WHAT CAN WE LEARN ABOUT THE EFFECTS OF FOOD STAMPS ON OBESITY
IN THE PRESENCE OF MISREPORTING?

Lorenzo Almada
Ian M. McCarthy
Rusty Tchernis

What Can We Learn About the Effects of Food Stamps on Obesity in the Presence of Misreporting?
Lorenzo Almada, Ian M. McCarthy, and Rusty Tchernis
NBER Working Paper No. 21596
September 2015
JEL No. C01,H4,I1,I28

ABSTRACT

There is an increasing perception among policy makers that food stamp benefits contribute positively to adult obesity rates. We show that these results are heavily dependent on one's assumptions regarding the accuracy of reported food stamp participation. When allowing for misreporting, we ″find no evidence that SNAP participation signifi″cantly increases the probability of being obese or overweight among adults. Our results also highlight the inherent bias and inconsistency of common point estimates when ignoring misreporting, with treatment e˙ffects from instrumental variable methods exceeding the non-parametric upper bounds by over 200% in some cases.

Lorenzo Almada
School of Social Work
1255 Amsterdam Ave
Room 718
la2587@columbia.edu

Ian M. McCarthy
Emory University
Department of Economics
1602 Fishburne Dr
Rich Memorial Building, Room 306
Atlanta, GA 30322
ianmccarthy.econ@gmail.com

Rusty Tchernis
Department of Economics
Andrew Young School of Policy Studies
Georgia State University
P.O. Box 3992
Atlanta, GA 30302-3992
and NBER
rtchernis@gsu.edu

The obesity rate among U.S. adults has more than doubled over the past few decades, from around 15% in the 70s to over 35% by 2010 (Ogden *et al.*, 2014). Obesity has a number of detrimental effects on a person's overall health and on the general health care system (Rosin, 2008; Finkelstein *et al.*, 2009; Cawley & Meyerhoefer, 2012), as well as decreased labor productivity and wage growth (Bhattacharya & Bundorf, 2009). The alarming obesity trend and associated health problems have spurred a flurry of economic research, much of which finds that obesity rates are more prevalent among lower income households (Rosin, 2008).

Subsequently, since lower income populations are more likely to rely on federal nutrition assistance, the relationship between obesity and federal nutrition assistance programs is of obvious interest to policy makers. The largest of these nutrition programs is the Supplemental Nutrition Assistance Program (SNAP), formerly known as the Food Stamp Program. In 2000, approximately 17.2 million individuals participated in SNAP at a total cost of approximately $23 billion (in 2014 dollars). By 2014, this figure increased to over 46.5 million individuals for a total cost estimated at nearly $74 billion.[1]

SNAP's dramatic growth and its role in determining diets among lower income households necessitates a thorough understanding of the relationship between SNAP participation and obesity, which several authors have analyzed. Although the literature in this area is mixed, a common finding is that SNAP participation tends to increase the probability of being obese or overweight among adult women with no significant effects for adult men (Townsend *et al.*, 2001; Gibson, 2003; Chen *et al.*, 2005; Kaushal, 2007; Ver Ploeg & Ralston, 2008; Meyerhoefer & Pylypchuk, 2008; Baum, 2011).[2]

---

[1]Data available at http://www.fns.usda.gov/pd/supplemental-nutrition-assistance-program-snap.

[2]This result does not persist among children. Here, the literature generally finds a reduction in childhood obesity following SNAP participation, but not necessarily for other nutrition assistance programs such as the National School Lunch Program (Schanzenbach, 2009; Schmeiser, 2012; Kimbro & Rigby, 2010; Burgstahler *et al.*, 2012). See Gundersen (2015) for a recent review of the literature both for adults and children.

Recently, however, authors have also been concerned with misclassification of SNAP participation, with misreporting rates of over 30% in some studies (Bollinger & David, 1997; Meyer *et al.*, 2009; Vassilopoulos *et al.*, 2011; Kreider *et al.*, 2012). Yet despite the growing awareness of misreporting of SNAP benefits, few studies have re-considered the relationship between SNAP participation and adult weight outcomes when allowing for such misclassification.

In this paper, we revisit the question of whether SNAP participation affects adult obesity when accounting for misreported participation status. We pursue a range of alternative procedures for estimating average treatment effects (ATE), each with different identification assumptions, including point estimation and nonparametric bounds following Manski & Pepper (2000), Kreider & Pepper (2007), Kreider *et al.* (2012), and Gundersen *et al.* (2012). The contribution of the paper is two-fold. First, we advance the policy discussion surrounding the effects of food stamps on adult weight outcomes, taking more seriously the impact of endogenous and misreported SNAP participation. Given the growing expenditure on SNAP and the broader policy debate surrounding the benefits of federal nutrition assistant programs, accurately estimating these effects is critical for current policy decisions. Second, our comparison of empirical methods reveals important practical differences in estimates when confronted with misreported treatment.

Based on data from the restricted-use National Longitudinal Survey of Youth - 1979 Cohort (NLSY79), we estimate that respondents under-report SNAP participation by between 30 and 50%. Standard IV estimates that fail to account for misreporting will then overestimate (in absolute value) the true effect by a factor of between 2 and 3.3.[3] Our results suggest that average treatment effects from a fixed effects IV model are similarly overestimated, with treatment effects exceeding the nonparametric upper bounds by more than 200% under reasonable assumptions on the selection

---

[3]Intuitively, the magnitude of overestimation follows directly from the linear IV estimator. This estimator can be written as a ratio of the second and first-stage estimators, where Hausman *et al.* (1998) finds that the first-stage estimate will be inconsistent by a factor of 1 minus the misreporting rates. Similar calculations are provided in Frazis & Loewenstein (2003).

process and misreporting rates. When accounting for misreporting, we find no evidence that SNAP participation significantly increases the probability of being obese or overweight (either overall or among men and women separately). These results therefore call into question many of the existing findings of a positive effect of SNAP participation on adult weight outcomes among women, some of which are based on our same NLSY79 data.

Literature Review

In addition to the growing empirical literature on the effects of misreporting, there is also a large literature concerning the role of nutrition assistance programs and health outcomes. In this section, we briefly discuss relevant aspects of these fields. We first review some of the existing literature on SNAP participation and adult weight outcomes; we then discuss the extent to which respondents may misreport their SNAP participation; and we conclude this section with a brief discussion of the available methods for treatment effects estimation with misreported treatment status.

SNAP Participation and Adult Weight Outcomes

There are several theoretical arguments as to why SNAP participation and obesity may be related, stemming largely from one of three general concepts: 1) the Food Stamp Cycle; 2) the propensity for lower-income households to purchase less healthy foods; and 3) the in-kind nature of SNAP benefits. The first hypothesis, the Food Stamp Cycle, describes a behavior whereby SNAP participants exhaust their benefits non-uniformly over the course of the month, with more benefits redeemed at the beginning of the month than the end of the month or vice versa. This results in periods of food surpluses and potential overeating followed by periods of food shortages and undereating, a pattern in which both children and adults tend to gain weight (Blackburn *et al.*, 1989; Polivy *et al.*, 1994; Dietz, 1995; Fisher & Birch, 1999). In research specifically related to the SNAP program, Wilde & Ranney (2000) and Hastings & Washington (2010) find evidence of such chronic dieting behavior

among SNAP participants.

A second theoretical explanation derives from the type of foods purchased by SNAP participants. In particular, lower income households are known to rely on diets relatively higher in calories and fat compared to healthier, more expensive fruits and vegetables (Dowler, 2003). Cole & Fox (2008) reached similar conclusions, finding that SNAP participants' diets contained more solid fats, alcoholic beverages, and sugars compared to SNAP-eligible non-participants and higher income individuals. However, there is little empirical evidence that SNAP benefits actually *cause* participants to purchase less healthy foods.

A third hypothesis linking SNAP participation and obesity involves the in-kind nature of SNAP benefits. Since SNAP benefits must be used on food, households with an excess allowance may either allow their benefits to go unused or spend more on food than they otherwise would. For example, Fox *et al.* (2004) estimated that households increase food expenditure by up to \$0.47 on average for every dollar of SNAP benefits received. Prior studies also found that SNAP recipients spent more on food than they would following changes in cash income (Devaney & Moffitt, 1991; Fraker *et al.*, 1995), although recent studies have found no such differences (Hoynes & Schanzenbach, 2009). To the extent that SNAP participants do spend more on food, this increased spending may be positively or negatively related to obesity, depending on the types of foods purchased. If households are already purchasing relatively unhealthy foods, then the increased spending would tend to increase obesity if individuals simply purchase more of the same types of foods. If, however, SNAP benefits lead households to purchase healthier (more expensive) foods, then SNAP participation would tend to decrease obesity. SNAP benefits may also allow participants to reduce work hours and spend more time preparing healthy meals.

Ultimately, the relationship between SNAP participation and obesity is theoretically ambiguous, with some mechanisms suggesting a positive relationship and others suggesting a negative relationship. The question of whether SNAP participation increases or decreases obesity is therefore an empirical question. To this end, Gibson (2003) and Baum (2011) examined the relationship between

SNAP participation and obesity with fixed effects models using data from the NLSY79, finding a 2-3 percentage point increase in the probability of being obese following SNAP participation for adult women. Adopting a difference-in-difference approach with propensity score matching, Fan (2010) also estimated the effects of SNAP participation on adult obesity using data from the NLSY79; however, he found no statistically significant effects of SNAP participation on BMI or probability of being overweight or obese. Almada & Tchernis (2015) reach a similar conclusion focusing on the intensive margin, based on the estimated effect of an increase in SNAP benefits among existing SNAP participants. Using data from the Medical Expenditure Panel Survey (MEPS), Meyerhoefer & Pylypchuk (2008) adopt an instrumental variables approach using state expenditures on food stamp outreach programs as their instrument for women, and for men their instruments are whether the state uses fingerprint scanning and whether the state requires 50% or more of food stamp participants to recertify their program eligibility within three months. The authors find that SNAP participation increases the probability of being obese by 6.7% among adult women with no statistically significant effects on adult men. As such, although the empirical results remain somewhat mixed, the general conclusion presented to policy makers is that SNAP participation tends to increase the probability of obesity among adult women with no significant effects among adult men (Ver Ploeg & Ralston, 2008).

Misreporting of SNAP Participation

As research on the empirical relationship between obesity and SNAP grows, so too grow concerns over the extent and effects of misclassification of SNAP participation. There are essentially two forms of misclassification: 1) false positives, which occur when individuals report participating in the program but in fact do not participate; and 2) false negatives, which result from individuals' failure to report their participation when in fact they do participate. When ignored, the presence of misreporting is known to substantially bias one's results (Aigner, 1973; Bollinger & David, 1997; Hausman *et al.*, 1998; Black *et al.*, 2000; Lewbel, 2007; Frazis & Loewenstein, 2003; Brachet, 2008;

Millimet *et al.*, 2010; Kreider *et al.*, 2012).[4]

Traditionally, false negatives are more prevalent than false positives in surveys of SNAP partic-ipation. Using data from 1984 Survey of Income and Program Participation (SIPP), Bollinger & David (1997) and Marquis & Moore (2010) find under-reporting of 12% at household level and 22% at individual level, respectively. More recent studies find increasingly prevalent under-reporting rates in several datasets, with a general theme that under-reporting is most prevalent in larger, national datasets and has increased over time (Meyer *et al.*, 2009). For example, Cody & Tuttle (2002) find that under-reporting rates in the Current Population Survey (CPS) increased from 21% in 1991 to 36% in 1999. More recently, Meyer *et al.* (2011) estimate under-reporting of 35% in the 2001 American Community Survey (ACS) and 50% in the 2002-2005 CPS. The authors also find that misreporting in general appears to be least problematic in the SIPP.

Estimates of over-reporting in food stamp data are often less than 1% (Bollinger & David, 1997). However, this varies across datasets and may also depend on the specific form of question asked. For example, Goerge *et al.* (2009) reports false positive rates as high as 18% among food stamp eligible households around Chicago. The authors also indicate that the time frame covered by the survey question may influence under- and over-reporting rates. For example, the question of "do you currently receive food stamp benefits," as posed in the SIPP, may be an easier question to accurately answer than "have you received food stamps in the previous 12 months."

Authors typically estimate misreporting either with aggregate or linked administrative data, or functional form assumptions. For example, Meyer *et al.* (2011) compare self-reported SNAP

---

[4]This bias results even in the case of purely random misreporting. Moreover, even small amounts of misreporting can significantly increase the uncertainty associated with point estimates. For ex-ample, Kreider (2010) finds that even with less than 2% misreporting, the confidence intervals based on the misreported data may not overlap with the confidence intervals from the true data. See Chen *et al.* (2011) for a recent review of the broader literature on nonlinear errors in variables models, and see Millimet (2011) for a review of measurement error in treatment effects models.

participation from the ACS and CPS to linked administrative data on SNAP recipients. In our case, we do not have access to any relevant supplemental data, and as such we estimate misreporting using functional form assumptions as in Hausman *et al.* (1998). Our nonparametric bounds analysis also considers a range of misreporting rates, the qualitative findings from which are not generally dependent on any specific misreporting estimates.

Existing Methods with Misreported Treatment

Although estimated misreporting rates clearly differ across datasets, the common theme is that misreporting is prevalent in surveys of SNAP participation and that, without accounting for such misreporting, standard point estimates may be substantially biased and inconsistent. Estimating coefficients with a mismeasured binary regressor is the subject of a relatively small but growing literature. Mahajan (2006) considered a nonparametric model with a mismeasured binary regressor. Under the appropriate assumptions, he shows that the conditional mean and marginal effects are identified and derives nonparametric and parametric estimators. Hu & Schennach (2008) propose an alternative estimation procedure that also allows for a more general discrete (not necessarily binary) mismeasured regressor. Lewbel (2007) further considers the estimation of average treatment effects with mismeasured treatment, relaxing some of the conditional independence assumptions but requiring the instrument to take on more unique values relative to Mahajan (2006).

Identification in models such as Mahajan (2006) and Lewbel (2007) is achieved through the presence of an "instrument-like variable" that is correlated with true treatment status but conditionally independent of the measurement error.[5] In the absence of such an instrument, methods for point-identification of treatment effects are sparse. Frazis & Loewenstein (2003) develop several

---

[5]Mahajan (2006) adopts the language "instrument-like variable" in order to distinguish the estimator from traditional IV methods, which are inconsistent in the case of a mismeasured binary regressor due to the inherent correlation between the true participation status and the measurement error (Black *et al.*, 2000; Frazis & Loewenstein, 2003).

moment conditions to allow consistent estimation of treatment effects and misreporting probabilities via GMM; however, their proposed estimator is under-identified when treatment is both endogenous and misreported. Another approach from Brachet (2008) is a two-stage estimator in which misreporting probabilities are estimated in the first stage (following Hausman *et al.* (1998)) and treatment effects are estimated in the second stage. McCarthy & Tchernis (2011) also propose a two-stage approach using Bayesian Markov Chain Monte Carlo methods. Alternatively, we can forgo point estimation and bound the estimated ATE using a range of assumptions on the selection process and extent of misreporting (Manski & Pepper, 2000; Kreider & Pepper, 2007; Kreider *et al.*, 2012; Gundersen *et al.*, 2012). Ultimately, our analysis considers a range of point estimates and bounds, the details of which are discussed in the methods section below.

Data

Our analysis is based on data from the National Longitudinal Survey of Youth - 1979 Cohort (NLSY79). The NLSY79 contains information from a nationally representative sample of 12,686 men and women who were between the ages of 14 and 22 at the time of their first survey in 1979. The dataset is comprised of three subsamples: 1) a cross-sectional sample of 6,111 youths designed to be representative of the non-institutionalized civilian youth population; 2) a supplemental sample of 5,295 youths designed to oversample the civilian Hispanic, black, and economically disadvantaged non-black/non-Hispanic youth population; and 3) a military sample of 1,280 youths. The cohort was interviewed annually through 1994 and subsequently every two years, with nearly 90% retention rates during most of the sample period.

Among other variables, the NLSY79 contains data on household size, number of children in the household, age of the respondent and other members of the household, as well as indicators for race, gender, marital status, and current employment status. The survey also collects information on the total net income of each respondent's household, the respondent's highest level of education,

the highest level of education for the respondent's mother, and an indicator if SNAP benefits were received during the past year. To calculate BMI and create our obesity and overweight indicators, we use measures of self-reported weight and height collected in multiple waves of the survey.[6] The respondent's weight in pounds is reported in the following survey years: 1981, 1982, 1985, 1986, 1988, 1989, 1990, 1992, 1993, 1994, 1996, 1998, 2000, 2002, 2004, 2006, 2008 and 2010. A measure of height in inches for each respondent is reported in 1981, 1982 and 1985. Following the existing literature, we use the 1985 height measure to calculate each respondent's BMI for all years with reported weight. For a very small fraction of respondents, we use 1982 measures of height if height was not reported in 1985.[7] We exclude from the analysis any observations with a missing BMI measure as well as females-year observations that report currently being pregnant.

The restricted-use NLSY79 further provide location identifiers for each respondent, which allows a linkage between state-level SNAP administration policies and respondent-level data. We incorporate state-level SNAP administration data from The Urban Institute's Food Stamp Program Rules Database Documentation (Finegold *et al.*, 2007) which includes data on whether the state uses biometric identification technology (i.e., fingerprint scanning) and the percentage of SNAP benefits issued by direct mail (i.e., coupons mailed to the participant's household). These policy variables are available for all 50 states from 1996 through 2004 and serve as potential instruments for our observed SNAP participation indicator. We discuss our instrument selection in more detail in the results section.

SNAP eligibility is formally restricted to those with gross household incomes at or below 130% of

---

[6]As highlighted in Cawley (2004), Courtemanche *et al.* (2015), and others, self-reported height and weight is subject to nonrandom measurement error. Our initial analysis uses these self-reported height and weight measures as given, without any adjustments; however, we examine the sensitivity of our results to potential misreporting of height and weight in the "Robustness and Sensitivity" subsection of our results.

[7]Our final sample includes only 91 individuals with height measures reported from the 1982 survey wave.

the federal poverty level (FPL) based on household size. However, households with *annual* incomes above the 130% FPL may still have periods of SNAP eligibility or participation due to within-year income fluctuations. As such, some authors have suggested alternative thresholds as high as 250% of the FPL (Mykerezi & Mills, 2010). Here, we focus our analysis on respondents at or below the 250% FPL, and we consider the sensitivity of our results to alternative FPL thresholds in the "Robustness and Sensitivity" subsection. This criterion captures over 99% of all reported SNAP participants in the data. Also, since part of our identification strategy uses state-level variation over time, we further restrict the analysis to respondents who are observed in at least two waves between 1996 through 2004. Our final dataset therefore consists of 3,603 individuals and 12,015 person-year observations with valid measures of height and weight, as summarized in table 1.

Four separate columns divide the 12,015 respondent-year observations into different SNAP participation groups. The first column of table 1 summarizes the characteristics of the respondents who reported participating in SNAP in the past year. The data show that, on average, current SNAP participants earn less income ($18,332 vs. $27,784), tend to have larger households (3.52 vs. 3.37) with more children (1.82 vs. 1.60), are less likely to have graduated high school (72.1% vs. 82.3%) or have mothers who graduated high school (38.7% vs 49%), and less likely to be employed (48.3% vs. 77.2%), married (24.9% vs. 42.2%), or be living in households with an elderly adult (15.9% vs. 21.1%) compared to non-SNAP participants (Column 2 of table 1). Additionally, those who report participating in SNAP are more likely to be black (34.6% vs. 22.1%), female (70.5% vs. 51.7%), participating in WIC (16.2% vs. 5.4%), and living in a household with a child under the age of 5 (26.4% vs. 19.6%). A simple comparison of the different weight measures reveals that SNAP participants, on average, are more likely to be obese and overweight compared to those not reporting any SNAP benefits. We see similar differences, albeit less pronounced, between respondents who reported participating in SNAP in at least one of the 5 waves (Column 3 of table 1) compared to those who never participated in SNAP (Column 4 of table 1) during this time period.

Methodology

Our overall goal is to estimate the average treatment effect (ATE) of SNAP participation on being obese or overweight in light of both self-selection into treatment as well as misreporting of treatment status.[8] We pursue both a point estimation of the ATE based on methods described in Hausman *et al.* (1998) and Brachet (2008), as well as partial (or set) identification of the ATE using a nonparametric bounds approach following Manski & Pepper (2000), Kreider & Pepper (2007), Kreider *et al.* (2012) and Gundersen *et al.* (2012). We present the details of both approaches throughout the remainder of this section.[9]

Parametric Approach

Consider the linear model

$$y_i = \gamma T_i + x_i \beta + \varepsilon_i, \tag{1}$$

where y_i denotes the outcome of interest (i.e., whether the respondent is obese or overweight), T_i denotes an indicator for SNAP participation, x_i is a vector of respondent and household characteristics, and ε_i is an error term. As has been documented in the literature, estimation of equation 1 by ordinary least squares (OLS) is subject to omitted variables bias (Gibson, 2003; Kaushal, 2007;

---

[8]As part of the larger debate on the value of BMI as a measure of health, there is evidence suggesting that overweight but not obese individuals (BMI between 25 and 30) are actually healthier than individuals with "normal" BMI between 20 and 25 (Flegal *et al.*, 2013). As such, an analysis of whether people are overweight may be of limited practical value. Nonetheless, for purposes of comparing results to the existing literature, we include this measure in our analysis. We are grateful to an anonymous referee for highlighting this important point.

[9]Our available instruments do not intuitively satisfy the conditions in Mahajan (2006) and Lewbel (2007), such as conditional independence with the measurement error or minimum support requirements not satisfied with a binary instrument. Along with concerns regarding endogenous treatment participation, we therefore do not pursue these estimators in our application.

Meyerhoefer & Pylypchuk, 2008; Baum, 2011). Moreover, if T_i is misreported, estimates of γ in equation 1 by both OLS and instrumental variables (IV) will be inconsistent (Aigner, 1973; Hausman *et al.*, 1998; Black *et al.*, 2000; Frazis & Loewenstein, 2003; Lewbel, 2007; Brachet, 2008). This inconsistency due to misreporting is particularly important given the extent of misreported SNAP participation in household surveys documented in the literature (Bollinger & David, 1997; Primus *et al.*, 1999; Bitler *et al.*, 2003; Marquis & Moore, 2010; Meyer *et al.*, 2011).

We denote observed treatment status by T_i and true (unobserved) treatment status by T_i^*, where $T_i^* = 1\,(z_i\delta + \nu_i > 0)$ and z_i is a vector of covariates including x_i as well as any instruments that are excluded from x_i and unrelated to ε_i (state-level SNAP policies in our application). Assuming misreporting probabilities are random and constant across individuals, we denote the conditional misreporting probabilities by

$$\alpha_0 \equiv \Pr\left(T_i = 1 | T_i^* = 0\right)$$

$$\alpha_1 \equiv \Pr\left(T_i = 0 | T_i^* = 1\right),$$

where α_0 denotes the probability of a false positive (reporting participation in SNAP when in fact the respondent does not participate) and α_1 denotes the probability of a false negative.[10] In the presence of such misreported treatment, Frazis & Loewenstein (2003) show that IV estimates of equation 1 are inconsistent and overestimated by a factor of $1/(1 - \alpha_0 - \alpha_1)$, similar to the derivations in Lewbel (2007) and Battistin & Sianesi (2011).

As an alternative to the linear IV model, we can potentially improve our point estimates by modeling the misreporting process directly. Specifically, with misreporting probabilities α_0 and α_1,

---

[10]The assumption of random and constant misreporting across individuals is strong and has been found to be violated in prior studies (Black *et al.*, 2003). Indeed, we also estimate large differences in misreporting rates when we split the sample between men and women. Violations of this assumption further support the use of alternative methods of treatment effect estimation, which we consider in detail throughout the remainder of the paper.

the conditional expectation for the observed treatment status is then

$$\mathrm{E}\left[T_i|z_i\right] = \Pr\left(T_i = 1|z_i, T_i^* = 0\right)\Pr\left(T_i^* = 0|z_i\right) + \Pr\left(T_i = 1|z_i, T_i^* = 1\right)\Pr\left(T_i^* = 1|z_i\right)$$

$$= \alpha_0 + (1 - \alpha_0 - \alpha_1)F_\nu(z_i\delta), \tag{2}$$

where $F_\nu(\cdot)$ is the cumulative distribution function (CDF) of ν_i. Assuming a functional form for $F_\nu(\cdot)$ and assuming $\alpha_0 + \alpha_1 < 1$, Hausman *et al.* (1998) propose the estimation of α_0 and α_1 using nonlinear least squares (NLS) by minimizing

$$\frac{1}{N}\sum_{i=1}^{N}\left\{T_i - \alpha_0 - (1 - \alpha_0 - \alpha_1)F_\nu(z_i\delta)\right\}^2. \tag{3}$$

In addition to the requirement that $\alpha_0 + \alpha_1 < 1$, note that the misreporting probabilities are identified based on the non-linearity of F_ν. In particular, as discussed in Hausman *et al.* (1998), misreporting probabilities would not be identified in a linear probability model where

$$\mathrm{E}\left[T_i|z_i\right] = \alpha_0 + (1 - \alpha_0 - \alpha_1)z_i\delta = (\alpha_0 + \delta_0) + z_i(1 - \alpha_0 - \alpha_1)\delta_1.$$

Extending Hausman *et al.* (1998), Brachet (2008) proposes an alternative estimator in which δ is estimated from equation 3 in the first stage, and the second stage then estimates equation 1 after replacing the observed treatment status with the predicted treatment status. The second stage therefore estimates the equation,

$$y_i = \gamma F_\nu(z_i\hat{\delta}) + x_i\beta + \varepsilon_i. \tag{4}$$

The methods described in Hausman *et al.* (1998) and Brachet (2008) apply directly to a cross-sectional setting; however, their appropriateness in a panel data setting is not clear. For example, to the extent that misreporting is time-invariant but (possibly) individual-specific, then misreported

treatment should not bias estimates in a fixed effects setting. If, however, misreporting varies over time, then some amount of misreporting will persist even after controlling for individual-level fixed effects. In the current analysis, we therefore extend Brachet's proposed estimator to a panel data setting as follows:

1. Separately for each wave, estimate δ by NLS estimation of equation 3, and denote the estimates in year t by $\hat{\delta}_t$. This estimation includes in z_{it} the state-level instruments along with all other variables in x_{it}.

2. Form the predicted true treatment status, $\hat{P}_{it} = F_\nu(z_{it}\hat{\delta}_t)$, for each year.

3. Estimate a standard linear fixed effects model, $y_{it} = \gamma\hat{P}_{it} + x_{it}\beta + \phi_i + \varepsilon_{it}$.

For convenience, we term this estimator the misreporting-adjusted fixed effects (MA-FE) estimator. As in Hausman *et al.* (1998) and Brachet (2008), identification in this case derives from the nonlinearity in F_ν as well as the requirement that $\alpha_0 + \alpha_1 < 1$ in each period.

Nonparametric Bounds

Our bounds estimates follow closely that of Kreider & Pepper (2007) and Kreider *et al.* (2012). Here, we present the intuition of our estimates and the alternative assumptions employed. Additional discussion is provided in the appendix. Generally, the bounds rely on observed sample averages and differences in these averages between the treated and untreated groups. Alternative formulas for the bounds then follow from two broad considerations.

The first is self-selection into treatment. As is always the case in the potential outcomes framework, the sampling process alone cannot identify the expected outcomes conditional on counterfactual treatment status. The assumption of exogenous treatment selection (random treatment assignment) avoids this problem altogether such that the ATE can be estimated purely from sample averages for the treated and untreated subsamples. Conversely, the worst-case selection bounds make no assumptions on the selection process. For example, denoting by $Y(1)$ the outcome of the

treated and by $Y(0)$ the outcome of the untreated, consider the counter-factual probability of being obese for the treatment group conditional on not receiving treatment, $P[Y(1) = 1|T = 0]$. Under the assumption of exogenous selection, this is the same as the observed conditional probability, $P[Y(1) = 1|T = 1]$, and the ATE is point identified assuming no misreporting. With a binary outcome, the counterfactual probabilities necessarily fall in the $[0, 1]$ interval, and the worst-case selection bounds exploit this fact to bound the overall ATE.

In addition to the exogenous selection and worst-case selection bounds, we also consider bounds under the monotone treatment selection (MTS) and monotone instrumental variable (MIV) assumptions. In our setting, MTS assumes that obese or overweight individuals are more likely to receive SNAP benefits. As such, the upper bound on the estimated ATE should be lower relative to the worst-case selection bounds, since some amount of the observed differences in probabilities of being obese or overweight are accounted for by self-selection. Finally, the MIV assumption implies that the probability of a good outcome is weakly monotonically increasing with an observed covariate. In our case, we follow Gundersen *et al.* (2012), (Kreider *et al.*, 2012), and others in using the income to poverty line ratio (IPR) as our MIV, where the MIV assumption is that an increase in IPR will tend to decrease the probability of being obese.[11] Combined with the MTS assumption, the mechanics of the MIV estimator are such that we form alternative MTS bounds for different values (or range of values) of the MIV, and we then take a weighted average of all such bounds to form the overall ATE bounds. The assumption is therefore that MTS holds conditionally for each MIV group.[12]

---

[11]Dividing the IPR distribution into five quantiles (IPR less than 1.5, 1.5 to 2.8, 2.8 to 3.9, 3.9 to 5.5, and above 5.5), the obesity rates within each quantile are 34%, 31%, 28%, 27%, and 21%, respectively. Obesity rates among women specifically are 38%, 34%, 27%, 25%, and 19%, while the rates among men are 28%, 28%, 29%, 29%, and 24%. Our figures are similar to those reported in Gundersen (2015) based on the NHANES, suggesting that the monotonicity assumption is appropriate in the full sample and separately for women, but may not be appropriate among men.

[12]The MIV estimator is known to suffer from finite sample bias that will tend to yield narrower bounds relative to the true bounds (Manski & Pepper, 2000, 2009). Kreider & Pepper (2007)

A second consideration is the potential misreporting of treatment status. Broadly, the presence of misreporting widens the ATE bounds relative to the assumption of no misreporting, and the extent to which the bounds widen depends on the assumed level of misreporting in the data. In our application, we assume maximum overall misreporting rates of 1%, 2%, 5%, and 10%. We also consider gender-specific misreporting rates of 20% and 30% among women and men, respectively, which provides a more direct comparison between the point estimates and the nonparametric bounds. Finally, for consistency with existing survey evidence on the nature of SNAP misreporting, we consider the case of no false positives. This assumes that no individual reports having received SNAP benefits when in fact they did not.

Results

We first discuss separately our results on the estimated misreporting rates, point estimates, and nonparametric bounds. Note that, consistent with the literature, the parametric methods and nonparametric bounds adopt different definitions of the misreporting probabilities (conditional versus unconditional, respectively). In the "Comparison of Methods" subsection, we therefore convert our estimated conditional misreporting probabilities from our parametric approach to unconditional misreporting probabilities so that we can more directly compare results across the two methods.

propose a bias correction; however, Hirano & Porter (2012) argue that such bias corrections may be inappropriate as they may not eliminate the bias and may substantially increase the variance. For this reason, we do not pursue a bias-corrected version of the MIV estimator in our analysis. Moreover, the Monte Carlo results in Manski & Pepper (2009) suggest that the bias will generally be small (less than 0.05) based on our sample sizes. A bias of this magnitude will not affect our qualitative findings.

Extent of Misreporting

Our analysis begins by first estimating the misreporting probabilities, α_0 and α_1, from equation 3. The results based on the full sample as well as specific estimates for each year and gender are summarized in table 2. Consistent with the existing literature, we find high rates of misreporting in the NLSY79, with false positives close to 0 and false negatives ranging from 23% to 45% in the full sample and 57% to 72% among men. We also find variable misreporting rates across time. Our misreporting estimates collectively suggest three important findings: 1) misreporting is prevalent in the NLSY79 data; 2) misreporting predominantly takes the form of underreporting of benefits; and 3) misreporting rates are highly variable across gender and across time.

From Frazis & Loewenstein (2003), the misreporting rates in table 2 imply that standard linear IV estimates are overstated by a factor of between 1.4 and 2.1, with ranges of 1.7 to 2.2 for women and 2.8 to 4.2 for men. Such high rates of misreporting and the subsequent overestimate in linear IV methods are critical to our understanding of the true effect of SNAP participation on weight outcomes, suggesting that existing estimates aimed at controlling for endogeneity via instrumental variables are overstating the true treatment effects by well over 100%.

Parametric Approach

Results for the effect of SNAP on the probability of being obese are summarized in table 3. Table 4 presents the analogous results for the estimated effect of SNAP on the probability of being overweight. In addition to the MA-FE estimator, we include for comparison purposes the results from a fixed effects (FE) model and fixed effects with instrumental variables (FE-IV). Standard errors in the FE and FE-IV estimators are clustered at the individual level, and the standard errors for the MA-FE estimator are derived from 200 bootstrap replications.

Our FE-IV and MA-FE estimators each adopt two instruments: 1) whether the state uses biometric identification technology; and 2) the percentage of SNAP benefits issued by direct mail in

each state.[13] Intuitively, both of these state policies may affect an individual's decision to participate in SNAP by varying the costs associated with participation, but have no effect on weight outcomes. For example, states that implement biometric identification technology are increasing the cost and burden to individuals considering participating in SNAP compared to states that do not. Similarly, states that issue a higher rate of SNAP benefits by direct mail via coupons are increasing the cost of using SNAP benefits compared to states that issue benefits via electronic bank transfer debit cards (EBT cards) or other direct means that do not require physical use of coupons.

To formally assess the appropriateness and strength of our instruments, we present the first-stage IV results in table 5.[14] Among the overall dataset as well as females specifically, these first stage regressions generally yield high and significant global F-statistics, and tests of overidentifying restrictions generally provide low and insignificant Hansen's J-statistics. The first-stage results therefore suggest that our instruments are highly correlated with reported SNAP participation and that our instrument set is appropriate for our analysis of the full sample and for a separate analysis of females. For males, however, the global F-statistics are low and insignificant suggesting that our instrument set is a weak predictor of SNAP participation among men.

In general, our three estimators produce mixed results on the effects of SNAP participation on the

---

[13]Arizona, California, Massachusetts, and Texas all adopted biometric technology at some point in our study period. Meanwhile, New York always required biometric identification during the time period and all other states did not require biometric ID. Direct mail issuance of SNAP benefits show relatively more variation over time. In 1996, there were 24 states that issued at least 50% of their benefits via mail. This drops to 17 in 1998, 8 in 2000, 6 in 2002, and zero in 2004. All states were required to issue 100% of benefits via EBT cards by 2004.

[14]Following Meyerhoefer & Pylypchuk (2008), we also considered as potential instruments the state expenditures on food stamp outreach programs as well as whether the state requires 50% or more of food stamp participants to recertify their program eligibility within three months. Other instruments we considered included state asset limit rules (e.g., whether vehicles are exempt from the asset eligibility test) and whether benefits were available to non-citizens. None of these additional instruments were significant predictors of SNAP participation.

probability of being obese, controlling for a full set of demographic characteristics and year dummies (table 3). The FE estimates suggest a positive effect of SNAP on the probability of being obese for the full sample and for men in particular, with SNAP participation increasing the probability of being obese for the full sample of adults by 2 percentage points, statistically significant at the 5% level. The effect for men is larger, predicting a 5.4 percentage point increase and significant at the 1% level. Meanwhile, SNAP participation has virtually no effect on women's probability of being obese. The FE-IV estimates tend to be larger in magnitude, but are not statically significant at conventional levels.

With the exception of men, the estimated effects vary in sign depending on which instruments are used, with the full-sample FE-IV estimates ranging from a decrease of 22.1 to an increase of 14.2 percentage points. Among women, the FE-IV estimates range from a decrease of 20.3 to an increase of 18.8 percentage points. Interestingly, regardless of which exclusion restriction is applied, the effects of SNAP for men are all negative ranging from a decrease of 8.9 to 29.9 percentage points. Finally, the MA-FE estimates more closely resemble the FE results, where we estimate small (and insignificant) positive effects of SNAP participation on the the probability of being obese.

The effects of SNAP participation on the probability of being overweight are similarly mixed between the three estimators (table 4). With the FE estimates, we find that SNAP participation has practically no effect for both the full sample and for men and women separately. The FE-IV results suggest a positive effect of SNAP participation for all three groups, although the estimates are insignificant when estimating separately among men or women. For the full sample of adults and with the biometric identification as an instrument, participation in SNAP increases the probability of being overweight by 41 percentage points and is statistically significant at the 10% level. When we incorporate both instruments, we estimate an increase in the probability of being overweight of 32.4 percentage points, again statistically significant at the 10% level. By contrast, our MA-FE estimates suggest a negative effect of SNAP participation on the probability of being overweight, with SNAP participation reducing the likelihood of being overweight for our full sample of adults

by 3.5 percentage points. The effects for men and women separately are also negative, but smaller and not significant at conventional levels.

In general, we do not find significant effects of SNAP on weight outcomes for low-income men or women as the point estimates from all three estimators are generally indistinguishable from zero when examining each gender separately. However, for the full sample of adults, FE estimates suggest a positive and significant effect of SNAP participation on obesity, and FE-IV estimates suggest a positive and significant effect of SNAP on the probability of being overweight. When also accounting for misreporting in our MA-FE estimator, we instead find no significant effect on obesity and a negative effect on the probability of being overweight.

Our parametric estimates therefore differ from some of the previous studies in the literature, where other authors have estimated a significant effect of SNAP participation on obesity among women with no significant effect for men. This appears to be due to differences in sample construction across studies. For example, we allow for an unbalanced panel consisting of all individuals eligible for SNAP participation in at least two waves from 1996 through 2004. Meanwhile, Baum (2011) also uses the NLSY79 data, but his sample consists of a balanced panel of individuals who are eligible for SNAP participation in every wave. We also allow for a higher income eligibility threshold since our annual income measures may not identify specific periods of eligibility within a given year. Although there are benefits to examining a balanced panel and to considering alternative eligibility thresholds, there may also be important differences between individuals who are persistently eligible for SNAP benefits over several years compared to individuals who may only be eligible intermittently. These differences are likely more salient when imposing tighter eligibility thresholds. In the "Comparison to Literature" subsection, we present alternative results based on a sample that more closely replicates estimates from the existing literature.

Nonparametric Bounds

The FE and FE-IV results are robust to time-invariant, individual-specific misreporting. However, the misreporting estimates by year in table 2 do not support a hypothesis of time-invariant misreporting rates. The misreporting estimates by gender further suggest correlation between misreporting probabilities and observed covariates. Therefore, the columns in tables 3 and 4, as well as other methods of point estimation proposed in the literature, are subject to some degree of bias and inconsistency. In the absence of appropriate instruments to control for misreporting, we adopt a nonparametric bounds analysis to provide a more complete picture of the range of possible effects of SNAP participation on weight outcomes.

Table 6 presents the full-sample nonparametric bounds on the estimated ATE of SNAP participation on obesity. Each range in brackets reflects the ATE bounds under alternative assumptions regarding the selection process and magnitude of misreporting.[15] The "arbitrary misreporting" column allows for misreporting in both directions (false positives and false negatives) bounded above by the assumed error rates in the "error rates" column, while the "no false positives" column assumes $P(T = 1, T^* = 0) = 0$ with false negatives still bounded above by the relevant error rate. Each of the four panels of the table refers to a different assumption on the selection process (exogenous selection, worst-case selection, MTS, and MIV). For example, assuming exogenous selection and no misreporting, the estimated ATE is a 6.3 percentage point increase in the probability of being obese. Conversely, assuming MTS with a misreporting probability of 10% and no false positives, the estimated ATE ranges from -0.486 to 0.358 (between a 48.6 percentage point decrease and 35.8 percentage point increase in the probability of being obese). Tables 7 and 8 present analogous results

---

[15]In all cases, we present results on the estimated bounds and do not provide standard errors or confidence intervals. We do this for two primary reasons: 1) the calculation of confidence intervals is computationally cumbersome and time consuming, particularly for the MIV estimates; and 2) given our focus on the sensitivity of common parametric estimates to misreported treatment, the confidence intervals provide little additional information relative to the bounds themselves.

specifically for women and men, respectively.

The results clearly highlight the potential impact of misreporting, where the ATE bounds include negative values even for very small levels of misreporting, and similarly for different assumptions on the selection process. The results for men are particularly sensitive to nearly any amount of misreporting, while the increase in obesity for women (assuming exogenous selection) persists with up to 1% misreporting. More generally, across almost all assumptions regarding selection and misreporting, our ATE bounds on the effect of SNAP participation on obesity tend to include 0 for both men and women. Allowing for selection, the lower ATE bound decreases much more for women than for men. For example, with 5% misreporting and worst-case selection, SNAP participation among men leads to between a 72 percentage point increase and a 38 percentage point decrease in the probability of being obese (table 8), while women incur between a 62 percentage point increase and 48 percentage point decrease in the probability of being obese (table 7). This suggests that existing studies finding an increase in obesity for female SNAP participants may not fully account for the selection process. These bounds significantly tighten if we further allow for IPR as a monotone instrumental variable. In this case (again with 5% misreporting), the ATE bounds reveal between a 19 percentage point increase and 30 percentage point decrease in probability of obesity among men, and between an 10 percentage point increase and 44 percentage point decrease in the probability of obesity among women.

As is often the case when bounding the ATE, we cannot definitively sign the effect of SNAP on adult obesity. However, the results are important from a program evaluation perspective. For example, if the additional costs from a 15 percentage point increase in adult obesity are such that the the benefits of SNAP (e.g., reduction in food insecurity) still exceed the costs, then the SNAP program would still pass the standard cost-benefit assessment. In this sense, our results offer a best-case and worst-case scenario in terms of the benefits or additional costs due to changes in adult obesity following SNAP participation.

Tables 9-11 present the results for SNAP participation on being overweight, analogous to the

estimates on obesity in tables 6-8. Here, under the MIV and MTS assumptions, we can sign the ATE even when allowing for some small amount of misreporting. If we further assume no false positives, then we see from the bottom panel of tables 10 and 11 that SNAP participation decreases the probability of being overweight for both men and women provided misreporting rates do not exceed 2%. But when allowing for misreporting rates more consistent with those estimated in table 2, the estimated bounds include zero, and we cannot definitively sign the effect of SNAP participation on the probability of being overweight.

Comparison of Methods

A direct comparison between the parametric estimates and the nonparametric estimates is complicated by the different definitions of misreporting adopted in the two approaches.[16] Namely, the misreporting probabilities discussed in Hausman *et al.* (1998), Frazis & Loewenstein (2003), and Brachet (2008) are conditional, defined as the probability of the observed outcome conditional on the true outcome. The misreporting probabilities in the bounds literature, however, are unconditional joint probabilities of observed and true outcomes. The conditional misreporting probabilities summarized in table 2 for the full sample, along with the observed rate of SNAP participation in our sample (21%), imply $P(T = 0, T^* = 1) = 0.115$ and $P(T = 1, T^* = 0) = 0.026$.

Therefore, for the full sample, the nonparametric bounds most consistent with our estimated misreporting rates are those with an assumed 10% misreporting and no false positives. For females and males specifically, the estimated conditional probability of a false negative is closer to 20% or

---

[16]Of course, other issues further complicate the comparison between our two approaches. For example, there are functional form assumptions inherent in the parametric approach and avoided in the non-parametric bounds. In addition, the parametric estimates introduce a richer set of conditioning variables compared to the nonparametric bounds. Examining the implications of these differences is beyond the scope of the current paper, and we focus here instead on obtaining an appropriate comparison with regard to misreporting rates. We thank an anonymous reviewer for highlighting this point.

30%, respectively, again with very small rates of false positives. The bottom row of each panel in tables 6-11 presents the estimated ATE bounds under each of these scenarios, which we can then more directly compare to the point estimates in tables 3 and 4.

Table 12 presents a more concise summary of our results, focusing only on the bounds most consistent with the estimated misreporting rates. Specifically, table 12 presents the point estimates and standard errors based on the FE, FE-IV, and MA-FE estimators, along with nonparametric bounds estimated under the MTS and MIV-MTS assumptions (both cases with the relevant misreporting probabilities discussed above and no false positives). Focusing on the estimated treatment effect of SNAP on the probability of being overweight, the FE-IV results consistently exceed the upper-bound in the MTS and MIV-MTS columns, often by more than 200%. Conversely, the FE and MA-FE estimators provide more reasonable estimates toward the midpoint of the relevant nonparametric bounds estimates.

The large differences observed in the FE-IV estimates relative to the FE estimates and nonparametric bounds could be due to fundamental differences in the underlying parameter being estimated. In particular, our FE, MA-FE, and nonparametric bounds analysis are all intended to estimate the ATE;[17] however, in the presence of heterogeneous treatment effects, the local average treatment effect (LATE) estimated using instrument-based methods may substantially differ from the ATE (Imbens & Angrist, 1994; Heckman *et al.*, 2006). As stated succinctly in Heckman & Vytlacil (2005), "Two economists analyzing the same dataset but using different valid instruments will estimate different parameters that have different economic interpretations...The choice of an instrument defines the treatment parameter being estimated." In this sense, misreported treatment participation becomes an even larger concern as not only will we overestimate the true parameter, but the true parameter for which we are estimating may not be of interest. Our results in table 12 illustrate how these issues combine in this case to provide estimated treatment effects far different from the feasible range of

---

[17]Although we adopt instruments in the first stage of our MA-FE, the treatment effect is technically identified without such instruments due to functional form assumptions.

estimates on the treatment effect of interest.

Comparison to Literature

As discussed previously, the general consensus in the literature (and the finding highlighted in policy briefs on the subject) is that SNAP participation tends to increase obesity among adult women but not men. Baum (2011) finds similar results also using the NLSY79 data. However, our FE results in table 3 instead suggest a significant positive effect in the overall sample and specifically among men, with no significant effect among women. There are likely several reasons for the differences in our results relative to Baum (2011) and others, including differences in time frames under study and underlying differences in sample construction. For example, Baum (2011) considers a balanced panel of income-eligible respondents across the 1985, 1986, 1988, 1989, 1990, 1992, 1993, 1994, 1996, 1998, and 2000 waves of the NLSY79. Therefore, in order to highlight the specific role of misreporting in the context of the existing literature, we reconstruct our analytic sample following Baum (2011) and re-estimate the effects of SNAP participation on obesity.

In this analysis, we maintain our focus on the 1996 through 2004 waves of the NSLY79; however, following Baum (2011), our sample now consists only of a balanced panel of respondents who were consistently at or below 250% percent of the federal poverty line in each wave. Summary statistics for this alternative sample are provided in table 13. The comparison among SNAP and non-SNAP participants is qualitatively similar to our original analysis, where we find that SNAP participants are again more likely to be obese compared to those not reporting any SNAP benefits. Similarly, SNAP participants earn less income, tend to have larger households with more children, are less likely to have graduated high school or have mothers who graduated high school, and are less likely to be employed, married, or be living in households with an elderly adult.

Table 14 summarizes our estimated treatment effects of SNAP participation on obesity based on this alternative dataset.[18] Our FE estimates show no statistically significant effect among men, with

---

[18]Similar to our previous analysis, we again estimate high rates of misreporting, with overall

a positive and statistically significant effect on obesity among women. In this sense, our results based on this alternative sample are more in-line with the existing consensus in the literature. Although imprecisely estimated, the FE-IV estimates are much larger in magnitude among the full sample and particularly among women. Meanwhile, the MA-FE estimates are much smaller in magnitude and also insignificant. In all cases, nonparametric bounds include 0 and do not support the hypothesis that SNAP participation increases obesity (either overall or among men or women specifically). The FE-IV results for women also exceed the respective upper-bound of the MIV-MTS nonparametric bounds, similar to our initial findings in table 12.

Sensitivity and Robustness

As noted in our discussion of the NLSY79 data, the height and weight data in the NLSY are subject to measurement error. As such, our obesity and overweight outcomes may be misclassified and therefore further bias our estimated effects of SNAP participation (Hausman *et al.*, 1998). Note that, to the extent misreported weight outcomes are individual-specific but time-invariant, our FE analysis will still provide consistent treatment effects estimates. Nonetheless, we examine the sensitivity of our results to this additional measurement error by adopting the BMI adjustment proposed in Cawley (2004). We then re-estimate our results for the effects of SNAP participation on our adjusted measure of obesity for the overall sample. The results, summarized in table 15, are similar to our initial results based on the raw NLSY height and weight data (tables 3 and 4).[19]

_____

under-reporting rates of 32% and over-reporting rates of 3.3%.

[19]For the nonparametric analysis, we provide updated bounds estimates based on the adjusted weight outcomes. It is possible, however, to explicitly introduce a misreported outcome alongside misreported treatment (Gundersen & Kreider, 2008). Doing so naturally widens the bounds on the estimated treatment effects. Allowing for a misreported outcome as well as misreported treatment is also computationally burdensome, as we must now perform a search over all combinations of the different error rates. Since our bounds already include 0 when allowing for misreported treatment assignment, explicitly introducing a misreported outcome similar to Gundersen & Kreider (2008)

In addition to potential mismeasurement of our outcome, we also considered the appropriateness of our SNAP eligibility criteria. Specifically, our initial sample consisted of all respondents at or below 250% of the FPL based on income and household size. Although SNAP eligibility is formally restricted to those at or below 130% of the FPL, Mykerezi & Mills (2010) discuss the value of using higher thresholds. The authors argue that some of the more common eligibility criteria of 130% or 185% of the FPL are too restrictive and tend to exclude households that report receiving SNAP benefits or who were eligible for some portion of the year. The authors further point out that it is the higher income households that are more likely to be the marginal SNAP participants, and as such, any study using instrumental variables will rely on this segment of the population in identifying their treatment effects. In light of these issues, the authors ultimately used eligibility criteria of 150%, 200%, and 250% of the FPL.

For robustness, we re-estimated our models using eligibility criteria of both 130% and 185% of the FPL, with results summarized in table 16. For comparison, the table also presents the results based on our original criteria of 250% of the FPL. For the FE estimates, we see few differences in our estimated treatment effects across the alternative eligibility criteria. Our estimators employing instruments are more sensitive to the different criteria. For example, our MA-FE estimate of the effect of SNAP participation on the probability of being overweight is positive using eligibility of 185% of the FPL but negative when using 130% and 250%. Similarly, the FE-IV estimates of the effect of SNAP participation on obesity increase from 0.001 to 0.190 when moving from the 250% FPL criterion to the 130% criterion. As discussed in Mykerezi & Mills (2010), the sensitivity of our instrument-based estimates to the eligibility thresholds is to be expected because the higher income individuals are more likely to be the marginal SNAP participants. Changes to the income criteria therefore fundamentally change the relevant segment of the sample that serves to identify the local average treatment effect. That said, comparing the nonparametric bounds to the estimated treatment effects, the qualitative findings are consistent across different eligibility criteria,

will only further widen the bounds with qualitatively no change in the results.

28

where we continue to see large differences in the FE-IV estimates relative to the FE estimates and nonparametric bounds.

Conclusion

The consequences of SNAP participation are increasingly important given the growth of this program in recent years, with much of the existing literature estimating a positive and significant effect of SNAP participation on female obesity; however, surveys of SNAP participation and many other programs are also known to suffer from high rates of misreporting. In this paper, we consider more carefully the role of misreporting in SNAP and examine empirically how the presence of misreporting may impact the estimated effect of SNAP on adult weight outcomes. We pursue a variety of estimation strategies, including several alternative parametric approaches as well as nonparametric bounds.

Our results are reflective of the conflicting evidence in the existing literature and highlight the impact of misreporting on estimated treatment effects in practice. We estimate under-reporting of SNAP participation of between 26% and 48% overall, with under-reporting nearing 70% among men. We also find variable misreporting rates across time. The magnitudes of misreporting suggest that linear IV methods will tend to overestimate the treatment effect of SNAP participation on adult obesity by over 100% (Frazis & Loewenstein, 2003). Relative to our nonparametric bounds estimates, we also find that fixed effects IV estimates are similarly overstated, with estimated treatment effects exceeding the nonparametric upper bounds sometimes by more than 200% under reasonable assumptions on the selection process and misreporting rates.

In the presence of misreporting, our comparison between parametric estimates and nonparametric bounds further suggests that instrument-based corrections, even when the proposed instruments pass standard IV tests, may perform worse than applying no correction at all.[20] The problems with

---

[20]These results are similar to Meyer & Mittag (2014), who find that misreporting corrections

IV methods intuitively derive from two sources: 1) such methods are inconsistent in the presence of misreporting; and 2) with treatment effect heterogeneity, the parameter actually being estimated may not reflect the true parameter of interest. The interaction of these two issues alongside endogenous treatment participation is an important area of future applied research.

Specific to our application, we find no evidence that SNAP participation significantly increases adult obesity when allowing for misreported treatment participation. This result is robust to a range of different samples and income-eligibility thresholds. To the contrary, we find some evidence that SNAP participation may decrease the probability of being overweight; however, due to the large amount of misreporting (particularly among men), these results are inconclusive when identifying effects specifically by gender. Our results broadly suggest that few conclusions can be drawn regarding the effect of SNAP on adult weight outcomes in light of misreporting, mirroring the findings of Gundersen & Kreider (2008) in their study of food security. As such, our results call into question many of the existing findings of a positive effect of SNAP participation on adult obesity among women.

The policy implications of our findings are two-fold: 1) attempts to combat the presumed increases in adult obesity due to SNAP participation are premature and should be reconsidered in light of more controlled studies or improved data collection; and 2) existing survey measures of SNAP participation should be improved with additional clarifying questions or perhaps by cross-validating reported participation with government records of SNAP participation. Methodologically, we can instead forego the estimation of treatment effects on the extensive margin (i.e., participation versus non-participation) and focus on treatment effects at the intensive margin (i.e., expansion or reduction in benefits among existing participants). These effects at the intensive margin are the subject of Almada & Tchernis (2015), who find that expansion of benefits among existing SNAP participants does not increase adult obesity and may decrease obesity among adults living in households

relying on functional form assumptions or exclusion restrictions may provide worse estimates relative to a näive estimator with no correction for misreporting.

with children under 5 years of age. More broadly, our results follow a growing empirical literature highlighting the difficulty in estimating treatment effects with misreported treatment participation. Consistent with this literature, our findings serve to caution policy makers when basing decisions on research using survey data with self-reported treatment.

References

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

Almada, Lorenzo, & Tchernis, Rusty. 2015. *Measuring Effects of SNAP on Obesity at the Intensive Margin.* Working Paper. Georgia State University.

Battistin, Erich, & Sianesi, Barbara. 2011. Misclassified treatment status and treatment effects: An application to returns to education in the United Kingdom. *Review of Economics and Statistics*, **93**(2), 495–509.

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

Bhattacharya, J., & Bundorf, M.K. 2009. The incidence of the healthcare costs of obesity. *Journal of health economics*, **28**(3), 649–658.

Bitler, Marianne P, Currie, Janet, & Scholz, John Karl. 2003. WIC eligibility and participation. *Journal of Human Resources*, 1139–1179.

Black, Dan, Sanders, Seth, & Taylor, Lowell. 2003. Measurement of higher education in the census and current population survey. *Journal of the American Statistical Association*, **98**(463).

Black, Dan A, Berger, Mark C, & Scott, Frank A. 2000. Bounding parameter estimates with nonclassical measurement error. *Journal of the American Statistical Association*, **95**(451), 739–748.

Blackburn, George L, Wilson, G Terence, Kanders, Beatrice S, Stein, Leslie J, Lavin, Philip T, Adler, Jamie, & Brownell, Kelly D. 1989. Weight cycling: the experience of human dieters. *The American journal of clinical nutrition*, **49**(5), 1105–1109.

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

Brachet, Tanguy. 2008. *Maternal smoking, misclassification, and infant health*. Working Paper. University of Pennsylvania.

Burgstahler, Rebecca, Gundersen, Craig, Garasky, Steven, *et al.* 2012. The supplemental nutrition assistance program, financial stress, and childhood obesity. *Agricultural and Resource Economics Review*, **41**(1), 29.

Cawley, John. 2004. The impact of obesity on wages. *Journal of Human Resources*, **39**(2), 451–474.

Cawley, John, & Meyerhoefer, Chad. 2012. The medical care costs of obesity: an instrumental variables approach. *Journal of health economics*, **31**(1), 219–230.

Chen, Xiaohong, Hong, Han, & Nekipelov, Denis. 2011. Nonlinear models of measurement errors. *Journal of Economic Literature*, **49**(4), 901–937.

Chen, Zhuo, Yen, Steven T, & Eastwood, David B. 2005. Effects of food stamp participation on body weight and obesity. *American Journal of Agricultural Economics*, **87**(5), 1167–1173.

Cody, Scott, & Tuttle, Christina. 2002. The Impact of Income Underreporting in CPS and SIPP on Microsimulation Models and Participating Rates. *Washington, DC: Mathematica Policy Research, Inc, July*, **24**.

Cole, Nancy, & Fox, Mary Kay. 2008. *Diet quality of Americans by food stamp participation status: Data from the National Health and Nutrition Examination Survey, 1999-2004*. US Department of Agriculture, Food and Nutrition Service, Office of Research, Nutrition and Analysis.

Courtemanche, Charles, Pinkston, Joshua C, & Stewart, Jay. 2015. Adjusting body mass for measurement error with invalid validation data. *Economics & Human Biology*.

Devaney, Barbara, & Moffitt, Robert. 1991. Dietary effects of the food stamp program. *American Journal of Agricultural Economics*, **73**(1), 202–211.

Dietz, William H. 1995. Does hunger cause obesity? *Pediatrics*, **95**(5), 766–767.

Dowler, Elizabeth. 2003. Food and poverty: insights from the North. *Development Policy Review*, **21**(5-6), 569–580.

Fan, Maoyong. 2010. Do food stamps contribute to obesity in low-income women? Evidence from the National Longitudinal Survey of Youth 1979. *American Journal of Agricultural Economics*, aaq047.

Finegold, Kenneth, Margrabe, William, & Ratcliffe, Caroline. 2007. *Food Stamp Program Rules Database Documentation*. Technical Report. The Urban Institute.

Finkelstein, Eric A, Trogdon, Justin G, Cohen, Joel W, & Dietz, William. 2009. Annual medical spending attributable to obesity: payer-and service-specific estimates. *Health affairs*, **28**(5), w822–w831.

Fisher, Jennifer Orlet, & Birch, Leann Lipps. 1999. Restricting access to palatable foods affects children's behavioral response, food selection, and intake. *The American journal of clinical nutrition*, **69**(6), 1264–1272.

Flegal, Katherine M, Kit, Brian K, Orpana, Heather, & Graubard, Barry I. 2013. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *Jama*, **309**(1), 71–82.

Fox, Mary Kay, Hamilton, William, & Lin, Biing-Hwan. 2004. Effects of food assistance and nutrition programs on nutrition and health. *Food Assistance and Nutrition Research Report*.

Fraker, Thomas M, Martini, Alberto P, & Ohls, James C. 1995. The effect of food stamp cashout on

food expenditures: An assessment of the findings from four demonstrations. *Journal of Human Resources*, 633–649.

Frazis, Harley, & Loewenstein, Mark A. 2003. Estimating linear regressions with mismeasured, possibly endogenous, binary explanatory variables. *Journal of Econometrics*, **117**(1), 151–178.

Gibson, Diane. 2003. Food stamp program participation is positively related to obesity in low income women. *The Journal of nutrition*, **133**(7), 2225–2231.

Goerge, Robert M, Franzetta, Kerry, & Dilts, John. 2009. *Food Stamp Program Eligibility and Participation in Chicago*. Working Paper 09-03.

Gundersen, Craig. 2015. SNAP and Obesity. *In:* Bartfield, J., Gunderson, C., Smeeting, T., & Ziliak, J. (eds), *SNAP Matters: How Food Stamps Affect Health and Well Being*. Redwood City, CA: Stanford University Press.

Gundersen, Craig, & Kreider, Brent. 2008. Food Stamps and Food Insecurity: What Can Be Learned in the Presence of Nonclassical Measurement Error? *Journal of Human Resources*, **43**(2), 352–382.

Gundersen, Craig, Kreider, Brent, & Pepper, John. 2012. The impact of the National School Lunch Program on child health: A nonparametric bounds analysis. *Journal of Econometrics*, **166**(1), 79–91.

Hastings, Justine, & Washington, Ebonya. 2010. The First of the Month Effect: Consumer Behavior and Store Responses. *American Economic Journal: Economic Policy*, **2**(2), 142–62.

Hausman, Jerry A, Abrevaya, Jason, & Scott-Morton, Fiona M. 1998. Misclassification of the dependent variable in a discrete-response setting. *Journal of Econometrics*, **87**(2), 239–269.

Heckman, James J, & Vytlacil, Edward. 2005. Structural equations, treatment effects, and econometric policy evaluation1. *Econometrica*, **73**(3), 669–738.

Heckman, James J, Urzua, Sergio, & Vytlacil, Edward. 2006. Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics*, **88**(3), 389–432.

Hirano, K., & Porter, J.R. 2012. Impossibility results for nondifferentiable functionals. *Econometrica*, **80**, 1769–1790.

Hoynes, Hilary W, & Schanzenbach, Diane Whitmore. 2009. Consumption responses to in-kind transfers: Evidence from the introduction of the food stamp program. *American Economic Journal: Applied Economics*, **1**(4), 109–139.

Hu, Yingyao, & Schennach, Susanne M. 2008. Instrumental variable treatment of nonclassical measurement error models. *Econometrica*, **76**(1), 195–216.

Imbens, Guido W, & Angrist, Joshua D. 1994. Identification and estimation of local average treatment effects. *Econometrica*, 467–475.

Kaushal, Neeraj. 2007. Do food stamps cause obesity?: Evidence from immigrant experience. *Journal of Health Economics*, **26**(5), 968–991.

Kimbro, Rachel Tolbert, & Rigby, Elizabeth. 2010. Federal food policy and childhood obesity: a solution or part of the problem? *Health Affairs*, **29**(3), 411–418.

Kreider, Brent. 2010. Regression coefficient identification decay in the presence of infrequent classification errors. *The Review of Economics and Statistics*, **92**(4), 1017–1023.

Kreider, Brent, & Pepper, John V. 2007. Disability and employment: reevaluating the evidence in light of reporting errors. *Journal of the American Statistical Association*, **102**(478), 432–441.

Kreider, Brent, Pepper, John V, Gundersen, Craig, & Jolliffe, Dean. 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.

Lewbel, Arthur. 2007. Estimation of average treatment effects with misclassification. *Econometrica*, **75**(2), 537–551.

Mahajan, Aprajit. 2006. Identification and estimation of regression models with misclassification. *Econometrica*, **74**(3), 631–665.

Manski, C. F., & Pepper, J. 2009. More on monotone instrumental variables. *Econometrics Journal*, **12 (Supplement)**, S200S216.

Manski, Charles F, & Pepper, John V. 2000. Monotone instrumental variables: with an application to the returns to schooling. *Econometrica*, **68**(4), 997–1010.

Marquis, Kent H, & Moore, Jeffrey C. 2010. Measurement errors in SIPP program reports. *Survey Methodology*, 01.

McCarthy, I., Millimet, D., & Roy, M. 2014. Bounding Treatment Effects: Stata Command for the Partial Identification of the Average Treatment Effect with Endogenous and Misreported Treatment Assignment. *The Stata Journal*, forthcoming.

McCarthy, I.M., & Tchernis, R. 2011. On the Estimation of Selection Models when Participation is Endogenous and Misclassified. *Advances in Econometrics*, **27a**.

Meyer, Bruce, & Mittag, Nikolas. 2014. *Misclassification in binary choice models*. Working Paper 20509. National Bureau of Economic Research.

Meyer, Bruce D, Mok, Wallace KC, & Sullivan, James X. 2009. *The under-reporting of transfers in household surveys: its nature and consequences*. Working Paper. National Bureau of Economic Research.

Meyer, Bruce D, Goerge, Robert M, & Assistance, Food. 2011. *Errors in survey reporting and imputation and their effects on estimates of food stamp program participation*. Food Assistance and Nutrition Research Program.

Meyerhoefer, Chad D, & Pylypchuk, Yuriy. 2008. Does participation in the food stamp program increase the prevalence of obesity and health care spending? *American Journal of Agricultural Economics*, **90**(2), 287–305.

Millimet, D. L. 2011. The elephant in the corner: A cautionary tale about measurement error in treatment effects models. *Advances in Econometrics: Missing-Data Methods*, **27A**, 1–39.

Millimet, D.L., Tchernis, R., & Husain, M. 2010. School Nutrition Programs and the Incidence of Childhood Obesity. *Journal of Human Resources*, **45**(3), 640–654.

Mykerezi, Elton, & Mills, Bradford. 2010. The impact of food stamp program participation on household food insecurity. *American Journal of Agricultural Economics*, aaq072.

Ogden, CL, Carrol, l MD, Kit, BK, & Flega, l KM. 2014. Prevalence of childhood and adult obesity in the united states, 2011-2012. *JAMA*, **311**(8), 806–814.

Polivy, Janet, Zeitlin, Sharon B, Herman, C Peter, & Beal, A Lynne. 1994. Food restriction and binge eating: a study of former prisoners of war. *Journal of Abnormal Psychology*, **103**(2), 409.

Primus, Wendell, Rawlings, Lynette, Larin, Kathy, & Porter, Kathryn. 1999. *The initial impacts of welfare reform on the incomes of single-mother families.* Policy Brief. Center on Budget and Policy Priorities.

Rosin, O. 2008. The economic causes of obesity: a survey. *Journal of Economic Surveys*, **22**(4), 617–647.

Schanzenbach, Diane Whitmore. 2009. Do school lunches contribute to childhood obesity? *Journal of Human Resources*, **44**(3), 684–709.

Schmeiser, Maximilian D. 2012. The impact of long-term participation in the supplemental nutrition assistance program on child obesity. *Health economics*, **21**(4), 386–404.

Townsend, Marilyn S, Peerson, Janet, Love, Bradley, Achterberg, Cheryl, & Murphy, Suzanne P. 2001. Food insecurity is positively related to overweight in women. *The Journal of nutrition*, **131**(6), 1738–1745.

Vassilopoulos, Achilleas, Drichoutis, Andreas, Nayga, Rodolfo, & Lazaridis, Panagiotis. 2011. *Does the Food Stamp Program Really Increase Obesity? The Importance of Accounting for Misclassification Errors.* Working Paper. MPRA Paper 28768.

Ver Ploeg, Michele L, & Ralston, Katherine L. 2008. Food Stamps and obesity: what do we know? *Economic Information Bulletin.*

Wilde, Parke E, & Ranney, Christine K. 2000. The Monthly Food Stamp Cycle: Shooping Frequency and Food Intake Decisions in an Endogenous Switching Regression Framework. *American Journal of Agricultural Economics*, **82**(1), 200–213.

Appendix: Nonparametric Bounds Estimates

We first introduce the relevant notation in the nonparametric bounds literature, following closely that of Kreider & Pepper (2007) and Kreider *et al.* (2012). In a potential outcomes framework with binary outcome Y, the ATE is given by

$$ATE(1,0) = P\left[Y(1) = 1|X\right] - P\left[Y(0) = 1|X\right] \tag{1}$$

where $Y(1)$ denotes the outcome with treatment, $Y(0)$ denotes the outcome without treatment, and X denotes a vector of observed covariates. Denoting treatment status by the indicator T^*, the observed outcome for a particular individual is given by $Y = T^*Y(1) + (1 - T^*)Y(0)$. We can then rewrite the elements of equation 1 as:[21]

$$P\left[Y(1) = 1\right] = P\left[Y(1) = 1|T^* = 1\right] P(T^* = 1) + P\left[Y(1) = 1|T^* = 0\right] P(T^* = 0), \text{ and} \tag{2}$$

$$P\left[Y(0) = 1\right] = P\left[Y(0) = 1|T^* = 1\right] P(T^* = 1) + P\left[Y(0) = 1|T^* = 0\right] P(T^* = 0). \tag{3}$$

Equations 2 and 3 highlight two fundamental problems in estimating the ATE. First, we have a selection problem in that the sampling process alone cannot identify the counterfactual probabilities, $P\left[Y(1) = 1|T^* = 0\right]$ and $P\left[Y(0) = 1|T^* = 1\right]$. Second, true treatment status may not be observed for all individuals. Specifically, instead of observing T^*, assume that the indicator T is observed. If $T \neq T^*$ for some individuals, then the sampling process again cannot identify the true treatment status T^*, and the probabilities on the right hand side of equations 2 and 3 are unknown.

Denote by Z^* an indicator for whether reported treatment assignment is accurate or not, where $Z^* = 1$ if $T^* = T$ and zero otherwise. We can then define the unconditional misreporting probabili-

---

[21]Conditioning on X is assumed throughout the remainder of this section in order to simplify notation.

ties as follows:

$$\theta_1^+ \equiv P(Y = 1, T = 1, Z^* = 0),$$

$$\theta_0^+ \equiv P(Y = 0, T = 1, Z^* = 0),$$

$$\theta_1^- \equiv P(Y = 1, T = 0, Z^* = 0), \text{ and}$$

$$\theta_0^- \equiv P(Y = 0, T = 0, Z^* = 0).$$

With this notation, equations 2 and 3 become

$$
\begin{aligned}
P\left[Y(1) = 1\right] &= \frac{P\left[Y = 1, T^* = 1\right]}{P(T^* = 1)} P(T^* = 1) + P\left[Y(1) = 1 | T^* = 0\right] P(T^* = 0) \\
&= \left[P(Y = 1, T = 1) - \theta_1^+ + \theta_1^-\right] \\
&\quad + P\left[Y(1) = 1 | T^* = 0\right] \left[P(T = 0) + (\theta_1^+ + \theta_0^+) - (\theta_1^- + \theta_0^-)\right], \text{ and} \qquad (4) \\
P\left[Y(0) = 1\right] &= P\left[Y(0) = 1 | T^* = 1\right] P(T^* = 1) + \frac{P\left[Y = 1, T^* = 0\right]}{P(T^* = 0)} P(T^* = 0) \\
&= P\left[Y(0) = 1 | T^* = 1\right] \left[P(T = 1) - (\theta_1^+ + \theta_0^+) + (\theta_1^- + \theta_0^-)\right] \\
&\quad + \left[P(Y = 1, T = 0) + \theta_1^+ - \theta_1^-\right]. \qquad (5)
\end{aligned}
$$

Bounds on the ATE can then be derived under varying assumptions regarding the extent of misclassification, $\theta_1^+, \theta_0^+, \theta_1^-, \theta_0^-$, and the nature of the selection process. Regarding misclassification, we consider two alternative scenarios: 1) the assumption of no false positives, $P(Z^* = 1 | T = 1) = 1$; and 2) an arbitrary upper bound on the extent of misreporting, $P(Z^* = 0) \leq Q$, for which we introduce the following notation:

$$0 \leq \theta_1^- \leq \min\{Q, P(Y = 1, T = 0)\} \equiv \theta_1^{UB-}, \qquad (6)$$

$$0 \leq \theta_0^- \leq \min\{Q, P(Y = 0, T = 0)\} \equiv \theta_0^{UB-}, \qquad (7)$$

$$0 \leq \theta_1^+ \leq \min\{Q, P(Y=1, T=1)\} \equiv \theta_1^{UB+}, \tag{8}$$

$$0 \leq \theta_0^+ \leq \min\{Q, P(Y=0, T=1)\} \equiv \theta_0^{UB+}, \text{ and} \tag{9}$$

$$\theta_1^+ + \theta_1^- + \theta_0^+ + \theta_0^- \leq Q. \tag{10}$$

Assumptions regarding the selection process are more involved and are discussed in the remainder of this section. Additional details on the derivation of bounds under these alternative misreporting and selection assumptions are available in Manski & Pepper (2000), Kreider & Pepper (2007), Kreider *et al.* (2012), Gundersen *et al.* (2012) and McCarthy *et al.* (2014).

Exogenous Selection

The assumption of exogenous selection implies

$$P[Y(1) = 1] = P[Y(1) = 1|T^* = 1] = P[Y(1) = 1|T^* = 0] \text{ and}$$

$$P[Y(0) = 1] = P[Y(0) = 1|T^* = 1] = P[Y(0) = 1|T^* = 0].$$

Allowing for measurement error, T^* is unobserved and these quantities can be rewritten as

$$
\begin{aligned}
P[Y = 1|T^* = 1] &= \frac{P[Y = 1, T^* = 1]}{P(T^* = 1)} \\
&= \frac{P(Y = 1, T = 1) - \theta_1^+ + \theta_1^-}{P(T = 1) - (\theta_1^+ + \theta_0^+) + (\theta_1^- + \theta_0^-)} \\
P[Y = 1|T^* = 0] &= \frac{P[Y = 1, T^* = 0]}{P(T^* = 0)} \\
&= \frac{P(Y = 1, T = 0) + \theta_1^+ - \theta_1^-}{P(T = 0) + (\theta_1^+ + \theta_0^+) - (\theta_1^- + \theta_0^-)}.
\end{aligned}
\tag{11}
$$
$$\tag{12}$$

Following Proposition 1 and A.1 in Kreider & Pepper (2007), we can then estimate the ATE

bounds as follows:

$$\inf_{b\in\left[0,\theta_1^{UB+}\right],\tilde{b}\in\left[0,\min\{Q-b,\theta_0^{UB-}\}\right]}\left[\frac{P\left[Y=1,T=1\right]-b}{P(T=1)-b+\tilde{b}}-\frac{P\left[Y=1,T=0\right]+b}{P(T=0)+b-\tilde{b}}\right],\ \text{and}\qquad(13)$$

$$\leq ATE$$

$$\leq\sup_{a\in\left[0,\theta_1^{UB-}\right],\tilde{a}\in\left[0,\min\{Q-a,\theta_0^{UB+}\}\right]}\left[\frac{P\left[Y=1,T=1\right]+a}{P(T=1)+a-\tilde{a}}-\frac{P\left[Y=1,T=0\right]-a}{P(T=0)-a+\tilde{a}}\right].$$

Estimation proceeds by performing separate two-way grid searches for $(b,\tilde{b})$ and $(a,\tilde{a})$ over the feasible region, where:

$$b\in[0,\min(Q,P\left[Y=1,T=1\right])],$$

$$\tilde{b}\in[0,\min(Q-b,P\left[Y=0,T=0\right])],$$

$$a\in[0,\min(Q,P\left[Y=1,T=0\right])],\ \text{and}$$

$$\tilde{a}\in[0,\min(Q-a,P\left[Y=0,T=1\right])].$$

Worst-Case Selection Bounds

With no assumptions on the selection process, and allowing for arbitrary measurement error, the ATE bounds are derived directly from:

$$P[Y(1)=1]\in[P[Y=1,T=1]-\theta_1^{UB+},P[Y=1,T=1]+P(T=0)+\theta_0^{UB+}],\ \text{and}$$

$$P[Y(0)=1]\in[P[Y=1,T=0]-\theta_1^{UB-},P[Y=1,T=0]+P(T=1)+\theta_0^{UB-}].$$

Monotone Treatment Selection (MTS)

The worst-case bounds may be tightened if we are willing to impose some assumptions on the nature of the selection process. The MTS assumption assumes that expected potential outcomes move in a

particular direction when comparing individuals in the treatment and control groups. We consider the case of negative selection (MTSn), where individuals in the treatment group are assumed to be more likely to experience a bad outcome conditional on treatment assignment. Denoting a good outcome by $Y = 1$ and allowing for measurement error, the ATE bounds with the MTSn assumption are:

$$P[Y(1) = 1] \in \left[\frac{P(Y = 1, T = 1) - \theta_1^+ + \theta_1^-}{P(T = 1) - (\theta_1^+ + \theta_0^+) + (\theta_1^- + \theta_0^-)}, P(T = 0) + P[Y = 1, T = 1] + \theta_0^+ - \theta_0^- \right], \text{ and}$$

$$P[Y(0) = 1] \in \left[P[Y = 1, T = 0] + \theta_1^+ - \theta_1^-, \frac{P[Y = 1, T = 0] + \theta_1^+ - \theta_1^-}{P(T = 0) + (\theta_1^+ + \theta_0^+) - (\theta_1^- + \theta_0^-)} \right].$$

Monotone Instrumental Variable (MIV) The MIV assumption implies that the latent probability of a good outcome conditional on treatment assignment – $P[Y(1) = 1]$ and $P[Y(0) = 1]$ – varies (weakly) monotonically with an observed covariate. The MIV assumption alone has no identifying power; therefore, following Kreider *et al.* (2012), we combine the MIV assumption with the MTSn assumption.

Denote by ν the monotone instrument and assume without loss of generality that $P[Y(1) = 1]$ and $P[Y(0) = 1]$ are non-decreasing in ν. Defining $u_1 < u < u_2$, the MIV assumption implies

$$P[Y(1) = 1 | \nu = u_1] \leq P[Y(1) = 1 | \nu = u] \leq P[Y(1) = 1 | \nu = u_2], \text{ and}$$

$$P[Y(0) = 1 | \nu = u_1] \leq P[Y(0) = 1 | \nu = u] \leq P[Y(0) = 1 | \nu = u_2].$$

Combining MIV and MTSn involves the following steps:

1. Split the sample into J cells, $j = 1, 2, ..., J$, based on values of ν, and let P_j denote the sample fraction in cell j.

2. Calculate the MTSn bounds for $P[Y(1) = 1]$ and $P[Y(0) = 1]$ for each cell under the relevant

misreporting assumptions (i.e., arbitrary measurement error or no false positives assumption). This yields UB_d^j and LB_d^j, $j = 1, 2, ..., J$ and $d = 0, 1$.

3. Calculate the overall lower bound for $P[Y(1) = 1]$, denoted by LB_1, as

$$LB_1 = \sum_j P_j \left(\sup_{j' \leq j} LB_1^{j'} \right).$$

LB_1 is therefore a weighted average of the appropriate lower bound estimates across all J cells, constructed by taking the relevant lower bound in each jth cell and averaging over the supremum of the individual lower bounds across all cells below and including the jth cell.

4. Calculate the overall UB for $P[Y(1) = 1]$, denoted by UB_1, as

$$UB_1 = \sum_j P_j \left(\inf_{j' \geq j} UB_1^{j'} \right).$$

This assumes $Y(1) = 1$ and $Y(0) = 1$ are good outcomes. If $Y(1) = 1$ and $Y(0) = 1$ are bad outcomes, then $UB_1 = \sum_j P_j \left(\inf_{j' \leq j} UB_1^{j'} \right)$. As with the lower bound in step (3), UB_1 is again a weighted average of the appropriate upper bound estimates across all J cells, constructed by taking relevant upper bound in each jth cell and averaging over the infimum of the individual upper bounds across all cells above and including the jth cell.[22]

5. Repeat steps 3 and 4 to obtain the overall LB and UB for $P[Y(0) = 1]$, denoted LB_0 and

---

[22]The MIV estimator is known to suffer from finite sample bias that will tend to yield narrower bounds relative to the true bounds (Manski & Pepper, 2000, 2009). Kreider & Pepper (2007) propose a bias correction; however, Hirano & Porter (2012) argue that such bias corrections may be inappropriate as they may not eliminate the bias and may substantially increase the variance. For this reason, we do not pursue a bias-corrected version of the MIV estimator in our analysis. Moreover, the Monte Carlo results in Manski & Pepper (2009) suggest that the bias will generally be small (less than 0.05) based on our sample sizes. A bias of this magnitude will not affect our qualitative findings.

UB_0, respectively.

6. Obtain bounds for the ATE given by

$$ATE \in [LB_1 - UB_0, UB_1 - LB_0].$$

Tables and Figures

Table 1: **Descriptive Statistics by SNAP Status**[a]

	SNAP Participants	Non-SNAP Participants	Ever Participated	Never Participated
Total Net Family Income[b]	16,132	24,450	17,140	25,962
	(14,405)	(15,637)	(14,375)	(15,586)
Age (years)	38.18	38.56	38.53	38.47
	(3.70)	(3.57)	(3.64)	(3.57)
Household Size	3.52	3.37	3.41	3.41
	(1.87)	(1.85)	(1.92)	(1.82)
Number of Children	1.82	1.60	1.72	1.60
	(1.55)	(1.48)	(1.56)	(1.47)
Education (% High School or More)	0.721	0.823	0.741	0.837
Mother's Education (% High School or More)	0.387	0.490	0.402	0.507
Currently Employed (%)	0.483	0.772	0.571	0.798
Married (%)	0.249	0.422	0.272	0.452
Urban (%)	0.696	0.659	0.690	0.653
Hispanic (%)	0.099	0.095	0.100	0.093
Black (%)	0.346	0.221	0.325	0.202
Female (%)	0.705	0.517	0.671	0.489
WIC (%)	0.162	0.054	0.123	0.048
Infant (%)[c]	0.264	0.196	0.230	0.197
Elderly (%)[d]	0.159	0.211	0.185	0.212
Obese (%)[e]	0.385	0.293	0.357	0.285
Overweight (%)[f]	0.677	0.640	0.664	0.638
Number of Observations	2,505	9,510	4,648	7,367

[a]Based on waves 1996-2004, excluding individuals or households with income greater than 250% of the federal poverty line. Statistics represent weighted sample means with standard deviations in parenthesis.

[b]Annual income in 2004 dollars adjusted by the CPI.

[c]Measured as the percent of individuals with at least one child under 5 years old in the household.

[d]Measured as the percent of individuals with at least one adult over 65 years old in the household.

[e]BMI 30 or more.

[f]BMI 25 or more.

Table 2: **Misclassification of SNAP Participation**[a]

Year(s)		Full Sample	Females	Males
All Years	$\hat{\alpha}_0$	0.043***	0.064***	0.056***
		(0.008)	(0.008)	(0.014)
	$\hat{\alpha}_1$	0.289***	0.397***	0.696***
		(0.096)	(0.097)	(0.121)
	N	12,015	6,832	5,183
1996	$\hat{\alpha}_0$	0.072***	0.108***	0.079**
		(0.025)	(0.032)	(0.034)
	$\hat{\alpha}_1$	0.389***	0.312***	0.568**
		(0.111)	(0.085)	(0.239)
	N	2,578	1,450	1,128
1998	$\hat{\alpha}_0$	0.036***	0.045***	0.042**
		(0.018)	(0.016)	(0.017)
	$\hat{\alpha}_1$	0.351***	0.439***	0.667***
		(0.126)	(0.111)	(0.240)
	N	2,666	1,530	1,136
2000	$\hat{\alpha}_0$	0.033	0.055**	0.038**
		(0.022)	(0.022)	(0.015)
	$\hat{\alpha}_1$	0.240	0.454***	0.657**
		(0.160)	(0.126)	(0.263)
	N	2,322	1,315	1,007
2002	$\hat{\alpha}_0$	0.052**	0.098***	0.041*
		(0.022)	(0.020)	(0.022)
	$\hat{\alpha}_1$	0.225	0.429***	0.719***
		(0.162)	(0.092)	(0.256)
	N	2,258	1,289	969
2004	$\hat{\alpha}_0$	0.064***	0.090***	0.077***
		(0.019)	(0.018)	(0.019)
	$\hat{\alpha}_1$	0.447***	0.465***	0.577***
		(0.124)	(0.074)	(0.187)
	N	2,191	1,248	943

[a]Based on waves 1996-2004, excluding individuals or households with income greater than 250% of the federal poverty line. Misreporting probabilities estimated using the methods in Hausman *et al.* (1998) with nonlinear least squares and Normal CDF. Boostrap standard errors in parentheses based on 200 iterations. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Table 3: **Regression Results for the Effect of SNAP Participation on P(Obese)**[a]

| | FE-IV | | | |
FE	Biometric	Direct Mail	Both IVs	MA-FE
Full Sample, $N = 12,015$				
0.020**	0.142	-0.221	0.001	0.008
(0.010)	(0.212)	(0.280)	(0.170)	(0.017)
Females, $N = 6,832$				
0.004	0.188	-0.203	0.070	-0.005
(0.012)	(0.194)	(0.289)	(0.162)	(0.017)
Males, $N = 5,183$				
0.054***	-0.089	-0.299	-0.210	0.004
(0.016)	(0.693)	(0.689)	(0.510)	(0.024)

[a]Based on waves 1996-2004, excluding individuals or households with income greater than 250% of the federal poverty line. Standard errors in parenthesis are clustered at the individual level. Additional covariates excluded from the table include household size, number of children, log income and log income squared, respondent's age, respondent education, and indicators for whether respondent is married, employed, living in an urban area, receives WIC benefits, has an elderly individual in the home, or has an infant in the home. Regressions for men exclude the WIC indicator, and year dummies are included in all regressions. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 4: **Regression Results for the Effect of SNAP Participation on P(Overweight)**[a]

| | FE-IV | | | |
FE	Biometric	Direct Mail	Both IVs	MA-FE
Full Sample, $N = 12,015$				
-0.006	0.410*	0.188	0.324*	-0.035*
(0.010)	(0.245)	(0.286)	(0.196)	(0.019)
Females, $N = 6,832$				
-0.005	0.301	0.164	0.259	-0.003
(0.012)	(0.206)	(0.302)	(0.181)	(0.017)
Males, $N = 5,183$				
-0.004	0.838	0.263	0.507	-0.020
(0.017)	(1.009)	(0.676)	(0.588)	(0.022)

[a]Based on waves 1996-2004, excluding individuals or households with income greater than 250% of the federal poverty line. Standard errors in parenthesis are clustered at the individual level. Additional covariates excluded from the table include household size, number of children, log income and log income squared, respondent's age, respondent education, and indicators for whether respondent is married, employed, living in an urban area, receives WIC benefits, has an elderly individual in the home, or has an infant in the home. Regressions for men exclude the WIC indicator, and year dummies are included in all regressions. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$

Table 5: **First-stage Results for IV Estimates**[a]

	Biometric Only	Direct Mail Only	Both
		Instrument Set	
Full Sample, $N = 12,015$			
Biometric	-0.059***		-0.053***
	(0.016)		(0.016)
Direct Mail		-0.054***	-0.047**
		(0.190)	(0.019)
Global F-statistic	13.18***	8.61***	9.47***
Hansen's J-statistic		Obese	1.273
		Overweight	0.392
Females, $N = 6,832$			
Biometric	-0.089***		-0.082***
	(0.023)		(0.023)
Direct Mail		-0.070**	-0.057**
		(0.277)	(0.028)
Global F-statistic	14.53***	6.32**	9.23***
Hansen's J-statistic		Obese	1.505
		Overweight	0.158
Males, $N = 5,183$			
Biometric	-0.025		-0.022
	(0.020)		(0.020)
Direct Mail		-0.034	-0.031
		(0.022)	(0.022)
Global F-statistic	1.58	2.34	1.68
Hansen's J-statistic		Obese	0.051
		Overweight	0.302

[a]Based on waves 1996-2004, excluding individuals or households with income greater than 250% of the federal poverty line. Standard errors in parenthesis are clustered at the individual level. Additional covariates in the first-stage regressions included household size, number of children, log income and log income squared, respondent's age, respondent education, and indicators for whether respondent is married, employed, living in an urban area, receives WIC benefits, has an elderly individual in the home, or has an infant in the home. Regressions for men exclude the WIC indicator, and year dummies are included in all regressions. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Table 6: **Nonparametric Bounds for Full Sample**, $y = \mathbf{1}(\text{Obese})^a$

Error Rates	Arbitrary Misreporting		No False Positives	
Exogenous Selection				
0	[0.063,	0.063]	[0.063,	0.063]
.01	[0.024,	0.100]	[0.041,	0.100]
.02	[-0.018,	0.134]	[0.020,	0.134]
.05	[-0.169,	0.226]	[-0.034,	0.226]
.1	[-0.397,	0.459]	[-0.111,	0.358]
Worst-case Selection				
0	[-0.386,	0.614]	[-0.386,	0.614]
.01	[-0.396,	0.624]	[-0.396,	0.624]
.02	[-0.406,	0.634]	[-0.406,	0.634]
.05	[-0.436,	0.664]	[-0.436,	0.664]
.1	[-0.486,	0.714]	[-0.486,	0.714]
Negative Monotone Treatment Selection (MTSn)				
0	[-0.386,	0.063]	[-0.386,	0.063]
.01	[-0.396,	0.100]	[-0.396,	0.100]
.02	[-0.406,	0.134]	[-0.406,	0.134]
.05	[-0.436,	0.226]	[-0.436,	0.226]
.1	[-0.486,	0.459]	[-0.486,	0.358]
Monotone Instrumental Variable with MTSn				
0	[-0.276,	-0.016]	[-0.276,	-0.016]
.01	[-0.296,	0.032]	[-0.286,	0.032]
.02	[-0.316,	0.072]	[-0.296,	0.073]
.05	[-0.368,	0.128]	[-0.326,	0.142]
.1	[-0.442,	0.225]	[-0.376,	0.235]

[a] Based on waves 1996-2004, excluding individuals or households with income greater than 250% of the federal poverty line. Terms in brackets reflect the lower and upper bounds of the ATE of SNAP on the probability of being obese (BMI $\geq$ 30). See Appendix for details on the estimation.

Table 7: **Nonparametric Bounds for Females**, $y = \mathbf{1}(\text{Obese})^a$

Error Rates	Arbitrary Misreporting		No False Positives	
	Exogenous Selection			
0	[0.042,	0.042]	[0.042,	0.042]
.01	[0.010,	0.071]	[0.022,	0.071]
.02	[-0.023,	0.100]	[0.003,	0.100]
.05	[-0.134,	0.180]	[-0.050,	0.180]
.1	[-0.384,	0.329]	[-0.128,	0.301]
.2	[-0.504,	0.729]	[-0.272,	0.531]
	Worst-case Selection			
0	[-0.428,	0.572]	[-0.428,	0.572]
.01	[-0.438,	0.582]	[-0.438,	0.582]
.02	[-0.448,	0.592]	[-0.448,	0.592]
.05	[-0.478,	0.622]	[-0.478,	0.622]
.1	[-0.528,	0.672]	[-0.528,	0.672]
.2	[-0.628,	0.772]	[-0.628,	0.772]
	Negative Monotone Treatment Selection (MTSn)			
0	[-0.428,	0.042]	[-0.428,	0.042]
.01	[-0.438,	0.071]	[-0.438,	0.071]
.02	[-0.448,	0.100]	[-0.448,	0.100]
.05	[-0.478,	0.180]	[-0.478,	0.180]
.1	[-0.528,	0.329]	[-0.528,	0.301]
.2	[-0.628,	0.729]	[-0.628,	0.531]
	Monotone Instrumental Variable with MTSn			
0	[-0.336,	-0.036]	[-0.336,	-0.036]
.01	[-0.356,	0.029]	[-0.346,	0.029]
.02	[-0.376,	0.047]	[-0.356,	0.048]
.05	[-0.431,	0.088]	[-0.386,	0.101]
.1	[-0.508,	0.161]	[-0.436,	0.186]
.2	[-0.628,	0.345]	[-0.536,	0.339]

[a]Based on waves 1996-2004 for females only, excluding individuals or households with income greater than 250% of the federal poverty line. Terms in brackets reflect the lower and upper bounds of the ATE of SNAP on the probability of being obese (BMI $\geq$ 30). See Appendix for details on the estimation.

Table 8: **Nonparametric Bounds for Males**, $y = \mathbf{1}(\text{Obese})^a$

Error Rates	Arbitrary Misreporting		No False Positives	
Exogenous Selection				
0	[0.054,	0.054]	[0.054,	0.054]
.01	[-0.010,	0.110]	[0.027,	0.110]
.02	[-0.084,	0.160]	[0.002,	0.160]
.05	[-0.317,	0.283]	[0.057,	0.283]
.1	[-0.335,	0.756]	[-0.128,	0.438]
.3	[-0.437,	0.959]	[-0.327,	0.769]
Worst-case Selection				
0	[-0.330,	0.670]	[-0.330,	0.670]
.01	[-0.340,	0.680]	[-0.340,	0.680]
.02	[-0.350,	0.690]	[-0.350,	0.690]
.05	[-0.380,	0.720]	[-0.380,	0.720]
.1	[-0.430,	0.770]	[-0.430,	0.770]
.3	[-0.630,	0.970]	[-0.630,	0.914]
Negative Monotone Treatment Selection (MTSn)				
0	[-0.330,	0.054]	[-0.330,	0.054]
.01	[-0.340,	0.110]	[-0.340,	0.110]
.02	[-0.350,	0.160]	[-0.350,	0.160]
.05	[-0.380,	0.283]	[-0.380,	0.283]
.1	[-0.430,	0.756]	[-0.430,	0.438]
.3	[-0.630,	0.959]	[-0.630,	0.769]
Monotone Instrumental Variable with MTSn				
0	[-0.253,	-0.054]	[-0.253,	-0.054]
.01	[-0.273,	0.030]	[-0.263,	0.030]
.02	[-0.292,	0.081]	[-0.273,	0.081]
.05	[-0.337,	0.211]	[-0.303,	0.194]
.1	[-0.404,	0.311]	[-0.353,	0.320]
.3	[-0.610,	0.895]	[-0.553,	0.643]

[a]Based on waves 1996-2004 for males only, excluding individuals or households with income greater than 250% of the federal poverty line. Terms in brackets reflect the lower and upper bounds of the ATE of SNAP on the probability of being obese (BMI $\geq$ 30). See Appendix for details on the estimation.

Table 9: **Nonparametric Bounds for Full Sample**, $y = \mathbf{1}(\text{Overweight})^a$

Error Rates	Arbitrary Misreporting		No False Positives	
Exogenous Selection				
0	[0.008,	0.008]	[0.008,	0.008]
.01	[-0.033,	0.051]	[-0.033,	0.026]
.02	[-0.071,	0.098]	[-0.071,	0.043]
.05	[-0.173,	0.268]	[-0.173,	0.088]
.1	[-0.317,	0.381]	[-0.317,	0.152]
Worst-case Selection				
0	[-0.607,	0.393]	[-0.607,	0.393]
.01	[-0.617,	0.403]	[-0.617,	0.403]
.02	[-0.627,	0.413]	[-0.627,	0.413]
.05	[-0.657,	0.443]	[-0.657,	0.443]
.1	[-0.707,	0.493]	[-0.707,	0.493]
Negative Monotone Treatment Selection (MTSn)				
0	[-0.607,	0.008]	[-0.607,	0.008]
.01	[-0.617,	0.051]	[-0.617,	0.026]
.02	[-0.627,	0.098]	[-0.627,	0.043]
.05	[-0.657,	0.268]	[-0.657,	0.088]
.1	[-0.707,	0.381]	[-0.707,	0.152]
Monotone Instrumental Variable with MTSn				
0	[-0.531,	-0.070]	[-0.531,	-0.070]
.01	[-0.551,	-0.023]	[-0.541,	-0.046]
.02	[-0.571,	0.029]	[-0.551,	-0.025]
.05	[-0.631,	0.105]	[-0.581,	0.030]
.1	[-0.707,	0.222]	[-0.631,	0.078]

^{a}Based on waves 1996-2004, excluding individuals or households with income greater than 250% of the federal poverty line. Terms in brackets reflect the lower and upper bounds of the ATE of SNAP on the probability of being overweight (BMI $\geq$ 25). See Appendix for details on the estimation.

Table 10: **Nonparametric Bounds for Females**, $y = \mathbf{1}$(Overweight)[a]

Error Rates	Arbitrary Misreporting		No False Positives	
Exogenous Selection				
0	[0.035,	0.035]	[0.035,	0.035]
.01	[0.000,	0.071]	[0.000,	0.050]
.02	[-0.032,	0.108]	[-0.032,	0.065]
.05	[-0.123,	0.236]	[-0.123,	0.107]
.1	[-0.259,	0.413]	[-0.259,	0.171]
.2	[-0.737,	0.473]	[-0.512,	0.292]
Worst-case Selection				
0	[-0.566,	0.434]	[-0.566,	0.434]
.01	[-0.576,	0.444]	[-0.576,	0.444]
.02	[-0.586,	0.454]	[-0.586,	0.454]
.05	[-0.616,	0.484]	[-0.616,	0.484]
.1	[-0.666,	0.534]	[-0.666,	0.534]
.2	[-0.766,	0.634]	[-0.766,	0.634]
Negative Monotone Treatment Selection (MTSn)				
0	[-0.566,	0.035]	[-0.566,	0.035]
.01	[-0.576,	0.071]	[-0.576,	0.050]
.02	[-0.586,	0.108]	[-0.586,	0.065]
.05	[-0.616,	0.236]	[-0.616,	0.107]
.1	[-0.666,	0.413]	[-0.666,	0.171]
.2	[-0.666,	0.473]	[-0.766,	0.292]
Monotone Instrumental Variable with MTSn				
0	[-0.491,	-0.041]	[-0.491,	-0.041]
.01	[-0.511,	-0.007]	[-0.501,	-0.025]
.02	[-0.531,	0.025]	[-0.511,	-0.009]
.05	[-0.591,	0.107]	[-0.541,	0.031]
.1	[-0.666,	0.227]	[-0.591,	0.088]
.2	[-0.766,	0.399]	[-0.691,	0.173]

[a]Based on waves 1996-2004 for females only, excluding individuals or households with income greater than 250% of the federal poverty line. Terms in brackets reflect the lower and upper bounds of the ATE of SNAP on the probability of being overweight (BMI $\geq$ 25). See Appendix for details on the estimation.

Table 11: **Nonparametric Bounds for Males**, $y = \mathbf{1}(\text{Overweight})^{a}$

Error Rates	Arbitrary Misreporting		No False Positives	
	Exogenous Selection			
0	[-0.026,	-0.026]	[-0.026,	-0.026]
.01	[-0.083,	0.040]	[-0.083,	0.000]
.02	[-0.135,	0.116]	[-0.135,	0.023]
.05	[-0.261,	0.321]	[-0.261,	0.079]
.1	[-0.749,	0.340]	[-0.418,	0.148]
.3	[-0.949,	0.444]	[-0.765,	0.343]
	Worst-case Selection			
0	[-0.662,	0.338]	[-0.662,	0.338]
.01	[-0.672,	0.348]	[-0.672,	0.348]
.02	[-0.682,	0.358]	[-0.682,	0.358]
.05	[-0.712,	0.388]	[-0.712,	0.388]
.1	[-0.762,	0.438]	[-0.762,	0.438]
.3	[-0.962,	0.638]	[-0.911,	0.638]
	Negative Monotone Treatment Selection (MTSn)			
0	[-0.662,	-0.026]	[-0.662,	-0.026]
.01	[-0.672,	0.040]	[-0.672,	0.000]
.02	[-0.682,	0.116]	[-0.682,	0.023]
.05	[-0.712,	0.321]	[-0.712,	0.079]
.1	[-0.762,	0.340]	[-0.762,	0.148]
.3	[-0.962,	0.444]	[-0.911,	0.343]
	Monotone Instrumental Variable with MTSn			
0	[-0.596,	-0.130]	[-0.596,	-0.130]
.01	[-0.616,	-0.025]	[-0.606,	-0.072]
.02	[-0.636,	0.032]	[-0.616,	-0.036]
.05	[-0.696,	0.166]	[-0.646,	0.022]
.1	[-0.762,	0.301]	[-0.696,	0.085]
.3	[-0.962,	0.388]	[-0.879,	0.251]

[a]Based on waves 1996-2004 for males only, excluding individuals or households with income greater than 250% of the federal poverty line. Terms in brackets reflect the lower and upper bounds of the ATE of SNAP on the probability of being overweight (BMI ≥ 25). See Appendix for details on the estimation.

Table 12: **Summary of Estimated Effects**[a]

	FE	FE-IV[b]	MA-FE	MTSn		MIV-MTSn	
				Obese - Full Sample			
	0.020**	0.001	0.008	[-0.486,	0.358]	[-0.376,	0.235]
	(0.010)	(0.170)	(0.017)				
				Obese - Females			
	0.004	0.070	-0.005	[-0.628,	0.531]	[-0.536,	0.339]
	(0.012)	(0.162)	(0.017)				
				Obese - Males			
	0.054***	-0.210	0.004	[-0.630,	0.769]	[-0.553,	0.643]
	(0.016)	(0.510)	(0.024)				
				Overweight - Full Sample			
	-0.006	0.324*	-0.035*	[-0.707,	0.152]	[-0.631,	0.078]
	(0.010)	(0.196)	(0.019)				
				Overweight - Females			
	-0.005	0.259	-0.003	[-0.766,	0.292]	[-0.691,	0.173]
	(0.012)	(0.181)	(0.017)				
				Overweight - Males			
	-0.004	0.507	-0.020	[-0.911,	0.343]	[-0.879,	0.251]
	(0.017)	(0.588)	(0.022)				

[a]Estimates taken from the detailed results in Tables 3 - 11. The bounds in columns MTSn and MIV-MTSn are based on misreporting rates of 10% for the full sample, 20% among females, and 30% among males (with no false positives in all cases).

[b]Results presented when using both instruments (direct mail and biometric identification) in the FE-IV estimation.

Table 13: **Descriptive Statistics by SNAP Status in Alternative Sample**[a]

	SNAP Participants	Non-SNAP Participants
Total Net Family Income[b]	12,316	20,397
	(8,265)	(12,699)
Age (years)	38.38	38.81
	(3.68)	(3.54)
Household Size	3.73	3.51
	(1.92)	(1.93)
Number of Children	2.14	1.83
	(1.60)	(1.58)
Education (% High School or More)	0.697	0.799
Mother's Education (% High School or More)	0.330	0.461
Currently Employed (%)	0.450	0.772
Married (%)	0.237	0.411
Urban (%)	0.658	0.625
Hispanic (%)	0.096	0.094
Black (%)	0.392	0.219
Female (%)	0.761	0.527
WIC (%)	0.171	0.068
Infant (%)[c]	0.274	0.192
Elderly (%)[d]	0.132	0.185
Obese (%)[e]	0.423	0.337
Number of Observations	978	2290

[a]Based on a balanced panel over waves 1996-2004, excluding individuals or households with income greater than 250% of the federal poverty line in any wave. Statistics represent weighted sample means with standard deviations in parenthesis.

[b]In 2008 dollars adjusted by the CPI

[c]Measured as the percent of individuals with at least one child under 5 years old in the household.

[d]Measured as the percent of individuals with at least one adult over 65 years old in the household.

[e]BMI 30 or more.

Table 14: **Summary of Estimated Effects with Alternative Sample**[a]

	FE	FE-IV[b]	MA-FE	MTSn		MIV-MTSn	
			Full Sample, $N = 3,268$				
	0.042**	0.073	0.015	[-0.541,	0.291]	[-0.423,	0.164]
	(0.018)	(0.276)	(0.018)				
			Females, $N = 2,030$				
	0.038*	0.346	0.017	[-0.680,	0.505]	[-0.578,	0.240]
	(0.021)	(0.364)	(0.027)				
			Males, $N = 1,238$				
	0.049	-0.305	-0.017	[-0.676,	0.744]	[-0.624,	0.564]
	(0.031)	(0.419)	(0.037)				

[a]Based on a balanced panel over waves 1996-2004, excluding individuals or households with income greater than 250% of the federal poverty line in any wave. The bounds in columns MTSn and MIV-MTSn are based on misreporting rates of 10% for the full sample, 20% among females, and 30% among males, with no false positives in all cases.

[b]Results presented when using both instruments (direct mail and biometric identification) in the FE-IV estimation.

Table 15: **Estimated Effects of SNAP Participation with Misreported Outcome**[a]

	FE	FE-IV	MA-FE	MTSn		MIV-MTSn[b]	
$y = 1(\text{Obese} = 1)$							
	0.014	0.043	0.006	[-0.507,	0.344]	[-0.401,	0.236]
	(0.009)	(0.178)	(0.018)				
$y = 1(\text{Overweight} = 1)$							
	0.002	0.217	-0.027	[-0.717,	0.150]	[-0.630,	0.091]
	(0.010)	(0.187)	(0.019)				

[a]Based on waves 1996-2004, excluding individuals or households with income greater than 250% of the federal poverty line. Self reported height and weight for our NLSY79 sample are adjusted by gender and race using similar aged respondents from the Third National Health and Nutrition Examination Survey (Cawley, 2004). Standard errors in parenthesis are clustered at the individual level. Additional covariates excluded from the table include household size, number of children, log income and log income squared, respondent's age, respondent education, and indicators for whether respondent is married, employed, living in an urban area, receives WIC benefits, has an elderly individual in the home, or has an infant in the home. Year dummies are included in all regressions. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

[b]Results based on 10% misreporting rates and no false positives, with income to poverty line ratio as the MIV.

Table 16: **Estimated Effect of SNAP Participation with Alternative Eligibility**[a]

	FE	FE-IV	MA-FE	MTSn		MIV-MTSn[b]	
$y = 1(\text{Obese} = 1)$							
250% of FPL	0.020**	0.001	0.008	[-0.486,	0.358]	[-0.376,	0.235]
	(0.010)	(0.170)	(0.017)				
185% of FPL	0.024**	0.094	0.024	[-0.506,	0.335]	[-0.306,	0.221]
	(0.011)	(0.144)	(0.019)				
130% of FPL	0.027**	0.190	0.014	[-0.522,	0.337]	[-0.272,	0.207]
	(0.013)	(0.200)	(0.022)				
$y = 1(\text{Overweight} = 1)$							
250% of FPL	-0.006	0.324*	-0.035*	[-0.707,	0.152]	[-0.631,	0.078]
	(0.010)	(0.196)	(0.019)				
185% of FPL	-0.004	0.195	0.009	[-0.677,	0.145]	[-0.561,	0.079]
	(0.011)	(0.160)	(0.019)				
130% of FPL	-0.004	0.195	-0.030	[-0.640,	0.145]	[-0.468,	0.057]
	(0.012)	(0.223)	(0.021)				

[a]Based on waves 1996-2004, excluding individuals or households with income greater than 250%, 185%, and 130% of the federal poverty line. Standard errors in parenthesis are clustered at the individual level. Additional covariates excluded from the table include household size, number of children, log income and log income squared, respondent's age, respondent education, and indicators for whether respondent is married, employed, living in an urban area, receives WIC benefits, has an elderly individual in the home, or has an infant in the home. FE-IV estimates include both the direct mail and biometric indicators as instruments, and year dummies are included in all regressions. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

[b]Results based on 10% misreporting rates and no false positives, with income to poverty line ratio as the MIV.