12-2009

The Impact of the National School Lunch Program on Child Health: A Nonparametric Bounds Analysis

Craig Gundersen
University of Illinois

Brent Kreider
Iowa State University, bkreider@iastate.edu

John Pepper
University of Virginia

The Impact of the National School Lunch Program on Child Health: A Nonparametric Bounds Analysis

Abstract
Children in households reporting the receipt of free or reduced price school meals through the National School Lunch Program (NSLP) are more likely to have negative health outcomes than eligible nonparticipants. Assessing the causal effects of the program is made difficult, however, by the presence of endogenous selection into the program and systematic misreporting of participation status. Using data from the National Health and Nutrition Examination Survey (NHANES), we extend and apply partial identification methods to account for these two identification problems in a single unifying framework. Similar to a regression discontinuity design, we introduce a new way to conceptualize the monotone instrumental variable (MIV) assumption using eligibility criteria as monotone instruments. Under relatively weak assumptions, we find evidence that receipt of free and reduced price lunches through the NSLP improves the health outcomes of children.

Keywords
partial identification, selection problem, classification error, monotone instrumental variable, regression discontinuity, National School Lunch Program, food insecurity, obesity

The Impact of the National School Lunch Program on Child Health: A Nonparametric Bounds Analysis

Craig Gundersen, Brent Kreider, and John Pepper

Working Paper No. 09027
December 2009

IOWA STATE UNIVERSITY
Department of Economics
Ames, Iowa, 50011-1070

The Impact of the National School Lunch Program on Child Health:
A Nonparametric Bounds Analysis*

Craig Gundersen
Department of Agricultural and Consumer Economics
University of Illinois
cggunder@illinois.edu

Brent Kreider
Department of Economics
Iowa State University
bkreider@iastate.edu

John Pepper
Department of Economics
University of Virginia
jvpepper@virginia.edu

December 2009

Abstract: Children in households reporting the receipt of free or reduced price school meals through the National School Lunch Program (NSLP) are more likely to have negative health outcomes than eligible nonparticipants. Assessing the causal effects of the program is made difficult, however, by the presence of endogenous selection into the program and systematic misreporting of participation status. Using data from the National Health and Nutrition Examination Survey (NHANES), we extend and apply partial identification methods to account for these two identification problems in a single unifying framework. Similar to a regression discontinuity design, we introduce a new way to conceptualize the monotone instrumental variable (MIV) assumption using eligibility criteria as monotone instruments. Under relatively weak assumptions, we find evidence that receipt of free and reduced price lunches through the NSLP improves the health outcomes of children.

Keywords: partial identification, selection problem, classification error, monotone instrumental variable, regression discontinuity, National School Lunch Program, food insecurity, obesity

JEL classification numbers: C14, C21, I38

* The authors received valuable comments from Helen Jensen, Peter Orazem, Nese Yildiz, and seminar participants at Iowa State University, Purdue University, the Southern Economic Association Meetings, and the conference on Identification and Decisions in honor of Charles Manski. Pepper acknowledges financial support from the Bankard Fund for Political Economy, and Gundersen acknowledges financial support from the US Department of Agriculture (USDA), Cooperative State Research, Education and Extension Service (CSREES), Hatch project no. ILLU-470-331.

1. Introduction

Every school day, more than 30 million children in the United States receive free or reduced-price lunches through the National School Lunch Program (NSLP).[1] With expenditures exceeding $9 billion in FY 2008, the NSLP is a large and important child nutrition program that is a vital component of the social safety net for low income children. As such, policymakers expect the program to have a positive impact on the health of this vulnerable population. Yet, the existing empirical literature reveals little supporting evidence and, in some cases, appears to find deleterious effects. In particular, the literature has found that children receiving free or reduced price lunches are more likely to have negative health outcomes than observationally similar eligible nonparticipants.

Despite these findings, the causal effects of the National School Lunch Program remain uncertain. Assessing the effects of the program is made difficult by the presence of two fundamental identification problems. First, self-selection of households into the program suggests that children receiving free or reduced price meals through the NSLP (hereafter referred to as free lunch recipients) are likely to differ from eligible nonparticipants in ways that are not observed in the data. Second, the association between participation in the NSLP and poor nutritional health may be at least partly an artifact of household misreporting of participation in the program. Meyer et al. (2009), for example, find evidence of aggregate NSLP underreporting rates of 45% in the Current Population Survey and 27% in the Panel Study of Income Dynamics.

While these identification problems have long been known to confound inferences on the impact of the free lunch program, credible solutions remain elusive. Most studies treat selection as exogenous and, to our knowledge, all studies assume that the classification of receipt is accurately reported.[2]

---

[1] The NSLP is based on three categories of payments: free, reduced price, and full price. The former two categories apply for children with household incomes below 185% of the poverty line. While those between 130% and 185% of the poverty line receive reduced price meals, the costs (no more than 40 cents per meal) are substantially less than the full price meals. See http://www.fns.usda.gov/cnd/lunch/ for administrative details about the program.

[2] Recent studies that address the possibility of endogenous selection include Schanzenbach (2007) and Millimet et al. (2008). Schanzenbach (2007) uses both a panel data design and a regression discontinuity design that exploits the income eligibility cutoff as an instrument. Millimet et al. (2008) use panel designs, propensity score matching, and sensitivity testing methods to assess the impact of inferences to different selection assumptions. Bhattacharya et al. (2006), who focus on the National School Breakfast Program (NSBP), use a difference-in-difference design that compares students' nutrient intakes during the school year with their intakes during the summer when school is not in session. They also compare outcomes of children attending schools that provide breakfast programs to outcomes of students enrolled in nonparticipating schools. In general, these studies conclude that the NSBP improves

Addressing these identification problems using observational data is always difficult but especially so when the classical prescriptions are untenable. State variation in program rules, which often serve as instruments to study the impact of other means-tested programs in the United States, are less useful as instrumental variables in this setting. Most program rules are set by the federal government and, except for some of the nutritional guidelines, have not changed substantially over time.[3] Moreover, as with much of the empirical literature on means-tested assistance programs, conventional parametric restrictions imposed to identify treatment effects – for example, the classical linear response model assumption – are difficult to justify when considering programs that are thought to have heterogeneous effects. Finally, classical measurement error models do not apply when the inaccurately measured covariate is discrete (see, for example, Bollinger, 1996).

In this paper, we evaluate the impact of the free lunch program in light of the ambiguity created by the selection and measurement problems. Extending a recent analysis of the Food Stamp Program in Kreider, Pepper, Gundersen, and Jolliffe (2009) (KPGJ hereafter), we apply partial identification bounding methods that allow one to consider weaker assumptions than required under conventional parametric approaches. We use these methods coupled with data from the 2001-2004 waves of the National Health and Nutrition Examination Survey (NHANES) to assess the impact of the free lunch program on the nutritional well-being of children.

After describing the data in Section 2, we formally define the empirical questions and identification problems in Section 3. The usual program evaluation literature formally acknowledges uncertainty associated with counterfactuals but not uncertainty associated with misreporting. As a departure from this literature, we account for both identification problems in a single framework. Given classification errors in the realized treatment variable, we extend Manski's (1995) basic selection bounds using ideas for making inferences from corrupt samples from Horowitz and Manski (1995) and for addressing missing treatments from Molinari (forthcoming and 2008). We also use related partial

---

nutritional outcomes while the NSLP seems to increase obesity rates. Evidence on selection is mixed. See Currie (2003) and Millimet et al. (2008) for thorough overviews of the literature.

[3] Although program rules are set by the federal government, school districts have some flexibility in designing the meal plan.

identification work from Manski and Pepper (2000), Pepper (2000), Kreider and Pepper (2007 and forthcoming), Kreider and Hill (2008), and KPJG.

After modifying the Manski selection bounds to account for classification error in the treatment, we introduce a number of monotonicity assumptions that help tighten inferences. In Section 4, we consider the identifying power of a *monotone instrumental variable* (MIV) assumption that certain observed covariates are known to be monotonically related to the latent response variable (Manski and Pepper, 2000). Requiring no *a priori* exclusion restriction, the MIV assumption can be plausible in many applications where the standard independence assumption is a matter of considerable controversy. In our application, we maintain the relatively weak assumption that the latent probability of a poor health outcome is nonincreasing with income.

In addition, we also consider two important variations of the MIV. The first variation, introduced in Manski and Pepper (2000), replaces the exogenous treatment selection assumption implicitly imposed in much of the literature with a weaker *monotone treatment selection* (MTS) restriction. This self-selection model formalizes the commonplace explanation for why recipients have higher food insecurity and obesity rates than nonrecipients – namely, that the decision to participate in the NSLP is presumed to be (weakly) monotonically related to poor latent health outcomes.

Our second variation introduces a new way to conceptualize the MIV assumption by using eligibility criteria as monotone instruments. There is a long history of using ineligible respondents to identify the impact of a wide array public policies, including several recent studies that assess the impacts of the NSLP and National School Breakfast Program (NSBP). Schanzenbach (2007), for example, uses a regression discontinuity design that exploits the income eligibility cutoff as an instrument, and Bhattacharya et al. (2006) use children attending schools that do not offer meal programs. While the basic idea of the discontinuity design is appealing, in practice there can be many limitations. Considerable disagreement often arises over the implicit assumption that ineligible respondents reveal the counterfactual outcome distribution for participants. Moreover, even if the comparison group is credible, these designs generally only identify the effect for persons near the eligibility cutoff and are not robust to classification error.

In contrast, the MIV assumption allows us to relax this traditional identifying assumption by holding that mean outcomes among subgroups of ineligible respondents bound instead of identify the counterfactual outcome distribution. In our application, we focus on three MIV ineligible subgroups: (1) children in households with incomes above the income eligibility threshold for free or reduced price lunches (185% of the federal poverty line), (2) children enrolled in schools that do not participate in the NSLP (primarily private schools), and (3) children who have dropped out of school. Children in the first two groups are presumed to have no worse latent health outcomes on average than eligible children, while children in the third group are presumed to have no better outcomes on average. Importantly, we show that this assumption provides substantial identifying information for the average treatment effect (ATE) even when there is classification error. Most notably, the latent outcome distribution under nonparticipation is point-identified for ineligible households since there is no selection or classification problem for this subgroup. As a result, this eligibility MIV assumption has substantial identifying power in this application, reducing the width of the Manski worst-case bounds (Manski, 1995) on the average treatment effect for the food insecurity rate by nearly 70 percent. Beyond the school lunch setting, the idea of using ineligible respondents as an MIV may prove to have wide applicability in the program evaluation literature.

In Section 5, we consider the additional identifying power of a *monotone treatment response* (MTR) assumption that participation in the free lunch program would not increase the prevalence of poor health. While free lunch recipients appear to be worse off than eligible recipients prior to accounting for selection, many have argued that the free lunch program cannot have negative casual impacts on food security or health (e.g., Currie, 2003). Combined with our MIV assumptions, the MTR assumption narrows the range of uncertainty about the effects of the free lunch program.

Section 6 draws conclusions. In this application, we find that the three MIV assumptions applied to ineligible groups have substantial identifying power. In particular, when we combine all three MIV assumptions, the estimated ATE bounds under the assumption of fully accurate reporting narrow to a few point range, and we can identify that the free lunch program has beneficial impacts on health-related outcomes. Most notably, in contrast to the previous literature, we find evidence that the NSLP reduces obesity.

2. Data

To study the impact of the NSLP on children's nutritional health, we use data from the 2001-2004 NHANES.[4] The NHANES, conducted by the National Center for Health Statistics, Centers for Disease Control (NCHS/CDC), is a program of surveys designed to assess the health and nutritional status of adults and children in the United States through interviews and direct physical examinations. The survey currently examines a national sample of about 5,000 persons each year, about half of whom are children. Vulnerable groups, including Hispanics and African Americans, are oversampled. The NHANES provides detailed and varied information on dietary and health-related outcomes collected from self-reports of health and nutritional well-being, medical and dental examinations, physiological measurements, and laboratory tests. Given the wealth of health-related information, the NHANES has been widely used in previous research on health- and nutrition-related child outcomes (e.g. KPGJ, 2009).

We restrict attention to households with children that are eligible to receive free or reduced price lunches through the NSLP. Specifically, we restrict the analysis to children between the ages of eight and 17 who are attending schools with lunch programs and residing in households with income less than 185% of the federal poverty line. This results in a sample of 2,219 eligible children. We also utilize information from three groups of children who are ineligible to receive free or reduced price lunches: income eligible children in this age range who are not enrolled in schools with lunch programs (N=65), income eligible children who are no longer in school (N=107), and children residing in households with incomes between 185% and 300% of the poverty line (N=778).

Table 1 displays the means and standard deviations for the variables used in the analysis for the entire sample of eligible children and by self-reported participation status. For each observation, we observe a limited set of socioeconomic and demographic information, including age and the poverty income ratio (PIR), the ratio of a family's income to the poverty threshold defined by the U.S. Census Bureau accounting for the family's composition. We also observe a self-reported (by the parent) indicator of participation in the NSLP and an indicator for whether the meals are free or reduced price. In this

---

[4] We pool the 2001-2002 and 2003-2004 two-year cycles of the NHANES. Weights are established within the NHANES for use when multiple cycles are combined.

survey, 73% of the eligible households claim to have children who received free or reduce price lunches through the NSLP during the school year.

We examine three outcomes: food insecurity, poor health, and obesity.[5] In our data, 24.4 percent of the respondents are food insecure, 18.7 percent are obese, and 7.4 percent report being in poor or fair health. These three measures are considered to be central indicators of the nutritional health and well-being of children and reflect a wide range of health related outcomes that might be impacted by the NSLP. All three outcomes are known to be associated with a range of negative physical, psychological, and social consequences that have current and future implications for health. Children in households suffering from food insecurity, for example, are more likely to have poor health, psychosocial problems, frequent stomachaches and headaches, increased odds of being hospitalized, greater propensities to have seen a psychologist, behavior problems, worse developmental outcomes, more chronic illnesses, less mental proficiency, and higher levels of iron deficiency with anemia (for a review, see Gundersen et al., 2009). Likewise, childhood obesity is known to have negative physical, psychological, and social consequences including reduced life expectancy (Fontaine et al., 2003).

While policymakers have expressed a great deal of interest in understanding the impact of assistance programs on food insecurity, no one, to the best of our knowledge, has examined the impact of the free lunch program on food insecurity or measures of poor general health.[6] There is, however, a growing body of literature that examines the program's impact on obesity, a recent concern among policymakers. This literature finds that the NSLP appears to lead to modest increases in obesity rates (see Millimet et al., 2008, and Schanzenbach, 2007).

---

[5] To calculate official food insecurity rates in the U.S., defined over a 12 month period, a series of 18 questions about food-related needs and resources in the household are posed in the Core Food Security Module (CFSM) for families with children. Examples include "I worried whether our food would run out before we got money to buy more" (the least severe outcome) and "Did a child in the household ever not eat for a full day because you couldn't afford enough food?" A child is considered to reside in a food insecure household if the respondent answers affirmatively to three or more of these questions. We measure obesity using measures based on a child's body mass index (BMI, kg/m^2) such that a child is classified as obese if his or her BMI is at or above the 95th percentile for age and gender. General health is based on a self-reported measure provide by the child's parent. A child's health under this measure is placed into one of five categories based on responses from the parent: excellent, very good, good, fair, or poor. In this paper, we combine these general health categories into an indicator of fair or poor health.

[6] Food insecurity and self-reported measures of health have been widely used as outcomes in studies of other food assistance programs such as the Food Stamp Program (e.g., see Currie, 2003, for a review of the literature).

In Table 1, we see that free lunch recipients appear to be somewhat worse off than eligible nonparticipants. Most striking are the results for food insecurity. The rates of food insecurity among self-reported recipients are nearly twelve points higher than for nonparticipants. The obesity rate is slightly higher for children in households claiming to participate, and the self-reported measures of poor health are nearly identical for recipients and nonrecipients. For both obesity and poor health, these differences are not statistically significant.

3. The Average Treatment Effect with Endogenous Selection and Classification Errors

Our interest is in learning the average effect of the free lunch program among eligible households:

$$\text{ATE}(1,0 \mid X \in \Omega) = E[Y(1) \mid X \in \Omega] - E[Y(0) \mid X \in \Omega] \tag{1}$$

where $Y(1)$ denotes the health of a child if participating in the NSLP, $Y(0)$ denotes the analogous outcome if not participating, and $X \in \Omega$ denotes conditioning on observed covariates whose values lie in the set Ω. Thus, the average treatment effect reveals how the mean outcome would differ if all eligible children would receive assistance versus the mean outcome if all eligible children would not receive assistance. Conditioning on *X* allows the researcher to focus on specific subpopulations of interest. In our analysis, we consider the full sample of NSLP-eligible households.[7] To simplify exposition, in what follows we suppress the conditioning on *X*.

Two forms of uncertainty arise when assessing the impact of the free lunch program on children's outcomes. First, even if participation were observed, the outcome $Y(1)$ is counterfactual for all children who did not receive assistance, and $Y(0)$ is counterfactual for all children who did receive assistance. This is referred to as the selection problem. Second, participation may not be accurately observed for all respondents. This is referred to as the classification error problem.

In Section 3.1, we formalize the identification problems that arise from the selection and classification error problems. Then, in Section 3.2 we focus on what can be inferred about the ATE in

[7] Note that there are no regression orthogonality conditions to be satisfied, and excluding other personal characteristics does not introduce any omitted variable bias into the analysis.

light of the classification error problem alone. In Section 3.3, we additionally consider the selection problem.

3.1 The Identification Problem

To disentangle the selection and classification problems, it is useful to introduce notation for accurate and inaccurate reports of participation. In particular, let $S^* = 1$ indicate that the child actually participates in the program, with $S^* = 0$ otherwise, such that the observed health outcome is $Y = Y(1)S^* + Y(0)(1 - S^*)$. The classification error problem arises because we only observe S, which indicates whether the household reports participating in the program, and not S^*.

Focusing on binary outcomes, we can highlight these two identification problems by writing

$$\begin{aligned} P[Y(1)=1] &= P[Y(1)=1 \mid S^*=1]P(S^*=1) + P[Y(1)=1 \mid S^*=0]P(S^*=0) \\ &= [P(Y=1,\ S=1) - \theta_1^+ + \theta_1^-] \ +\ P[Y(1)=1 \mid S^*=0]\,[P(S=0)\ +\ (\theta_1^+ + \theta_0^+) - (\theta_1^- + \theta_0^-)] \end{aligned} \tag{2}$$

where $\theta_i^+ = P(Y=i,\ S=1,\ S^*=0)$ and $\theta_i^- = P(Y=i,\ S=0,\ S^*=1)$ denote the fraction of false positive and false negative classifications of free lunch participation, respectively, for children realizing a health outcome of i = 0 or 1. Notice that the first term $P[Y(1)=1 \mid S^*=1]\ P(S^*=1)$ is not identified because of the classification error problem. If we rule out these errors such that $\theta_1^+ = \theta_1^- = 0,$ then $P[Y(1)=1,\ S^*=1]$ is revealed by the data. The second term is not identified because of both the selection and classification error problems. The data cannot reveal the counterfactual outcome distribution, $P[Y(1)=1 \mid S^*=0]$, regardless of whether participation is measured accurately, and, in the presence of classification errors, the sampling process does not reveal which respondents received assistance, $P(S^*)$.

3.2 Exogenous Selection Bounds

To isolate the identification problem created by classification errors, we first focus on the important special case that arises when the selection process is exogenous: $P[Y(t)=1 \mid S^*] = P[Y(t)=1]$. In this case, the average treatment effect is simply the difference in mean health outcomes:

$$\text{ATE}(1,0) = E[Y(1) \mid S^* = 1] - E[Y(0) \mid S^* = 0].$$

While Table 1 reveals the differences in mean outcomes based on self-reports of receipt, there is evidence that these reports are often misclassified. In this case, the outcome probabilities are not identified. To see this, it is helpful to write

$$E[Y(1) \mid S^* = 1] = P(Y = 1 \mid S^* = 1) = \frac{P(Y = 1, S^* = 1)}{P(S^* = 1)} = \frac{P(Y = 1, S = 1) + \theta_1^- - \theta_1^+}{P(S = 1) + \theta_1^- + \theta_0^- - \theta_1^+ - \theta_0^+}. \tag{3}$$

Mean outcomes among participants, $P(Y = 1 \mid S^* = 0)$, can be decomposed in a similar fashion. Since the probabilities, θ, are unknown, neither the numerator nor the denominator is identified and, in the absence of further assumptions, the data are uninformative. However, using the basic conceptualization of corrupt sampling developed by Horowitz and Manski (1995), we have found that imposing weak restrictions on the pattern of classification errors can lead to informative bounds on the conditional probability in Equation (3).

In particular, to address the classification error problem we maintain the following two assumptions:

(A1) *Upper Bound Error Rate Assumption:* $P(S^* \neq S) \leq \lambda$

(A2) *No False Positives Assumption: If* $S = 1$, *then* $S^* = 1$.

Assumption A1, which is maintained throughout our analysis, places an upper bound on the potential degree of data corruption. We assess the sensitivity of inferences to NSLP participation errors by varying λ between 0 and 0.25. If $\lambda = 0.1$, for example, then up to 10% of households might misreport their NSLP participation status. The existing literature evaluating the efficacy of the free lunch program has uniformly maintained the assumption of accurate reporting, thus implicitly setting $\lambda = 0$. Assumption A2, rules out false positive reports. That is, respondents reporting to have received free or reduced price lunch are verified to provide accurate reports.

The restrictions in (A1) and (A2) are generally consistent with the limited evidence on misreporting of free school lunch participation and the more detailed information on misreporting of participation in other assistance programs. In particular, validation studies for other food assistance programs reveal negligible errors of commission (see Marquis and Moore, 1990, and Bollinger and David, 1997) and substantial error rates among households reporting to have not received assistance (Meyer et al., 2009).

These two assumptions imply informative restrictions on the unknown probabilities, θ. From Assumption 1, we know for i = 0, 1 that

$$0 \le \theta_i^- \le \min[\lambda,\ P(Y=i,\ S=0)] \equiv \theta_i^{UB-},$$

$$0 \le \theta_i^+ \le \min[\lambda,\ P(Y=i,\ S=1)] \equiv \theta_i^{UB+},$$

$$\text{and } \theta_1^- + \theta_0^- + \theta_1^+ + \theta_0^+ \le \lambda. \qquad (4)$$

Under Assumption 2, we know that $\theta_1^+ = \theta_0^+ = 0$ so that $\theta_i^{UB+} = 0$.

Thus, under Assumptions 1 and 2, it follows that (see Gundersen and Kreider, 2008, Prop. 1):

$$\frac{P(Y=1,S=1)}{P(S=1)+\ \theta_0^{\mathrm{UB}-}} \le P(Y=1 \mid S^*=1) \le \frac{P(Y=1,S=1)+\theta_1^{UB-}}{P(S=1)+\theta_1^{UB-}} \qquad (5)$$

and

$$\frac{P(Y=1,Z=0)-\theta_1^{UB-}}{P(S=0)-\theta_1^{UB-}} \le P(Y=1 \mid S^*=0) \le \frac{P(Y=1,S=0)}{P(Y=0)-\theta_0^{UB-}}. \qquad (6)$$

Kreider and Pepper (2007, Prop. 1) derive the analogous bounds under arbitrary errors for the case that Assumption 2 is not imposed.

Given the results in Equations (5) and (6), a naïve bound on the difference in mean outcomes between recipients and nonrecipients can be found by subtracting the upper (lower) bound on $P(Y=1 \mid S^*=0)$ from the lower (upper) bound on $P(Y=1 \mid S^*=1)$. These naïve bounds on the ATE, however, do not account for the fact that the two conditional probabilities are linked by the unknown probabilities, θ_1^- and θ_0^-. We compute sharp bounds on mean differences in health outcomes using numerical methods that impose these constraints (see Gundersen and Kreider, 2008).

Figures 1A-1C trace out point-estimates of the bounds on the ATE for the three outcomes of interest as λ varies between 0 and 0.25. The accompanying tables reproduce these results for selected values of λ (0, 0.05, 0.10, and 0.25) along with Imbens-Manski (2004) confidence intervals that cover the ATE with 90% probability. For now, we focus on the case of exogenous selection.

When λ = 0 such that reporting errors are ruled out by assumption, the ATE is point-identified as the self-reported gap (matching the descriptive statistics reported in Table 1). For example, free lunch recipients are found to be significantly more likely to be food insecure. Specifically, the difference in food insecurity rates between participants and nonparticipants is estimated to be 0.397 - 0.282 = 0.115 for households (Figure 1A). Since the confidence intervals do not include 0, we can identify these ATE values as strictly positive even after accounting for sampling variability (again, consistent with Table 1). At λ = 0, free lunch participants are estimated to be in slightly better general health (Figure 1B) but also more likely to be obese (Figure 1C). For both of these outcomes, however, the accompanying 90% confidence intervals include both positive and negative values.

These figures reveal that small degrees of uncertainty about the reliability of NSLP classifications translate into large degrees of uncertainty about the ATE. For each outcome, identification of the ATE deteriorates rapidly as λ departs from 0. When λ = 0.10, for example, the estimated bounds on the food insecurity gap when imposing Assumption A1 but not A2 widen to [-0.177, 0.467], and the gaps for poor general health and obesity are estimated to lie within the wide ranges [-0.273, 0.110] and [-0.329, 0.258], respectively.

Thus, small degrees of potential classification error are sufficient to overturn conclusions from the data that free lunch recipients are more likely to be food insecure and obese than eligible nonrecipients. In fact, if even 6.2 percent of participating households failed to report receiving benefits, we cannot identify that food insecurity is more prevalent among children receiving free lunches than in nonparticipating households (see Figure 1A). For obesity, the analogous critical value for being unable to identify the sign of the ATE is only 2.8 percent. That is, once we allow for NSLP classification errors, it is difficult to conclude that poor health outcomes are more prevalent among free lunch recipients than among eligible nonrecipients. Such conclusions would require a large degree of confidence in self-reported free lunch participation status, an assumption not supported by validation studies.

3.3 Relaxing the Exogeneity Assumption

In the absence of assumptions on the selection process, Manski (1995) shows that the bounds on the ATE given fully accurate data always have a width of 1 and always include 0. That is, the data cannot reveal the sign of the treatment effect. Adding ambiguity created by the reporting error problem will naturally widen these treatment effect bounds. In this section, we show how the Manski worst-case selection bounds can be generalized to account for the possibility of classification errors (see KPGJ). As above, we consider what can be learned under Assumptions A1 and A2.

Consider the problem of drawing inferences on the probability of a negative health outcome if all eligible children were to receive free lunches, $P[Y(1)=1]$. As noted above, the fact that we cannot observe this latent outcome for children who did not receive assistance is the selection problem. In particular, we cannot identify the latent outcome probability, $P[Y(1)=1 \mid S^*=0]$. The fact that we may not observe which children received assistance is the classification error problem. That is, we do not know the fraction of false positive and negative reports, θ.

Still, we do know that $P[Y(1)=1 \mid S^*=0]$ must lie within the unit interval, and Assumptions A1 and A2 place restrictions on the form of the response error. Thus, we can derive informative bounds on $P[Y(1)=1]$ and $P[Y(0)=1]$. In particular, given Assumption 1 alone, we have

$$P(Y=1,\ S=1)-\theta_1^{UB+} \le P[Y(1)=1] \le P(Y=1,\ S=1)+P(S=0)+\theta_0^{UB+}, \tag{7}$$

$$P(Y=1,\ S=0)-\theta_1^{UB-} \le P[Y(0)=1] \le P(Y=1,\ S=0)+P(S=1)+\theta_0^{UB-}. \tag{8}$$

Returning to Figures 1A-1C, these worst-case selection bounds that allow for endogeneity (labeled endogenous selection bounds in the figures) can be compared with the exogenous selection bounds discussed in the previous section. Clearly, much less can be known about the average treatment effects when the exogeneity assumption is relaxed.

For $\lambda = 0$, receipt of free lunches is known to be accurately reported. In this case, the bounds in Equations (7) and (8) simplify to the well-known worst-case selection bounds reported in Manski (1995). Classification errors increase uncertainty about the ATE. When λ rises from 0 to 0.1 under A1 and A2, for example, the ATE bounds on the food insecurity rate for households expand from [-0.518, 0.482], with a width of 1, to [-0.618, 0.557], with a width of 1.18. Under A1 alone, the bounds widen further to [-0.718, 0.657], with a width of 1.38.

Interestingly, under the no-false positive assumption, A2, the bounds on $P[Y(1)=1]$ in Equation (7) are identical to the worst-case bounds without reporting errors. Even in this environment, however, we cannot rule out the possibility that the free lunch program has a large positive or negative effect on the likelihood of poor health. To narrow these bounds, we consider a number of additional identifying assumptions in the next section.

4. Monotone Instrumental Variable Assumptions: Variations on a Theme

Many observed variables are thought to be monotonically related to the latent health outcomes in Equation (1). In this section, we formalize and assess the identifying power of several different types of monotone instrumental variable assumptions. We begin by considering the seemingly innocuous assumption that the latent probability of negative health outcomes weakly decreases with income adjusted for family composition, as in KPGJ. To formalize the notion that the latent probability of a negative health outcome $P[Y(t)=1]$ is known to vary monotonically with an observed covariate, let *v* be a monotone instrumental variable such that

(A3) *Income MIV*

$$u_1 \le u \le u_2 \Rightarrow P[Y(t)=1 \mid v=u_2] \le P[Y(t)=1 \mid v=u] \le P[Y(t)=1 \mid v=u_1].$$

While these conditional probabilities are not identified, they can be bounded. Let LB(*u*) and UB(*u*) be the known lower and upper bounds evaluated at *v* = *u*, respectively, given the available information. Then the monotone instrumental variable restriction (MIV) presented in Manski and Pepper (2000, Proposition 1) implies:

$$\sup_{u \le u_2} LB(u_2) \le P[Y(t)=1 \mid v=u] \le \inf_{u \ge u_1} UB(u_1).$$

In the absence of other information, this bound on $P[Y(t)=1 \mid v=u]$ is sharp.

The MIV bound on the unconditional latent probability, $P[Y(t)=1]$ can then be obtained using the law of total probability. If the latent probability is weakly decreasing with the MIV income, then

$$LB_{MIV} \equiv \sum_u P(v=u)\left\{\sup_{u \le u_2} LB(u_2)\right\} \le P[Y(t)=1] \le \sum_u P(v=u)\left\{\inf_{u \ge u_1} UB(u_1)\right\} \equiv UB_{MIV} \qquad (9)$$

To estimate the MIV bounds on the rates of poor health, we take the appropriate weighted average of the plug-in estimators of lower and upper bounds across the different values of the instrument.[8] As discussed in Manski and Pepper (2000; 2009) and Kreider and Pepper (2007), this MIV estimator is consistent but biased in finite samples. We employ Kreider and Pepper's (2007) modified MIV estimator that accounts for the finite sample bias using a nonparametric bootstrap correction method.

In practice, the income MIV assumption has some identifying power but does not substantially narrow the worst case selection bounds in our application. Thus, rather than present bounds under the income MIV assumption alone, we combine this assumption with two other distinct but related instrumental variable restrictions. In Section 4.1, we introduce and assess the idea that eligibility criteria for the NSLP might be monotonically related to the latent outcomes. For example, income ineligible children – i.e., children residing in households with income greater than 185% of the federal poverty line – are likely to have better average health outcomes the income eligible children. In Section 4.2, we apply the monotone treatment selection assumption that participation in the program is (weakly) negatively related to expected health outcomes.

[8] To estimate these income MIV bounds, we divide the sample into 10 PIR groups.

4.1 Ineligible Comparison Groups as an MIV

Many program evaluations rely on ineligible respondents to reveal the counterfactual outcome distribution under nonparticipation. This, for example, is the central idea of the regression discontinuity design. Schanzenbach (2007) and Bhattacharya et al. (2006) use ineligible children to identify the impact of school meal programs on health.

In our application, we observe three groups of ineligible respondents: income eligible children who have dropped out of school (v_2 = dropped out), income eligible children attending schools without the NSLP (v_2 = no lunch program), and children whose household income is between 185% and 300% of the poverty line (v_2 = income ineligible). Table 2 displays the means and standard deviations of the variables used in this analysis for these three ineligible groups of children as well as for the group of eligible children. Children in schools without the NSLP are similar to eligible children, dropouts are much less likely to be obese but are more likely to be food insecure and in poor health, and income ineligible children are better off with respect to food insecurity and general health.

These comparison groups are unlikely to satisfy the standard instrumental variable restriction that the latent health outcomes are mean independent of eligibility status. However, the MIV assumption holding that mean response varies monotonically across these subgroups seems credible, especially for the food insecurity and poor health outcomes. Children in households with incomes above the eligibility cutoff for the NSLP (i.e., above 185% of the poverty line), for example, are likely to have no worse average latent health outcomes than children below this line. Likewise, children attending schools without the NSLP– which are primarily private schools – are thought to have better outcomes, and dropouts might be assumed to have relatively poor latent health outcomes.[9] We apply these monotonicity restrictions to the food insecurity and poor health outcomes, but not obesity where the relationships between the three subgroups and latent measures of obesity are less certain.

[9] The NSLP offers lunches in 99 percent of U.S. public schools and in 83 percent of private and public schools combined (USDA/ERS, 2004).

To formalize the notion that the latent probability of poor health, $P[Y(t) = 1]$, is known to be monotonically related to these observed ineligible subgroups, let v_2 be the monotone instrumental variable such that

(A4) *Ineligible Comparison Group MIV*

i. $P[Y(t)=1] \geq P[Y(t)=1 \mid v_2 = \text{income ineligible}]$,

ii. $P[Y(t)=1] \geq P[Y(t)=1 \mid v_2 = \text{no school lunch program}]$, and

iii. $P[Y(t)=1] \leq P[Y(t)=1 \mid v_2 = \text{dropped out}]$.

For these ineligible subgroups defined by v_2, we assume $S^* = 0$, and there is no selection or classification error problem.[10] The data identify $P[Y(0)=1 \mid v_2]$ but provide no information on $P[Y(1)=1 \mid v_2]$. Thus, the monotone instrumental variable restriction (MIV) in Assumption 4 implies

$$\begin{gathered} \max\left\{P(Y=1 \mid v_2 = \text{income ineligible}),\ P(Y=1 \mid v_2 = \text{no school lunch program})\right\} \\ \leq P[Y(0)=1] \leq \\ P(Y=1 \mid v_2 = \text{dropped out}). \end{gathered} \tag{10}$$

To estimate these MIV bounds on $P[Y(0)=1]$, we plug the observed sample probabilities into Equation (10). For the food insecurity rate outcome, for example, $P[Y(0)=1]$ is constrained to lie within [0.361, 0.451]. Otherwise, in the absence of classification error $P[Y(0)=1]$ can lie anywhere within [0.075, 0.809], and these bounds will be still wider if we allow for reporting errors. Thus, this MIV

[10] The assumption that $S^*= 0$ (similar to a sharp discontinuity design) may not be valid for the observed income threshold where income measures used to determine eligibility may reflect different time periods than measures collected in the NHANES. A household whose eligibility was established in one period may have income that exceeds the threshold when the survey is conducted. With a "fuzzy" threshold where $S^* = 1$ for some "ineligible" respondents, the methods can be adapted to allow for selection and measurement error within "ineligible" subgroups. In this case, the data would provide informative bounds on both latent outcome probabilities.

assumption has substantial identifying power in this application, reducing the width of the bounds under fully accurate reporting from 0.734 to 0.090.[11]

Figures 2A-2C trace out sharp bounds on the ATE under the MIV assumptions and different degrees of misreporting. For each outcome except obesity, we apply the joint income MIV (A3) and ineligible MIV (A4) assumptions. For obesity, we only apply the income MIV Assumption (A3). For ease of comparison, the “no MIV” bounds presented in these figures reproduce the endogenous selection bounds presented in Figures 1A-1C.

The MIV assumptions, and especially the A4 restriction on ineligible respondents, substantially reduce the uncertainly created by the selection problem. In the case with no classification errors ($\lambda = 0$), the width of the bounds on the ATE for household food insecurity declines from 1 to 0.261, a 74 percent reduction. For the poor health rate, the width declines by about four-fifths from 1 to 0.207. The estimated MIV bound of [-0.613, 0.164] for the ATE on the obesity rate, where Assumption 4 is not applied, has a width of 0.777.

While these MIV assumptions have considerable identifying power, there remains much uncertainty. Most notably, we cannot identify the sign of the treatment effect even if we rule out reporting errors. These bounds highlight our inability to make strong inferences about the efficacy of the free lunch program without making stronger assumptions about unknown counterfactuals. In the absence of further restrictions, the estimated bounds identified using the ineligible and income MIV assumptions alone cannot rule out the possibility that the NSLP has a large positive or negative effect on the likelihood of poor health outcomes.

Moreover, the estimated bounds generally expand with the degree of potential misreporting. For example, as λ increases from 0 to 0.10 under arbitrary patterns of misreporting, the estimated bounds on the ATE for household food insecurity expand from [-0.151, 0.110] to [-0.251, 0.210]. In contrast, under the no false positives assumption (A2), the bounds on food insecurity rates do not vary with λ. Under this assumption, classification errors do not contaminate inference on either $P[Y(1) =1]$ or $P[Y(0) = 1 \mid v_2]$. Thus, as long as Assumption A4 is binding, the bounds on the ATE will not be sensitive to classification

[11] Notice that for obesity, Assumption 4 is rejected by the data: the estimated MIV lower bound of 0.190 exceeds the upper bound of 0.074. As discussed above, we do not apply Assumption 4 to the obesity outcome.

error. Notice that for obesity, where we do not apply Assumption A4, the bounds are sensitive to the upper bound misreporting rate, λ.

4.2 Monotone Treatment Selection

Self-selection into the NSLP is the most common explanation for the positive correlation between participation and poor health. Unobserved factors associated with poor health are thought to be positively associated with the decision to take-up the program, S^*. The *monotone treatment selection* (MTS) assumption, which replaces the strong exogenous selection assumption implicit in much of the literature, specifies that children receiving free lunches are likely to have worse latent health outcomes on average than nonparticipants:

(A5) *Monotone Treatment Selection (MTS)*:

$$P[Y(t)=1 \mid S^*=0] \leq P[Y(t)=1 \mid S^*=1] \text{ for } t=0,1.$$

Under this MTS assumption, it follows that $P[Y(1)=1] \leq P[Y(1)=1 \mid S^*=1]$ and $P[Y(0)=1 \mid S^*=0] \leq P[Y(0)=1]$ (Manski and Pepper, 2000). Notice that this assumption is a special case of the A3 MIV assumption discussed above.

Figures 3A-3C trace out bounds on the ATE when combining the previous MIV assumptions with the MTS assumption. For all three outcomes, the most striking finding is that the joint MIV-MTS model identifies the ATE as strictly negative and substantial as long as the degree of misreporting is small. If participation is accurately reported, these estimates suggest that free lunch program reduces food insecurity by at least 5.9 points (from 0.405 to 0.346), or 15%, poor health by at least 3.2 points (from 0.108 to 0.072), or 33%, and obesity by at least 9.4 points (from 0.244 to 0.150), or 38%.

While these findings seem to indicate that the NSLP plays an important role in improving children's health, there are a number of reasons to temper conclusions based on this evidence. First, identification of the sign of the ATE is precluded under even mall degrees of classification error. Second, even with fully accurate reporting, the 90 percent confidence intervals include zero. Thus, we cannot reject the hypothesis that the program is ineffective in promoting healthy outcomes.

Still, while these findings do not clearly identify the sign of the ATE, the ambiguity created by the selection and measurement problems is notably reduced under this joint MIV-MTS model. Under low rates of misreporting, the ATE is estimated to lie within a few point range such that, at worst, the free lunch program appears to have negligible impacts on food insecurity, poor general health rates, and child obesity. At best, the program may substantially reduce the prevalence of negative health outcomes including obesity. This finding is in contrast to the previous literature that generally finds that the NSLP leads to modest increases in obesity.

5. Monotone Treatment Response

The free lunch program seems likely to decrease food insecurity and poor health. To formalize this idea, we consider the identifying power of the *monotone treatment response* (MTR) assumption introduced by Manski (1995 and 1997):

(A6) *Monotone Treatment Response (MTR):* $Y(0) \geq Y(1)$.

The MTR assumption implies that the ATE must be nonnegative: the free lunch program, by assumption, cannot increase the probability of being food insecure or in poor health[12]

The validity of this assumption is likely to be sensitive to the particular outcome in question. In particular, the MTR restriction seems innocuous when studying the impact of the NSLP on food security and general health. As Currie (2003) suggests, it is difficult to imagine that providing a free lunch would

---

[12] To see this, notice that observations where $S^* \neq t$ may now be informative about the latent outcome of interest. For instance, children receiving assistance would have no better health outcomes if they had not received assistance. Thus, the observed outcome $Y(1)$ provides a lower bound for the unobserved outcome $Y(0)$. Given this monotone treatment response (MTR) assumption, we know that $P[Y(0)=1 \mid S^*=0] \geq P[Y(1)=1 \mid S^*=0]$ and $P[Y(1)=1 \mid S^*=1] \leq P[Y(0)=1 \mid S^*=1]$. Thus, $P[Y(1)=1] \leq P(Y=1)$ and $P(Y=1) \leq P[Y(0)=1]$ (Manski, 1995; 1997).

increase the prevalence of food insecurity. In contrast, however, there is much debate about whether the NSLP might increase or decrease obesity. Because NSLP administrators must adhere to nutritional guidelines, one might expect the free lunch program to reduce obesity. Yet, the evidence provides a mixed picture suggesting that school lunches lead to some improved nutrient intake but also a higher portion of fat-related calories associated with obesity (see, e.g., Millimet et al., 2008). Moreover, the receipt of free meals through the NSLP allows families to purchase more food, which in turn could contribute to obesity, but also better quality food, which might lead to reductions in obesity.

Both the MTS and MTR assumptions reduce the upper bound on $P[Y(1)=1]$ and the lower bound on $P[Y(0)=1]$. Thus, under the joint MTR-MTS assumption, we have

$$P[Y(1)=1] \leq \min\{P[Y(1)=1 \mid S^{*}=1],\ P(Y=1)\}$$

and

$$P[Y(0)=1] \geq \max\{P[Y(0)=1 \mid S^{*}=0],\ P(Y=1)\}.$$

Under this joint assumption, the upper bound on the ATE is nonpositive whereas the upper bound under the MTS assumption alone may be positive or negative. The lower bound on the average treatment effect is unaffected by these assumptions.

Figures 3A-3C trace out bounds on the ATE when combining the MIV assumptions with the MTS and MTR assumptions. This MIV-MTS-MTR model allows us to identify the ATE for food insecurity as strictly negative across all degrees of potential misreporting. For example, the estimated bounds on ATE for the food insecurity rates (Figure 3A) under arbitrary misreporting vary from the 8-point range [-0.151, -0.076] when λ = 0 to [-0.398, -0.076] when λ = 0.25. In the latter case, the lower bound improves to -0.151 under the no false positives restriction. The ATE on the poor health probability (Figure 3B) is nearly point-identified in the absence of reporting error, lying in the range [-0.040, -0.032]. The sign cannot be identified, however, after allowing for the possibility of small degrees of reporting error. For obesity, where the eligibility criteria MIV assumption is not applied, much more uncertainty remains. Still, we can identify a strictly beneficial effect of the program on child obesity under the MIV-MTS-MTR assumption, even accounting for sampling variability, as long as the data are reported accurately. Under small degrees of misreporting, we cannot reject the hypothesis of no beneficial effects on child obesity.

6. Conclusion

Children receiving free or reduced-price school lunches through the National School Lunch Program often have worse health outcomes on average than observationally similar children who do not participate, especially in the case of food insecurity. Whether these puzzling correlations reflect causal impacts of the program has become a matter of considerable debate among researchers and policymakers. Much of the empirical literature maintains the untenable exogenous selection assumption, and the classification error problem appears to have been completely ignored. Reviewing the general literature on the causal impacts of food assistance programs, Currie (2003) goes so far as to conclude that "many studies have … simply 'punted' on the issue of identification."[13]

Our analysis considered the impact of the National School Lunch Program on health-related outcomes using nonparametric methods that allow us to simultaneously account for the selection and misclassification problems in a single unifying framework. The partial identification approach is well-suited for this application where conventional methods of addressing these selection and measurement problems (e.g., standard instrumental variable methods) are often viewed as non-credible and there remains much uncertainty about even the qualitative impacts of the program. By successively layering stronger identifying assumptions into the model, the approach makes transparent how assumptions on the selection and reporting error processes shape inferences about the causal impacts of the program.

Our general approach involves replacing common identifying independence assumptions with weaker monotonicity assumptions. In particular, we apply a series of monotone instrumental variables including the income MIV, used by KPGJ, and the MTS assumption suggested in the literature and

[13] Bhattacharya et al. (2006), in studying the National School Breakfast Program, suggest that "no study has dealt convincingly with endogenous participation."

formalized by Manski and Pepper (2000). Similar to a regression discontinuity design, we also introduce a new way to conceptualize the MIV assumption using eligibility criteria as monotone instruments. The eligibility MIV assumption has substantial identifying power in this application and may prove to have much wider applicability to other settings.

Under these plausible monotonicity restrictions, the ambiguity reflected in the bounds is narrowed substantially and in some cases suggests that the free lunch program improves health. For example, without classification errors, the joint MIV-MTS model reveals that the NSLP reduces poor health by at least 3.5 points, reduces food insecurity by at least 7.6 points, and reduces obesity by at least 9.4 points. These lower bound estimates are substantial. For example, the estimated upper bound for obesity of -9.4 implies that the NSLP reduces obesity by at least 38 percent. When we add in the MTR assumption, the estimated bounds narrow further and, except for the poor general health outcome, remain strictly negative even with substantial degrees of classification error. While in many cases we cannot reject the hypothesis that the NSLP has no impact of children's health, the results imply that at worst the NSLP has little impact on health but may lead to substantial reductions in food insecurity, poor health, and obesity.

A final contribution is our focus on outcomes besides obesity. While obesity is clearly an important health outcome, the NSLP literature has paid little attention to food insecurity and self-reported measures of health. In contrast, the literature assessing the efficacy of the Food Stamp Program (now known as the Supplemental Nutrition Assistance Program (SNAP)) has examined a wide range of outcomes, with a particular focus on food insecurity (e.g., DePolt et al., 2009; Gundersen and Kreider, 2008; Gundersen and Oliveira, 2001). In this application, we introduce powerful assumptions that can be applied credibly to the analysis of food insecurity and general health outcomes.

References

Bhattacharya, J., J. Currie, and S. Haider, 2006, Breakfast of Champions? The School Breakfast Program and the Nutrition of Children and Families. Journal of Human Resources 41, 445-466.

Bollinger, C.R., 1996, Bounding Mean Regressions When a Binary Variable is Mismeasured. Journal of Econometrics 73(2), 387-399.

Bollinger, C. and M. David,1997, Modeling Discrete Choice with Response Error: Food Stamp Participation. Journal of the American Statistical Association 92 (439), 827-835.

Currie, J., 2003, U.S. Food and Nutrition Programs, in: R. Moffitt, (Eds.), Means Tested Transfer Programs in the US, Chapter 4. University of Chicago Press.

DePolt, R.A., R.A. Moffitt, D.C.Ribar, 2009, Food Stamps, Temporary Assistance for Needy Families and Food Hardships in Three American Cities. Pacific Economic Review 14(4), 445-4473.

Fontaine, K., D. Redden, C. Wang, A. Westfall, and D. Allison, 2003, Years of Life Lost Due to Obesity. Journal of the American Medical Association 289(2), 187-193.

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), 352-382.

Gundersen, C. and V. Oliveira, 2001, The Food Stamp Program and Food Insufficiency. American Journal of Agricultural Economics, 84(3), 875-887.

Gundersen, C., S. Garasky, and B. Lohman B., 2009, Food Insecurity is not Associated with Childhood Obesity as Assessed Using Multiple Measures of Obesity. Journal of Nutrition 139, 1173-1178.

Horowitz, J.L., and C.F. Manski, 1995, Identification and Robustness With Contaminated and Corrupted Data. Econometrica 63(2), 281-02.

Horowitz, J.L., and C.F. Manski, 1998, Censoring of Outcomes and Regressors Due to Survey Nonresponse: Identification and Estimation Using Weights and Imputations. Journal of Econometrics 84, 37-58.

Imbens, G. and C. Manski, 2004, Confidence Intervals for Partially Identified Parameters. Econometrica 72(6), 1845-1857.

Kreider, B. and J. V. Pepper, forthcoming, Identification of Expected Outcomes in a Data Error Mixing Model with Multiplicative Mean Independence. Journal of Business and Economic Statistics.

Kreider, B., 2007, Disability and Employment: Reevaluating the Evidence in Light of Reporting Errors. Journal of the American Statistical Association 102(478), 432-441.

Kreider, B., and S.C. Hill, 2008, Partially Identifying Treatment Effects with an Application to Covering the Uninsured. Journal of Human Resources 44(2), 409-449.

Kreider, B., J. Pepper, C. Gundersen, and D. Jolliffe, 2009, Identifying the Effects of Food Stamps on Child Health Outcomes. Working Paper.

Manski, C.F., 1995, Identification Problems in the Social Sciences. Cambridge, MA: Harvard University Press.

Manski, C.F, 1997, Monotone Treatment Response. Econometrica 65(6), 1311-1334.

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

Manski, C.F. and J.V. Pepper, 2009, More on Monotone Instrumental Variable. The Econometrics Journal, 200-216.

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, 721-745.

Meyer, B., D.W. Mok, and J.X. Sullivan, 2009, The Under-Reporting of Transfers in Household Surveys: Its Nature and Consequences. Working Paper.

Millimet, D.L., R. Tchernis, and M. Husain, 2008, School Nutrition Programs and the Incidence of Childhood Obesity. NBER WP 14297.

Molinari, F, forthcoming, Missing Treatments. Journal of Economic and Business Statistics.

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

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

Schanzenbach, D., forthcoming, Does the Federal School Lunch Program Contribute to Childhood Obesity? Journal of Human Resources.

Table 1. Means by National School Lunch Program participation

	Income eligible children†	Recipients	Nonrecipients
Age in years	12.0	11.6*	13.1
	(0.087)	(0.091)	(0.194)
Income to Poverty Ratio (PIR)	0.982	0.923*	1.144
	(0.015)	(0.016)	(0.028)
NSLP recipient	0.734	1.000	0.000
	(0.014)	(0.000)	(0.000)
Outcomes			
Food Insecure (Household)	0.366	0.397*	0.282
	(0.015)	(0.018)	(0.027)
Obese (BMI ≥95th percentile)	0.187	0.194	0.168
	(0.012)	(0.015)	(0.022)
Poor or Fair Health	0.074	0.072	0.080
	(0.008)	(0.010)	(0.015)
N	2219	1683	536

Notes: Sample estimates weighted with the medical exam weight. The estimated means for the NSLP recipient population are superscripted with * or ** to indicate that they are statistically significantly different from the means for the non-recipient population (with p-values less than 0.01 and 0.05, respectively, based on Wald statistics corrected for complex design).

† Includes all children residing in households with income less than 185% of the poverty line and attending schools that offer a lunch provided by the National School Lunch Program.

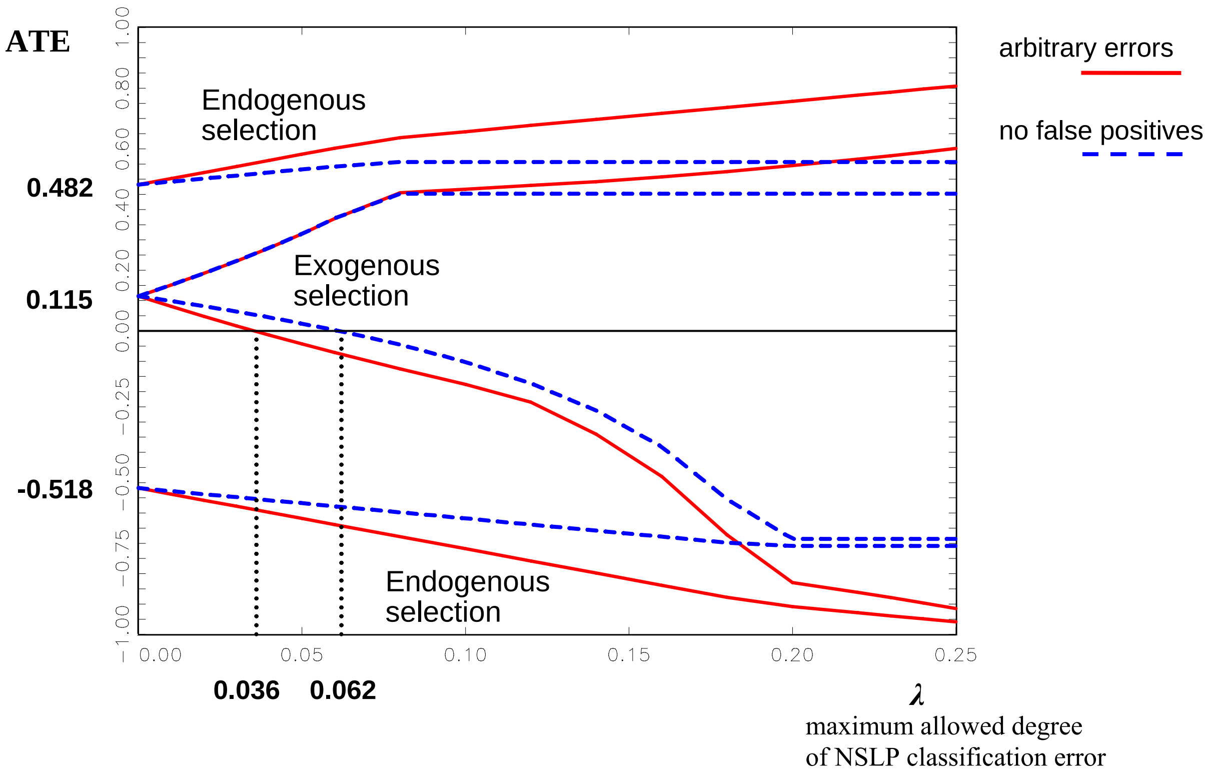
Table 2. Means by National School Lunch Program participation

	Income Eligible Children Attending NSLP Schools†	Income Eligible Children Attending non-NSLP Schools†	Income Eligible Dropouts†	Income Ineligible+
Age in years	12.0	13.3*	15.1*	12.2
	(0.087)	(0.489)	(0.350)	(0.134)
PIR	0.982	1.068**	1.003	2.421**
	(0.015)	(0.069)	(0.060)	(0.016)
NSLP recipient	0.734	0.000	0.000	0.226**
	(0.014)	(0.000)	(0.000)	(0.018)
Outcomes				
Food insecure (household)	0.366	0.361	0.451	0.132**
	(0.015)	(0.088)	(0.066)	(0.016)
Obese (BMI $\geq 95^{th}$ percentile)	0.187	0.190	0.074*	0.186
	(0.012)	(0.072)	(0.026)	(0.019)
Poor or Fair Health	0.074	0.071	0.093	0.018**
	(0.008)	(0.058)	(0.028)	(0.005)
N	2219	65	107	778

Notes: Sample estimates weighted with the medical exam weight. The estimated means for Columns (2) through (4) are superscripted with * or ** to indicate that they are statistically significantly different from the means for the income eligible population in Column (1) with p-values less than 0.01 and 0.05, respectively (based on Wald statistics corrected for complex design).

† Children residing in households with income less than 185 percent of the federal poverty line (FPL) are classified as income eligible, whereas those with income between 185-300% of the FPL are classified as ineligible. Among the income eligible households, some attend schools without the NSLP and some have dropped out of school.

Figure 1A. Sharp Bounds on the ATE for **Household Food Insecurity**:
Endogenous vs. Exogenous NSLP Participation†



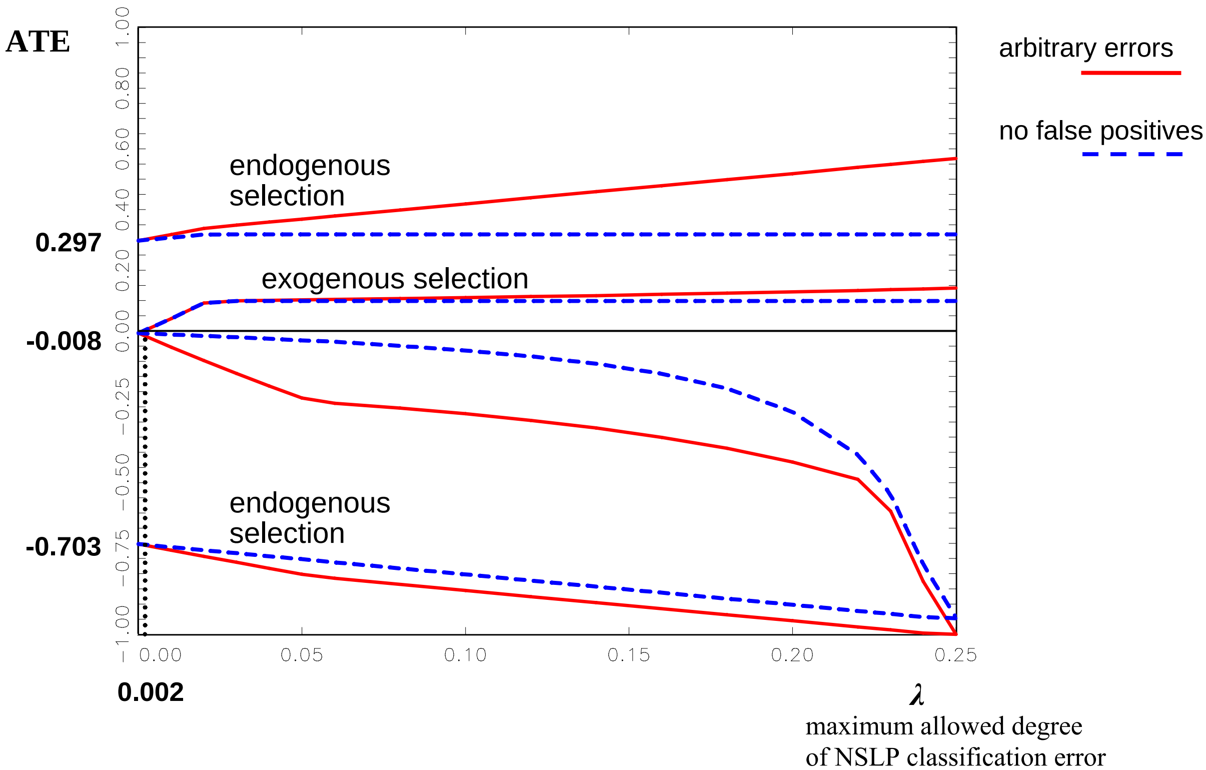
Selected values of λ

	Exogenous Selection		Endogenous Selection	
	Arbitrary Errors	No False Positives	Arbitrary Errors	No False Positives
λ=0	[0.115, 0.115]†	[0.115, 0.115]	[-0.518, 0.482]	[-0.518, 0.482]
	[0.061 0.169]‡	[0.061 0.169]	[-0.538 0.503]	[-0.538 0.503]
λ=0.05	[-0.043, 0.320]	[0.025, 0.320]	[-0.618, 0.582]	[-0.568, 0.532]
	[-0.082 0.370]	[-0.024 0.370]	[-0.638 0.603]	[-0.588 0.553]
λ=0.10	[-0.177, 0.467]	[-0.102, 0.452]	[-0.718, 0.657]	[-0.618, 0.557]
	[-0.219 0.489]	[-0.165 0.475]	[-0.738 0.677]	[-0.638 0.577]
λ=0.25	[-0.915, 0.601]	[-0.685, 0.452]	[-0.959, 0.807]	[-0.709, 0.557]
	[-0.952 0.637]	[-0.705 0.475]	[-0.977 0.827]	[-0.727 0.577]

† Point estimates of the population bounds

‡ 90% confidence intervals around ATE are calculated using methods from Imbens-Manski (2004) with 1,000 pseudosamples

Figure 1B. Sharp Bounds on the ATE for **Child Poor Health**: Endogenous vs. Exogenous NSLP Participation†



Selected values of λ

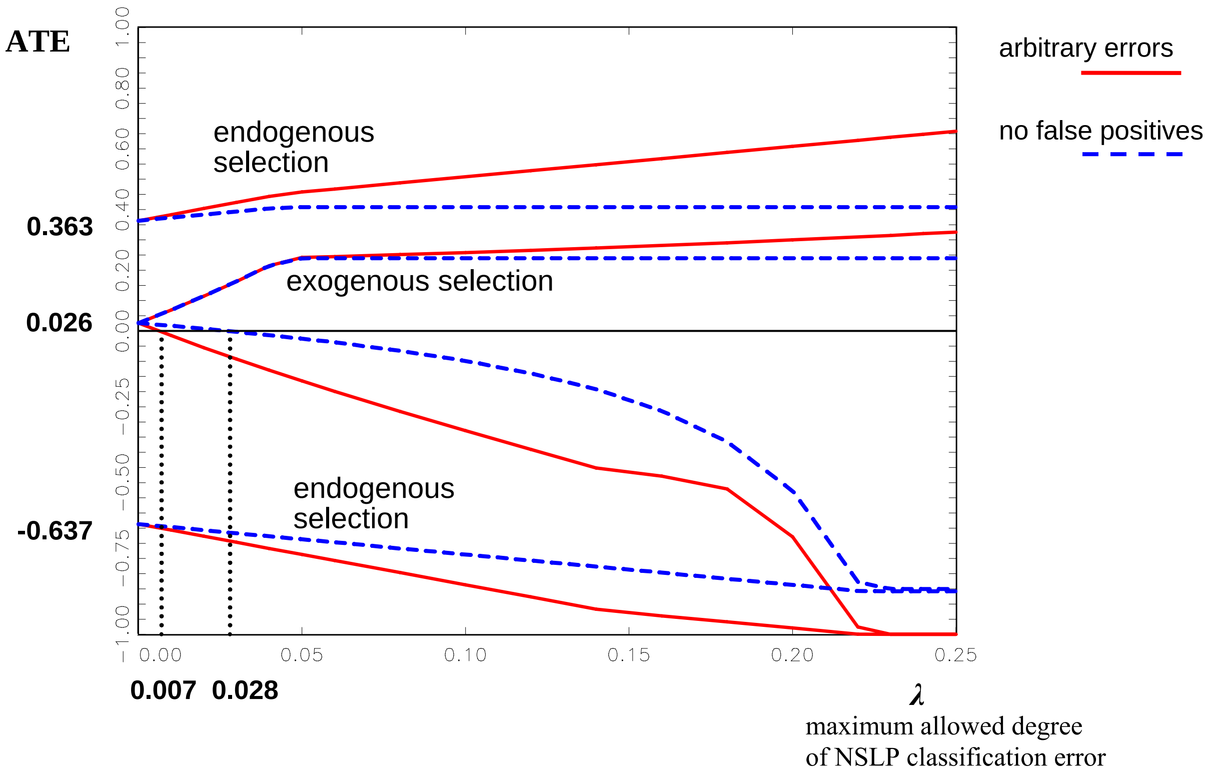
	Exogenous Selection		Endogenous Selection	
	Arbitrary Errors	No False Positives	Arbitrary Errors	No False Positives
λ=0	[-0.008, -0.008]†	[-0.008, -0.008]	[-0.703, 0.297]	[-0.703, 0.297]
	[-0.035 0.020]‡	[-0.035 0.020]	[-0.722 0.317]	[-0.722 0.317]
λ=0.05	[-0.221, 0.102]	[-0.031, 0.098]	[-0.803, 0.369]	[-0.703, 0.297]
	[-0.242 0.116]	[-0.056 0.112]	[-0.821 0.388]	[-0.722 0.317]
λ=0.10	[-0.273, 0.110]	[-0.064, 0.098]	[-0.855, 0.419]	[-0.803, 0.319]
	[-0.303 0.125]	[-0.097 0.112]	[-0.874 0.438]	[-0.822 0.338]
λ=0.25	[-1.000, 0.141]	[-0.946, 0.098]	[-1.000, 0.569]	[-0.947, 0.319]
	[-1.000 0.160]	[-1.000 0.112]	[-1.000 0.588]	[-0.958 0.338]

---

† Point estimates of the population bounds

‡ 90% confidence intervals around ATE are calculated using methods from Imbens-Manski (2004) with 1,000 pseudosamples

Figure 1C. Sharp Bounds on the ATE for **Child Obesity**:
Endogenous vs. Exogenous NSLP Participation†



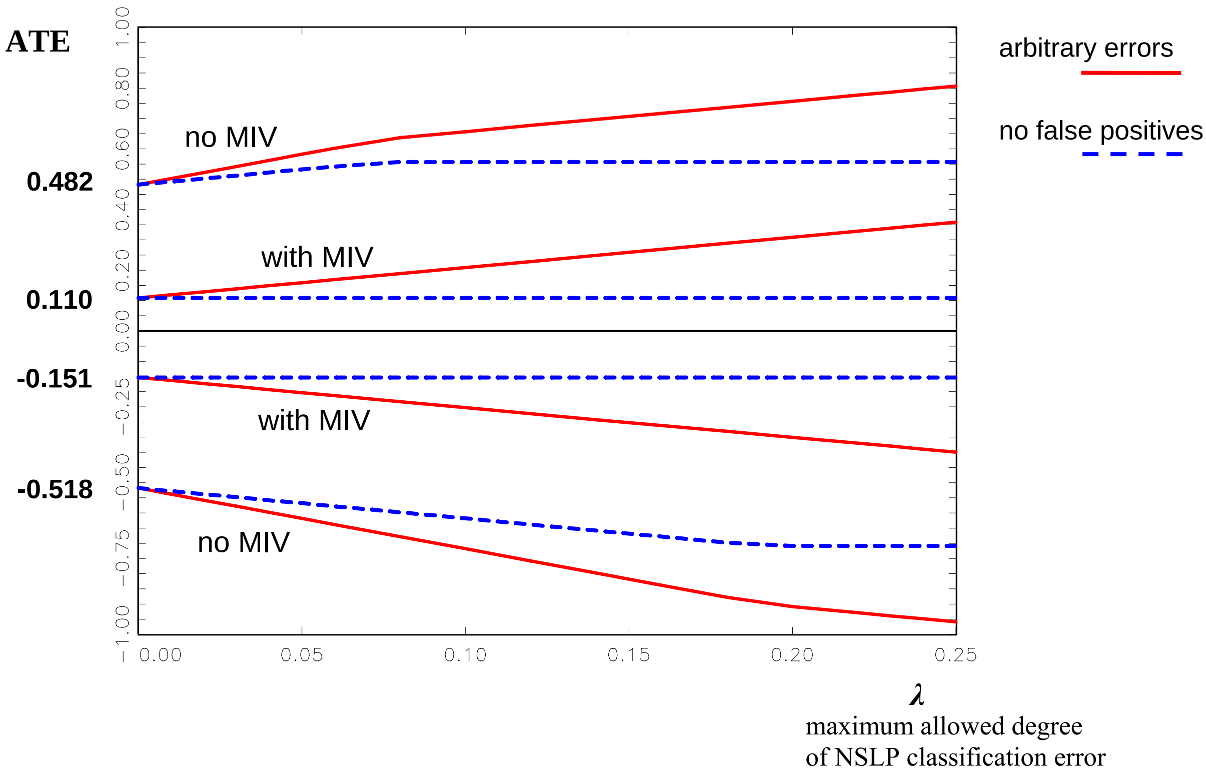
Selected values of λ

	Exogenous Selection		Endogenous Selection	
	Arbitrary Errors	No False Positives	Arbitrary Errors	No False Positives
λ=0	[0.026, 0.026]† [-0.018 0.069]‡	[0.026, 0.026] [-0.018 0.069]	[-0.637, 0.363] [-0.657 0.384]	[-0.637, 0.363] [-0.657 0.384]
λ=0.05	[-0.165, 0.242] [-0.197 0.263]	[-0.025, 0.240] [-0.065 0.261]	[-0.737, 0.458] [-0.757 0.479]	[-0.687, 0.408] [-0.707 0.429]
λ=0.10	[-0.329, 0.258] [-0.361 0.279]	[-0.099, 0.240] [-0.149 0.260]	[-0.837, 0.508] [-0.857 0.529]	[-0.737, 0.408] [-0.757 0.429]
λ=0.25	[-1.000, 0.326] [-1.000 0.351]	[-0.851, 0.240] [-0.874 0.260]	[-1.000, 0.658] [-1.000 0.679]	[-0.858, 0.408] [-0.872 0.429]

† Point estimates of the population bounds

‡ 90% confidence intervals around ATE are calculated using methods from Imbens-Manski (2004) with 1,000 pseudosamples

Figure 2A. Sharp Bounds on the ATE for **Household Food Insecurity**: Endogenous NSLP Participation, with vs. without MIV[†]



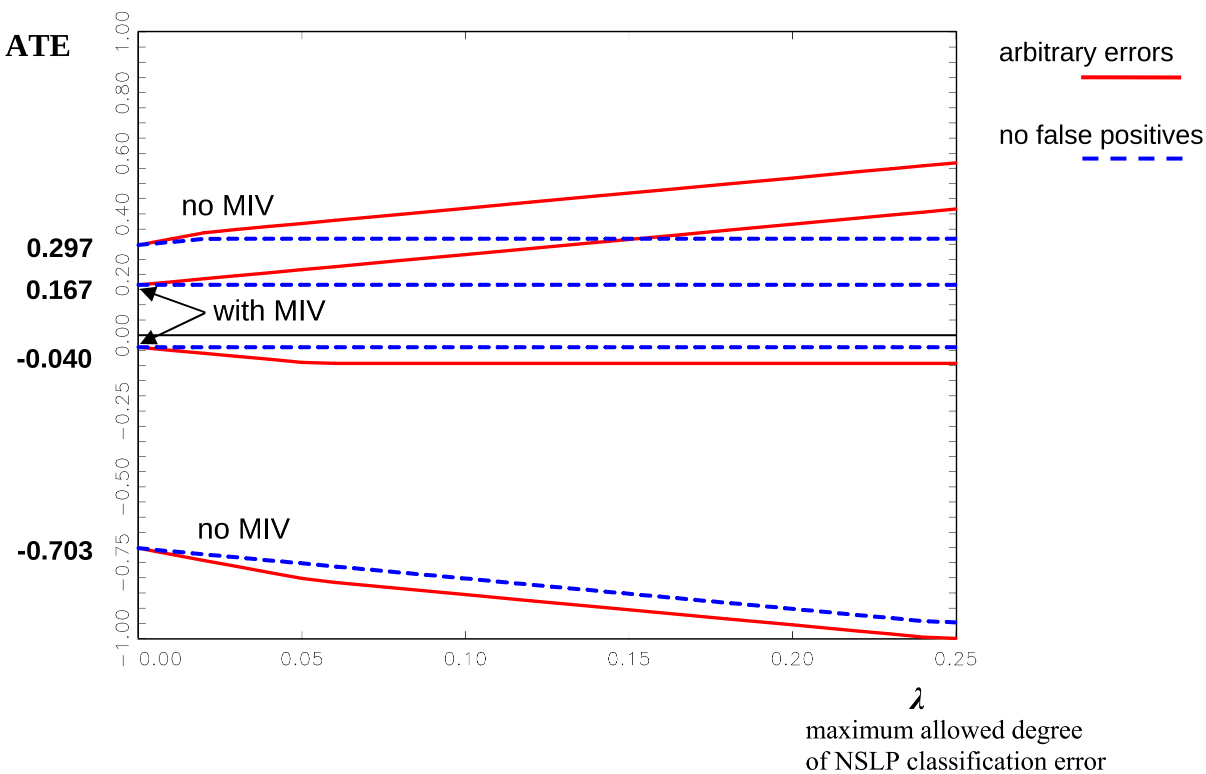
Selected values of λ

	No MIV		With MIV	
	Arbitrary Errors	No False Positives	Arbitrary Errors	No False Positives
λ=0	[-0.518, 0.482][†] [-0.538 0.503][‡]	[-0.518, 0.482] [-0.538 0.503]	[-0.151, 0.110] [-0.250 0.260]	[-0.151, 0.110] [-0.250 0.260]
λ=0.05	[-0.618, 0.582] [-0.638 0.603]	[-0.568, 0.532] [-0.588 0.553]	[-0.201, 0.160] [-0.300 0.311]	[-0.151, 0.110] [-0.250 0.260]
λ=0.10	[-0.718, 0.657] [-0.738 0.677]	[-0.618, 0.557] [-0.638 0.577]	[-0.251, 0.210] [-0.350 0.361]	[-0.151, 0.110] [-0.250 0.260]
λ=0.25	[-0.959, 0.807] [-0.977 0.827]	[-0.709, 0.557] [-0.727 0.577]	[-0.398, 0.359] [-0.498 0.510]	[-0.151, 0.110] [-0.250 0.260]

[†] Point estimates of the population bounds

[‡] 90% confidence intervals around ATE are calculated using methods from Imbens-Manski (2004) with 1,000 pseudosamples

Figure 2B. Sharp Bounds on the ATE for **Child Poor Health**: Endogenous NSLP Participation, with vs. without MIV†



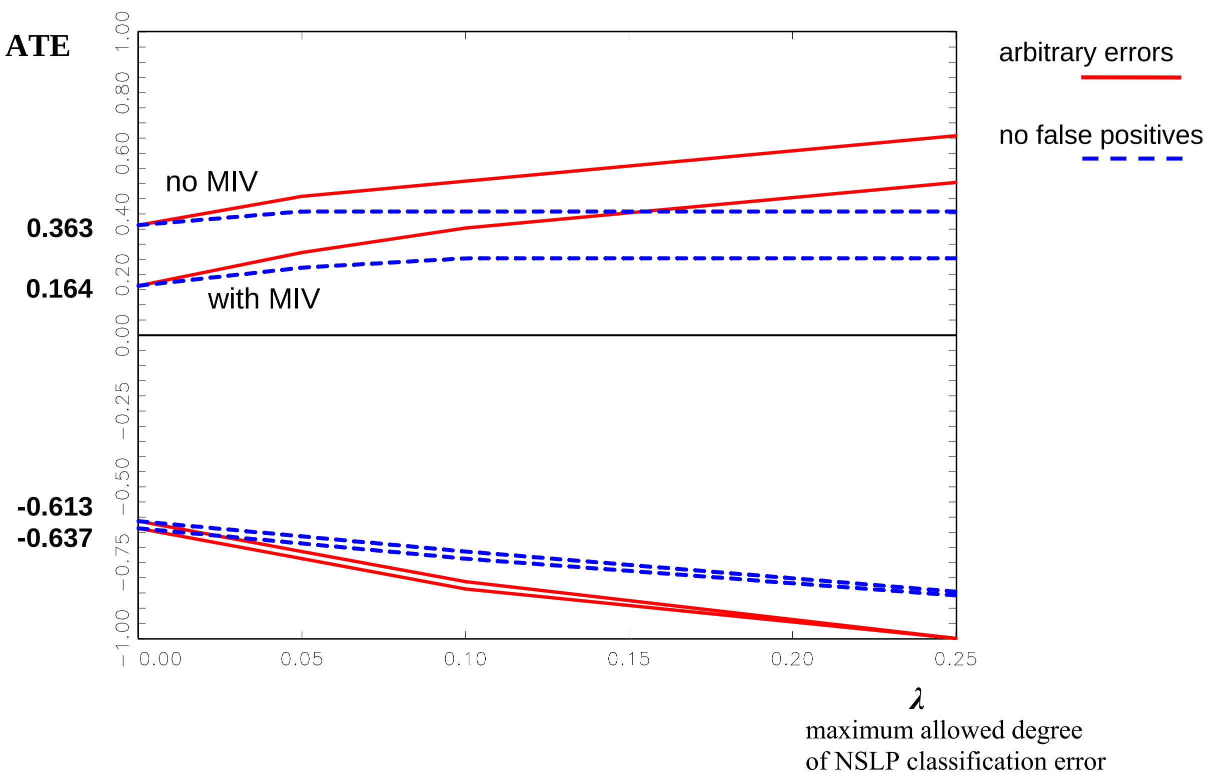
Selected values of λ

	No MIV		With MIV	
	Arbitrary Errors	No False Positives	Arbitrary Errors	No False Positives
λ=0	[-0.703, 0.297]†	[-0.703, 0.297]	[-0.040, 0.167]	[-0.040, 0.167]
	[-0.722 0.317]‡	[-0.722 0.317]	[-0.073 0.256]	[-0.073 0.256]
λ=0.05	[-0.803, 0.369]	[-0.703, 0.297]	[-0.090, 0.217]	[-0.040, 0.167]
	[-0.821 0.388]	[-0.722 0.317]	[-0.123 0.309]	[-0.073 0.256]
λ=0.10	[-0.855, 0.419]	[-0.803, 0.319]	[-0.093, 0.417]	[-0.040, 0.167]
	[-0.874 0.438]	[-0.822 0.338]	[-0.125 0.509]	[-0.073 0.256]
λ=0.25	[-1.000, 0.569]	[-0.947, 0.319]	[-0.093, 0.417]	[-0.040, 0.167]
	[-1.000 0.588]	[-0.958 0.338]	[-0.125 0.509]	[-0.073 0.256]

† Point estimates of the population bounds

‡ 90% confidence intervals around ATE are calculated using methods from Imbens-Manski (2004) with 1,000 pseudosamples

Figure 2C. Sharp Bounds on the ATE for **Child Obesity**: Endogenous NSLP Participation, with vs. without MIV†



Selected values of λ

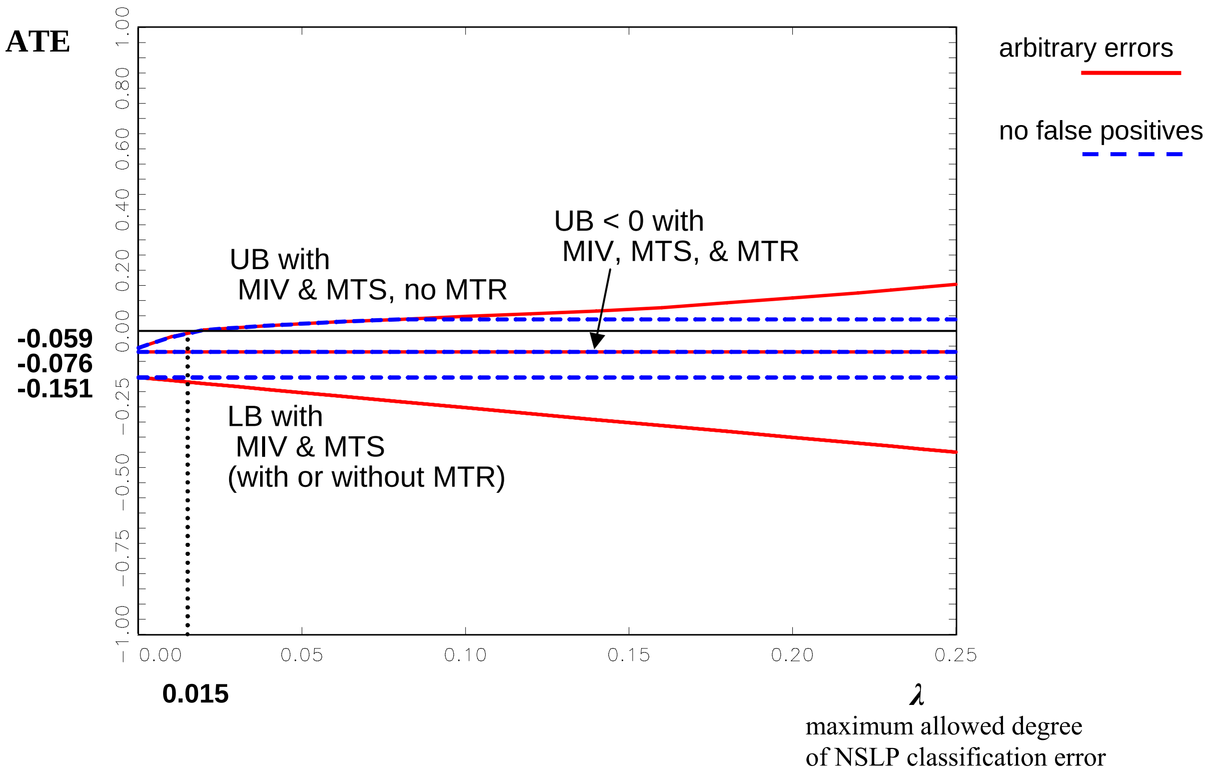
	No MIV		With MIV*	
	Arbitrary Errors	No False Positives	Arbitrary Errors	No False Positives
λ=0	[-0.637, 0.363]†	[-0.637, 0.363]	[-0.613, 0.164]	[-0.613, 0.164]
	[-0.657 0.384]‡	[-0.657 0.384]	[-0.656 0.256]	[-0.656 0.256]
λ=0.05	[-0.737, 0.458]	[-0.687, 0.408]	[-0.713, 0.273]	[-0.663, 0.223]
	[-0.757 0.479]	[-0.707 0.429]	[-0.756 0.364]	[-0.706 0.314]
λ=0.10	[-0.837, 0.508]	[-0.737, 0.408]	[-0.813, 0.354]	[-0.713, 0.254]
	[-0.857 0.529]	[-0.757 0.429]	[-0.856 0.436]	[-0.756 0.336]
λ=0.25	[-1.000, 0.658]	[-0.858, 0.408]	[-1.000, 0.504]	[-0.846, 0.254]
	[-1.000 0.679]	[-0.872 0.429]	[-1.000 0.582]	[-0.871 0.332]

† Point estimates of the population bounds

‡ 90% confidence intervals around ATE are calculated using methods from Imbens-Manski (2004) with 1,000 pseudosamples

* For the obesity outcome, only the income-to-poverty ratio MIV assumption is imposed; the MIV ineligibles assumptions are not imposed.

Figure 3A. Sharp Bounds on the ATE for **Household Food Insecurity**: Endogenous NSLP Participation with MIV, MTS, and MTR[†]



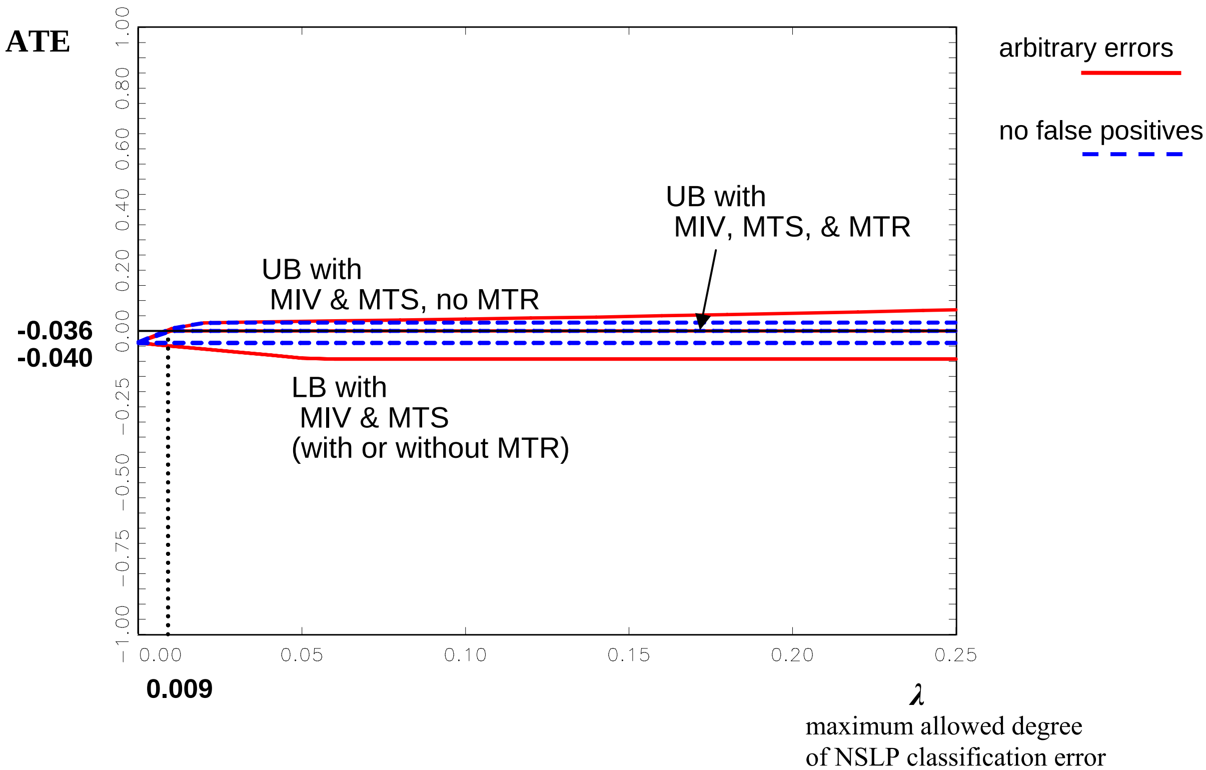
Selected values of λ

	MIV & MTS		MIV, MTS, & MTR	
	Arbitrary Errors	No False Positives	Arbitrary Errors	No False Positives
λ=0	[-0.151, -0.059][†] [-0.250 0.060][‡]	[-0.151, -0.059] [-0.250 0.060]	[-0.151, -0.076] [-0.250 0.000]	[-0.151, -0.076] [-0.250 0.000]
λ=0.05	[-0.201, 0.020] [-0.300 0.153]	[-0.151, 0.020] [-0.250 0.153]	[-0.201, -0.076] [-0.300 0.000]	[-0.151, -0.076] [-0.250 0.000]
λ=0.10	[-0.251, 0.044] [-0.350 0.184]	[-0.151, 0.031] [-0.250 0.171]	[-0.251, -0.076] [-0.350 0.000]	[-0.151, -0.076] [-0.250 0.000]
λ=0.25	[-0.398, 0.150] [-0.498 0.314]	[-0.151, 0.031] [-0.250 0.173]	[-0.398, -0.076] [-0.498 0.000]	[-0.151, -0.076] [-0.250 0.000]

† Point estimates of the population bounds

‡ 90% confidence intervals around ATE are calculated using methods from Imbens-Manski (2004) with 1,000 pseudosamples

Figure 3B. Sharp Bounds on the ATE for **Child Poor Health**:
Endogenous NSLP Participation with MIV, MTS, and MTR†



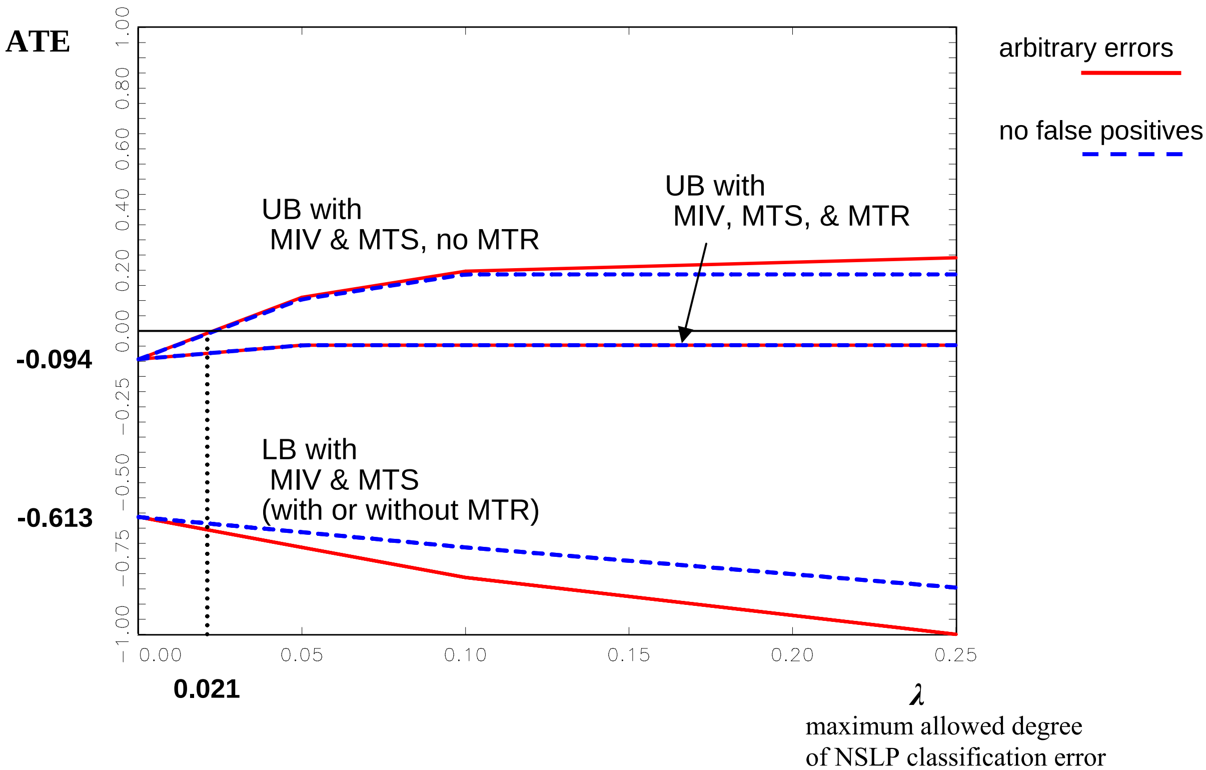
Selected values of λ

	MIV & MTS		MIV, MTS, & MTR	
	Arbitrary Errors	No False Positives	Arbitrary Errors	No False Positives
λ=0	[-0.040, -0.036]† [-0.075 0.039]‡	[-0.040, -0.036] [-0.075 0.039]	[-0.040, -0.036] [-0.075 0.000]	[-0.040, -0.036] [-0.075 0.000]
λ=0.05	[-0.090, 0.031] [-0.123 0.098]	[-0.040, 0.028] [-0.074 0.095]	[-0.090, 0.000] [-0.123 0.000]	[-0.040, 0.000] [-0.074 0.000]
λ=0.10	[-0.093, 0.039] [-0.125 0.105]	[-0.040, 0.028] [-0.074 0.095]	[-0.093, 0.000] [-0.125 0.000]	[-0.040, 0.000] [-0.074 0.000]
λ=0.25	[-0.093, 0.070] [-0.125 0.137]	[-0.040, 0.028] [-0.074 0.095]	[-0.093, 0.000] [-0.125 0.000]	[-0.040, 0.000] [-0.074 0.000]

† Point estimates of the population bounds

‡ 90% confidence intervals around ATE are calculated using methods from Imbens-Manski (2004) with 1,000 pseudosamples

Figure 3C. Sharp Bounds on the ATE for **Child Obesity**:
Endogenous NSLP Participation with MIV, MTS, and MTR†



Selected values of λ

	MIV* & MTS		MIV*, MTS, & MTR	
	Arbitrary Errors	No False Positives	Arbitrary Errors	No False Positives
λ=0	[-0.613, -0.094]† [-0.656 0.012]‡	[-0.613, -0.094] [-0.656 0.012]	[-0.613, -0.094] [-0.656 -0.001]	[-0.613, -0.094] [-0.656 -0.001]
λ=0.05	[-0.713, 0.111] [-0.756 0.217]	[-0.663, 0.105] [-0.706 0.211]	[-0.713, -0.047] [-0.756 0.000]	[-0.663, -0.047] [-0.706 0.000]
λ=0.10	[-0.813, 0.197] [-0.856 0.276]	[-0.713, 0.186] [-0.756 0.258]	[-0.813, -0.047] [-0.856 0.000]	[-0.713, -0.047] [-0.756 0.000]
λ=0.25	[-1.000, 0.241] [-1.000 0.329]	[-0.846, 0.186] [-0.871 0.249]	[-1.000, -0.047] [-1.000 0.000]	[-0.846, -0.047] [-0.871 0.000]

† Point estimates of the population bounds

‡ 90% confidence intervals around ATE are calculated using methods from Imbens-Manski (2004) with 1,000 pseudosamples

*For the obesity outcome, only the income-to-poverty ratio MIV assumption is imposed; the MIV ineligibles assumptions are not imposed.