Partial Identification Methods for Evaluating Food Assistance Programs: A Case Study of the Causal Impact of SNAP on Food Insecurity

Craig Gundersen, Brent Kreider, and John V. Pepper

We illustrate how partial identification methods can be used to provide credible inferences on the causal impacts of food assistance programs, focusing on the impact that the Supplemental Nutrition Assistance Program (SNAP, formerly known as the Food Stamp Program) has on food insecurity among households with children. Recent research applies these methods to address two key issues confounding identification: missing counterfactuals and nonrandomly misclassified treatment status. In this paper, we illustrate and extend the recent literature by using data from the Survey of Income and Program Participation (SIPP) to study the robustness of prior conclusions. The SIPP confers important advantages: the detailed information about income and eligibility allows us to apply a modified discontinuity design to sharpen inferences, and the panel nature allows us to reduce uncertainty about true participation status. We find that SNAP reduces the prevalence of food insecurity in households with children by at least six percentage points.

Key words: Supplemental Nutrition Assistance Program, Food Stamp Program, food insecurity, partial identification, treatment effects, nonparametric bounds, classification error, Survey of Income and Program Participation.

JEL codes: C14, C21, I38.

Evaluations of food assistance programs must address fundamental identification problems that limit what can be directly inferred from observed data on program participation and outcomes of interest such as food security. Survey data are typically measured with substantial error, and even perfectly measured data cannot reveal counterfactual outcomes. To address these identification problems, data must be combined with economic and statistical assumptions.

The conventional approach to assessing the impact of food assistance programs has been to impose assumptions strong enough to yield a definitive finding, in which case the parameter of interest—for example, an average treatment effect—is point-identified. Researchers may, for example, treat selection into a government program as (conditionally) exogenous or, in some cases, apply a linear response model coupled with an orthogonality assumption that an instrumental variable (IV) affects the outcome only indirectly through program participation. Errors in participation status, if allowed, are classically treated as independent of unobserved true participation status, the outcome conditional on true participation status, and household attributes. Strong identifying assumptions, however, often reflect “convenience rather than conviction,” (Bound, Brown, and Mathiowetz 2001). When such assumptions do not hold, they may yield flawed, counterintuitive, and conflicting conclusions (Manski 2007; Manski and Pepper 2015).

Craig Gundersen is a professor and Soybean Industry Endowed Professor in Agricultural Strategy in the Department of Agricultural and Consumer Economics, University of Illinois. Brent Kreider is a professor in the Department of Economics, Iowa State University. John V. Pepper is a professor in the Department of Economics, University of Virginia. The authors acknowledge financial support from the Survey of Income and Program Participation Analytic Research Small Grants Competition of the National Poverty Center. Pepper also acknowledges financial support from the Bankard Fund for Political Economy. The views expressed in this paper are solely those of the authors. Correspondence may be sent to: cggunder@illinois.edu.

Amer. J. Agr. Econ. 99(4): 875–893; doi: 10.1093/ajae/aax026
Published online April 25, 2017

These concerns are especially salient in the literature evaluating food assistance programs. Consider, for example, the problem of evaluating the impact that the Supplemental Nutritional Assistance Program (SNAP, formerly known as the Food Stamp Program) has on food insecurity, the focus of this paper. As the cornerstone of the federal food assistance system in the United States, SNAP is intended to serve as the first line of defense against hunger, especially for children, who constitute 60% of all recipients.[1] Yet much of the literature finds that children residing in households receiving SNAP are more likely to be food insecure and suffer from an array of health-related problems than observationally similar nonparticipating households with children (e.g., Currie 2003; Coleman-Jensen, Gregory, and Singh 2014).

Two key identification problems are known to confound inferences of the causal impacts of SNAP: endogenous selection into the program and systematic underreporting of participation status. First, a high proportion of eligible households with children do not participate in the program, and a household's decision to participate in SNAP is not random. Second, there is extensive nonrandom underreporting of SNAP participation status in national surveys (see, e.g., Bollinger and David 1997; Meyer, Mock, and Sullivan 2009; Meyer, Goerge, and Mittag 2015).

Credible solutions to these selection and classification error problems remain elusive (e.g., Kreider et al. 2012). For some assistance programs, there is substantial variation across time, space, or both. In contrast, SNAP has remained relatively unchanged since the 1980s and the rules are primarily established at the federal level. States do have some leeway (e.g., setting gross income cutoffs, whether or not to have asset tests), but policy actions to change SNAP rules may endogenously reflect state-level food insecurity rates and, hence, may not be valid instruments. And even if an instrument is technically valid, it may suffer from a weak instruments problem: it is difficult to find plausibly exogenous measures that are highly correlated with SNAP participation. Finally, addressing the problem of classification errors is difficult. The classical model assumption of non-mean-reverting errors cannot apply with binary variables, and the systematic underreporting of SNAP participation violates the classical assumption that measurement error arises independently of the true value of the underlying variable. While prior research has considered the endogenous selection problem using a wide array of instruments, almost no research has assessed how household reporting errors may affect inferences about the impacts of the program.[2]

In this setting, we have found it useful to consider inference under a spectrum of assumptions of varying identifying power regarding self-selection and reporting error. Analysis of partial identification of treatment effects has developed and been applied for nearly three decades, beginning with Manski (1990).[3] The basic insight of partial identification analysis is that identification need not be an all or nothing undertaking (Manski 2007). Rather, the approach addresses the tension between the strength of assumptions and their credibility. Available data and credible assumptions may not point-identify average treatment effects, but they do partially identify them, yielding bounds rather than point estimates. In some cases, the results imply little uncertainty, whereas in others the data and credible assumptions may not be able to identify even the direction of the causal impact of a program. As usual, sampling variability represents an additional layer of uncertainty that can be reflected through confidence intervals.

Recently, there has been a growing interest in applying these methods to study the impact of food assistance programs in the United States. Kreider et al. (2012), for example, develop new partial identification methods to examine the impact of SNAP on food insecurity and other outcomes (see also, e.g., Gundersen, Kreider, and Pepper 2012;

[1] By far the largest food assistance program in the United States, SNAP provided assistance to almost 19 million children in 2011, the most recent year for which data are available (Eslami and Cunnyngham 2014). Nearly half of all children will receive assistance at some point during their childhood (Rank and Hirschl 2009).

[2] A number of papers address the selection problem using instrumental variables and nonlinear selection models. Borjas (2004), Gundersen and Oliveira (2001), Jensen (2002), Hoynes and Schanzenbach (2009), Meyerhoefer and Pylypchuk (2008), Mykerezi and Mills (2010), Van Hook and Balistreri (2006), and Yen et al. (2008), for example, address the selection problem using instrumental variables within a linear response model, and Devaney and Moffitt (1991) and Ratcliffe, McKernan, and Zhang (2011) use nonlinear selection models.

[3] See Manski (2007) for textbook expositions and, e.g., Manski and Nagin (1998), Manski and Pepper (2000), Pepper (2000), Blundell et al. (2007), Molinari (2008), and Kreider et al. (2012) for applications.

Almada, McCarthy, and Tchernis 2016). These methods do not require the linear response model, classical measurement error assumptions, a traditional IV exclusion, or any distributional assumptions. Instead, Kreider et al. (2012) apply weaker models that are straightforward to motivate in practice and result in informative bounds on the average effect of SNAP on food security.

In this paper, we build on Kreider et al. (2012) to illustrate how partial identification methods can be used to evaluate the impact of SNAP on the food security of children using data from the Survey of Income and Program Participation (SIPP). Kreider et al. (2012)'s estimates are obtained using data from cross-sections taken from the National Health and Nutrition Examination Survey (NHANES) and Current Population Survey (CPS). The SIPP confers important advantages that allow us to update and improve upon the earlier analyses for food insecurity outcomes.[4] First, the panel nature of the SIPP allows us to use information across multiple waves to reduce the degree of uncertainty about the reliability of self-reported program participation status. Second, the refined income and asset information in the SIPP allows us to sharpen inferences by applying a modified regression discontinuity design (see Gundersen, Kreider, and Pepper 2012).[5] In this framework, we construct a monotone instrumental variable (MIV) based on SNAP eligibility that does not require an orthogonality restriction. Unlike the usual discontinuity design, we do not achieve point identification of an average treatment effect at the discontinuity. As described below, however, the income-eligibility cutoff does allow us to bound a counterfactual distribution. Finally, we formally highlight the identifying power of the exogenous selection assumption imposed in some of the literature. Under this assumption, the average treatment effect is point-identified if SNAP receipt is accurately classified, but partially identified otherwise.

Data

Our analysis uses data from the 2004 SIPP which surveyed 46,500 households eight times over a two-and-a-half year period. Conducted by the U.S. Census Bureau, the SIPP is comprised of a series of national panels designed to measure the effectiveness of government programs such as SNAP on the well-being of households in the United States.[6]

We focus on households with children that are income- and asset-eligible for SNAP.[7] To be eligible for assistance, a household's gross income before taxes in the previous month cannot exceed 130% of the poverty line, and assets generally must be less than $2,000. The central variable of interest in this paper is food insecurity, and it is only available in the Supplement in Wave 5.[8] We therefore use information on household structure and income from the Wave 5 Core Module to identify households with children who are eligible for SNAP based on the gross income criterion.[9,10]

[4] The NHANES data were especially useful in Kreider et al. (2012) because that analysis studied a wide array of health outcomes, including obesity and anemia. The primary strength of the NHANES is the wealth of health-related information collected from medical examinations, physiological measurements, and laboratory tests. For example, data on height and weight were collected by trained personnel, and anemia was measured using a blood test. The CPS was used in order to conduct robustness checks using state-level instruments.

[5] A central goal of the SIPP is a more comprehensive and accurate portrayal of income receipt in the United States; please see http://www.census.gov/sipp/intro.html. As such, there is less misreporting on income, assets, SNAP eligibility, and SNAP participation in the SIPP than in other commonly used datasets (Meyer, Goerge, and Mitta 2015). The SIPP measures income as a continuous variable and includes information on assets, while the NHANES measures income as a categorical variable and does not include detailed measures of assets. Also, the SIPP uses a four-month recall window for household responses in contrast to the annual window used in the NHANES and CPS. This four-month window has been shown to mitigate recall errors (Goerge, Franzetta, and Dilts 2009; Almada, McCarthy, and Tchernis 2016), is better aligned with the income used to measure SNAP eligibility, and mitigates potential time inconsistencies between SNAP participation and food insecurity (e.g., a respondent may be food insecure in March and take up SNAP in September).

[6] Other research on SNAP that uses data from the SIPP includes, for example, Bollinger and David (2001, 2005), Gundersen and Oliveira (2001), and Ribar and Hamrick (2003).

[7] When this survey was conducted, SNAP was called the Food Stamp Program. We use the current name SNAP to avoid confusion.

[8] Because the food insecurity questions are only asked in one wave, panel data methods such as fixed effects models are not feasible.

[9] Households must also meet a net income test, but nearly all gross income-eligible households with children are also net income-eligible (Gundersen and Offutt 2005). In years after these data were collected in the SIPP, states were allowed to set higher gross-income eligibility thresholds and waive the asset test (although the net income threshold would still be binding). Neither of these cases hold for the time period of this analysis. We do not include higher-income households that may have become adjunctively eligible for SNAP by virtue of participating in another assistance program such as Supplemental Security Income (SSI) or Temporary Assistance for Needy Families (TANF).

[10] Measurement error may lead to some contamination of our eligibility indicator (see Gundersen, Kreider, and Pepper 2012). In addition to the standard problems of accurately measuring income in self-reported surveys, there may be a mismatch between the time periods used in the SIPP to determine eligibility and the actual time periods used to determine eligibility. This mismatch

Table 1. Weighted Means and Standard Deviations by Reported Supplemental Nutrition Assistance Program (SNAP) Participation

	Income-eligible children	Recipients	Nonrecipients	Difference
Income to poverty ratio	0.715 (0.393)	0.620* (0.364)	0.797 (0.400)	0.177*
SNAP Recipient	0.466 (0.499)			
Ever received SNAP	0.626 (0.484)		0.300 (0.458)	
Food insufficient household	0.081 (0.273)	0.094 (0.292)	0.069 (0.254)	0.024*
Child not eating enough	0.115 (0.319)	0.134 (0.341)	0.099 (0.299)	0.035*
Either measure	0.150 (0.357)	0.177 (0.382)	0.126 (0.332)	0.051*
N	2,589	1,288	1,301	

Note: The difference column displays the difference between the recipient and nonrecipient estimates. Asterisk * indicates statistically significantly different means between recipients and nonrecipients with p-values less than 0.01 based on Wald statistics corrected for the sample design.

In total, 2,936 children reside in income-eligible households. Detailed information from Topical Modules in Waves 3 and 6 also allows us to further restrict the sample to asset-eligible households with children.[11] Of the 2,936 children residing in income-eligible households, not quite 10%, 283, are asset-ineligible. Thus, our primary sample includes information on 2,653 children who reside in income- and asset-eligible households. While our sample is restricted to Wave 5 respondents, SNAP participation is reported in all waves. We use this information from the other waves to construct our "ever reported receiving SNAP" indicator described below.

Table 1 displays means and standard deviations for the key variables used in this study. The estimates in this table and elsewhere are weighted to account for the survey design. For each respondent, we observe information on household income relative to the poverty line. Our sample has an average household income level equal to 71.5% of the poverty line, with respondents claiming to receive SNAP having notably less income than those claiming to have not received benefits from SNAP.[12] Two measures of SNAP receipt are presented. First, 46.6% of the eligible households with children report receiving benefits in Wave 5. Second, 62.6% of eligible households with children report receiving SNAP in at least one of the eight waves covered by the 2004 SIPP.[13]

The participation rate found in Wave 5 is similar to the contemporaneous rates found in other surveys (e.g., the CPS and NHANES) but lower than analogous rates found when administrative data are used to establish the number of participants (i.e., the numerator in the calculation of participation rates). Wolkwitz (2008) finds, for example, that just over half of all eligible households and 80% of eligible households with children participate. Differences between the participation rates from administrative and self-reported surveys are thought to largely reflect classification errors in the self-reported survey data (Trippe, Doyle, and Asher 1992; Bitler, Currie, and Scholz 2003; Taeuber et al. 2004, Meyer, Goerge, and Mittag 2015).[14] Bollinger and David (1997) provide direct evidence of misreporting in the SIPP by comparing individual reports of food stamp participation status with matched reports from administrative data (also see Marquis and Moore 1990; Meyer, Goerge, and Mittag 2015). They find that 12.0% of responses in the 1984 SIPP involve errors of omission (i.e., cases where households report that they do not receive SNAP when, in

is a common problem in evaluating eligibility for assistance programs using survey data. Still, regardless of the potential for errors in classifying eligibility, this sample restriction generates a well-defined subpopulation of interest—namely households with children reporting income less than 130% of the poverty line and assets less than $2,000.

[11] The value of a vehicle above a certain level may be considered a part of assets unless it is used for work or for the transportation of disabled persons. Our analysis does not include the value of respondents' vehicles when defining asset eligibility.

[12] To assess the characteristics of our sample relative to other national estimates, we pool data from six rounds of the 2001–2006 CPS, March Supplement. These data indicate that during this same time period, income-eligible children lived in families with an average income equal to 70% of the poverty line.

[13] While benefit amounts are sometimes imputed conditional on reported receipt, indicators of receipt themselves are not imputed.

[14] Differences between administrative and self-reported data might also reflect different time periods used to measure receipt. Administrative participation rates are often calculated on an annual basis, while the SIPP measures participation over a four-month window.

reality, they do) while only 0.3% involve errors of commission (i.e., cases where households report that they receive SNAP when, in reality, they do not).

Using data from the Wave 5 Topical Module, we observe two measures of food insecurity. First, the survey includes the food insufficiency question that has been included in numerous surveys since 1977. This question asks respondents to describe their food intake in terms of the following: "Which of these statements best describe the food eaten in your household in the last month?" Respondents have four choices: "Enough of the kinds of food we want to eat"; "Enough but not always the kinds of food we want to eat"; "Sometimes not enough to eat"; or "Often not enough to eat." Households reporting that they sometimes or often do not get enough to eat are defined as "food insufficient."[15] Second, we also focus on a variable indicating whether the child is not eating enough—the 13th question in the Core Food Security Module (CFSM; Coleman-Jensen, Gregory, and Singh 2014; appendix table A1).

Consistent with previous work on this topic, SNAP recipients in the data have worse food insecurity outcomes than eligible nonparticipants. For example, food insufficiency rates are 9.4% for households reported as SNAP recipients, 2.4 percentage points higher than the food insufficiency rate of 6.9% among eligible nonparticipants. Likewise, the fraction of children not eating enough is about 13.4% for children in household reporting to receive SNAP but 9.9% in households reporting not to have received benefits.

Inference on the Average Treatment Effect

Our interest is in learning about the average treatment effect (ATE) of SNAP participation on food insecurity among eligible households with children. Let $S^* = 1$ denote that the household receives SNAP benefits and $S^* = 0$ denote that the household does not receive SNAP benefits. Focusing on binary outcomes, the ATE can be expressed as

$$(1) \quad ATE(1,0|X \in \Omega) = P[FI(1) = 1|X \in \Omega] - P[FI(0) = 1|X \in \Omega]$$

where FI is the realized food insecurity outcome, $FI(1)$ denotes the potential food insecurity outcome if the household were to receive SNAP, $FI(0)$ denotes the analogous outcome if the household were not to receive SNAP, and $X \in \Omega$ denotes conditioning on observed covariates whose values lie in the set Ω. Thus, the ATE reveals how the food insecurity rate would differ if all eligible households with children received SNAP versus the food insecurity rate if all eligible households with children did not receive SNAP.

In what follows, we simplify notation by suppressing the conditioning on subpopulations of interest captured in X. For our main analysis, we condition on eligible households with children. We also assess how the results differ by marital status, educational level, race, and ethnicity. In a usual regression framework, the inclusion of additional observed covariates is motivated as a means of controlling for other factors that may influence food security outcomes; omitting relevant explanatory variables could lead to biased estimates. Instead, in our framework there are no regression disturbance orthogonality conditions to be met; conditioning on covariates serves only to define subpopulations of interest, and our problem is well-defined regardless of how the subpopulations are specified (Pepper 2000).

Two key identification problems confound inferences on the ATE. First, the outcome $FI(1)$ is counterfactual for all children who did not receive SNAP, while $FI(0)$ is counterfactual for all children who did receive SNAP. The data cannot reveal $FI(1)$ for nonparticipants or $FI(0)$ for participants. This is referred to as the selection problem. The realized food insecurity outcome is $FI = FI(1)S^* + FI(0)(1 - S^*)$. Second, true participation status, S^*, is not observed if participation may be misreported. Instead, we observe a self-reported counterpart, S, which may differ from S^*.

Nonparametric Models to Address the Selection Problem

We begin by considering the case in which reports of SNAP receipt are assumed to be accurate: $S = S^*$. To address the selection problem, we consider a number of different

[15] This food insufficiency measure has been used in other studies such as Bitler, Gundersen, and Marquis (2005), Gundersen and Oliveira (2001), and Gundersen and Ribar (2011).

monotonicity restrictions. A natural starting point is to assume nothing about the counterfactual probabilities: what do the data alone reveal?

Using the Law of Total Probability, we can decompose the first term in equation (1) into the following components:

$$P[FI(1)=1] = P[FI(1)=1|S^*=1]P(S^*=1) + P[FI(1)=1|S^*=0]P(S^*=0).$$

For SNAP participants, $S^*=1$, the potential outcome under participation, $FI(1)$ is observed in the data: $P[FI(1)=1|S^*=1] = P(FI=1|S^*=1)$. The latent probability $P[FI(1)=1|S^*=0]$ is unknown. For households that did not participate in SNAP, the data cannot reveal the counterfactual food insecurity outcome had they participated. All we know is that $P[FI(1)=1|S^*=0]$, as a probability, must lie within [0,1]. Taking the two polar extremes, it follows that

$$\text{(2a)}\quad P(FI=1, S^*=1) \leq P[FI(1)=1] \leq P(FI=1, S^*=1) + P(S^*=0).$$

Thus, the worst-case lower bound is derived by supposing that all nonrecipients would have been food secure had they received SNAP, and the upper bound is derived by supposing that all nonrecipients would have been food insecure had they received SNAP. Analogously, we obtain bounds on $P[FI(0)=1]$, the potential outcome under nonparticipation:

$$\text{(2b)}\quad P(FI=1, S^*=0) \leq P[FI(0)=1] \leq P(FI=1, S^*=0) + P(S^*=1).$$

Returning to equation (1), an upper bound on the ATE is obtained by subtracting the lower bound on $P[FI(0)=1]$ from the upper bound on $P[FI(1)=1]$. Similarly, a lower bound on the ATE is obtained by subtracting the upper bound on $P[FI(0)=1]$ from the lower bound on $P[FI(1)=1]$. These bounds are sharp (i.e., logically as tight as possible under the maintained assumptions) for the case of accurately measured data (Manski 2007).

At the other end of the spectrum, the literature examining the impact of SNAP on food insecurity often assumes that selection is exogenous such that

$$\text{(3)}\quad P[FI(j)=1|S^*=j] = P[FI(j)=1|S^*\neq j] = P[FI(j)=1], j=1,0.^{16}$$

Under this exogenous selection assumption, the average treatment effect can be written as

$$\text{(4)}\quad ATE(1,0) = P[FI(1)=1|S^*=1] - P[FI(0)=1|S^*=0]$$

which in turn can be expressed as the difference in conditional means:

$$\text{(5)}\ \beta = P(FI=1|S^*=1) - P(FI=1|S^*=0).$$

The appeal of the exogenous selection assumption is obvious: if selection is exogenous and SNAP receipt S^* is observed, then the average treatment effect is identified by the sampling process. The sample means displayed in table 1, for example, suggest that SNAP is associated with a higher probability of being food insufficient (difference = 0.024) and a higher probability that a child is not eating enough (difference = 0.035). In the absence of random assignment, however, this parameter typically does not reflect causal effects of the program. Still, this mean outcome gap β—the difference in expected outcomes among SNAP recipients and nonrecipients—is an important descriptive measure of the association between SNAP participation and food insecurity.

A number of middle-ground assumptions narrow the endogenous selection bounds by restricting relationships between SNAP participation, food insecurity outcomes, and observed covariates. In particular, we apply three common monotonicity assumptions.

First, the Monotone Treatment Selection (MTS) assumption (Manski and Pepper 2000) places structure on the selection mechanism through which households with children become SNAP recipients. A common theme in

[16] Kreider et al. (2012) did not study this identification assumption.

the literature on the impact that SNAP has on food insecurity is that unobserved factors associated with food insecurity are likely to be positively associated with a household's decision to take up food assistance (see Kreider et al. 2012; Gundersen and Oliveira, 2001; and Currie 2003). Currie (2003) makes this point directly: "… one suspects that these results are driven by negative selection into the Food Stamp Program (FSP) program. That is, those who participate may be less likely to eat a healthy diet [or be food secure] for reasons that have not been controlled for in the regression models estimated by these researchers."[17] In this case, recipients have worse latent food security outcomes than nonrecipients, on average. The MTS assumption is formalized as follows:

$$(6)\quad P[FI(j)=1|S^*=0] \leq P[FI(j)=1|S^*=1],\ j=1,0.$$

That is, conditional on either treatment, $j=1$ or 0, eligible households with children that receive SNAP, $S^*=1$, have a higher latent prevalence of food insecurity than eligible households with children that did not receive benefits.

Second, the Monotone Instrumental Variable (MIV) assumption (Manski and Pepper 2000) formalizes the notion that the latent probability of food insecurity, $P[FI(j)=1]$, varies monotonically with certain observed covariates. Kreider et al. (2012), for example, specify that this probability weakly declines with the ratio of a household's reported income to the poverty threshold associated with the household's family composition.[18] To formalize this idea, let v be the monotone instrumental variable such that

$$(7)\quad u_1 \leq u \leq u_2 \Rightarrow P[FI(j)=1|v=u_2] \leq P[FI(j)=1|v=u] \leq P[FI(j)=1|v=u_1]\ \text{for}\ j=1,0.$$

This mean monotonicity condition is weaker than the standard mean independence instrumental variables (IV) assumption. Under mean independence, these weak inequalities would be replaced with strict equalities. In our application, it is difficult to find instruments that would satisfy mean independence, so we impose the weaker MIV assumption.

Across forty ordered income categories, the estimation procedure enforces a restriction that the derived lower bound on $P[FI(j)=1]$ (based on the particular set of maintained assumptions) for any income group must be no smaller than the lower bound derived for any higher-income group. Similarly, any derived upper bound on $P[FI(j)=1]$ for any income group must be no larger than the upper bound derived for any lower income group. Then, to find the aggregate MIV bounds on the potential rates of food insufficiency, one takes the appropriate weighted average of the plug-in estimators (weighted to account for the survey design) of the lower and upper bounds across the different income groups observed in the data. Estimation details are provided in Kreider and Pepper (2007).[19]

Beyond using the income to poverty line ratio as an MIV, we sharpen inferences compared with Kreider et al. (2012) by incorporating the modified regression discontinuity design (RDD) developed in Gundersen, Kreider, and Pepper (2012). Unlike a standard RDD model, we do not identify an average treatment effect at the discontinuity. Instead, rather than assuming that ineligible respondents reveal the counterfactual outcome, as is the case in a RDD, we apply MIV restrictions that latent food insecurity outcomes vary monotonically with the benefit eligibility criteria. The detailed income and asset

[17] For information on differences between SNAP recipients and nonrecipients over commonly-observed covariates, see Cunnyngham (2005). SNAP recipients tend to have fewer resources than eligible nonrecipients.

[18] Table 3 in Coleman-Jensen, Gregory, and Singh (2014) shows that food insecurity rates in the United States fall as income increases at the following rates: 45.6% for those under the poverty line; 44.2% for those under 130% of the poverty line; and 40.3% for those under 185% of the poverty line, versus 7.7% for those over 185% of the poverty line. One concern with this assumption might be that some families have very low incomes (including negative incomes) as a result of a "business loss" component. These families may have more resources than those with higher incomes. While the SIPP does not allow us to identify households with substantial business loss income, we do know that only 195 respondents in our sample of 2,653 households have zero or negative income. As part of our sensitivity analysis, we slightly modify the MIV approach by removing the constraint that this zero or negative income group has no better food security outcomes, on average, than groups with higher reported income; that is, we allow this group to have either better or worse outcomes. As described in the results section, relaxing this constraint has little effect on the estimates.

[19] As discussed in Manski and Pepper (2000), the unadjusted MIV estimator is consistent but biased in finite samples. Kreider and Pepper's (2007) modified MIV estimator accounts for the finite sample bias using a nonparametric bootstrap correction method. Under the joint MTS-MIV assumption, the MTS assumption is assumed to hold at each value of the instrument, v.

information available in the SIPP make this eligibles-MIV design feasible in this paper.

To implement this modified RDD, we use two observed groups of ineligible households with children: those who are income-eligible but fail the asset test ($v' \equiv$ asset ineligible) and those whose household income falls between 130% and 150% of the poverty line ($v'' \equiv$ income ineligible). While these comparison groups are unlikely to satisfy the standard IV restriction that the latent food insecurity outcomes are mean independent of eligibility status, the MIV assumption holding that mean response varies weakly monotonically across these subgroups seems credible.[20] Children in households with assets valued at $2,000 or more, or with incomes above the SNAP eligibility cutoff (i.e., above 130% of the poverty line), are likely to have no worse average food insecurity outcomes than children living in eligible households. That is,

$$(8)\quad P[FI(j)=1] \geq \max\{P[FI(j)=1|v'=1], P[FI(j)=1|v''=1]\}.$$

Assume that $P(S^*=0)=1$ among these ineligible groups. Then the sampling process identifies $P[FI(0)=1|v'=1]$ and $P[FI(0)=1|v''=1]$ as $P(FI=1|v'=1)$ and $P(FI=1|v''=1)$, respectively.[21] In this case, these MIV ineligibles restrictions imply that $P[FI(0)=1] \geq \max\{P(FI=1|v'=1), P(FI=1|v''=1)\}$, which in turn can improve the upper bound on ATE. This restriction provides no information on $P[FI(1)=1]$.

Finally, despite the observed correlations in the data, providing someone with food assistance seems unlikely to decrease their level of food security. The *Monotone Treatment Response* (MTR) assumption (Manski 1997) formalizes this idea:

$$(9)\quad FI(1) \leq FI(0).$$

Instead of imposing this restriction for all households, we consider a weaker variant in which the restriction holds, on average, conditional on the true value of the treatment, S^*. It is difficult to imagine that SNAP is harmful to food security, on average. While the MTR assumption precludes a strictly positive ATE (thus assuming away the question of whether SNAP is detrimental, on average), it leaves open the question of whether the program has strong beneficial effects, mild beneficial effects, or no effects.

Nonparametric Models to Address Both the Selection and Classification Error Problems

We now consider implications of both the selection and classification error problems. To formalize these identification problems, let the latent variable Z^* indicate whether a report of SNAP participation is accurate, with $Z^*=1$ if $S^*=S$, and $Z^*=0$ otherwise. Using this indicator, we decompose the first term on the right side in equation (1) as

$$\begin{aligned}(10)\quad P[FI(1)=1] &= P(FI=1,S^*=1) \\ &\quad + P[FI(1)=1|S^*=0]P(S^*=0) \\ &= P(FI=1,S=1)+\theta_1^- - \theta_1^+ \\ &\quad + P[FI(1)=1|S^*=0][P(S=0) \\ &\quad + (\theta_1^+ + \theta_0^+) - (\theta_1^- + \theta_0^-)]\end{aligned}$$

where $\theta_k^+ \equiv P(FI=k,S=1,Z^*=0)$ and $\theta_k^- \equiv P(FI=k,S=0,Z^*=0)$ denote the fraction of false positive and false negative classifications of SNAP recipients, respectively, for the food insecurity outcome $k=0,1$. The first term, $P(FI=1,S^*=1)$, is not identified because of the classification error problem. The second term is not identified because of both the selection and classification error problems. The data cannot reveal the counterfactual outcome distribution, $P[FI(1)=1|S^*=0]$, regardless of whether participation is measured accurately, and, in the presence of classification errors, the sampling process does not reveal the proportion of respondents that received assistance, $P(S^*=1)$.

Without imposing assumptions on the measurement error probablities, θ, the

[20] Traditionally, the RDD identifies the average treatment effect for those close to the cutoff. We are focused on drawing inferences across the whole eligible population, not for just those with income near the eligibility thresholds. Our ineligible-MIV assumption, however, seems credible. Moreover, using the law of total probability, we can easily extend our MIV approach to apply to only those with incomes near the threshold. In this case, the MIV assumption in equation (8) might narrow the bounds for some households (e.g., those with incomes near the threshold) but not others (e.g., those with very low incomes).

[21] The assumption that $S^*=0$ for this group may not be valid for all households if income used by an administrator to determine SNAP eligibility was taken from a different month than income reported in the SIPP (see footnote 10). Likewise, adjunctive eligibility conferred to a TANF and SSI recipient might violate the assumption that $S^*=0$. Nevertheless, the eligibility-MIV inequality restriction in equation (8) should remain valid. In this case, the model is similar to a fuzzy discontinuity design.

probability $P[FI(1) = 1|S^* = 0]$ in equation (5) can lie anywhere between zero and one. Thus, in the absence of assumptions on the classification error problem, the data alone are uninformative. Given these identification problems, Kreider et al. (2012) show that the ATE can be sharply bounded as follows:

$$(11)\ -P(FI = 0, S = 1) - P(FI = 1, S = 0) + \Theta \leq ATE \leq P(FI = 1, S = 1) + P(FI = 0, S = 0) + \Theta$$

where $\Theta \equiv (\theta_1^- + \theta_0^+) - (\theta_0^- + \theta_1^+)$. While the observed joint probabilities $P(FI, S)$ are identified by the data, Θ is not identified since the error components $\{\theta\}$ are unknown.

To address the classification error problem, we restrict the range of Θ in equation (11) using two sources of information. First, Kreider et al. (2012) show that knowledge of the true participation rate, $P^* \equiv P(S^* = 1)$, confines the range of Θ. For example, the net fraction of false negative reports, $(\theta_0^- + \theta_1^-) - (\theta_0^+ + \theta_1^+)$—which is not the same as Θ—must equal the difference between the true and self-reported participation rates $\Delta \equiv P(S^* = 1) - P(S = 1)$. Using administrative data on caseloads from the USDA, Kreider et al. (2012) estimate that $P(S^* = 1)$ is about 0.50 for the period 2001–2006 among eligible households. The analogous self-reported rate in the SIPP is $P \equiv P(S = 1) = 0.466$ (see table 1), suggesting a net underreporting rate of 3.4%.[22] Our sensitivity analysis considers higher misreporting rates as well.

Second, we further restrict the feasible ranges of Θ by allowing for the possibility that some self-reports of participation status can be treated as accurate. Specifically, let $V = 1$ indicate that the report is assumed to be accurate (i.e., $Z^* = 1$, which in turn implies $S^* = S$), while $V = 0$ indicates a lack of evidence on the true value of S^* (i.e., Z^* may be one or zero). Kreider et al. (2012) considered a "no false positives" assumption that $S = 1$ implies $S^* = 1$ (though not the converse). This assumption is consistent with evidence from SNAP validation studies that find errors of commission to be negligible. As discussed earlier, Bollinger and David (1997) find that only 0.3% of SIPP responses involve errors of commission (see also Marquis and Moore 1990). Thus, classification error can be described as mostly an underreporting problem.

Furthermore, Bollinger and David (2005) find that those who report SNAP receipt in one period are likely to be accurate reporters of participation status in other periods as well. Exploiting the panel nature of the SIPP allows us to formalize this notion:

Verification Assumption: A response $S = 1$ in any wave implies $V = 1$.

The idea is that a household willing to report SNAP participation in any wave has revealed itself to be willing to acknowledge the receipt of benefits despite any perceived social stigma. We treat such households as reliable responders, even in waves they do not report benefits. In our sample, 46.6% of the sample reports receiving SNAP in the current wave (wave 5, see table 1), while 66.2% of SNAP-eligible households with children report receiving SNAP in at least one wave of the 2004 SIPP.[23] For other households, we remain agnostic about their reporting reliability. Kreider, Pepper, and Roy (2016) derive analytic worst-case bounds, as well as exogenous selection and MTS bounds, on average treatment effects under this type of partial verification assumption.[24] As discussed below, this verification assumption has strong identifying power in some contexts (e.g., when combined with the MTS assumption alone), but little power in others (e.g., when combined with no monotonicity assumptions or when combined with several monotonicity assumptions).

[22] While this value is consistent with the results of Kreider et al. (2012) and of Meyer, Mock, and Sullivan (2009) when evaluating misreporting of the SIPP, it is much smaller than found in the CPS and Panel Survey of Income Dynamics (PSID). Using the CPS, for example, the participation rate is estimated to be around 70% (Cunnyngham 2005).

[23] In some instances, information on SNAP participation may be missing in other waves. Rather than dropping those observations, we conservatively treat them as noncorroborative: a $V = 1$ designation requires specific information that the respondent reported receiving SNAP benefits in some wave.

[24] Verification also places restrictions on the set of feasible values of P^*. Under the no false positives assumption, P^* can be no smaller than P. More generally, P^* is confined to the range $\max\{P - P(S = 1, V = 0),\ P(Y = 1, S = 1) - \theta_1^{+\max}\} \leq P^* \leq \min\{P + P(S = 0, V = 0),\ 1 - P(Y = 1, S = 0) + \theta_1^{-\max}\}$ where $\theta_1^{+\max}$ and $\theta_1^{-\max}$ are derived upper bounds on these error components.

Table 2. Summary Analysis with No Classification Errors, Food Insufficiency Measure

Observed moments

Selection probability: $P(S=1)=0.466$

Food insufficiency rate: $P(FI=1)=0.081$

Food insufficiency rate among those receiving SNAP: $P[FI(1)=1|S=1]=0.094$

Food insufficiency rate among those not receiving SNAP: $P[FI(0)=1|S=0]=0.069$

Treatment response probabilities: $P[FI(j)=1],\ j=0,1$

I. Worst-case: $P[FI(j)=1|S\neq j]\in[0,1]$

$$(0.093*0.466)\ =\ 0.044\ \leq\ P[FI(1)=1]\ \leq\ 0.578\ =\ (0.094*0.466+0.534)$$
$$(0.069*0.534)\ =\ 0.037\ \leq\ P[FI(0)=1]\ \leq\ 0.503\ =\ (0.069*0.534+0.466)$$

II. MTS: $P[FI(j)=1|S=0]\ \leq\ P[FI(j)=1|S=1]$

$$(0.094*0.466)\ =\ 0.044\ \leq\ P[FI(1)=1]\ \leq\ 0.094\ =\ (0.094*0.466+0.097*0.534)$$
$$(0.069*0.534+0.069*0.466)\ =\ 0.069\ \leq\ P[FI(0)=1]\ \leq\ 0.503\ =\ (0.069*0.534+0.466)$$

III. Exogenous selection: $P[FI(j)=1|S=j]=P[FI(j)=1|S\neq j]=P[FI(j)=1]$

$$P[FI(1)=1]\ =0.094$$
$$P[FI(0)=1]=0.069$$

ATE: $P[FI(1)=1]-P[FI(0)=1]$

I. Worst-case: $(0.044-0.503)\ =\ -0.459\ \leq\ ATE\ \leq\ 0.541\ =\ (0.578-0.037)$

II. MTS: $(0.044-0.503)\ =\ -0.459\ \leq\ ATE\ \leq\ 0.024\ =\ (0.094-0.069)$

MTS + MIV:[†] $-0.286\ \leq\ ATE\ \leq\ -0.061$

MTS + MIV + MTR:[a,*] $-0.286\ \leq\ ATE\ \leq\ -0.070$

III. Exogenous selection:

Nonparametric (no covariates)* $ATE=0.024\ =\ (0.094-0.069)$

Linear probability model (with covariates)[b] $ATE=0.014$

IV. Regression discontinuity[c] $ATE=0.018$

Note: Asterisk * indicates that ATE is statistically significantly different than zero at the 10% significance level. (a) Estimation includes the income and eligibility monotone instruments. (b) The linear model results are almost identical to those estimated using a probit model. The covariates are income-to-poverty ratio, age, and indictors for whether the respondent graduated from high school, is married, is African-American, is Hispanic, and is a homeowner. (c) The estimated model uses the discontinuity generated by the income eligibility threshold, but not the asset eligibility threshold, as an instrument. We report the Wald estimator.

Results

We present empirical results in two parts. To make ideas concrete, we first consider drawing inferences under the assumption that SNAP receipt is reported accurately, and focus exclusively on the selection problem. Results are presented for all eligible households with children as well as for different subsamples. This no-classification-error model is consistent with most of the literature that assumes accurate reporting of SNAP participation. Initial focus on this case also allows us to illustrate the basic methodological approach that researchers can use to study questions in different literatures. We also contrast the results from this partial identification analysis to those found using standard point-identified models under the exogenous selection assumption (with and without covariates) and under a parametric regression discontinuity design where the income eligibility cutoff (130% of the FPL) is used as an instrumental variable. Finally, after evaluating the results found under the accurate reporting assumption, we then present results that account for both the selection and classification error problems.

Estimates of the ATE Given the Selection Problem Without Classification Error

Table 2 presents worst-case bounds on the ATE of SNAP on food insufficiency for the case of no classification errors, along with successively narrower bounds derived under the MTS, MTS+MIV, MTS+MIV+MTR, and point-identified models. The bounds are estimated by replacing population probabilities with the corresponding sample probabilities. To illustrate the identification problem arising from unobservable counterfactual outcomes, in this table we display only the estimates of the bounds and not confidence intervals. We do, however, indicate when the estimates of the ATE are statistically significantly different from zero. For other tables in the paper, we present both estimates of the bounds that reflect identification uncertainty and confidence intervals that reflect sampling variability.

The data reveal that 46.6% of the eligible households with children report receiving SNAP. Food insufficiency rates are 9.4% for children residing in households that receive SNAP and 6.9% for households that do not receive SNAP. Under no misreporting, it follows that the food insufficiency rate would lie within [0.044, 0.578] if all households received SNAP, and within [0.037, 0.503] if all households did not receive SNAP. Thus, abstracting from classification error and sampling variability, the data alone reveal that the ATE of SNAP on food insufficiency lies within the range of [−0.459,0.541] (using 0.044–0.503 = −0.459 and 0.578–0.037 = 0.541 as shown near the bottom of table 2). As formalized in Manski (1990), these worst-case bounds have a width of one and do not identify the sign of the ATE.[25]

Rather than focusing on the worst case model, it is useful to apply middle ground assumptions. Continuing with table 2, we assess what can be identified with no misreporting under the MTS, MIV, and MTR assumptions described above. Under the MTS assumption alone, the upper bound on the ATE falls to 0.024. When MTS is additionally combined with the income and eligibility MIV assumptions, the bounds on the ATE shrink to [−0.286, −0.061]. Thus, given these two monotonicity restrictions, SNAP reduces food insufficiency rates by at least 6.1 percentage points, and perhaps much more—up to 29 percentage points. Finally, if we add the MTR assumption, the estimated bounds imply that SNAP reduces food insufficiency rates by at least 7.0 percentage points among households with children.

These partial identification results can be usefully compared to results from conventional models that point-identify the effect of SNAP on food insecurity. We consider three point-identified models: first, the basic exogenous selection assumption in equation (3); second, a linear probability model, where the exogenous selection assumption is assumed to hold conditional on a set of covariates; and third, a regression discontinuity design, where the income eligibility cutoff (130% of the FPL) is used to define the discontinuity. The linear probability model controls for the income-to-poverty ratio, age, and indictors for whether the respondent graduated from high school, is married, is African-American, is Hispanic, and is a homeowner.

The results from these three models suggests that SNAP leads to a small increase in the rate of food insufficiency, a result that is inconsistent with the bounds estimated under the MTS-MIV assumptions. For example, if selection is exogenous, then the ATE is point-identified and estimated to equal 0.024, and if conditionally exogenous, the estimate from the linear probability model implies that the ATE is 0.014. Finally, the estimates from the regression discontinuity model imply that SNAP increases food insecurity by 0.018. The latter two estimates are not statistically significant at the 10% significance level.

These inconsistencies imply that some (or possibly all) of the assumptions in either the point-identified or partially-identified models must be invalid. In fact, as we have noted above, there are good reasons to question the validity of the point-identified models. Even with a rich set of covariates, the decision to take up SNAP is likely to be related to unobserved factors associated with food insecurity (Currie 2003) and food insecurity is almost certainly related to income and thus the income eligibility cutoff.

Estimates of the ATE Allowing for Both Self-Selection and Classification Errors

Classification errors introduce additional uncertainty. Figures 1–3 trace out bounds on the ATE (vertical axis) as a function of the unknown true participation rate, P^* (horizontal axis). Negative values of *ATE* correspond with beneficial expected impacts of SNAP. Figure 1 traces out bounds for the food insufficiency outcome under five different sets of assumptions about the selection process: (*a*) exogenous selection; (*b*) worst-case selection; (*c*) worst-case selection combined with the MIV-ineligibles assumption; (*d*) monotone treatment selection (MTS); and (*e*) MTS combined with the MIV-ineligibles assumption. Figure 2 additionally imposes the income-MIV and MTR assumptions. Figure 3 is identical to figure 2, except the outcome "food insufficient

[25] While these worst-case assumption bounds leave much uncertainty about the ATE, note that the data alone (if accurately measured) have already eliminated half the uncertainty compared with the pre-data knowledge that the ATE can lie anywhere within $[-1, 1]$.

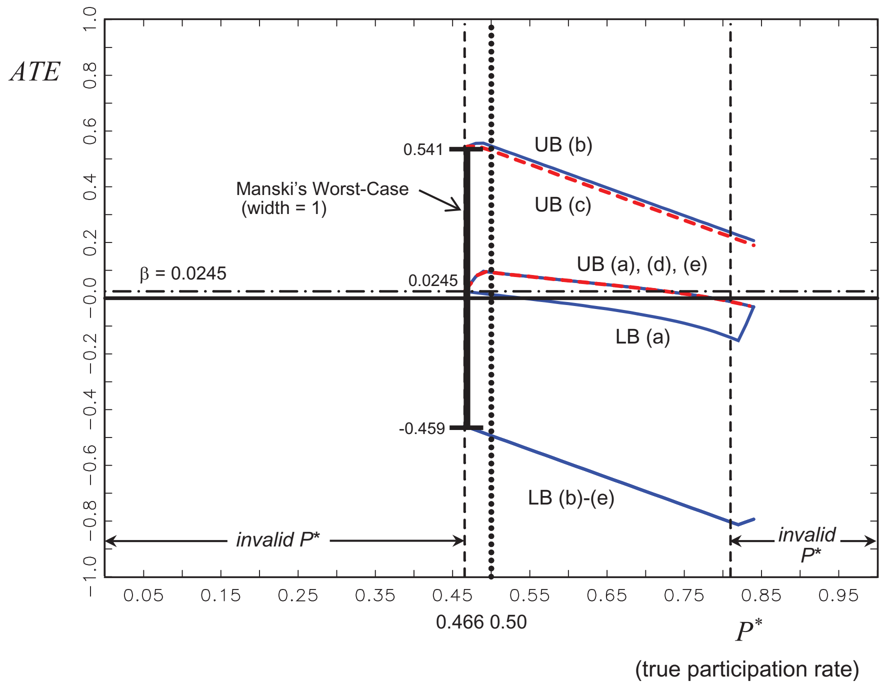

Figure 1. Bounds on *ATE* for "food insufficient household" as a function of P* exogenous selection, worst-case, MTS, and ineligibles MIV

Note: Letter (a) denotes exogenous selection; (b) denotes worst-case scenario; (c) denotes worst-case scenario plus ineligibles MIV; (d) denotes MTS; and (e) denotes MTS plus ineligibles MIV.

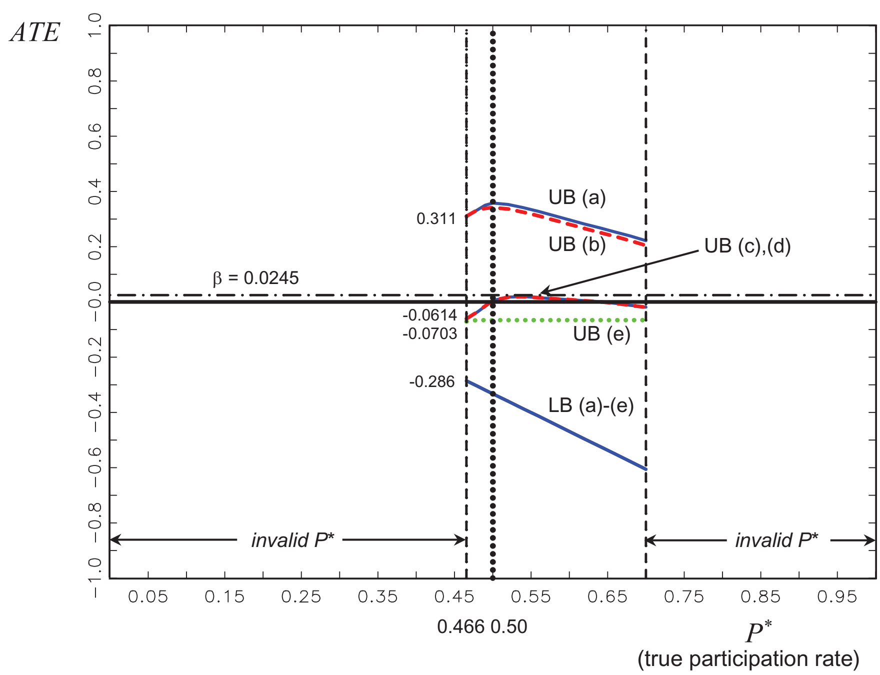

Figure 2. Bounds on *ATE* for "food insufficient household" as a function of P* MTS, MTR, ineligibles MIV, and income MIV

Note: Letter (a) indicates worst-case plus income MIV; (b) indicates ineligibles MIV plus income MIV; (c) denotes MTS plus income MIV; (d) denotes MTS, ineligibles MIV, plus income MIV; and (e) denotes MTS, ineligibles MIV, income MIV, plus MTR.

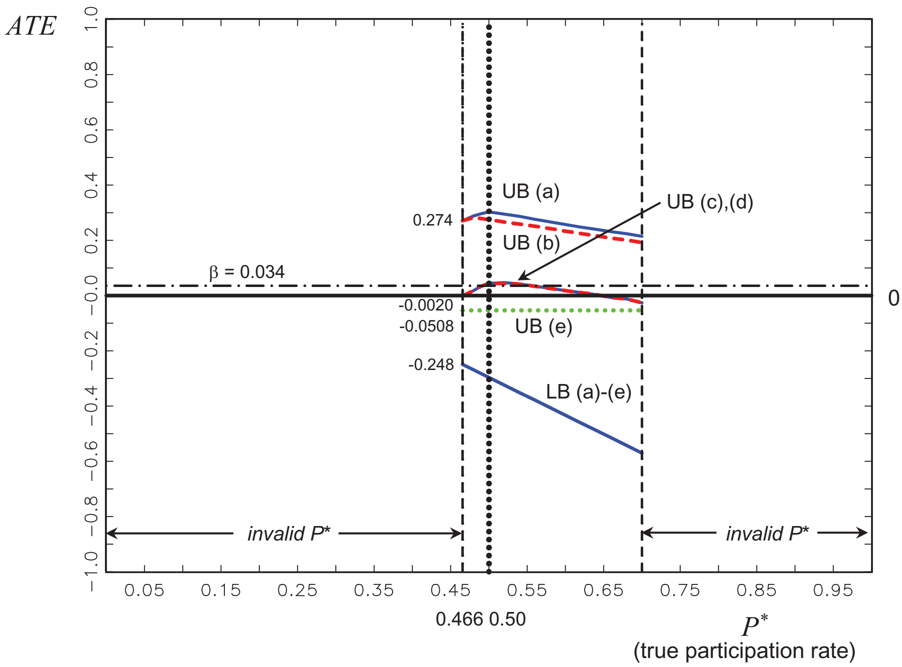

Figure 3. Bounds on *ATE* for "kids not eating enough" as a function of *P: MTS, MTR, ineligibles MIV, and income MIV**

Note: Letter (a) indicates worst case plus income MIV; (b) denotes ineligibles MIV plus income MIV; (c) denotes MTS plus income MIV; (d) denotes MTS, ineligibles MIV, plus income MIV; and (e) denotes MTS, ineligibles MIV, income MIV, plus MTR.

household" is replaced with "kids not eating enough" to examine another food insecurity measure.

For each figure, there is an accompanying table (3, 4, and 6) that highlights key results for the special cases in which the true participation rate, P^*, equals either the observed rate, $P^* = P = 0.466$, the administrative rate reported in Kreider et al. (2012), $P^* = P^o = 0.50$, or a higher rate, 0.70, to study the sensitivity of results to large-scale misreporting. These tables present Imbens-Manski (2004) confidence intervals (CI) that cover the true value of the ATE with 90% probability. Instead of covering the entire identification region (i.e., range between the lower bound and upper bound) with fixed probability, the Imbens-Manski approach is designed to cover the true value of the parameter (in our case unknown ATE) with this probability.[26] For cases when these confidence intervals are strictly negative, we also report a p-value.

Throughout, we impose the verification assumption that respondents who report SNAP participation in any wave of the SIPP accurately report participation status in Wave 5.[27] In addition to narrowing the bounds compared with the Kreider et al. (2012)

[26] Given standard confidence intervals on the lower and upper bounds, a straightforward confidence interval on the parameter can be found using the lower confidence interval on the lower bound and the upper interval on the upper bound. This naïve confidence interval is valid but generally conservative. Imbens and Manski (2004) modify this approach by making two adjustments to the critical value. First, they choose the critical value to ensure proper coverage. For example, to provide a 95% interval for a univariate expectation when there is missing data, it is easy to show that the one-sided critical value of 1.645 provides proper coverage while the traditional critical value of 1.96 provides a 97.5% interval. Second, they adjust the critical value to account for the discontinuity that arises when the model is (almost) point-identified. In the example above, one would use a 1.96 critical value if there were no missing data.

[27] In our application, this verification assumption plays a strong role in reducing the degree of uncertainty about the ATE under our middle-ground models, but it provides less identification power under the weakest and strongest models. Under the MTS assumption alone at $P^* = P^o = 0.50$, for example, verification narrows the bounds from [−0.581, 0.162] to [−0.493, 0.093], a 21% reduction in the width. The upper bound is reduced by nearly seven points. Under the worst-case bounds with no monotonicity assumptions, verification narrows the bounds from [−0.581, 0.581] to [−0.493, 0.546], only a 10% reduction in the width. Under our strongest model that combines knowledge of P^* with all of the monotonicity assumptions, verification narrows the bounds from [−0.389, −0.066] to [−0.333, −0.066]. While the lower bound is somewhat improved, the upper bound in this application is unaffected by the verification. These findings may naturally vary in other applications.

Table 3. Bounds on ATE for "Food Insufficient Household" as a Function of P^*: Exogenous Selection, Worst-Case, MTS, and Ineligibles MIV

		Self-reported rate $P^* = P = 0.466$		Administrative rate $P^* = P^o = 0.50$		Sensitivity $P^* = 0.70$	
			width		width		width
(a) Exogenous selection	p.e.[a]	[0.025, 0.025]	0.000	[0.013, 0.093]	0.080	[-0.062, 0.034]	0.095
	CI[b]	[0.005 0.044]		[-0.002 0.108]		[-0.081 0.049]	
(b) Worst case	p.e.	[-0.459, 0.541]	1.000	[-0.493, 0.546]	1.040	[-0.693, 0.346]	1.040
	CI	[-0.473 0.555]		[-0.508 0.562]		[-0.708 0.362]	
(c) Worst case + ineligibles MIV	p.e.	[-0.459, 0.541]	1.000	[-0.493, 0.530]	1.024	[-0.693, 0.330]	1.024
	CI	[-0.473 0.558]		[-0.508 0.552]		[-0.708 0.352]	
(d) MTS	p.e.	[-0.459, 0.025]	0.483	[-0.493, 0.093]	0.586	[-0.693, 0.034]	0.727
	CI	[-0.473 0.034]		[-0.508 0.108]		[-0.708 0.049]	
(e) MTS + ineligibles MIV	p.e.	[-0.459, 0.025]	0.483	[-0.493, 0.093]	0.586	[-0.693, 0.034]	0.727
	CI	[-0.473 0.040]		[-0.508 0.110]		[-0.708 0.049]	

Note: (a) Point estimates (p.e.) of the population bounds; (b) Imbens-Manski 5th and 95th percentile bounds (CI) based on 1,000 pseudosamples.

framework, this verification restriction confines P^* to lie sufficiently close to the self-reported rate, P (see footnote 24). In the figures, logically valid values of P^* are constrained to lie between the two vertical dashed lines.[28] The dashed horizontal line reflects the value of ATE under the benchmark case of exogenous selection and no misclassification, 0.024, consistent with this value in tables 1 and 2.

As seen in figure 1 and table 3, the worst-case endogenous selection bounds range from [−0.459,0.541] (a width of 1) at $P^* = P$, the self-reported SNAP participation rate, to [−0.493,0.546] at $P^* = P^o$, the "true" participation rate, to [−0.693,0.346] at $P^* = 0.70$. Adding the ineligibles-MIV assumption improves the upper bounds to 0.541, 0.530, and 0.330, respectively. The MTS assumption has substantial identifying power. Under the MTS assumption alone, the upper bounds fall to 0.025, 0.093, and 0.034. Adding the MIV-ineligibles assumption does not change the estimated bounds in this case. Regardless of the true value of P^*, these assumptions alone are not strong enough to identify the sign of the ATE—an important negative result. Nevertheless, the MTS assumption has reduced the upper bound at $P^* = P^o$ by nearly fifty points (from 0.546 to 0.093) compared with the worst-case upper bound. Notice that the upper bounds first rise but then fall with the degree of misclassification. In fact, for a large enough value of P^*, the upper bound on the ATE under misreporting may actually be lower than the upper bound under the no misreporting case. For example, the worst-case upper bound in figure 1 rises from 0.541 at $P^* = 0.466$ under no misreporting to 0.546 at $P^* = 0.50$, then soon declines to a value below 0.541 as the degree of underreporting rises.[29]

Using figure 2 and table 4, we assess what can be identified when we additionally impose the income MIV assumption (latent outcomes weakly improve with income) and the MTR assumption (SNAP participation may not be beneficial, but it does no harm, on average). Compared with the joint MTS + ineligibles-MIV assumption at $P^* = P^o = 0.50$ in row (d) of table 3, adding income-MIV narrows the bounds from ([−0.493,0.093]) to ([−0.333,0.005]), a 42% reduction in the width of the bounds (table 4). Additionally, adding the MTR assumption

[28] Outside this range, either P^* is invalid or the verification assumption is invalid. In Kreider et al. (2012), all values of P^* exceeding P, the self-reported rate, were treated as valid because they had no information on the extent of false negative responses.

[29] A bound can change direction with P^* as particular constraints on the error components $\{\theta\}$ become binding. For example, the upper bound in equation (11) under the no false positives assumption is rising with $\Theta \equiv \theta_1^- - \theta_0^-$. If these two false negative components were unconstrained, we would simply set $\theta_0^- \equiv P(Y=0, S=0, Z^*=0) = 0$ and as a worst case scenario assume that any classification error arises through $\theta_1^- \equiv P(Y=1, S=0, Z^*=0)$, the fraction of households that are food insecure and falsely claim not to participate in SNAP (and vice-versa for the lower bound). However, we know that $P(Y=1, S=0, Z^*=0)$ cannot exceed $P(Y=1, S=0)$: the fraction of false claims of nonparticipation cannot exceed the fraction of all claims of nonparticipation. For a large enough P^*, θ_1^- will become maxed out. Any further increase in P^* will need to involve an increase in θ_0^-—which improves the upper bound—to continue reconciling the gap between P^* and P.

Table 4. Bounds on *ATE* for "Food Insufficient Household" as a Function of P^*: MTS, MTR, Ineligibles MIV, and Income MIV

		Self-reported rate $P^* = P = 0.466$	width	Administrative rate $P^* = P^o = 0.50$	width	Sensitivity $P^* = 0.70$	width
(a) Worst case + income MIV	p.e.[a]	[-0.286, 0.311]	0.597	[-0.333, 0.358]	0.691	[-0.606, 0.224]	0.829
	CI[b]	[-0.384 0.415]		[-0.431 0.455]		[-0.677 0.297]	
(b) Ineligibles MIV + income MIV	p.e.	[-0.286, 0.311]	0.597	[-0.333, 0.342]	0.675	[-0.606, 0.208]	0.813
	CI	[-0.392 0.414]		[-0.433 0.438]		[-0.677 0.272]	
(c) MTS + income MIV	p.e.	[-0.286, -0.061]	0.225	[-0.333, 0.005]	0.338	[-0.606, -0.025]	0.581
	CI	[-0.382 0.022]		[-0.430 0.092]		[-0.677 0.049]	
(d) MTS + ineligibles MIV+ income MIV	p.e.	[-0.286, -0.061]	0.225	[-0.333, 0.005]	0.338	[-0.606, -0.025]	0.581
	CI	[-0.390 0.024]		[-0.433 0.062]		[-0.677 0.046]	
(e) MTS + ineligibles MIV+ income MIV + MTR	p.e.	[-0.286, -0.070]	0.216	[-0.333, -0.066]	0.267	[-0.606, -0.066]	0.540
	CI	[-0.390 -0.012][c]		[-0.433 -0.013] [d]		[-0.677 -0.016][e]	

Note: (a) Bias-corrected point estimates of the population bounds; (b) Imbens-Manski 5th and 95th percentile bounds (1,000 pseudosamples); (c) p-value for a strictly negative ATE is 0.06; (d) p-value = 0.06; (e) p-value = 0.05.

in row (e) in table 4, the ATE is confined to be no greater than −0.066, or a 6.6 percentage point reduction in food insecurity. As shown, this favorable impact of the program is statistically significant at the 10% level (the p-value is 0.06) and robust to higher misclassification rates. This finding is comparable to the −0.089 upper bound estimated in Kreider et al. (2012) using NHANES data.[30]

Qualitatively similar results are found across a variety of different subgroups of income-eligible households with children. Table 5 displays estimates and confidence intervals under the MIV-MTS-MTR model by marital status, educational level, race, and ethnicity.[31] Results are provided when the true participation rate, P^*, is assumed to equal either the observed rate, $P^* = P$, or an administrative rate, $P^* = P^o$, imputed under the assumption that the percentage degree of underreporting is constant across subsamples. The results imply that SNAP reduces food insufficiency across the different subgroups, with the estimated upper bounds (at $P^* = P^o$) ranging from −0.021 for Hispanics to −0.083 for households in which the head is not married. Reflecting the relatively small sample sizes, however, these estimates are statistically insignificant except for unmarried households (p-value = 0.03), whites (p-value = 0.08), and for the full sample (p-value = 0.06). Finally, notice that although there is much variation in estimates across the different groups, all of the different bounds overlap. Thus, we cannot reject the hypothesis that the effect of SNAP is the same across the different subgroups.

The percentage reduction in the prevalence of food insufficiency under SNAP compared with food insufficiency in the absence of the program is given by $-\{P[FI(1) = 1] - P[FI(0) = 1]\}/P[FI(0) = 1]$, or $-ATE/P[FI(0) = 1]$. Under the combined MTS-MIV-MTR assumptions, the estimated lower bound on $P[FI(0) = 1]$ for the full sample is 0.104. Thus, our estimates imply that SNAP reduces the prevalence of food insecurity in households with children by at least $-ATE^{UB}/P[FI(0) = 1]^{LB} = 0.066/0.104$, or by at least 63%. This sizable impact of SNAP in part reflects the relatively small base of food-insufficient households with children.

[30] These results are nearly identical if we do not impose the income MIV constraint on the group reporting zero or negative income (see footnote 18). Specifically, the row (e) estimates of the bounds at $P^* = P = 0.466$ change slightly from [−0.286, −0.070] to [−0.286, −0.069], with confidence intervals changing from [−0.390, −0.012] to [−0.389, −0.012]. At $P^* = P^o = 0.50$, the point estimates change slightly from [−0.333, −0.066] to [−0.332, −0.062], with the confidence intervals changing from [−0.433, −0.013] to [−0.431, −0.015].

[31] The number of MIV groups is reduced proportionally with the reduction in sample size.

Table 5. Bounds on *ATE* for "Food Insufficient Household" as a Function of P^* for Subsamples: MTS + ineligibles MIV + income MIV[c] + MTR

		Self-reported rate		Administrative rate[b]	
		$P^* = P = 0.466$		$P^* = P^o = 0.50$	
			width		width
Main sample (N = 2589)	p.e. [a]	[−0.286, −0.070]	0.216	[−0.333, −0.066]	0.267
	CI [b]	[−0.390 −0.012]*	p-value: 0.06 [e]	[−0.433 −0.013] *	p-value: 0.06
Married (N = 999)	$P^* = P = 0.282$			$P^* = P^o = 0.30$	
	p.e.	[−0.188, −0.036]	0.152	[−0.212, −0.023]	0.184
	CI	[−0.274 0.000]	p-value: 0.26	[−0.299 0.000]	p-value: 0.26
Not married (N = 1590)	$P^* = P = 0.592$			$P^* = P^o = 0.64$	
	p.e.	[−0.560, −0.083]	0.476	[−0.609, −0.083]	0.526
	CI	[−0.588 −0.027] **	p-value: 0.03	[−0.632 −0.027] **	p-value: 0.03
High school graduates (N = 1994)	$P^* = P = 0.449$			$P^* = P^o = 0.48$	
	p.e.	[−0.251, −0.037]	0.215	[−0.296, −0.031]	0.265
	CI	[−0.354 0.000]	p-value: 0.24	[−0.394 0.000]	p-value: 0.26
Nongraduates (N = 595)	$P^* = P = 0.519$			$P^* = P^o = 0.56$	
	p.e.	[−0.476, −0.037]	0.438	[−0.516, −0.037]	0.478
	CI	[−0.523 0.000]	p-value: 0.19	[−0.562 0.000]	p-value: 0.20
White (N = 1260)	$P^* = P = 0.442$			$P^* = P^o = 0.47$	
	p.e.	[−0.281, −0.071]	0.211	[−0.324, −0.071]	0.253
	CI	[−0.390 −0.008] *	p-value: 0.08	[−0.428 −0.009]*	p-value: 0.08
African American (N = 720)	$P^* = P = 0.621$			$P^* = P^o = 0.67$	
	p.e.	[−0.564, −0.051]	0.512	[−0.614, −0.051]	0.563
	CI	[−0.611 0.000]	p-value: 0.20	[−0.657 0.000]	p-value: 0.21
Hispanic (N = 560)	$P^* = P = 0.370$			$P^* = P^o = 0.40$	
	p.e.	[−0.365, −0.032]	0.333	[−0.395, −0.021]	0.374
	CI	[−0.412 0.000]	p-value: 0.30	[−0.439 0.000]	p-value: 0.30

Note: (a) Point estimates (p.e.) of the population bounds; (b) Imbens-Manski 5th and 95th percentile bounds (CI) based on 1,000 pseudosamples; (c) Number of MIV groups is reduced proportionally with the reduction in sample size; (d) True participation rate under the assumption that the percentage degree of underreporting is constant across subsamples; (e) p-value for a strictly negative ATE. Asterisk *indicates significance at the 10% level; **indicates significance at the 5% level.

Alternatively focusing on the larger base of food-sufficient households with children, we find that SNAP increases the prevalence of food security by at least 7.3%. The results from the subgroup analysis reveal similarly large effects of SNAP. These estimates, which as noted above are often statistically insignificant, imply that SNAP reduces the prevalence of food insecurity by at least 69% when the household head is not married (43% when married), 37% when the household head graduated from high school (40% when a drop out), 66% when the household head is white, 40% when African American, and 26% when Hispanic.

Kids Not Eating Enough

Figure 3 and table 6 displays analogous findings for the "kids not eating enough" outcome. Focusing on $P^* = P^o$, the ATE is again identified as strictly negative (beneficial) under the joint MTR, MTS, and MIV assumptions. In particular, the ATE is estimated to lie in the range of −0.295 to −0.051. Based on these estimates, SNAP is found to decrease the fraction of kids not eating enough by at least 5.1 percentage points, or 41%, relative to the baseline case that SNAP did not exist. Alternatively focusing on the base of households in which children are eating enough, SNAP increases this fraction by at least 5.8%. Unlike the case for food insufficiency, however, these results are not statistically significant at the 10% level.

Importantly, these results make especially transparent that even large degrees of classification error need not hinder inference, depending on the research question. While the total degree of uncertainty about the magnitude of the ATE (width of the bounds) increases with the degree of misreporting, note that the upper bound itself is improving with the degree of underreporting (see earlier discussion).

Table 6. Bounds on *ATE* for "Kids not Eating Enough" as a Function of P^*: MTS, MTR, Ineligibles MIV, and Income MIV

		Self-reported rate $P^* = P = 0.466$		Administrative rate $P^* = P^o = 0.50$		Sensitivity $P^* = 0.70$	
			width		width		width
(a) Worst case + income MIV	p.e.[a]	[−0.248, 0.274]	0.522	[−0.295, 0.305]	0.600	[−0.568, 0.218]	0.786
	CI[b]	[−0.337 0.385]		[−0.387 0.411]		[−0.639 0.292]	
(b) Ineligibles MIV+ income MIV	p.e.	[−0.248, 0.274]	0.522	[−0.295, 0.277]	0.572	[−0.568, 0.195]	0.763
	CI	[−0.352 0.385]		[−0.394 0.382]		[−0.639 0.267]	
(c) MTS+ income MIV	p.e.	[−0.248, −0.002]	0.246	[−0.295, 0.045]	0.339	[−0.568, −0.030]	0.538
	CI	[−0.336 0.053]		[−0.387 0.124]		[−0.639 0.076]	
(d) MTS + ineligibles MIV+ income MIV	p.e.	[−0.248, −0.002]	0.246	[−0.295, 0.045]	0.339	[−0.568, −0.030]	0.538
	CI	[−0.349 0.053]		[−0.393 0.123]		[−0.639 0.076]	
(e) MTS + ineligibles MIV+ income MIV + MTR	p.e.	[−0.248, −0.051]	0.197	[−0.295, −0.051]	0.244	[−0.568, −0.053]	0.516
	CI	[−0.349 0.000] [c]		[−0.393 0.000] [d]		[−0.639 0.000][e]	

Note: (a) Bias-corrected point estimates of the population bounds; (b) Imbens-Manski 5th and 95th percentile bounds (1,000 pseudosamples); (c) p-value for a strictly negative ATE is 0.17; (d) p-value = 0.17; (e) p-value of 0.16.

Conclusion

Without strong assumptions on the selection and measurement problems, the causal effects of food assistance programs cannot be identified. Kreider et al. (2012) recently developed a nonparametric partial identification approach that addresses these issues in a single unifying framework. In this paper, we illustrate how to apply these partial identification methods using data from the SIPP to study the robustness of prior conclusions.

While this econometric approach cannot fully resolve the uncertainty about the causal effects of SNAP, it serves to bound the truth and make transparent how assumptions on the selection and reporting error processes shape inferences. In general, the results are similar to those found in Kreider et al. (2012). For some models, the data reveal very little about the impact of SNAP on food insecurity in households with children. Under our strongest nonparametric assumptions, we find that SNAP reduces the fraction of food-insufficient households with children by at least six to eleven percentage points depending on the degree of SNAP misreporting and how food security is measured.

References

Almada, L., I. McCarthy, and R. Tchernis. 2016. What Can We Learn about the Effects of Food Stamps on Obesity in the Presence of Misreporting? *American Journal of Agricultural Economics* 98 (4): 997–1017.

Bitler, M., J. Currie, and J. Scholz. 2003. WIC Eligibility and Participation. *Journal of Human Resources* 38: 1139–79.

Bitler, M., C. Gundersen, and G. Marquis. 2005. Are WIC Non-Recipients at Less Nutritional Risk than Recipients? An Application of the Food Security Measure. *Review of Agricultural Economics* 27 (3): 433–8.

Blundell, R., A. Gosling, H. Ichimura, and C. Meghir. 2007. Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds. *Econometrica* 75 323–63.

Bollinger, C., and M. David. 1997. Modeling Discrete Choice with Response Error: Food Stamp Participation. *Journal of the American Statistical Association* 92 (439): 827–35.

———. 2001. Estimation with Response Error and Nonresponse: Food Stamp

Participation in the SIPP. *Journal of Business and Economic Statistics* 19 (2): 129–41.

———. 2005. I Didn't Tell, and I Won't Tell: Dynamic Response Error in the SIPP. *Journal of Applied Econometrics* 20: 563–9.

Borjas, G. 2004. Food Insecurity and Public Assistance. *Journal of Public Economics* 88: 1421–43.

Bound, J., C. Brown, and N. Mathiowetz. 2001. Measurement Error in Survey Data, In *Handbook of Econometrics*, ed. J. Heckman and E. Leamer, 3705–843. Amsterdam: ScienceDirect.

Coleman-Jensen, A, C. Gregory, and A. Singh. 2014. *Household Food Security in the United States in 2013*. Washington DC: U.S. Department of Agriculture, Economic Research Report No. 173.

Cunnyngham, K. 2005. *Food Stamp Program Participation Rates: 2003*. Washington DC: U.S. Department of Agriculture, Food, and Nutrition Service.

Currie, J. 2003. U.S. Food and Nutrition Programs. In *Means Tested Transfer Programs in the U.S.* ed. Robert Moffitt, 199–290. University of Chicago Press.

Devaney, B., and R. Moffitt. 1991. Dietary Effects of the Food Stamp Program. *American Journal of Agricultural Economics* 73 (1): 202–11.

Eslami, E., and K. Cunnyngham. 2014. *Supplemental Nutrition Assistance Program Participation Rates: Fiscal Years 2010 and 2011*. Washington DC: Prepared by Mathematica Policy Research, Inc., for the U.S. Department of Agriculture, Food and Nutrition Service.

Goerge, R., K. Franzetta, and J. Dilts. 2009. Food Stamp Program Eligibility and Participation in Chicago. Working Paper. Chicago, IL: Chapin Hall Center for Children at the University of Chicago.

Gundersen, C., and B. Kreider. 2008. Food Stamps and Food Insecurity: What Can be Learned in the Presence of Nonclassical Measurement Error? *Journal of Human Resources* 43 (2): 352–82.

Gundersen, C., B., Kreider, and J.V. Pepper. 2012. The Impact of the National School Lunch Program on Child Health: A Nonparametric Bounds Analysis. *Journal of Econometrics* 166 (1): 79–91.

Gundersen, C., and S. Offutt. 2005. Farm Poverty and Safety Nets. *American Journal of Agricultural Economics* 87 (4): 885–99.

Gundersen, C., and V. Oliveira. 2001. The Food Stamp Program and Food Insufficiency. *American Journal of Agricultural Economics* 84 (3): 875–87.

Gundersen, C., and D. Ribar. 2011. Food Insecurity and Insufficiency at Low Levels of Food Expenditures. *Review of Income and Wealth* 57 (4): 704–26.

Hoynes, H., and D. Schanzenbach. 2009. Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program. *American Economic Journal: Applied Economics* 1 (4): 109–39.

Imbens, G., and C. Manski. 2004. Confidence Intervals for Partially Identified Parameters. *Econometrica* 72 (6): 1845–57.

Jensen, H. 2002. Food Insecurity and the Food Stamp Program. *American Journal of Agricultural Economics* 84 (5): 1215–28.

Kreider, B., and J.V. Pepper. 2007. Disability and Employment: Reevaluating the Evidence in Light of Reporting Errors. Journal of the American Statistical Association 102 (478): 432–41.

Kreider, B., J.V. Pepper, C. Gundersen, and D. Jolliffe. 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): 958–75.

Kreider, B., J.V. Pepper, and M. Roy. 2016. Does the Women, Infants, and Children Program (WIC) Improve Infant Health Outcomes? Working paper. Ames, IA: Iowa State University

Manski, C. 1997. Monotone Treatment Response. *Econometrica* 65 (6): 1311–34.

——— 2007. *Identification for Prediction and Decision*. Cambridge, MA: Harvard University Press.

——— 1990. Nonparametric Bounds on Treatment Effects *American Economic Review Papers and Proceedings*. 80 (2): 319–323.

Manski, C., and D. Nagin. Bounding Disagreements About Treatment Effects: A Case Study of Sentencing and Recidivism. *Sociological Methodology* 28: 99–137.

Manski, C., and J.V. Pepper. 2000. Monotone Instrumental Variables: With an Application to the Returns to Schooling. *Econometrica* 68 (4): 997–1010.

———. 2015. How do Right-to-Carry Laws Affect Crime Rates? Coping with Ambiguity Using Bounded-Variation Assumptions, NBER # 21701.

Marquis, K., and J. Moore. 1990. Measurement Errors in SIPP Program Reports. *Proceedings of the Bureau of the Census Annual Research Conference*. Washington DC: Bureau of the Census, 721–45.

Meyer, B.D., W.K.C. Mock, and J.X. Sullivan. 2009. The Under-Reporting of Transfers in Household Surveys: Its Nature and Consequences. NBER Working Paper no. 15181.

Meyer, B., R. Goerge, and N. Mittag. 2015. Errors in Survey Reporting and Imputation and their Effects on Estimates of Food Stamp Program Participation. Unpublished Manuscript.

Meyerhoefer, C., and V. Pylypchuk. 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.

Molinari, F. 2008. Partial Identification of Probability Distributions with Misclassified Data. *Journal of Econometrics* 144 (1): 81–117.

Mykerezi E., and B. Mills. 2010. The Impact of Food Stamp Program Participation on Household Food Insecurity. *American Journal of Agricultural Economics* 92 (5): 1376–91.

Pepper, J.V. 2000. The Intergenerational Transmission of Welfare Receipt: A Nonparametric Bounds Analysis. *Review of Economics and Statistics* 82 (3): 472–88.

Rank, M., and T. Hirschl. 2009. Estimating the Risk of Food Stamp Use and Impoverishment During Childhood. *Archives of Pediatrics and Adolescent Medicine* 163 (11): 994–9.

Ratcliffe, C., S. McKernan, and S. Zhang. 2011. How Much Does the Supplemental Nutrition Assistance Program Reduce Food Insecurity? *American Journal of Agricultural Economics* 93 (4): 1082–98.

Ribar, D., and K. Hamrick. 2003. *Dynamics of Poverty and Food Sufficiency.* Washington DC: U.S. Department of Agriculture, Food Assistance and Nutrition Research Report No. 33.

Taeuber, C., D.M. Resnick, S.P. Love, J. Stavely, P. Wilde, and R. Larson. 2004. Differences in Estimates of Food Stamp Program Participation between Surveys and Administrative Records. Working paper, U.S. Census Bureau.

Trippe, C., P. Doyle, and A. Asher. 1992. *Trends in Food Stamp Program Participation Rates, 1976 to 1990.* Washington DC: U.S. Department of Agriculture, Food and Nutrition Service.

U.S. Department of Agriculture. 1999. *Annual Historical Review: Fiscal Year 1997.* Food and Nutrition Service.

Van Hook, J., and K. Stamper Balistreri. 2006. Ineligible Parents, Eligible Children: Food Stamps Receipt, Allotments and Food Insecurity among Children of Immigrants. *Social Science Research* 35 (1): 228–51.

Wolkwitz, K. 2008. *Trends in Food Stamp Program Participation Rates: 2000–2006.* Washington DC: Prepared by Mathematica Policy Research, Inc. for the U.S. Department of Agriculture, Food and Nutrition Service.

Yen, S., M. Andrews, Z. Chen, and D. Eastwood. 2008. Food Stamp Program Participation and Food Insecurity: An Instrumental Variables Approach. *American Journal of Agricultural Economics* 90: 117–32.