, with rates of false positives ranging from 3.5% to 5.5% in different samples. However, the false positives and false negatives do appear to be correlated with the outcome variables examined (food security, diet quality, and food spending) but are otherwise not far from classical in the sense that they are not highly correlated with other socioeconomic characteristics which are presumed to be exogenous. Nevertheless, the endogenous nature of the reporting error should be expected to affect the degree and direction of bias. Following these analyses, the paper presents its main results based on a comparison of the estimated effects of SNAP participation on food-related outcomes using, in one case, the reported SNAP participation and, in the other, SNAP participation as revealed in the administrative data. Estimation is conducted on several different samples. Both OLS and IV estimates are obtained and compared.

The results show that the degree of misreporting is very small, both false positives and false negatives. In addition, while there are some biases present from a correlation of misreporting with food-related outcomes, this bias is also small. Estimates of SNAP participation impacts on food insecurity, diet quality, and food expenditure show approximately equivalent estimates whether reported participation or participation based on administrative data are used. A final, brief analysis of the School Breakfast Program and the National School Lunch Program impacts show no significant effects of those programs on health-related measures.

1 A Brief Theoretical Analysis of Bias from Measurement Error in Analyses of the Impact of Participation in a Program

Let y be an outcome variable of interest, P^* be a binary variable for true program participation, and P^r be a binary variable for reported program participation in a survey. Let the true model be

$$y = \alpha + \beta P^* + \varepsilon \tag{1}$$

where subscripts $i = 1, ...N$ are omitted for notational simplicity. The true effect of participation on the outcome is β. Define the false negative error rate as $\pi_{01} = \Pr(P^r = 0|P^* = 1)$ and the false positive rate as $\pi_{10} = \Pr(P^r = 1|P^* = 0)$.

The following results are proven in the Appendix. If the two measurement errors are independent of ε, then the OLS estimator of β is, in the limit,

$$\beta[1 - \frac{\pi_{01}(1-q^*)}{q^r} - \frac{\pi_{10}q^*}{1-q^r}] \tag{2}$$

where $q^* = \Pr(P^* = 1)$ and $q^r = \Pr(P^r = 1)$. The two ratios inside the brackets can be shown to be less than 1, implying that the OLS estimator provides an estimate of β that is biased toward zero.[1] If the measurement error is correlated with ε, the OLS estimator of β has the limit

$$\beta[1 - \frac{\pi_{01}(1-q^*)}{q^r} - \frac{\pi_{10}q^*}{1-q^r}] + A, \tag{3}$$

where

$$\begin{aligned} A =& q^r[E(\varepsilon|P^r = 1, P^* = 1) - E(\varepsilon|P^r = 0, P^* = 1)] \\ &+ (1-q^r)[E(\varepsilon|P^r = 1, P^* = 0) - E(\varepsilon|P^r = 0, P^* = 0)]. \end{aligned} \tag{4}$$

In this expression, the first term in brackets is the difference in y for those who are on the program in the administrative records who do and do not report participation in the survey (and hence is a measure of endogeneity related to false negatives) and the second term in brackets is the difference in y for those who are not on the program in the administrative records but who do and do not report participation in the survey (and hence is a measure of endogeneity related to false positives). The bias term A is a weighted average of the two biases in brackets. If the presumed true participation variable, P^*, is correlated with ε, all OLS estimators are additionally biased. If an instrumental variable z is available which is

[1] Kane et al. (1999) also show this result.

independently distributed of ε and if IV is applied, the IV estimator of β has the limit

$$\frac{\beta}{1 - \pi_{10} - \pi_{01}} \tag{5}$$

and since the denominator is less than one, the estimator is biased upwards. [2] However, in this case, if the probability of a false positive or the probability of a false negative is correlated with ε, the limit of the IV estimator remains the same as long as those probabilities are distributed independently of z.

In our work, we will provide numerical estimates of all these bias terms.

2 The FoodAPS Data and Its Implied Response Errors

2.1 The Data and Our Main Outcome Variables

The FoodAPS data is a nationally representative survey of 4,826 households surveyed between April 2012 and January 2013, collecting data on general socioeconomic characteristics as well as specific survey questions on food security, diet quality, and food spending. The primary respondent of each sampled household was interviewed twice in person and up to three times by telephone. To concentrate on a SNAP-eligible population, we select those households in the survey who had income less than 200 percent of the federal poverty line, had assets less than $3,000, and who had at least one child. We also restrict our sample to those who gave consent to have the SNAP administrative records matched to their survey data. The resulting sample has 1,283 households.

The outcome measures we use from the FoodAPS data are as follows.

Household food insecurity To determine how SNAP contributes to food insecurity, we use a binary variable indicating whether a household was food insecure. The survey asks

[2] Kane et al. (1999) also have this result.

10 questions based on USDA's 30-day Adult Food Security Scale, and categorizes households that affirmed three (six) or more items as having low (very low) food security. We classify households with low or very low food security as food insecure.

Diet Quality To measure the effect of SNAP on diet quality, we use household-level data on Healthy Eating Index (HEI) scores as our primary outcome variable. The index is a measure of diet quality based on the 2010 Dietary Guidelines for Americans that consists of scores from 12 dietary elements: total vegetables, greens and beans, total fruit, whole fruit, whole grain, dairy, total protein foods, seafood and plant proteins, fatty acids, sodium, refined grains, and SOFAAS (solid fats, alcoholic beverages, and added sugar). We use the sum of the scores (HEI-2010) as well as all of the scores from each of the components.

Food Expenditure The FoodAPS survey asks each household to provide food expenditures for each event of food acquisition over the 7-day survey period. We use the sum of these expenditure over the week to create our measure of total food expenditure. We create three main outcome measures of food expenditures: food-at-home expenditures, food-away-from-home expenditures, and total food expenditures.

Table 1 shows the means of the outcome variables for samples A, B, and C which we describe momentarily.

2.2 The SNAP Survey and Administrative Data

The survey question about SNAP participation in the FoodAPS asks if a household had participated in the last 30-31 days and is therefore quite simple. However, the collection of SNAP administrative data was complex. In most, but not all, of the states, SNAP administrative records (so-called ADMIN data) from caseload records were available to match individuals probabilistically to the survey data using name, address, and other characteristics. This was not possible in states that did not provide ADMIN (i.e., caseload) data. In addition, states provided an administrative data set recording recent SNAP transactions at food outlets, the so-called ALERT data. In states with no caseload data, FoodAPS respondents were matched probabilistically to the ALERT data. In states with caseload data and with a caseload identifier in those data, cases in the ADMIN data were matched to those in the ALERT data on the basis of that caseload identifier; but if the ADMIN data did not have case identifiers or the identifiers did not match, the match to the ALERT data was made probabilistically.

The issue with this complex administrative data design is that households in different states had different chances of being matched to administrative data, depending on the constellation and nature of the administrative data that happened to be available in the state of residence. Because there is no way of being sure which configuration of state administrative data give the most accurate responses, we simply conduct our estimates on three different samples. Sample A is the sample of all states, which is presumably the least accurate because it includes states with no ADMIN data and states with ADMIN data and no caseload identifiers. Sample B is the sample of all states except those which provided no caseload, ADMIN data, and is therefore presumably somewhat more accurate. Sample C is a sample only of states who provided ADMIN data with caseload identifiers, which should be the most accurate sample. However, the tradeoff is that sample sizes fall, the more restrictive the sample.

For each sample, we use a binary SNAP participation from the survey data and one from the administrative data. The variable in the FoodAPS database which is entirely drawn from survey questions is labeled SNAPNOWREPORT. A variable in the FoodAPS database which incorporates the administrative data, labeled SNAPNOWHH, revises SNAPNOWREPORT by setting it equal to 1 if the household is found in the administrative data for SNAP receipt or transactions within the last 30-31 days (i.e., the time period covered by the FoodAPS survey question) and equal to 0 if SNAP receipt or transactions in the administrative data were more than 32 days before. SNAPNOWHH is set equal to SNAPNOWREPORT if a match could not be made.

Table 2 gives the distributions of SNAPNOWREPORT crossed with SNAPNOWHH for our three samples. In sample A, about 54% of households participate in SNAP according to SNAPNOWHH, whereas the percentage according to SNAPNOWREPORT is about slightly above 50%. The SNAP participation rate as measured by either variable falls as the sample becomes more accurate: the respective percentages are 52% and 48.6% for sample B, and 49.3 and 45.7% for sample C, somewhat surprisingly. Across sample groups, the fractions of people who misreport are relatively small and do not vary much across samples. Across the groups, about 4 percent of the sample have mismatch in the two variables, as can be seen by summing the off-diagonal elements. The difference in participation rates by the two variables within each sample can be mostly explained by the fraction of false negatives (those with $SNAPNOWREPORT = 0, SNAPNOWHH = 1$). The fractions of false positives are 0.004, 0.005, and 0.002, respectively.

Table 3 shows the degree of bias in the OLS and IV estimators as discussed in section 1. The row labeled "OLS bias" reports the factor loading on OLS coefficient estimates when the measurement errors are independent of ϵ, as discussed in equation (2). OLS coefficients are biased toward 0 as the theory predicts, and in all three samples OLS estimates would underestimate the true effect of SNAP on outcome variables by 7 to 9 percent. The second row reports estimates of the factor loading on the bias of an IV estimator in equation (5) if an instrumental variable is available and independently distributed from the error term in the main equation. The IV estimator is biased upward by about 8 to 9 percent. The factor loading on bias differs slightly across the samples, ranging from 8.1 in sample A to 9.4 in sample B. Note that the IV bias is not the smallest in the most accurate sample, sample C.

The estimates given in table 3 arise only from misreporting when the measurement error is independent of the error term in the outcome equation, and affect estimates by the same factor regardless of the outcome variable used. If measurement error is correlated with ϵ, the OLS estimates are biased by an additional term that depends on the dependent variable, as shown in equation (3). Table 4 reports the estimates of equation (4) for all the outcome variables used. Comparing the magnitudes of the biases to the means of the outcomes variable in Table 1. While the biases are small relative to the means for some outcomes variables (e.g., LFS), they are sizable (e.g., more than 10 percent of the mean values) for others (sea plant, fatty acid, FAFH, Total Expenditure). In addition, contrary to results in table 3, both the direction and the magnitude of the bias due to endogeneity of reporting error differs across the three samples and is often largest for the most accurate sample, Sample C.

Having shown the existence of measurement error in SNAP participation and the degree to which such error would lead introduce bias in both OLS and IV estimators, we next study determinants of misreporting. Columns (1), (2), and (3) of Table 5 report OLS coefficients of false negatives on a vector of covariates for the three samples, conditional on SNAP non-participation. A household is less likely to falsely report as not receiving SNAP benefits if the primary respondent of the household is hispanic in sample A and B, but not in C. Similarly, the primary respondent's marital status is a significant determinant of false negatives in the first two groups. None of the covariates show significant association with false negatives in group C, which can be partly attributed to its relatively smaller sample size. Even in samples A and B, however, the the independent variable do not meaningfully predict misreporting, as shown by the low values of coefficients of determination.

Table 6 reports OLS coefficients estimate of a binary variable equal to 1 if falsely reporting SNAP participation on the same set of covariates. Column (1) shows that among SNAP participants, that being male, living in a metro area, and being in a marriage is positively associated with false positives, while having elderly present in the household is negatively associated with misreporting for sample A. These variables are also significant determinants of misreporting in sample B and sample C in columns (2) and (3), with the exception of living in metro for sample B, and male in sample C. R^2 values are higher for the regressions of false positive compared to those on false negatives but still low, signifying little explanatory power for misreporting by observable characteristics.

3 Results for the Effect of SNAP on Outcomes

We estimate two econometric models to show the effect of SNAP on a given food-related outcome using the two measures of SNAP participation for the three different samples. First, we present the results of the ordinary least squares model of the form :

$$y_i = \beta SNAP_i + X_i\delta + \epsilon_i \tag{6}$$

where y is one of our ourcome variables discussed in section 2.1, i indexes an individual, X_i is a vector of individual- or household-level covariates, and ϵ_i is an error term with mean 0. $SNAP_i$ is one of the SNAP participation variables in the main analysis, and we also consider SNAP benefit amount in some specifications. Table 7 reports the summary statistics for the control variables used.

We begin with food insecurity outcomes for sample A. Columns (1) and (2) of Table 8 report simple bivariate regressions of LFS on the reported and administrative SNAP participation variables, respectively.The point estimates are nearly identical, negative, and statistically insignificant. [3] When selection into SNAP is ignored, the point estimates cannot be given causal interpretations since there are many omitted factors that jointly predict LFS and SNAP participation. In the subsequent columns, we attempt to reduce the omitted variable bias by controlling for several observable factors. In column (3) and (4), we add the primary respondent's characters as controls. This moves the estimates of β further away from 0, though the estimate remains to be negative and insignificant. Adding household-level controls including family size, number of children, number of elderly present does not change the estimates much, as is reported in columns (6) and (7). Subsequent columns add household's economic variables. The point estimates are again nearly identical for the two SNAP variables, but both becoming more negative yet remaining insignificant. The estimates suggest

[3] The similarity of the two estimates is the result of offsetting biases, the Table 3 bias that makes the estimate using reported SNAP status smaller in absolute value combined with the Table 4 bias that moves it the other way.

that conditional on observed primary respondent and household characteristics including sex, race, age, educational attainment, household composition, work status and household income, the average marginal effect of participating in SNAP decreases the likelihood of experiencing food insecurity by about 3 percentage points although very imprecisely estimated. Table 9 reports estimates with the presumably more accurate Samples B and C. The results show that the point estimates are again quite similar across the two SNAP participation measures (the differences are slightly larger than in Table 8 bias the Table 4 biases are larger) but the SNAP impacts are not estimated to be positive without controls and negative and small with demographic characteristics controls. These again show the importance of omitted factors, however.

While conditional association between SNAP receipt and LFS can be studied using the OLS model, the point estimates are biased if SNAP participation is endogenous in equation (6), which is likely to occur because of selection bias. The selection bias in SNAP participation in the literature has been controlled largely in two ways. In one strand, the historical introduction of SNAP and over-time policy variation in the SNAP programs is used to identify the effect of SNAP (Hoynes et al. (2016); Hoynes and Schanzenbach (2012); Currie and Moretti (2013)). Because of the cross-sectional nature of the FoodAPS data set, this method cannot be used for the present analysis. Instead, we use variation in state level SNAP policies as instruments to address selection, and compare IV coefficients across different samples to address the effect of measurement error on these estimates. Explicitly, we estimate the following model:

$$\begin{aligned} y_i &= \beta SNAP_i + X_i\delta + \epsilon_i \\ SNAP_i &= Z_i\gamma + \nu_i \\ \nu_i \perp \epsilon_i&,\ E(Z_i'\epsilon_i) = 0,\ E(Z_i SNAP_i) \neq 0 \end{aligned} \tag{7}$$

That is, we need variables that are correlated with SNAP variables but uncorrelated with the outcome variable, except through their effects on SNAP participation. Table 10 shows past studies using IVs and the instruments used; we attempt to use similar instruments. We merge the FoodAPS data set with information on State-level SNAP policies in the SNAP Policy Database to create relevant instruments. We use indicator variables for whether or not the State requires fingerprinting of SNAP applicants and whether or not the State uses uses a reporting option that reduces requirements for reporting changes in household circumstances, and per capita outreach spending defined by the sum of Federal, State, and grant outreach spending divided by the state's population. The use of biometric technology is negatively associated with SNAP participation, while the latter two variables are positively related. The first-stage F-statistics for the instruments are 7.8, 10.9, and 9.6 for Samples A, B, and C, respectively.

Even after selection has been controlled for, SNAP has no effect on food security as measured by LFS, as shown in table 11. While all the coefficients are insignificant, several patterns different from OLS estimates emerge. First, the point estimates are positive for some of the specifications, namely for both measures of SNAP participation for sample A, and using the administrative SNAP variable for sample C. Also, both the within and across sample difference in SNAP coefficients is larger, although IV coefficients are less precisely estimated. Standard errors of β coefficients for sample C are about twice as large as than those of other samples, likely because of the significantly smaller sample size.

Next, we extend OLS and IV analysis above to study the effect of SNAP on diet quality and food expenditures across samples with differing accuracy on administrative record of SNAP receipt. OLS estimates of the effect of SNAP on HEI total and component scores are reported in table 12. SNAP has no significant association with diet quality as measured by HEI scores in samples A and B using SNAPNOWREPORT. The effect of SNAP on seafood and plant proteins, however, is large, significant, and negative. Larger differences in OLS coefficients appear across the two measures of SNAP participation, consistent with Table 4. Coefficients on SNAPNOWHH on total fruit consumption are large, significant, and negative for both sample group A and B, whereas the point estimate using SNAPNOWREPORT on wholefruit is positive for group A and about half the magnitude for group B.

IV estimates that account for selection differ significantly from OLS estimates, as reported in Table 13. Estimates across the two SNAP participation variables are relatively closer than OLS estimates, with the exception of the estimated effect of SNAP on Total HEI for sample C. However, the point estimates using IV show that SNAP has large, negative effects on most aspects of food consumption, including vegetable, greens and beans, total fruit, fruit, and seafood and plant proteins. Thus, the results show, contrary to expectation, negative effects on diet quality but mostly insignificant in OLS and much more frequently significant in IV. quality.

One possibility for these results is that the effect of SNAP on HEI differs by the amount of the benefit. To examine this possibility, we regress HEI scores on two measures of benefit amount received. The first measure is created by interacting the reported last received amount with reported SNAP participation; the second uses the SNAP receipt amount interacted with SNAPNOWHH.[4] However, the effects of SNAP benefit amounts on HEI scores are still large, negative, and significant for the total HEI scores, vegetables, greens and beens, total fruits, whole fruits, and SOFAAS across the samples, as reported in Table 14. Also, while the signs of coefficients on the two benefit amount variables align, the magnitude is generally larger using the reported SNAP amounts.

The last part of Table 12 shows the association of SNAP participation with food expenditures. The coefficients on FAFH across sample groups and across SNAP variables are negative and insignificant, while the analogous point estimates on FAH and total expenditures and large, positive, and significant. The estimates on total food expenditures are similar across the two SNAP participation variables, ranging from 21 in sample A to 35 in sample C. However, once we control for selection into SNAP, the results are quite different,

[4] We define the benefits amount in this way since while SNAPNOWHH and SNAPNOWREPORT measure SNAP participation in the last month, the SNAP amount variables in the FoodAPS include benefits measures received more than 32 days ago.

as shown by the IV estimates in Table 13[5]. While similarity in coefficients estimates across SNAPNOWHH and SNAPNOWREPORT remains, the positive association between FAH and total expenditures and SNAP participation disappears. The effect of SNAP on food expenditures is larger for sample C compared to the other samples, with the magnitudes being larger by a factor of 1.5 to 2.

4 A Brief Analysis of School Programs

The Analysis of the SBP and NSLP using the FoodAPS data set requires the construction of the two program participation variables, as the survey does not ask its respondents of participation status in the school programs. Instead, FoodAPS asks whether child's school breakfasts are free, at a reduced price, or full price and ask the same question for child's school lunches. We define binary variables indicating SBP and NSLP equal to 1 if breakfasts and lunches, respectively, were free or at a reduced price, and 0 otherwise. We restrict the analytic sample for school programs to children aged 5 to 18 years old who are enrolled in primary or secondary schools with whom school was in session at the time of the interviews, and whose household falls in sample A, B, or C above. The literature on school nutrition programs primarily studies impacts on child achievement (Ribar and Haldeman (2013); Frisvold (2015)), and dietary effects including obesity (Hofferth and Curtin (2005); Millimet et al. (2010)). The identification strategies used in the literature include instruments for variations across states in the threshold for SBP participation, and instrumental variable methods using public school attendance or school's food characteristics as exlusion restrictions.

Here, we provide estimates of the school nutrition programs on obesity and bmi scores as the main outcomes. The FoodAPS has a constructed variable "BMICAT" that categorizes individuals into 3 categories: not overweight, overweight, and obese. For children, those whose BMI-for-age percentile falls between 85 to 95th percentile are categorized as over-

[5]IV estimates using SNAP benefits amount show similar pattern, as reported in Table 15

weight, and those at or above the 95th percentile are categorized as obese. Here, we define overweight as those who are categorized as either overweight or obese in BMICAT.[6]

Columns (1)-(3) of Table 16 report OLS regressions of overweight, obesity, and bmi scores on the SBP participation variable, respectively. The coefficient estimates are insignificant in all three samples but the direction and the magnitude of the point estimates differ slightly across the samples. For instance, all three coefficients are negative in sample A, but they are all positive in sample C. These differences, however, are not statistically significant. Next, we also include SNAP participation dummies to control for the correlation between SNAP participation and BMI scores [7]. Columns (4)-(6) report estimates using SNAPNOWHH and columns (7)-(9) report analogous point estimates using SNAPNOWREPORT as the measure of SNAP participation. Again, none of the point estimates are significant, and point estimates for the three outcomes are almost identical across the two definitions of SNAP participation. Including SNAP participation as controls have insignificant effects on the outcome variables and the sample defined.

Table 17 reports similar estimates using NSLP as the key explanatory variable. Except for BMI in sample C, participation in NSLP has no significant association with BMI or being obese. Also, while coefficient estimates for sample C are more likely significant, they are not always of the expected sign. Again, point estimates differ more across sample than across definitions of SNAP participation. A key difference, however, is that the magnitude of coefficients using SNAPNOWREPORT is slightly larger compared to estimates using SNAPNOWHH, and that the association between NSLP and BMI is negative and significant at 10% for sample C. Overall, Tables 16 and 17 shows that school nutrition programs do not have discernible effects on overweight and obesity. BMI is negatively correlated with NSLP, although its significance depends on the sample, and exhibits no statistical association with SBP.

[6] We report OLS regression coefficients only, as instruments used in the literature also available in FoodAPS dataset suffer from a weak instruments problem.

[7] There is a reasonably large literature studying the relationship between SNAP and obesity: for a recent survey, see Gundersen (2015)

Next, we briefly examine the relationship between the two school nutrition programs and SNAP. In particular, we present estimates of linear regressions of the form

$$y_i = \gamma SNAP_i + X_i\beta + \epsilon_i \tag{8}$$

where SNAP denotes the program particiupation variable, y is either SBP or NSLP participation, the vector X includes a set of statistical controls, and i indexes students. The first part of Table 18 report OLS estimates of γ for SBP and NLSP using the two measures of SNAP participation. SNAP participants are more likely to participate in the school lunch program, as shown by the point estimates for NLSP which are around .2 for all 6 cases and significant. The association between SBP and SNAP participation is positive, but with smaller coefficients. The significance of the coefficients depends on the measure of SNAP, with reported SNAP participation showing smaller and insignificant associations between the two. Of course, the coefficient estimates suggest the correlation and cannot be taken as causal. The second part of Table 18 shows IV estimates using the same sets of instruments used in the previous section. The point estimates again show increase in NSLP participation among SNAP participants.

5 Summary and Conclusions

This study has used the FoodAPS data set to examine the effect of reporting error in survey measures of SNAP participation in gauging the impact of that participation on food insecurity, diet quality, and food expenditure. The FoodAPS data are partly matched with SNAP administrative data, allowing an examination of the issue. The administrative data were matched in different ways in different states, so we test our models on different samples of states with different rates of matching. The results show that the degree of misreporting is very small, both false positives and false negatives. False positive rates are generally less than 1 percent and false negative rates are less than 4 percent. These imply downward biases in OLS estimates of less than 10 percent and upward biases in IV estimates of about the same amount. We find that misreporting is correlated with food insecurity, diet quality, and food expenditures, with the consequent implied biases are sometimes large and sometimes small. We also find that the probabilities of having a false negative or a false positive are uncorrelated with most standard socioeconomic characteristics. Our main results, which compare OLS and IV estimates of the impact of SNAP on food-related outcomes, using reported SNAP participation and administrative-based SNAP participation, differ little. This is because the degree of misreporting is small, and the biases from endogenous misreporting are small. Consequently, the FoodAPS data set we employ suggests that misreporting of SNAP participation in surveys may not be a serious problem. Our analyses using either measure of SNAP participation reveals few significant effects in the expected direction with either OLS or IV. In fact, the IV estimates more frequently yield estimates in the unexpected direction. This may be because of weakness in the instruments, or because of the cross-sectional nature of the FoodAPS, which does not permit causal estimation using over-time variation commonly used elsewhere in the literature. A final, brief analysis of the School Breakfast Program and the National School Lunch Program impacts show no significant effects of those programs on health-related measures, although participation in the programs is positively affected by SNAP receipt, whether survey-based or administrative-based, as expected.

References

Baum, C. (2007). The effects of food stamps on obesity.

Bitler, M. (2016). The health and nutrition effects of snap: selection into the program and a review of the literature on its effects. In J. Bartfeld, C. Gundersen, and T. Ziliak (Eds.), *SNAP matters: how food stamps affect health and well-being*. Stanford, CA: Stanford University Press.

Currie, J. and E. Moretti (2013). Did the introduction of food stamps affect birth outcomes in california? making americans healthier: social and economic policy as health policy. 2008.

Deb, P. and C. A. Gregory (2016). Who benefits from the supplemental nutrition assistance program? an analysis of food insecurity and spending.

Frisvold, D. E. (2015). Nutrition and cognitive achievement: An evaluation of the school breakfast program. *Journal of public economics 124*, 91–104.

Greenhalgh-Stanley, N. and K. Fitzpatrick (2013). *Food Stamps, Food Sufficiency, and Diet-Related Disease among the Elderly*. Institute for Research on Poverty, University of Wisconsin-Madison.

Gregory, C., M. P. Rabbitt, and D. C. Ribar (2015). The supplemental nutrition assistance program. In J. Bartfeld, C. Gundersen, and T. Ziliak (Eds.), *SNAP matters: how food stamps affect health and well-being*. Stanford, CA: Stanford University Press.

Gregory, C. A., M. Ver Ploeg, M. Andrews, and A. Coleman-Jensen (2013). Supplemental nutrition assistance program (snap) participation leads to modest changes in diet quality.

Gundersen, C. (2015). Snap and obesity. *SNAP Matters: How Food Stamps Affect Health and Well Being*, 161–185.

Hofferth, S. L. and S. Curtin (2005). Poverty, food programs, and childhood obesity. *Journal of Policy Analysis and Management 24*(4), 703–726.

Hoynes, H. and D. W. Schanzenbach (2016). U.s. food and nutrition programs. In R. A. Moffitt (Ed.), *Economics of Means-Tested Transfers Programs*. Chicago: Univ. Chicago Press.

Hoynes, H., D. W. Schanzenbach, and D. Almond (2016). Long-run impacts of childhood access to the safety net. *The American Economic Review 106*(4), 903–934.

Hoynes, H. W., L. McGranahan, and D. W. Schanzenbach (2015). Snap and food consumption. In J. Bartfeld, C. Gundersen, and T. Ziliak (Eds.), *SNAP matters: how food stamps affect health and well-being*. Stanford, CA: Stanford University Press.

Hoynes, H. W. and D. W. Schanzenbach (2012). Work incentives and the food stamp program. *Journal of Public Economics 96*(1), 151–162.

Kabbani, N. S. and M. Yazbeck (2004). *The role of food assistance programs and employment circumstances in helping households with children avoid hunger*. University of Wisconsin–Madison, Institute for Research on Poverty.

Kane, T. J., C. E. Rouse, and D. Staiger (1999). Estimating returns to schooling when schooling is misreported. Technical report, National bureau of economic research.

Millimet, D. L., R. Tchernis, and M. Husain (2010). School nutrition programs and the incidence of childhood obesity. *Journal of Human Resources 45*(3), 640–654.

Ratcliffe, C. and S.-M. McKernan (2010). How much does snap reduce food insecurity? washington. *DC: The Urban Institute*.

Ribar, D. C. and L. A. Haldeman (2013). Changes in meal participation, attendance, and test scores associated with the availability of universal free school breakfasts. *Social Service Review 87*(2), 354–385.

Shaefer, H. L. and I. A. Gutierrez (2013). The supplemental nutrition assistance program and material hardships among low-income households with children. *Social Service Review 87*(4), 753–779.

Tables

Table 1: Summary Statistics

	Sample A	Sample B	Sample C
LFS	0.135	0.137	0.148
	(0.013)	(0.013)	(0.020)
HEI-2010	48.154	48.099	48.255
	(0.439)	(0.490)	(0.652)
Vegetables	2.603	2.614	2.587
	(0.055)	(0.047)	(0.051)
Greens and Beans	1.396	1.378	1.392
	(0.060)	(0.072)	(0.111)
Total Fruit	1.958	1.949	1.933
	(0.065)	(0.059)	(0.075)
Whole Fruit	2.259	2.251	2.25
	(0.080)	(0.081)	(0.079)
Whole Grain	1.729	1.74	1.741
	(0.133)	(0.122)	(0.121)
Dairy	5.777	5.735	5.614
	(0.133)	(0.159)	(0.165)
Total Protein	4.058	4.094	4.137
	(0.051)	(0.050)	(0.072)
Seafood, Plant Protein	1.672	1.71	1.768
	(0.066)	(0.074)	(0.111)
Fatty Acid	4.895	4.966	5.081
	(0.156)	(0.155)	(0.147)
Sodium	5.956	5.817	5.754
	(0.161)	(0.176)	(0.272)
Refined Grain	5.812	5.651	5.37
	(0.177)	(0.192)	(0.244)
SOFAAS	10.222	10.233	10.567
	(0.286)	(0.299)	(0.429)
FAFH Expenditure	50.244	52.252	54.437
	(2.093)	(2.214)	(3.073)
FAH Expenditure	117.635	122.053	120.13
	(5.630)	(4.891)	(5.592)
Total Expenditure	167.879	174.305	174.568
	(6.176)	(5.738)	(8.039)

LFS is low food security. FAFH is food away from home, FAH is food at home. Vegetables, greens and beans, total fruit, whole fruit, whole grain, dairy, total protein, seafood, plant protein, fatty acid, sodium, refined grain, and SOFAAS each represent HEI component scores.

Table 2: Distribution of SNAPNOWREPORT and SNAPNOWHH

	SNAPNOWREPORT					
	Sample A		Sample B		Sample C	
SNAPNOWHH	0	1	0	1	0	1
0	0.457	0.004	0.475	0.005	0.506	0.002
1	0.036	0.503	0.039	0.481	0.038	0.455
N	1282		1109		687	

The values reported are fraction of households in the category. All values are weighted by FoodAPS survey weights. In each SNAPNOWHH and SNAPNOWREPORT, 0 indicates non-participation and 1 indications participation.

Table 3: Estimated OLS and IV Biases

	Sample A	Sample B	Sample C
OLS Bias	0.919	0.914	0.926
IV Bias	1.081	1.094	1.088

OLS bias reports the sample estimate of $[1 - \frac{\pi_{01}(1-q^*)}{q^r} - \frac{\pi_{10}q^*}{1-q^r}]$ in equation (2), and IV bias is shows the estimate of $\frac{1}{1-\pi_{10}-\pi_{01}}$ in equation (5)

Table 4: Estimates of magnitude of Bias

	LFS	Total HEI	Vegetables	Greens and Beans	Fruit	Whole Fruit	Whole Grain
Sample A	0.004	0.632	0.286	-0.351	0.324	0.466	-0.207
Sample B	0.018	0.876	0.276	-0.315	0.277	0.372	-0.273
Sample C	0.043	-4.6	-0.634	-0.702	0.023	-0.157	-0.068
	Dairy	Protein	Sea Plant	Fatty Acid	Sodium	Refined Grain	SOFAAS
Sample A	0.738	0.05	0.257	-1.442	-0.578	0.769	0.322
Sample B	0.708	0.105	0.255	-1.365	-0.384	0.986	0.235
Sample C	0.432	-0.462	-0.918	-1.566	1.672	0.766	-2.984
	FAFH	FAH	Tot. Exp				
Sample A	-18.934	-4.077	-19.071				
Sample B	-20.346	0.218	-16.564				
Sample C	-2.157	10.756	8.872				

LFS is low food security, VLFS is very low food security. Total HEI is HEI-2010 total score, and vegetables, greens and beans, whole fruit, whole grain, dariy, protein, sea plant, fatty acid, sodium, refined grain, and SOFAAS each represent HEI-2010 component scores. The values reported are estimates of of equation (4) for the key variables

Table 5: Determinants of False Negatives

	Sample A	Sample B	Sample C
Male	0.013	0.018	0.011
	(0.014)	(0.018)	(0.009)
Age	0.003	0.004	0.001
	(0.003)	(0.003)	(0.001)
Age Squared	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Black	-0.011	-0.013	-0.008
	(0.011)	(0.014)	(0.005)
Hispanic	-0.016*	-0.019**	-0.006
	(0.008)	(0.009)	(0.007)
Married	-0.009	-0.012	-0.004
	(0.006)	(0.007)	(0.006)
Widowed	-0.018*	-0.020*	-0.009
	(0.010)	(0.011)	(0.007)
Less than HS	0.004	0.004	0.007
	(0.011)	(0.014)	(0.006)
HS only	0.000	0.002	0.001
	(0.004)	(0.006)	(0.006)
Household size	0.001	0.001	-0.001
	(0.002)	(0.003)	(0.003)
Children, aged 4 to 10	0.006	0.006	0.004
	(0.011)	(0.014)	(0.014)
Elder present	-0.003	-0.003	-0.006
	(0.009)	(0.011)	(0.004)
Metro	0.014	0.014	0.007
	(0.009)	(0.009)	(0.005)
Health	0.006	0.013	0.021
	(0.010)	(0.019)	(0.024)
Working	-0.003	-0.004	-0.004
	(0.006)	(0.008)	(0.008)
log income	-0.022	-0.022	0.001
	(0.021)	(0.024)	(0.006)
Own Housing	-0.001	0.000	-0.004
	(0.006)	(0.008)	(0.003)
Constant	0.109	0.101	-0.027
	(0.118)	(0.133)	(0.036)
N	462	394	261
R-squared	0.030	0.035	0.018

The dependent variable is a binary indicator equal to 1 if $SNAPNOWHH = 0$ and $SNAPNOWREPORT = 1$. FoodAPS sampling weights are used in all regressions.

Table 6: Determinants of False Positives

	Sample A	Sample B	Sample C
Male	0.079	0.131**	0.110
	(0.047)	(0.062)	(0.113)
Age	-0.010	-0.006	-0.004
	(0.008)	(0.009)	(0.012)
Age Squared	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Black	0.023	0.011	-0.037
	(0.022)	(0.028)	(0.037)
Hispanic	0.005	0.001	-0.005
	(0.023)	(0.023)	(0.033)
Married	0.076**	0.066*	0.086**
	(0.035)	(0.035)	(0.036)
Widowed	0.013	0.050	0.099
	(0.072)	(0.104)	(0.111)
Less than HS	0.011	-0.024	0.026
	(0.031)	(0.034)	(0.030)
HS only	-0.016	-0.040	-0.008
	(0.028)	(0.034)	(0.047)
Household size	0.003	0.006	-0.003
	(0.010)	(0.011)	(0.011)
Children, aged 4 to 10	0.006	0.012	-0.036
	(0.052)	(0.056)	(0.047)
Elder present	-0.105***	-0.127***	-0.110**
	(0.031)	(0.033)	(0.044)
Metro	0.066**	0.045	0.080**
	(0.027)	(0.033)	(0.036)
Health	-0.053	-0.067	-0.080
	(0.049)	(0.050)	(0.064)
Working	0.036	0.026	0.045
	(0.025)	(0.024)	(0.034)
log income	0.011	0.018	0.007
	(0.013)	(0.016)	(0.026)
Own Housing	0.032	0.045	0.075*
	(0.025)	(0.029)	(0.040)
Constant	0.041	-0.037	0.009
	(0.188)	(0.210)	(0.358)
N	766	668	398
R-squared	0.054	0.075	0.088

Note: $*p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$. The dependent variable is a binary indicator equal to 1 if $SNAPNOWHH = 1$ and $SNAPNOWREPORT = 0$. FoodAPS sampling weights are used in all regressions.

Table 7: Summary Statistics: Covariates

	Sample A	Sample B	Sample C
Male	0.156	0.141	0.122
	(0.015)	(0.014)	(0.017)
Age	38.281	38.226	38.506
	(0.465)	(0.499)	(0.744)
Black	0.268	0.245	0.149
	(0.049)	(0.049)	(0.027)
Hispanic	0.298	0.33	0.401
	(0.050)	(0.060)	(0.073)
Married	0.405	0.443	0.482
	(0.033)	(0.035)	(0.028)
Widowed	0.026	0.022	0.033
	(0.004)	(0.005)	(0.008)
Less than HS	0.26	0.266	0.277
	(0.017)	(0.017)	(0.025)
HS only	0.293	0.28	0.289
	(0.017)	(0.018)	(0.023)
Household Size	4.314	4.353	4.401
	(0.082)	(0.095)	(0.132)
Children, 4 to 10	0.145	0.153	0.145
	(0.012)	(0.015)	(0.018)
Elder Present	0.051	0.057	0.066
	(0.010)	(0.010)	(0.013)
Metro	0.873	0.885	0.896
	(0.044)	(0.050)	(0.047)
Health	0.085	0.081	0.06
	(0.013)	(0.011)	(0.013)
Working	0.472	0.469	0.465
	(0.021)	(0.020)	(0.031)
log income	7.55	7.559	7.618
	(0.038)	(0.036)	(0.044)
Own Housing	0.324	0.318	0.331
	(0.027)	(0.030)	(0.040)
N	1230	1064	659

All means and standard errors are weighted by FoodAPS sample weights. HS is highschool. Sex, race, and education measures are in fractions and report those of primary respondents.

Table 8: OLS Estimates of the Effect of SNAP on Households' Low Food Security Status: Sample A

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SNAPNOWREPORT	-0.007		-0.019		-0.019		-0.030	
	(0.031)		(0.033)		(0.033)		(0.028)	
SNAPNOWHH		-0.007		-0.018		-0.019		-0.031
		(0.030)		(0.032)		(0.034)		(0.030)
Male			0.041	0.040	0.049	0.049	0.047	0.048
			(0.054)	(0.055)	(0.056)	(0.056)	(0.051)	(0.051)
Age			0.007	0.007	0.014**	0.013*	0.012	0.012
			(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)
Age Squared			0.000	0.000	-0.0001**	0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Black			-0.013	-0.013	-0.010	-0.010	-0.009	-0.009
			(0.044)	(0.044)	(0.043)	(0.043)	(0.042)	(0.041)
Hispanic			0.048	0.050	0.044	0.046	0.034	0.035
			(0.039)	(0.039)	(0.036)	(0.036)	(0.036)	(0.035)
Married			-0.044	-0.042	-0.055	-0.054	-0.038	-0.036
			(0.036)	(0.035)	(0.035)	(0.035)	(0.034)	(0.034)
Widowed			-0.065	-0.069	-0.058	-0.062	-0.051	-0.054
			(0.056)	(0.056)	(0.054)	(0.054)	(0.059)	(0.059)
Less than HS			0.096**	0.099**	0.096**	0.099**	0.079**	0.081**
			(0.036)	(0.037)	(0.036)	(0.037)	(0.035)	(0.036)
HS only			0.051	0.051	0.050	0.049	0.028	0.028
			(0.047)	(0.047)	(0.048)	(0.048)	(0.039)	(0.039)
Household size					0.019	0.021*	0.027**	0.029**
					(0.012)	(0.011)	(0.012)	(0.011)
Children, aged 4 to 10					-0.098**	-0.104**	-0.103**	-0.109***
					(0.046)	(0.043)	(0.040)	(0.037)
Elder present					0.127	0.120	0.105	0.097
					(0.087)	(0.091)	(0.096)	(0.099)
Metro					-0.009	-0.007	-0.040	-0.039
					(0.041)	(0.042)	(0.039)	(0.039)
Health					-0.047	-0.050	-0.028	-0.030
					(0.045)	(0.044)	(0.039)	(0.038)
Working							-0.076**	-0.077**
							(0.036)	(0.036)
log income							-0.022	-0.023
							(0.020)	(0.021)
Own Housing							-0.062	-0.065
							(0.039)	(0.039)
Constant	0.151***	0.152***	-0.023	-0.019	-0.183	-0.182	0.061	0.070
	(0.023)	(0.025)	(0.147)	(0.148)	(0.168)	(0.166)	(0.186)	(0.191)
N	1282	1283	1276	1277	1276	1277	1228	1229
R-squared	0.000	0.000	0.027	0.027	0.040	0.042	0.059	0.061

Note: $*p < 0.10$, $** p < 0.05$, $*** p < 0.01$. Parameter estimates of equation (6) for sample A are reported. The dependent variable is a binary indicator equal to 1 if the household reports LFS and 0 otherwise. Sampling weights are used in all regressions.

Table 9: OLS Estimates of the Effect of SNAP on Households' Low Food Security Status, sample B and C

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
sample:	B	C	B	C	B	C	B	C
SNAPNOWREPORT	0.018	0.028	-0.006	-0.006				
	(0.035)	(0.042)	(0.030)	(0.037)				
SNAPNOWHH					0.020	0.030	-0.005	-0.004
					(0.034)	(0.042)	(0.030)	(0.038)
Controls	N	N	Y	Y	N	N	Y	Y
R-squared	0	0.002	0.059	0.074	0.001	0.002	0.061	0.074

Note: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Controls include indicators for whether primary respondent is male, black, hispanic, married, widowed, did not graduate from highschool, graduated from highschool but no postsecondary school enrollment, live in metro area, report having excellent health, currently working, and whether the household owns the residential unit. Continuous covariates include age and age squared and log household income.

Table 10: Summary of Key Studies

Study	Subsample	Dataset	IV	Dependent Variable
Greenhalgh-Stanley et al (2013) [8]	≥ 60 years-old $3,000 and $5,000 asset limit monthly income ≤ PL+ $500	Health and Retirement Study (HRS) 2000-2008 recertification period	Outreach, CAPs, EBT implementation,	SNAP participation Food Insecurity
Shaefer and Gutierrez (2012) [9]	HH w/ at least one adult and one child, <150 PL	SIPP 1996,2001,2004	short recertification period	FSP participation Food Insecurity
Ratcliffe et al (2010) [10]	HH w/ <150PL, $4000 or $5000 asset limit	SIPP 1996,2001,2004	biometric technology outreach spending partial/full immigrant eligibility	FSP participation Food Insecurity
Kabbani and Yazbeck (2004) [11]	< 185 PL	April 1995,97,99,01 CPS Food Security Supplements	short recertification monthly reporting EBT implementation outreach spending	FSP participation Food Insecurity
Gregory et al (2013) [12]	HH with 200 PL at least 19	2003-2008 NHANES	use of BBCE, exempt one vehicle from asset test	HEI, macro-nutrients intake
Gregory et al (2016) [13]	HH with <130 PL	2009-2011 CPS-FSS	citizenship, certification interval	Food Insecurity
Deb and Gregory (2016) [14]	HH with <185 PL	2006-2013 Dec. CPS -FSS	Outreach spending	Food Spending Food Insecurity
Baum (2007) [15]	Individuals with no more than a HS education	NLSY79 NLSY79	value of vehicles elderly	Obesity

Table 11: IV Estimates of the Effect of SNAP on Households' Low Food Security Status

	(1)	(2)	(3)	(4)	(5)	(6)
	Sample A		Sample B		Sample C	
SNAPNOWREPORT	0.058		-0.088		-0.014	
	(0.188)		(0.160)		(0.298)	
SNAPNOWHH		0.040		-0.094		0.023
		(0.182)		(0.158)		(0.334)

Note: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Coefficients estimates on SNAP participation variables for equation (7) are reported. All regressions control for race, age, sex, marital status, education, health and work status of the primary respondent, and housing status and income of the family.

Table 12: OLS Estimates of the Effect of SNAP on Food Outcomes

		HEI-2010	Vegetables	Beans	Fruit	Whole Fruit	Whole Grain	Dairy	Protein
SNAPNOWREPORT	sample A	-0.702	-0.063	-0.009	-0.124	-0.089	-0.095	0.116	-0.081
		(0.798)	(0.097)	(0.144)	(0.145)	(0.190)	(0.323)	(0.269)	(0.103)
	sample B	-0.762	0.014	-0.04	-0.175	-0.211	-0.232	0.135	-0.03
		(0.755)	(0.094)	(0.159)	(0.135)	(0.196)	(0.314)	(0.272)	(0.101)
	sample C	0.005	0.046	-0.034	0.13	0.111	0.106	0.028	-0.116
		(0.737)	(0.130)	(0.208)	(0.139)	(0.150)	(0.286)	(0.366)	(0.142)
SNAPNOWHH	sample A	-1.095	-0.072	-0.012	-0.265*	-0.226	-0.172	0.1	-0.026
		(0.880)	(0.120)	(0.143)	(0.148)	(0.197)	(0.330)	(0.271)	(0.101)
	sample B	-1.211	0.041	-0.051	-0.296**	-0.31	-0.332	0.131	0.008
		(0.908)	(0.114)	(0.155)	(0.140)	(0.217)	(0.318)	(0.290)	(0.100)
	sample C	-0.3	0.121	-0.003	0.066	0.079	0.018	0.215	-0.073
		(0.901)	(0.136)	(0.200)	(0.152)	(0.163)	(0.295)	(0.436)	(0.136)
		Seafood	Fatty Acid	Sodium	Refined Grain	SOFAAS	FAFH	FAH	Tot. Expenditure
SNAPNOWREPORT	sample A	-0.077	-0.194	-0.215	0.203	-0.073	-8.479	26.080***	21.674*
		(0.134)	(0.270)	(0.317)	(0.281)	(0.467)	(5.392)	(8.397)	(11.229)
	sample B	-0.176	-0.095	-0.064	0.189	-0.077	-8.624	39.669***	35.046***
		(0.143)	(0.300)	(0.299)	(0.338)	(0.404)	(6.316)	(8.969)	(10.715)
	sample C	-0.377**	-0.173	-0.16	0.252	0.191	-8.656	33.792***	28.404**
		(0.165)	(0.396)	(0.278)	(0.415)	(0.506)	(9.228)	(8.472)	(13.211)
SNAPNOWHH	sample A	-(0.056)	-0.05	-0.257	0.063	-0.124	-6.467	25.962***	21.572*
		(0.144)	(0.285)	(0.325)	(0.292)	(0.497)	(5.606)	(8.488)	(11.689)
	sample B	-(0.168)	0.031	-0.131	-0.032	-0.102	-6.647	36.987***	31.698***
		(0.138)	(0.330)	(0.309)	(0.331)	(0.433)	(6.661)	(8.224)	(10.789)
	sample C	-(0.244)	-0.213	-0.454	-0.017	0.205	-7.829	33.689***	26.545*
		(0.154)	(0.455)	(0.275)	(0.436)	(0.586)	(9.230)	(9.308)	(14.437)

Note: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Coefficients estimates on SNAP participation variables for equation (6) are reported. All regressions control for race, age, sex, marital status, education, health and work status of the primary respondent, and housing status and income of the family.

Table 13: IV Estimates of the Effect of SNAP on Food Outcomes

		HEI-2010	Vegetables	Beans	Fruit	Whole Fruit	Whole Grain	Dairy	Protein
SNAPNOWREPORT	sample A	-21.414***	-1.544*	-3.767*	-3.243***	-5.137***	-1.893	-1.191	0.299
		(7.808)	(0.905)	(1.930)	(1.009)	(1.450)	(1.486)	(1.814)	(0.596)
	sample B	-23.377***	-2.062**	-3.626**	-3.565***	-4.476**	-0.549	-1.091	0.357
		(8.175)	(0.841)	(1.602)	(0.978)	(1.074)	(1.815)	(1.661)	(0.509)
	sample C	-16.304	-2.737**	-2.355	-1.143	-0.726	2.216	3.55	-1.189
		(11.504)	(1.149)	(2.002)	(0.878)	(1.077)	(2.214)	(2.248)	(0.823)
SNAPNOWHH	sample A	-21.691**	-1.546	-3.769*	-3.349**	-5.215***	-1.981	-1.231	0.318
		(8.221)	(0.942)	(2.077)	(1.142)	(1.498)	(1.329)	(1.738)	(0.589)
	sample B	-23.189***	-1.999**	-3.627**	-3.626***	-4.627	-0.789	-1.201	0.413
		(8.354)	(0.884)	(1.677)	(1.116)	(1.051)	(1.667)	(1.626)	(0.500)
	sample C	18.532	-2.905**	-2.447	-1.407	-0.976	2.018	3.804	-1.169
		(12.911)	(1.273)	(2.110)	(0.990)	(1.159)	(2.627)	(2.390)	(0.871)
		Seafood	Fatty Acid	Sodium	Refined Grain	SOFAAS	FAFH	FAH	Tot. Expenditure
SNAPNOWREPORT	sample A	-2.419**	0.073	-1.734	3.72	-4.577	-60.185	-28.703	-60.455
		(1.096)	(1.992)	(2.237)	(2.900)	(3.738)	(35.806)	(52.291)	(64.601)
	sample B	-1.731*	-0.238	-3.126	2.635	-5.906*	-35.404	-32.685	-56.986
		(0.963)	(1.914)	(2.407)	(2.507)	(3.446)	(22.777)	(64.809)	(72.522)
	sample C	-3.454*	-4.221	-3.005	2.744	-5.984	-67.827*	-54.403	-99.695
		(1.896)	(2.571)	(2.941)	(3.800)	(5.639)	(36.250)	(95.393)	(112.753)
SNAPNOWHH	sample A	-2.394**	0.216	-1.773	3.696	-4.663	-57.494	-33.755	-65.372
		(1.081)	(1.850)	(2.351)	(2.998)	(3.787)	(38.798)	(56.712)	(74.116)
	sample B	-1.751*	0.017	-3.054	2.663	-5.607	-36.366	-35.308	-60.792
		(0.952)	(1.766)	(2.523)	(2.547)	(3.534)	(26.318)	(67.373)	(82.522)
	sample C	-3.944*	-4.617	-3.731	2.556	-5.714	-77.635*	-66.644	-117.955
		(1.976)	(2.920)	(3.413)	(4.478)	(6.919)	(42.031)	(110.363)	(132.834)

Note: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Coefficients estimates on SNAP participation variables for equation (7) are reported. F-statistics for test of weak instrument is 7.750 for sample A, 10.90 for sample B, and 9.583 for sample C.

Table 14: OLS Estimates of the Effect of SNAP Benefit amounts on Food Outcomes

		HEI-2010	Vegetables	Beans	Fruit	Whole Fruit	Whole Grain	Dairy	Protein
Amount,	sample A	-2.732	-0.134	-0.269	-0.133	-0.191	0.101	0.359	-0.060
Reported		(1.709)	(0.224)	(0.275)	(0.373)	(0.471)	(0.626)	(0.445)	(0.184)
	sample B	-2.104	0.074	-0.422	-0.175	-0.422	-0.179	0.208	0.091
		(1.535)	(0.184)	(0.308)	(0.326)	(0.404)	(0.577)	(0.453)	(0.167)
	sample C	-0.219	0.258	-0.192	0.501	0.224	0.091	0.305	-0.017
		(1.654)	(0.256)	(0.335)	(0.333)	(0.453)	(0.442)	(0.685)	(0.229)
Amount,	sample A	-0.182	-0.294	-0.533	-0.480	-0.653*	-0.038	0.464	-0.001
Matched		(1.850)	(0.189)	(0.322)	(0.311)	(0.365)	(0.615)	(0.487)	(0.165)
	sample B	-0.096	-0.049	-0.571*	-0.355	-0.682*	-0.133	0.431	-0.014
		(1.543)	(0.177)	(0.322)	(0.278)	(0.348)	(0.612)	(0.460)	(0.177)
	sample C	1.796	0.044	-0.538	0.233	-0.097	0.013	0.947	-0.116
		(1.479)	(0.225)	(0.397)	(0.311)	(0.404)	(0.434)	(0.789)	(0.290)
		Seafood	Fatty Acid	Sodium	Refined Grain	SOFAAS	FAFH	FAH	Tot. Expenditure
Amount,	sample A	-0.292	-0.127	-0.318	1.160	-0.118	-28.302**	87.727***	68.967**
Reported		(0.269)	(0.595)	(0.642)	(0.692)	(0.958)	(11.359)	(22.590)	(27.543)
	sample B	-0.389	0.059	-0.174	1.132	0.226	-32.369**	109.806***	87.286**
		(0.313)	(0.590)	(0.556)	(0.773)	(0.956)	(12.504)	(28.669)	(33.301)
	sample C	-0.759*	-0.129	-0.257	0.873	1.112	-30.628*	87.917**	63.497
		(0.373)	(0.926)	(0.688)	(0.852)	(1.176)	(17.223)	(32.284)	(40.335)
Amount,	sample A	-0.110	0.186	-0.922	0.742	-1.254	-23.671**	99.198***	80.812***
Matched		(0.277)	(0.483)	(0.631)	(0.782)	(0.849)	(9.694)	(27.797)	(27.489)
	sample B	-0.295	0.141	-0.671	0.800	-0.830	-27.812**	105.820***	83.902***
		(0.295)	(0.469)	(0.631)	(0.798)	(0.784)	(10.608)	(28.002)	(26.975)
	sample C	-0.500	-0.452	-0.828	0.329	0.531	-33.955**	87.828**	54.989
		(0.330)	(0.902)	(0.727)	(0.784)	(0.930)	(15.608)	(35.428)	(32.364)

Note: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Coefficients estimates on SNAP participation variables for equation (6) are reported. All regressions control for race, age, sex, marital status, education, health and work status of the primary respondent, and housing status and income of the family.

Table 15: IV Estimates of the Effect of SNAP Benefit amounts on Food Outcomes

		HEI-2010	Vegetables	Beans	Fruit	Whole Fruit	Whole Grain	Dairy	Protein
Amount,	sample A	-52.772***	-3.119	-8.390**	-8.604***	-13.158***	-5.623	-2.807	1.047
Reported		(15.731)	(1.934)	(3.909)	(2.343)	(3.323)	(3.526)	(4.255)	(1.395)
	sample B	-58.864***	-4.485**	-8.951**	-9.639***	-12.783***	-4.025	-3.820	1.620
		(17.570)	(1.972)	(3.424)	(2.553)	(2.914)	(3.949)	(4.383)	(1.307)
	sample C	-58.980**	-7.547**	-8.336	-6.732***	-7.675***	-0.962	5.094	1.642
		(24.637)	(3.427)	(4.873)	(2.375)	(2.247)	(5.170)	(5.834)	(2.438)
Amount,	sample A	-44.647***	-2.580	-8.575***	-8.813***	-11.930***	-5.598	-3.133	2.337
Matched		(14.853)	(1.823)	(2.904)	(2.161)	(2.817)	(3.663)	(4.675)	(1.797)
	sample B	-39.098***	-3.066**	-7.038***	-7.313***	-9.139***	-3.144	-1.987	1.414
		(11.885)	(1.466)	(2.264)	(1.508)	(1.923)	(3.333)	(3.736)	(1.254)
	sample C	-40.963**	-6.561**	-7.640**	-4.433**	-4.959**	2.917	8.334*	0.589
		(18.486)	(2.627)	(2.943)	(1.680)	(1.807)	(3.728)	(4.560)	(2.052)
		Seafood	Fatty Acid	Sodium	Refined Grain	SOFAAS	FAFH	FAH	Tot. Expenditure
Amount,	sample A	-5.782**	1.115	-4.468	6.893	-9.876	-142.607*	-61.429	-141.292
Reported		(2.390)	(4.472)	(4.933)	(6.115)	(8.669)	(78.630)	(113.749)	(147.141)
	sample B	-4.362*	1.105	-6.300	5.892	-13.116	-98.327	-30.406	-103.290
		(2.226)	(4.408)	(5.419)	(5.637)	(8.067)	(63.932)	(129.730)	(161.264)
	sample C	-9.534**	-7.760	-11.134	7.795	-13.832	-174.741	-67.464	-172.865
		(4.237)	(6.007)	(9.067)	(10.334)	(17.709)	(117.628)	(223.986)	(306.610)
Amount,	sample A	-2.604	4.832	-2.049	8.526	-15.058*	-98.559	-2.728	-70.880
Matched		(2.404)	(4.761)	(5.080)	(5.273)	(7.954)	(68.378)	(111.015)	(139.050)
	sample B	-2.171	2.637	-3.500	6.760	-12.551*	-71.789	-20.070	-72.936
		(1.873)	(3.912)	(4.682)	(4.592)	(6.360)	(47.768)	(109.662)	(126.689)
	sample C	-6.326*	-5.379	-7.643	12.039	-21.900	-93.821	-6.039	-43.255
		(3.359)	(4.106)	(6.370)	(7.547)	(13.273)	(71.460)	(157.019)	(195.595)

Note: $*p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$. Coefficients estimates on SNAP participation variables for equation (7) are reported. F-statistics for test of weak instrument is 7.750 for sample A, 10.90 for sample B, and 9.583 for sample C.

Table 16: OLS Estimates of the Effect of SBP

	(1) Overweight	(2) Obese	(3) BMI	(4) Overweight	(5) Obese	(6) BMI	(7) Overweight	(8) Obese	(9) BMI
Sample A	-0.037	-0.056	-0.418	-0.030	-0.036	0.007	-0.031	-0.038	-0.015
	(0.039)	(0.042)	(0.578)	(0.038)	(0.043)	(0.651)	(0.038)	(0.043)	(0.661)
Sample B	0.005	-0.057	-0.379	0.002	-0.037	-0.156	0.001	-0.040	-0.184
	(0.037)	(0.057)	(0.756)	(0.039)	(0.057)	(0.791)	(0.039)	(0.058)	(0.802)
Sample C	0.028	0.040	0.291	0.023	0.093	0.835	0.020	0.089	0.768
	(0.049)	(0.073)	(1.053)	(0.054)	(0.072)	(1.061)	(0.055)	(0.074)	(1.089)
SNAPNOWHH	N	N	N	Y	Y	Y	N	N	N
SNAPNOWREPORT	N	N	N	N	N	N	Y	Y	Y

Note: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. OLS coefficients estimates on SBP variables are reported. Overweight and Obese are binary indicators equal to 1 if a person is categorized as being overweight/obese and 0 otherwise. BMI is a continous measure. All regressions control for race, age, sex, marital status, education, health and work status of the primary respondent, age of child, and housing status and income of the family.

Table 17: OLS Estimates of the Effect of NSLP

	(1) Overweight	(2) Obese	(3) BMI	(4) Overweight	(5) Obese	(6) BMI	(7) Overweight	(8) Obese	(9) BMI
Sample A	-0.003	-0.016	-0.594	0.030	-0.020	-0.540	0.025	-0.031	-0.652
	(0.031)	(0.033)	(0.452)	(0.031)	(0.033)	(0.509)	(0.031)	(0.034)	(0.495)
Sample B	-0.003	-0.006	-0.551	0.026	-0.015	-0.484	0.021	-0.027	-0.591
	(0.031)	(0.036)	(0.444)	(0.032)	(0.033)	(0.452)	(0.032)	(0.034)	(0.459)
Sample C	0.008	-0.070	-1.251**	0.031	-0.081*	-1.172	0.024	-0.092*	-1.325*
	(0.037)	(0.041)	(0.586)	(0.029)	(0.043)	(0.767)	(0.029)	(0.046)	(0.747)
SNAPNOWHH	N	N	N	Y	Y	Y	N	N	N
SNAPNOWREPORT	N	N	N	N	N	N	Y	Y	Y

Note: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. OLS coefficients estimates on NSLP variables are reported. Overweight and Obese are binary indicators equal to 1 if a person is categorized as being overweight/obese and 0 otherwise. BMI is a continous measure. All regressions control for race, age, sex, marital status, education, health and work status of the primary respondent, age of child, and housing status and income of the family.

Table 18: OLS and IV, Substitutability of SNAP and School Nutrition Programs

	SBP		NLSP	
OLS	SNAPNOWREPORT	SNAPNOWHH	SNAPNOWREPORT	SNAPNOWHH
Group A	0.043	0.060***	0.226***	0.223***
	(0.029)	(0.020)	(0.031)	(0.025)
Group B	0.049	0.068***	0.230***	0.194***
	(0.030)	(0.021)	(0.034)	(0.028)
Group C	0.05	0.047	0.185***	0.204***
	(0.044)	(0.037)	(0.039)	(0.045)
	SBP		NLSP	
IV	SNAPNOWREPORT	SNAPNOWHH	SNAPNOWREPORT	SNAPNOWHH
Group A	0.279*	0.215*	0.501**	0.423**
	(0.142)	(0.126)	(0.202)	(0.171)
Group B	0.266**	0.242**	0.348**	0.355***
	(0.106)	(0.091)	(0.129)	(0.121)
Group C	0.155	0.126	0.073	0.093
	(0.137)	(0.123)	(0.206)	(0.228)

Note: *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. OLS and IV coefficients estimates on SBP variables are reported. Overweight and Obese are binary indicators equal to 1 if a person is categorized as being overweight/obese and 0 otherwise. BMI is a continous measure. All regressions control for race, age, sex, marital status, education, health and work status of the primary respondent, age of child, and housing status and income of the family.

Appendix

Let y be a continuous dependent variable, P^* be true SNAP participation, and P^r be reported SNAP participation. Let the measurement errors be $\pi_{01} = \Pr(P^r = 1|P^* = 0) =$ probability of a false positive and $\pi_{10} = \Pr(P^r = 0|P^* = 1) =$ probability of a false negative. Assume the linear model

$$y_i = \alpha + \beta P^* + \varepsilon_i \tag{9}$$

1. Case 1: P^* exogenous and measurement error independent of ε. The OLS estimate of β is

$$\Delta = E(y|P^r = 1) - E(y|P^r = 0) \tag{10}$$

and hence

$$
\begin{aligned}
E(y|P^r &= 1) = \alpha + \beta E(P^*|P^r = 1) && (11)\\
&= \alpha + \beta \Pr(P^* = 1|P^r = 1) && (12)\\
&= \alpha + \beta[1 - \Pr(P^* = 0|P^r = 1)] && (13)\\
&= \alpha + \beta[1 - \frac{\Pr(P^* = 0, P^r = 1)}{\Pr(P^r = 1)}] && (14)\\
&= \alpha + \beta[1 - \frac{\Pr(P^r = 1|P^* = 0)\Pr(P^* = 0)}{\Pr(P^r = 1)}] && (15)\\
&= \alpha + \beta[1 - \frac{\pi_{01}(1 - q^*)}{q^r}] && (16)
\end{aligned}
$$

where $q^* = \Pr(P^* = 1)$ and $q^r = \Pr(P^r = 1)$. Following a similar derivation,

$$E(y|P^r = 0) = \alpha + \beta[\frac{\pi_{10}q^*}{1 - q^r}] \tag{17}$$

Therefore

$$\Delta = \beta[1 - \frac{\pi_{01}(1 - q^*)}{q^r} - \frac{\pi_{10}q^*}{1 - q^r}] \tag{18}$$

Now, since $q^r = \Pr(P^r = 1) = \pi_{01}(1 - q^*) + (1 - \pi_{10})q^*$, the two ratios in this expression must be positive and less than 1. Therefore OLS is biased toward zero.

2. Case 2: P^* exogenous but measurement error is correlated with ε.First, continue to let π_{01} and π_{10} be the overall false positive and false negative error rates, but recognizing that they are determined by the integral of individual error rates (which are a function of ε) over the distribution of ε. Second, the OLS estimator of β is now

$$\Delta = E(y|P^r = 1) - E(y|P^r = 0) + E(\varepsilon|P^r = 1) - E(\varepsilon|P^r = 0) \tag{19}$$

While $E(\varepsilon|P^*) = 0$, this is not the case for $E(\varepsilon|P^r)$ if measurement error is correlated with ε. The first part of the expression, $E(y|P^r = 1) - E(y|P^r = 0)$, is the same as before and equals the expression in (XX). But the sign of the second part depends on the sign of the correlation of ε with the occurrence of false positives and false negatives. Since

$$E(\varepsilon|P^r = 1) = E(\varepsilon|P^r = 1, P^* = 0)(1 - q^*) + E(\varepsilon|P^r = 1, P^* = 1)q^* \tag{20}$$

$$E(\varepsilon|P^r = 0) = E(\varepsilon|P^r = 0, P^* = 0)(1 - q^*) + E(\varepsilon|P^r = 0, P^* = 1)q^* \tag{21}$$

the second bias expression in (XX) above is

$$E(\varepsilon|P^r = 1) - E(\varepsilon|P^r = 0) = q^*[E(\varepsilon|P^r = 1, P^* = 1) - E(\varepsilon|P^r = 0, P^* = 1] \tag{22}$$

$$+(1 - q^*)[E(\varepsilon|P^r = 1, P^* = 0) - E(\varepsilon|P^r = 0, P^* = 0]$$

The first part of this expression represents the difference in y values among those on SNAP who do and do not report it correctly (because of false negative reports) and the second half represents the difference in y among those off SNAP who do and do not report accurately (because of false positive reports). This expression could be either positive or negative. Hence the direction of bias in Δ is indeterminant.

Case 3. P^* endogenous, measurement error independent of ε. Assume we have an instrument z which is correlated with P^* but independent of ε and also independent of measurement error. To show the bias in the IV estimator of β, we shall express the IV estimator as the ratio of the coefficients on z in reduced form regressions of y and P on z (alternatively,one could express it in 2SLS form). The true equation is still (XX). The reduced forms can be denoted as $E(y|z)$ and $E(P|z)$. We have

$$E(y|z) = \alpha + \beta E(P^*|z) = \alpha + \beta q^*(z) \tag{23}$$

where $q^*(z) = \Pr(P^* = 1|z)$. We also have

$$\begin{aligned} E(P^r|z) &= \Pr(P^r = 1|z) = \Pr(P^r = 1|P^* = 1, z)q^*(z) + \Pr(P^r = 1|P^* = 0.z)(1 - q^*(z)) && (24) \\ &= (1 - \pi_{10})q^*(z) + \pi_{01}(1 - q^*(z)) && (25) \\ &= \pi_{01} + q^*(z)(1 - \pi_{10} - \pi_{01}) && (26) \end{aligned}$$

The coefficients on z in these reduced form equations are those on $\partial q^*(z)/\partial z$ and their ratio is

$$\frac{\beta}{1 - \pi_{10} - \pi_{01}} \tag{27}$$

Since the denominator of this expression is less than 1, the IV estimator of β is biased away from zero.

Case 4. P^* endogenous, measurement error correlated with ε. This case is the same as the previous case if z is still independent of ε and of measurement error. The error term ε does not enter because it is independent of z.