Does the Women, Infants, and Children Program (WIC) Improve Infant Health Outcomes?*

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

John V. Pepper  
Department of Economics  
University of Virginia  
jvp3m@virginia.edu

Manan Roy  
IMPAQ International, LLC  
mroy@impaqint.com

July 18, 2014

Abstract: We derive new nonparametric bounds on average treatment effects to evaluate causal impacts of prenatal WIC participation on healthy birth outcomes. These methods allow for “partial verification” of treatment status given potentially misreported WIC receipt and introduce the notion of a partially observed monotone instrument in a modified regression discontinuity design. The estimated bounds are compared with point estimates from a parametric approach developed by Lewbel (2012). Combining survey data from the Early Childhood Longitudinal Study – Birth cohort of 2001 (ECLS-B) with administrative data from the USDA, we estimate that WIC reduces the prevalence of unhealthy birth weights by at least 20 percent and unhealthy gestation duration by at least 6 percent.

JEL: C14, C21, I12, I18

Keywords: Special Supplemental Nutrition Program for Women, Infants, and Children, WIC, Children, Treatment Effects, Health Outcomes, Instrumental Variables, Partial Identification, Nonparametric Bounds, Classification Error

*The authors received valuable comments from Craig Gundersen, James Ziliak, and participants at the Research Program on Childhood Hunger Progress Report Conferences. This project was supported with a grant from the University of Kentucky Center for Poverty Research through funding by the U.S. Department of Agriculture, Food and Nutrition Service, contract number AG-3198-B-10-0028. Pepper also acknowledges financial support from the Bankard Fund for Political Economy. The opinions and conclusions are solely those of the authors and do not necessarily reflect the opinions or policies of the UKCPR or any agency of the Federal Government.
1. INTRODUCTION

The Supplemental Nutrition Program for Women, Infants, and Children (WIC) is an early intervention food assistance program that provides benefits to about nine million recipients per year – and nearly half of all pregnant women in the United States – at an annual cost of about $7 billion.\(^1\) Driven in part by concerns over discretionary government spending and in part by the growing perception that early life conditions have long term impacts on adult life outcomes (Heckman and Carneiro, 2003; and Almond and Currie, 2011), there has been a renewed interest in understanding the impact of WIC on infant health.

Evaluating the efficacy of WIC, however, is complicated by two distinct identification problems. First, a selection problem arises because the decision to take up WIC is likely to be endogenous; unobservable factors associated with both program participation and health outcomes confound inference. Second, a measurement problem arises because households are known to systematically underreport the receipt of food assistance in national surveys (e.g., Bollinger and David, 1997; Bitler et al., 2003; and Meyer et al., 2009), and the propensity to misreport may vary across households based on unobserved characteristics.\(^2\) The classical measurement error assumption of non-mean-reverting errors is automatically violated with binary variables (e.g., Bollinger 1996).

To evaluate causal impacts of WIC on birth outcomes, we use a partial identification framework that simultaneously addresses these two identification problems. To do so, we extend the nonparametric bounding strategy proposed in Kreider, Pepper, Gundersen, and Jolliffe (2012, hereafter KPGJ) to allow for “partial verification” of potentially misreported treatment status and a modified discontinuity design with partially observed monotone instruments. In this setting, we derive sharp bounds on causal treatment effects under both exogenous and endogenous selection into the program.

We estimate these nonparametric models using data from the Early Childhood Longitudinal Study-Birth cohort of 2001 (ECLS-B) and auxiliary information from USDA administrative data. We focus

\(^1\) See http://www.fns.usda.gov/sites/default/files/WIC-Fact-Sheet.pdf

\(^2\) Potential stigma, for example, has long been thought to play a major role in underreported benefits (e.g., Moffitt, 1983).
on evaluating average treatment effects of a mother’s prenatal WIC participation on indicators of
normal birth weight (between 2500 grams and 4000 grams) and full term pregnancy (gestation age
between 38 and 42 weeks).

Several recent analyses address the selection problem using instrumental variable (IV)
models in narrowly defined samples (e.g., Bitler and Currie, 2005; Joyce et al., 2005, 2008; and
Figlio et al., 2009). For example, Bitler and Currie (2005) evaluate the impact of WIC on Medicaid-
eligible women, and Figlio et al. (2009) evaluate the impact on infants with older school aged
siblings. More recently, Hoynes et al. (2011) evaluate the performance of WIC at the time of its
establishment as a pilot program in 1972, exploiting the plausibly exogenous variation in
participation due to the staggered introduction of the program in the 1970s. These studies tend to
find that WIC leads to improved infant health, though evidence is mixed. We contribute to this
literature by relaxing traditional IV assumptions and by explicitly modeling the potential
underreporting of WIC benefits. Our framework does not require the traditional instrumental variable
or measurement error assumptions, or the linear response model.3

After describing the data in Section 2, we formalize the empirical question and identification
problems in Section 3. In Section 4, we derive sharp bounds on gaps in healthy birth probabilities
between WIC recipients and nonrecipients when true participation status may be mismeasured. With
no classification errors, these descriptive statistics point-identify average treatment effects (ATE)
under exogenous selection. With classification errors, we can only bound these differences in
conditional outcome probabilities.

In Section 5, we turn our attention to drawing inferences on the ATE given the endogenous
selection problem, abstracting away from measurement issues. Manski’s (1995) classic worst-case
ATE bounds impose no assumptions on the selection process. To tighten these bounds, we consider
the identifying power of several types of monotone instrumental variables (MIV) (Manski and

3 The classical linear response model assumption is difficult to justify when considering programs that are thought to
have heterogeneous effects (Moffitt 2005).
Pepper, 2000) in which we posit monotonic relationships between the latent probabilities of healthy birth outcomes and certain observed covariates, such as household income. Compared with standard IVs, these less restrictive MIVs are attractive in that they require no priori exclusion restriction (or mean independence assumption). We require only that healthy birth probabilities vary monotonically with the instrument. Modifying a discontinuity design, we also introduce the notion of a partially observed monotone instrument based on a federal law that confers adjunctive eligibility into WIC through participation in other assistance programs. Finally, we consider the identifying power of a monotone treatment response (MTR) assumption (Manski, 1997).

In Section 6, we derive sharp bounds on average treatment effects that simultaneously account for both the classification error and selection problems. Combining endogenous selection with measurement error naturally weakens what we can conclude about the average treatment effects. Nevertheless, we can still identify beneficial impacts of WIC on healthy birth outcomes under relatively mild assumptions. Perhaps counterintuitively, in some cases the estimated lower bounds on the average treatment effects rise, not fall, with the degree of misclassification. To complement our partial identification results, we present point estimates using Lewbel’s (2012) parametric model that circumvents the need for traditional instruments by exploiting conditional second moment restrictions. Section 7 concludes.

2. DATA

To study the impact of WIC on infant health, we use data from the first wave (2001) of the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), a nationally representative cohort of children born in 2001 through first grade. Assembled by the U.S. Department of Education, the ECLS-B focuses on children’s early environmental characteristics like health care and in- and out-of-home experiences that play a crucial role in the overall development of children and the first brush with the demands of formal school. The survey collects information directly from the children’s

---

4 ECLS-B case-level data are available to researchers who are granted a restricted-use data license. Information about receiving a restricted-use data license can be found at http://nces.ed.gov/pubsearch/licenses.asp.
parents, video-tapes parent-child interactions, and assesses child care settings. Data were collected between Fall 2001 and Fall 2002. The parents of 10,700 children born in 2001 participated in the first wave of the study when the children were approximately nine months old. As is common in the literature, we restrict our analysis to households that are income-eligible for WIC: family gross income cannot exceed 185 percent of the U.S. Poverty Income Guidelines ($44,123 in 2014 for a family of four). The sample is further restricted to infants who are singletons without any missing information on age. Our primary sample includes 4,350 nine-month-old infants from WIC income-eligible families. For parts of our analysis, we also use information on infants residing in households with incomes above the income threshold but below the fourth quintile of the SES distribution (N ≈ 1,250) to estimate a modified regression discontinuity design similar to the one applied in Gundersen et al. (2012) and Schanzenbach (2009).

Table 1 displays means and standard deviations for the variables used in our analysis for the main sample of infants classified as income-eligible to receive WIC (N=4,350). The last two columns report the differences in means between the treated (prenatal WIC recipients) and the untreated (income eligible nonparticipants) with p-values.

2.1. Covariates

For each respondent, we observe measures of the infant’s race (white, black, Asian, and Hispanic), gender, household socioeconomic status (SES), mother’s age, father’s age, dummy variables for whether the mother has a high school (HS) degree or less, a corresponding indicator for father’s education, an indicator for whether the parents are married, region (Northeast, Midwest,
South, and West), and city type (urban cluster, urban area, and rural). These covariates are not required for the partial identification analyses but are used when implementing Lewbel’s (2012) model in Section 6.

Table 1 reveals differences in the demographic characteristics by self-reported WIC status. Infants whose mothers report prenatal WIC receipt are less likely to be white or Asian, and more likely to be black or Hispanic. These infants tend to be from households with lower SES and fewer family members, and they are more likely to have younger unmarried parents. The mothers are less likely to report inadequate care (consistent with WIC’s focus on nutritional and health care education), more likely to be overweight, and less likely to have more than a high school degree.

### 2.2. WIC Participation Indicator and Misclassification

In addition to these covariates, we also observe a self-reported indicator of prenatal participation in WIC. This binary treatment variable takes a value of one if the mother reports receiving WIC benefits during pregnancy, and zero otherwise. In the survey, 68.7 percent of the eligible households report prenatal WIC receipt. This participation rate is similar to those found in other surveys (e.g., the CPS and SIPP) but lower than analogous rates found using administrative data (Bitler et al., 2003; USDA, 2006). As discussed above, Bitler et al. (2003) and Meyer et al. (2009) find evidence of substantial underreporting of WIC participation in the CPS, SIPP, and PSID.

To directly assess the nature and degree of classification errors in the ECLS-B, we use administrative data on the size of the WIC caseload to estimate the size of the eligible population (USDA, 2014). This calculation reveals a true participation rate of 0.74, notably higher than the self-reported rate of 0.687.

Additional insight into the nature and degree of classification errors can be found using the Hausman et al. (1998) parametric model of asymmetric misclassification. Estimates from this model imply a true participation rate of 0.838 with fewer than one percent of the eligible population falsely reporting WIC participation (see Section 4, footnote 13 for additional details). Thus, consistent with
the related SNAP literature, estimates suggest large degrees of underreporting of WIC participation in the ECLS-B but negligible rates of false positive reporting.

In the classification error models applied below, we show how knowledge of the true participation rate and the self-reported rate can be used to provide informative restrictions on the measurement error problem. Given uncertainty about the true participation rate, we assess how the results vary with different conjectured values of the true participation rate.

2.3. Birth Outcomes

Finally, we observe a number of different measures of infant health related to birth weight and gestation age. We focus mostly on indicators of normal birth weight (between 2500 grams and 4000 grams) and normal gestation length (between 38 and 42 weeks). Descriptive statistics in Table 1 reveal that infants whose parents report having received prenatal WIC have slightly better birth weight outcomes on average but worse gestation length outcomes than income-eligible nonparticipants. For example, the gap in the probability of a normal birth weight is 1.9 percentage points while the gap in the probability of a normal gestation length is −2.2 percentage points. We also evaluate other measures of favorable birth outcomes including birth weight of at least 1500 grams (not very low birth weight), at least 2500 grams (not low birth weight), no more than 4000 grams (not macrosomic), and indicators for near-term pregnancy – gestation age of at least 33 weeks (not very premature) or at least 37 weeks (not premature).10

A large literature documents the importance of early health outcomes, even as early as in utero, in influencing future adult outcomes (see Almond, 2006). That birth weight and gestation length affect future outcomes has been widely documented. Breslau et al. (1994), Brooks-Gunn et al. (1996), and Currie and Hyson (1999), for example, link birth weight to average scores on several

---

10 About 7 percent of birth weights in the sample are clinically macrosomic. Boulet et al. (2004) find that macrosomia is related to fetal injury, perinatal asphyxia, and fetal death, as well as complications for the mother like increasing the probability of caesarean delivery. Results for these outcomes follow similar patterns to those in the main analysis and are available from the authors upon request.
different tests of intellectual and social development. Goldenberg and Cullhane (2007) find that low birth weight is strongly associated with later adult chronic medical conditions like diabetes, hypertension, and heart disease. Using birth weight, gestation age, and Apgar score as metrics for infant health, Oreopoulos et al. (2008) conclude that poor infant health is a predictor of mortality within one year as well as mortality until age 17.\footnote{\textsuperscript{11} Other studies exploiting sibling comparisons include Conley and Bennett (2000), Johnson and Schoeni (2007), Lawlor et al. (2006), Black et al. (2007), Royer (2009), and Currie and Moretti (2007).}

Boyle and Boyle (2013) review and summarize the current available literature on infants born at moderate preterm (32-33 weeks) and late preterm (34-36 weeks) gestations and conclude that preterm infants face significantly greater risks of morbidity and mortality than previously believed. Goldenberg et al. (2008) also identify preterm birth as the leading cause of perinatal morbidity and mortality in developed countries. For even longer term outcomes, Crump et al. (2011) find that shorter gestation is most significantly associated with increased mortality in early childhood and mortality related to congenital anomalies and respiratory, endocrine, and cardiovascular disorders in young adults. Morse et al. (2009) report a positive association between preterm birth and risks of developmental delay and school-related problems such as risk of suspension and disability in prekindergarten at age three and four, among others.

3. THE SELECTION AND MEASUREMENT PROBLEMS

Our interest is in learning about average treatment effects (ATE) of prenatal maternal WIC participation on infant health among income-eligible households. For binary outcomes, these treatment effects can be expressed as

\[
ATE(1, 0 \mid X \in \Omega) = E[H(1) \mid X \in \Omega] - E[H(0) \mid X \in \Omega]
\]

\[
= P[H(1) = 1 \mid X \in \Omega] - P[H(0) = 1 \mid X \in \Omega]
\]

where \(H\) is the realized health outcome, \(H(1)\) denotes the infant’s health if he or she were to receive WIC, \(H(0)\) denotes the analogous outcome if the infant were not to receive WIC, and \(X \in \Omega\) denotes

\textsuperscript{11} Other studies exploiting sibling comparisons include Conley and Bennett (2000), Johnson and Schoeni (2007), Lawlor et al. (2006), Black et al. (2007), Royer (2009), and Currie and Moretti (2007).
conditioning on observed covariates whose values lie in the set Ω. Thus, the average treatment effect reveals the mean health effect of prenatal WIC participation (compared with nonparticipation) for an infant chosen randomly from the underlying population. In what follows, we will simplify notation by suppressing the conditioning on subpopulations of interest captured in X. For this analysis, we focus on infants who are income-eligible for WIC.

Two identification problems arise when assessing the impact of WIC on infant health. First, even if WIC participation were observed for all eligible households, the potential outcome \( H(1) \) is counterfactual for all infants who did not receive WIC, while \( H(0) \) is counterfactual for all infants who did receive WIC. This is referred to as the selection problem. Using the Law of Total Probability, this identification problem can be highlighted by writing the first term of Equation (1) as

\[
P[H(1) = 1] = P[H(1) = 1 | W^* = 1]P(W^* = 1) + P[H(1) = 1 | W^* = 0]P(W^* = 0)
\]

where \( W^* = 1 \) denotes that an infant resides in a household that truly receives WIC and \( W^* = 0 \) otherwise. If WIC receipt is observed, the sampling process identifies the selection probability \( P(W^* = 1) \), the censoring probability \( P(W^* = 0) \), and the expectation of outcomes when the outcome is observed, \( P[H(1) = 1 | W^* = 1] \). Still, the sampling process cannot reveal the mean outcome conditional on censoring, \( P[H(1) = 1 | W^* = 0] \). Given this censoring, \( P[H(1) = 1] \) is not point-identified by the sampling process alone. Analogously, the second term in Equation (1), \( P[H(0) = 1] \), is not identified.

Second, true participation status may not be observed for all respondents. This is referred to as the measurement or classification error problem. Instead of observing \( W^* \), we observe a self-reported indicator, \( W \), where \( W = 1 \) if an infant resides in a household that reported receiving prenatal WIC and 0 otherwise. Without assumptions restricting the nature or degree of classification errors, the sampling process does not reveal useful information on WIC receipt, \( W^* \), and thus all of the probabilities on the right hand side of Equation (2) are unknown.
To highlight this measurement problem, let the latent variable $Z^*$ indicate whether a report is accurate, where $Z^* = 1$ if $W^* = W$ and $Z^* = 0$ otherwise. Using this variable, we can further decompose the first term of Equation (1) as

$$P[H(1) = 1] = [P(H = 1, W = 1) - \theta_1^+ + \theta_1^-] + P[H(1) = 1 | W^* = 0][P(W = 0) + (\theta_1^+ + \theta_0^-) - (\theta_1^- + \theta_0^+)]$$  \hspace{1cm} (2')$$

where $\theta_j^+ = P(H = j, W = 1, Z^* = 0)$ and $\theta_j^- = P(H = j, W = 0, Z^* = 0)$ denote the fraction of false positive and false negative classifications of WIC recipients, respectively, for infants realizing health outcome $j = 1, 0$. Given this decomposition, it is easy to see that the first term in Equation (2), $P[H(1) = 1 | W^* = 1]P(W^* = 1)$, is not identified because of the classification error problem while the second term is not identified because of both the selection and classification error problems. The data cannot reveal the counterfactual outcome distribution, $P[H(1) = 1 | W^* = 0]$, 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, $P(W^* = 1)$.

To address these two identification problems, we extend the approach developed in KPGJ. We proceed in three steps. First, we focus on the implications of the measurement error problem alone when we have information on the true participation rate, $P^*$, and the self-reported rate, $P$. Second, assuming reports of WIC participation are accurate, we focus on the selection problem. Finally, we assess what can be inferred when accounting for both identification problems simultaneously.
4. EXOGENOUS SELECTION BOUNDS

Much of the literature examining the impact of WIC on health assumes that selection is exogenous so that $P[H(j) \mid W^*] = P[H(j)]$, $j = 1, 0$. Under this assumption, the average treatment effect can be written as

$$ATE(1, 0 \mid X \in \Omega) = P[H(1) = 1 \mid W^* = 1] - P[H(0) = 1 \mid W^* = 0],$$

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

$$\beta \equiv P(H = 1 \mid W^* = 1) - P(H = 1 \mid W^* = 0). \quad (3)$$

The appeal of the exogenous selection assumption is obvious: if selection is exogenous and WIC receipt $W^*$ is observed, then the average treatment effect is identified by the sampling process.

Though the exogenous selection assumption is unlikely to hold in our setting, the difference in expected outcomes among WIC recipients and nonrecipients – the mean outcome gap, $\beta$ – is an important descriptive measure of the association between WIC participation and birth outcomes. The sample means displayed in Table 1, for example, suggest that WIC is associated with a slightly higher probability of a normal birth weight but lower probability of a normal gestation age.

If one allows for the possibility of classification errors in $W^*$, however, this outcome gap is not identified. To make progress in partially identifying this mean outcome gap, we decompose the first term in Equation (3) into identified and unidentified quantities:

$$P(H = 1 \mid W^* = 1) = \frac{P(H = 1, W^* = 1)}{P(W^* = 1)} = \frac{P(H = 1, W = 1) + \theta^- - \theta^+}{P(W^* = 1)} \quad (4)$$

where $P(H = 1, W = 1)$ is identified by the data. In the numerator, $\theta^- - \theta^+$ reflects the unobserved excess of false negative versus false positive classifications among those with a favorable birth outcome. The quantity $P(H = 1 \mid W^* = 0)$ can be decomposed analogously. Clearly, without assumptions restricting the nature or degree of classification errors, the data are uninformative.

To address the classification error problem, we combine auxiliary data on the size of the caseload from the USDA (2014) with survey data from the ECLS-B to estimate the true and self-
reported participation rates, $P^* = P(W^* = 1)$ and $P = P(W = 1)$. As noted above, these data reveal that $P^* = 0.74$ and $P = 0.687$. Thus, following KPGJ, the auxiliary information identifies the difference in false negative and positive reporting rates:

$$
\Delta = P^* - P = (\theta_1^- + \theta_0^-) - (\theta_1^+ - \theta_0^+).
$$

In our application, $\Delta$ is estimated to equal $0.053 (= 0.74 - 0.687)$; the fraction of false negative reports must exceed the fraction of false positive reports by this quantity.

We consider two additional restrictions on the classification error problem:

**ME1:** Maximum error rate: $P(Z^* = 0) \leq Q_u$  

**ME2:** Verification: $V = 1 \Rightarrow W^* = W$ (i.e., $Z^* = 1$)

where $Q_u$ is a known upper bound on the degree of data corruption in the spirit of Horowitz and Manski (1995) and $V$ is an indicator of whether a self-report of WIC participation is treated as accurate.

For Assumption ME1, the value of $Q_u$ must logically lie within the range $[|\Delta|, 1]$. In the polar case that $Q_u$ equals 1, ME1 is uninformative. We refer to the case that $Q_u = 1$ (and no responses are verified to be accurate) as the “arbitrary errors model.” In the other polar case that $Q_u$ equals $|\Delta|$, the researcher is imposing a “no excess errors” assumption. In the case of net underreporting ($\Delta > 0$), this assumption is equivalent to imposing a “no false positives” assumption that respondents do not falsely claim to participate in WIC. This no excess errors restriction serves as a useful benchmark for the receipt of WIC in our application. Validation data suggest very few instances of households falsely claiming to receive food assistance (e.g., Bollinger and David 1997), and the estimates obtained using Hausman et al.’s (1998) parametric model suggest that less than one percent of the

---

12 See Hotz, Mullins, and Sanders (1997) for an illustration of how auxiliary data can be used to address similar measurement problems.
eligible WIC population reports benefits not actually received: \( \theta^* < 0.01 \). Middle-ground positions are obtained by setting \( Q_u \) between \( |P^* - P| \) and 1.

Assumption ME2 allows for the possibility that some respondents are known to provide accurate reports, denoted \( V = 1 \). For \( V = 0 \) respondents, a report may be either accurate or inaccurate. In the no false positives model, for example, respondents reporting the receipt of WIC are validated to provide accurate reports. We also consider a model that verifies a WIC response for anyone willing to report the receipt of benefits from any social welfare program asked about in the ECLS-B (WIC, SNAP, TANF, or Medicaid). Under this verification assumption, 46 percent of the WIC responses are treated as accurate.

These assumptions also imply informative upper bounds on the false reporting probabilities \( \theta^*, \theta^+, \theta^0 \) and \( \theta^* \). Specifically, it follows that

\[
\theta^* \leq \min \left\{ P(H = i, W = 0, V = 0), P^*, \frac{1}{2}(Q_u + \Delta) \right\} \equiv \theta^*_{UB}, \quad i = 1, 0 \tag{7a}
\]
\[
\theta^+ \leq \min \left\{ P(H = i, W = 1, V = 0), 1 - P^*, \frac{1}{2}(Q_u - \Delta) \right\} \equiv \theta^+_{UB}, \quad i = 1, 0 \tag{7b}
\]

For example, the fraction of false negative responses among households with a particular health outcome cannot exceed the fraction of all negative responses for that outcome. Nor can it exceed the fraction of households participating in WIC. The third argument goes to 0 as \( Q_u \) and \( \Delta \) go to 0.

With verification, \( P^* \) is logically required to lie within the range

\[
\Lambda = [\max \{ P - P(W = 1, V = 0), P(H = 1, W = 1) - \theta^*_{UB} \}, \min \{ P + P(W = 0, V = 0), 1 - P(H = 1, W = 0) + \theta^+_{UB} \}].
\]

---

\(^{13}\) Using Hausman et al.’s (1998) model, the true participation rate is specified as \( P^* = F(X' \gamma) \), where \( F(\cdot) \) is the standard normal CDF and \( X \) is the vector of covariates reported in Table 1. Given data on the self-reported rate, \( P = P(W = 1) \), the model identifies the conditional false negative reporting rate, \( P(W = 0 \mid W^* = 1) \), and the conditional false positive reporting rate, \( P(W = 1 \mid W^* = 0) \). We estimate these rates to be 0.19 and 0.05, respectively. Based on these estimates, we are able to back out an estimate of \( P^* \) and the unconditional misreporting rates using \( P^* = [P - P(W = 1 \mid W^* = 0)}/[1 - P(W = 1 \mid W^* = 0) - P(W = 0 \mid W^* = 1)] \). Our estimate of \( \theta^* = (\theta^* + \theta^*_0) \) is 0.159, and our estimate of \( \theta^+ = (\theta^+ + \theta^+_0) \) is 0.008. Full estimation results are available from the authors.

\(^{14}\) We do not presume that any particular response is inaccurate.
Intuitively, $P^*$ cannot be too far away from $P$ when some households are known to provide valid responses. Under no false positives (and no further verification), this range reduces to $P^* \in [P, 1]$.

Given these restrictions, we derive the following bounds on the outcome gap or, equivalently, the ATE under the exogenous selection assumption:

**Proposition 1** Given the classification error model restrictions in Equations (7a) and (7b), the outcome gap, $\beta$, is bounded as follows:

$$\frac{P(H = 1, W = 1) - P^* P(H = 1) - \min \{ \theta_1^{UB}, \theta_0^{UB} - \Delta \}}{P^*(1 - P^*)} \leq \beta \leq \frac{P(H = 1, W = 1) - P^* P(H = 1) + \min \{ \theta_1^{UB}, \theta_0^{UB} + \Delta \}}{P^*(1 - P^*)}$$

for $P^* \in \Lambda \cap (0, 1)$. If the valid range $\Lambda$ includes 1, the bounds converge to $[-P(H = 0), P(H = 1)]$ as $P^*$ approaches 1. Similarly, if the valid range includes 0, the bounds converge to $[-P(H = 1), P(H = 0)]$ as $P^*$ approaches 0.

These bounds simplify in an arbitrary errors model:

**Corollary 1** Under arbitrary errors, the Proposition 1 bounds reduce to:

$$-\min \left\{ \frac{P(H = 1)}{1 - P^*}, \frac{P(H = 0)}{P^*} \right\} \leq \beta \leq \min \left\{ \frac{P(H = 0)}{1 - P^*}, \frac{P(H = 1)}{P^*} \right\}.$$  

\(^{15}\) The arguments involving $\theta_1^{UB}$ and $\theta_0^{UB}$ ensure that the conditional probabilities $P(H = 1 | W^* = 1)$ and $P(H = 1 | W^* = 0)$ do not exceed 1.
See the appendix for a proof of these results.

All of the probabilities in Proposition 1 and the corollary are consistently estimated using data from the ECSL-B except $P^*$, the true WIC participation rate. We estimate $P^* = 0.74$ using administrative data on the size of the WIC caseload as described above, but we also assess the sensitivity of the bounds to variation in $P^*$.

### 4.1. Sharp ATE Bounds When $P^*$ is Not Known

If Assumptions ME1 and ME2 hold under arbitrary errors but the researcher does not know the value of $P^*$, sharp bounds on $\beta$ are given as follows:

$$
\beta_{UB} = \inf_{\theta_1^*, \theta_0^*} \left\{ \frac{P(H = 1, W = 1) - \theta_1^*}{P(W = 1) - \theta_1^* + \theta_0^*} - \frac{P(H = 1, W = 0) + \theta_1^*}{P(W = 0) + \theta_1^* - \theta_0^*} \right\}
$$

$$
\beta_{LB} = \sup_{\theta_1^*, \theta_0^*} \left\{ \frac{P(H = 1, W = 1) + \theta_1^*}{P(W = 1) + \theta_1^* - \theta_0^*} - \frac{P(H = 1, W = 0) - \theta_1^*}{P(W = 0) - \theta_1^* + \theta_0^*} \right\}
$$

These bounds are obtained by minimizing or maximizing $P(H = 1 | W^* = 1) - P(H = 1 | W^* = 0)$ using Equation (4) and its analogous counterpart for $P(H = 1 | W^* = 0)$. They can be estimated by conducting separate grid searches over $\{\theta_1^*, \theta_0^*\}$ and $\{\theta_1^*, \theta_0^*\}$ in the feasible regions to minimize $\beta_{LB}$ and maximize $\beta_{UB}$ subject to the constraint that none of the conditional probabilities (given by the ratios) exceeds 1. Intuitively, these bounds when $P^*$ is unknown are wider than the Proposition 1 bounds when $P^*$ is known. Under the no false positives assumption, these bounds simplify to:

$$
ATE_{LB} = \inf_{\theta_0^*} \left\{ \frac{P(H = 1, W = 1)}{P(W = 1) + \theta_0^*} - \frac{P(H = 1, W = 0)}{P(W = 0) - \theta_0^*} \right\}
$$

$$
ATE_{UB} = \sup_{\theta_1^*} \left\{ \frac{P(H = 1, W = 1) + \theta_1^*}{P(W = 1) + \theta_1^*} - \frac{P(H = 1, W = 0) - \theta_1^*}{P(W = 0) - \theta_1^*} \right\}.
$$
4.2. Results Under Exogenous Selection

Figures 1A and 1B trace out the Proposition 1 bounds on the outcome gap as the true participation rate $P^*$ varies between 0 and 1 for the normal birth weight outcome (Figure 1A) and normal gestation length outcome (Figure 1B) under (a) arbitrary classification errors (no verification), (b) no false positive reports of WIC receipt, and (c) verified WIC status responses for households that reported benefits from any government program. The bounds traced out in these figures account only for identification uncertainty and abstract away from the additional layer of uncertainty associated with sampling variability. The associated tables under the figures display the bounds at the self-reported participation rate, $P^* = P = 0.687$, and the true participation rate, $P^* = P^* = 0.74$, along with Imbens-Manski (2004) confidence intervals that cover the true value of the ATE with 90% probability.\(^\text{16}\)

If all WIC responses are known to be accurate such that $P^* = P$, then the mean outcome gap is point-identified as $\beta = 0.8565 - 0.8375 = 0.019$ for normal birth weight and as $\beta = 0.7416 - 0.7632 = -0.022$ for normal gestation duration (consistent with the descriptive statistics in Table 1). Otherwise, the difference in mean outcomes between recipients and nonrecipients can only be partially identified. Consider, for example, the case where $P^* = 0.74$ and there are no false positive reports of participation (dashed curves). In this case, $\beta$ can lie anywhere in the range $[-0.202, 0.062]$ for the normal birth weight gap and $[-0.232, 0.046]$ for the normal gestation duration gap. Under the stronger verification assumption (dotted curves), these ranges narrow to $[-0.077, 0.062]$, a 47% reduction in the width of the bounds, and $[-0.120, 0.046]$, a 41% reduction

\(^\text{16}\) Complete sets of estimation results for the other outcomes (not low birth weight, not very low birth weight, not macrosomic weight, not premature, not very premature) are available from the authors upon request.
in the width, respectively. In none of these cases, however, can the sign of \( \beta \) be identified at our preferred value \( P^* = 0.74 \).\(^{17}\)

Under arbitrary errors, there is a great deal of uncertainty about the true value of \( \beta \) even if it is known that \( P = P^* = 0.687 \) such that there is no net misreporting. While the fraction of false negative reports must exactly offset the fraction of false positive reports, there still could be a substantial degree of misclassification. Even ignoring sampling variability, \( \beta \) could lie anywhere between -0.217 and 0.478 (a 69 percentage point gap) for the normal birth weight outcome and anywhere between -0.367 and 0.803 (a 117 point gap) for the normal gestation period outcome.

One key result in Figure 1 is that identification of the expected birth outcome gaps deteriorates sufficiently rapidly that small degrees of classification error preclude us from identifying the sign of either the normal birth weight or normal gestation duration gap. The conclusion that normal birth outcomes are more (or less) prevalent among WIC recipients than among eligible nonrecipients requires a large degree of confidence in self-reported WIC participation status, an assumption not supported by validation studies.

Another interesting result from Figure 1 is that the introduction of classification errors in the treatment variable improves the estimated lower bound on \( \beta \) when we know the direction of misreporting. Specifically, notice that the estimated lower bounds under arbitrary errors are continuously rising with the degree of underreporting \((P^* - P)\) for all conjectured values of \( P^* \) within \([P, 1]\). The lower bounds under the verification assumptions are also rising with the degree of underreporting after some initial deterioration around \( P \). In Section 6, these patterns will carry through to the case of endogenous program participation.

\(^{17}\) Under the stronger verification assumption for the normal birth weight outcome, \( \beta \) is point identified as 0.019 at \( P^* = P \) and as 0.002 at \( P^* = P + P(H = 1, V = 0) \), the two endpoints for the valid range of \( P^* \) (see the vertical dashed lines in Figure 1A). The corresponding point estimates for normal gestation are -0.022 and 0.005.
5. ENDOGENOUS SELECTION BOUNDS

5.1. Models

In this section, we focus on what can be learned about the ATE when selection is endogenous for the special case of perfectly accurate reporting \( (P^* = P \text{ with full verification}) \). A natural starting point is to ask what can be learned in the absence of any assumptions invoked to address the selection problem (see Manski, 1995 and Pepper, 2000). Since the counterfactual latent probability \( P[H(1) = 1 | W^* = 0] \) must lie within \([0, 1]\), it follows from Equation (2) that

\[
P[H(1) = 1 | W^* = 1] P(W^* = 1) \leq P[H(1) = 1] \leq P[H(1) = 1 | W^* = 1] P(W^* = 1) + P(W^* = 0).
\]

Intuitively, the width of this bound on \( P[H(1) = 1] \) equals the censoring probability, \( P(W^* = 0) \). Thus, if a large fraction of mothers receive WIC, the bounds on \( P[H(1) = 1] \) are relatively narrow.

Taking the difference between the upper bound on \( P[H(1) = 1] \) and the lower bound on \( P[H(0) = 1] \) obtains a sharp upper bound on \( ATE \), and analogously a sharp lower bound (Manski, 1995). As a result, the width of the bounds on the average treatment effect always equals 1 and, in the absence of identifying restrictions, the data cannot reveal the sign of the effect of WIC on birth outcomes.

To narrow the bounds, prior information must be brought to bear. While the exogenous selection assumption seems untenable, a number of middle ground assumptions restrict the relationship between WIC participation, birth outcomes, and observed covariates. We consider the identifying power of several monotonicity assumptions: one on treatment selection, two using instruments, and one on treatment response.

5.2. Monotone Instrumental Variables

The Monotone Instrumental Variable (MIV) assumption (Manski and Pepper, 2000) formalizes the notion that the latent probability of a negative health outcome, \( P[H(j) = 1] \), varies monotonically with certain observed covariates. Arguably, for example, this probability decreases
with household income.\textsuperscript{18} To formalize this idea, let $v$ be the monotone instrumental variable such that

$$u_1 \leq u \leq u_2 \Rightarrow P[H(j) = 1 \mid v = u_1] \leq P[H(j) = 1 \mid v = u] \leq P[H(j) = 1 \mid v = u_2].$$

That is, the latent positive birth outcome probabilities weakly increase with income.

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 MIV assumption formalized in Manski and Pepper (2000, Proposition 1) implies:

$$\sup_{u \in [u_1, u_2]} LB(u) \leq P[H(j) = 1 \mid v = u] \leq \inf_{u \in [u_1, u_2]} UB(u).$$

These bounds on $P[H(j) = 1 \mid v = u]$ are sharp. Bounds on the unconditional latent probability, $P[H(j) = 1]$, can then be obtained using the law of total probability.\textsuperscript{19}

\subsection*{5.3. Monotone Treatment Selection}

A special case of the MIV assumption arises when the realized treatment is a monotone instrument. This Monotone Treatment Selection (MTS) assumption (Manski and Pepper, 2000) places structure on the selection mechanism through which mothers become WIC recipients. The literature suggests that unobserved factors associated with poor health are likely to be positively associated with the decision to take up the program, and WIC recipients are known to have unfavorable demographic, socioeconomic, and health characteristics (e.g., Bitler and Currie, 2005;

---

\textsuperscript{18} Chen et al. (2002) report that child health improves monotonically with socioeconomic status, and Deaton (2002) provides evidence of a negative income gradient in realized health outcomes.

\textsuperscript{19} To estimate these MIV bounds, we first divide the sample into nine income categories provided in the ECLS-B. We assume that the ratio of actual to potential net underreporting does not vary across MIV groups. To find the MIV bounds on the rates of poor health, one takes the appropriate weighted average of the plug-in estimators of lower and upper bounds across the groups. As discussed in Manski and Pepper (2000), 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.
Gundersen, 2005). In this case, recipients have worse latent health outcomes than nonrecipients on average.

We formalize the MTS assumption as follows:

\[
P[H(j) = 1 | W^* = 1] \leq P[H(j) = 1 | W^* = 0] \quad \text{for } j = 1, 0.
\]

That is, on average, eligible households that chose to receive WIC, \( W^* = 1 \), have worse latent birth outcomes than eligible households that did not take up WIC, \( W^* = 0 \).

5.4. Ineligibles MIV

We also use eligibility criteria to construct a monotone instrument by extending an approach in Gundersen et al. (2012). In particular, we focus on the subgroup of households with socioeconomic status (SES) in the third or lower quintile yet ineligible for WIC because (a) their income exceeded 185 percent of the Poverty Income Guidelines (denoted \( \text{inc} > 185\% \)) and (b) they were not adjunctively eligible for WIC through the receipt of TANF, SNAP, or Medicaid benefits. About 56 percent of our sample in this income range reported no participation in these programs.

While latent birth outcomes are unlikely to be mean independent of eligibility status, an MIV assumption holding that mean response varies monotonically across these groups seems credible. In particular, suppose the latent probability of a healthy birth outcome among WIC-eligible households is no better than among higher-income ineligible households:

\[
P[H(j) = 1] \leq P[H(j) = 1 | v = \text{ineligible}].
\]

This ineligibility bound on \( P[H(j) = 1] \) has the same structure as the income-MIV restriction in Equation (9), but it has the unique property that \( W^* = 0 \) among ineligible households. If eligibility is observed, the data point-identify \( P[H(0) = 1 | v = \text{ineligible}] \) as \( P(H = 1 | v = \text{ineligible}) \). This
quantity then serves as an upper bound on the potential outcome $P[H(0) = 1]$ for our primary population of income-eligible households. We learn nothing new about $P[H(1) = 1]$.

We extend the MIV ineligibles model in Gundersen et al. (2012) to account for the problem that adjunctive eligibility into WIC through participation in other programs may be measured with error. Let $B^\ast = 0$ indicate a true lack of benefits from these programs and $B = 0$ the self-reported counterpart, where the ECLS-B data are used classify a household reporting no prenatal WIC benefits as ineligible for WIC if $inc > 185\%$ and the household reports no benefits from Medicaid, SNAP, or TANF since the child was born.

It follows that $P[H(0) = 1]$ cannot exceed $P(H = 1 | inc > 185\%, B^\ast = 0)$. This conditional probability can be written as

$$
\frac{P(H = 1, B^\ast = 0 | inc > 185\%)}{P(B^\ast = 0 | inc > 185\%)} = \frac{P(H = 1, B = 0 | inc > 185\%)}{P(B = 0 | inc > 185\%)} + \tilde{\theta}_1 - \tilde{\theta}_0
$$

where $\tilde{\theta}_j = P(H = j, B = 1, B^\ast = 0 | inc > 185\%)$ and $\tilde{\theta}_j = P(H = j, B = 0, B^\ast = 1 | inc > 185\%)$ denote false positive and false negative classifications for $j = 1, 0$ among this relatively higher income group.

Ruling out false positive responses to the participation questions, we obtain a feasible upper bound given by

$$
P[H(0) = 1] \leq \frac{P(H = 1, B = 0 | inc > 185\%)}{P(B = 0 | inc > 185\%)} - \tilde{\theta}_0^{UB}
$$

where $\tilde{\theta}_0^{UB} = \min \{q, P(H = 0, B = 0)\}$ and $q \in [0, 1]$ is any imposed maximum allowed error rate in classifying $B^\ast$. More generally, we can apply Kreider and Pepper’s (2007) Proposition 1 for misclassified conditional probabilities to obtain the following arbitrary errors upper bound:

$$
P[H(0) = 1] \leq \frac{P(H = 1, B = 0 | inc > 185\%) + \gamma}{P(B = 0 | inc > 185\%) + 2\gamma - q}
$$

(12)
\[
\min \{q, P(H = 1, B = 1 | inc > 1.85)\}
\]

where \( \gamma = \begin{cases} P(H = 1, B = 0 | inc > 185\%) - P(H = 0, B = 0 | inc > 185\%) + q \leq 0 \\ \max \{0, q - P(H = 0, B = 0 | inc > 185\%)\} \end{cases} \) o.w.

With no misclassification of adjunctive eligibility \( (q = 0) \), the upper bound on \( P[H(0) = 1] \) equals the observed probability of a positive birth outcome among ineligible households, \( P(H = 1 | inc > 185\%, B = 0) \), which equals 0.820 for the normal birth weight outcome and 0.795 for the normal gestation duration outcome. We assess the sensitivity of the bounds in Equations (11) and (12) to variation in \( q \).

### 5.5. Monotone Treatment Response

There is a general consensus among policymakers and researchers that prenatal WIC participation should not lead to worse birth outcomes (Currie, 2003). Given this consensus, we consider the identifying power of the **Monotone Treatment Response** (MTR) assumption (Manski, 1995 and 1997; Pepper 2000) that formalizes the idea that WIC participation would not lead to a reduction in health status:

\[
H(1) \geq H(0). 
\]

While this assumption weakly signs the ATE as nonnegative, it seems relatively innocuous in the context of WIC that restricts purchased food products to be nutritionally sound. It is difficult to imagine how receiving WIC would lead to worse birth outcomes. The assumption does not provide any information on the magnitude of the ATE and does not rule out a value of 0.

### 5.6. Results for the No Errors Case

For each of the outcomes, Table 2 displays the estimated bounds and confidence intervals under a variety of different models for the no errors case. In row (i), we make no assumptions about how eligible households select into the WIC program. The width of these worst case selection bounds always equals 1, and the bounds always include zero. These bounds highlight a researcher’s
inability to make strong inferences about the efficacy of WIC without making assumptions that address the problem of unknown counterfactuals. In the absence of restrictions that address the selection problem, we cannot rule out the possibility that WIC has a large positive or negative impact on the likelihood on positive birth outcomes.

The bounds, however, are narrowed substantially under common monotonicity assumptions on treatment selection (MTS) and relationships between the latent outcome and observed instrumental variables (MIV). In row (ii), the MTS assumption alone identifies the sign of the impact of WIC on the probability of normal birth weight. In particular, we estimate that WIC increases the probability of normal birth weight by at least 1.9 percentage points and by as much as 63.9 points. Similar results are found for the ‘not low’ and ‘not very low’ birth weight outcomes. For the other outcomes, the MTS assumption notably increases the estimated lower bound but it does not identify the sign of the effect of WIC. For example, the estimated lower bound on the probability of avoiding a macrosomic birth weight increases from $-0.333$ to $-0.018$. The bounds narrow further when we combine the MTS assumption with the two MIV assumptions that the probability of a favorable birth outcome weakly increases with family resources (row iii) and is no worse among eligibles households (row iv). In fact, the MTS-MIV model (row iii) identifies the ATE as positive for six of the seven outcomes (the one exception being the probability of a delivery being premature), though the confidence interval includes zero in many of the cases. For example, the estimates imply that WIC increases the probability of a normal birth weight by at least 6.5 points, and the confidence interval does not include zero. In this no measurement error setting, layering on the MIV eligibles assumption (row iv) has no additional identifying power except for the case of avoiding a macrosomic birth. As discussed in Section 6, however, the eligibles MIV can have substantial identifying power when the true participation rate $P^*$ exceeds the reported rate $P$.

The last row in Table 2 adds the MTR assumption. In most cases, the MTR assumption does not notably improve the estimated lower bounds on the ATE relative to the estimates derived under
the MTS-MIV assumptions alone. Like the ineligibles MIV assumption, the MTR assumption does turn out to have substantial identifying power when WIC participation is underreported (Section 6).

Overall, the no-errors results indicate that prenatal WIC may lead to substantial increases in the favorable birth outcomes and, at worst, have slightly deleterious effects. Consistent with the previous literature, the findings provide stronger evidence that prenatal WIC participation improves birth weight outcomes in general than that it specifically prevents premature births. In the next section, we study what can be learned about the effects of the program under these monotonicity assumptions on the selection problem when WIC participation may be measured with error.

6. INFECTION WITH MEASUREMENT AND SELECTION PROBLEMS

The selection bounds estimated in Section 5 presume that everyone reports WIC participation accurately. With reporting errors, however, there is uncertainty not only about counterfactuals but also about the reliability of the data on WIC participation, \( W^* \). In this section, we merge the classification error models from Section 4 with the selection models in Section 5.

6.1. Models

KPGJ show that the worst-case selection bounds on the ATE with classification errors can be written as

\[
\begin{align*}
[-P(H = 1,W = 0) - P(H = 0,W = 1)] + \Theta \\
\leq ATE(1,0) \leq \\
[P(H = 1,W = 1) + P(H = 0,W = 0)] + \Theta.
\end{align*}
\]  

(14)

where \( \Theta = (\theta_1^- + \theta_0^-) - (\theta_0^+ + \theta_1^+) \). In the absence of classification errors, \( \Theta = 0 \), and Equation (14) simplifies to Manski’s (1995) worst-case selection bounds. With classification errors, the models from Section 4 – see Equations (6) and (7) – allow us to place informative restrictions on \( \Theta \).

Extending Proposition 1 in KPGJ to incorporate verification, we have:
Proposition 2  Given restrictions (6), (7), and (14) under endogenous selection, the ATE is bounded as follows:

\[
-2 \theta_i^{UB} - \Delta, -2 \theta_0^{UB} + \Delta \]

\[
\leq ATE(1,0) \leq
\]

\[
[2 \theta_i^{UB} - \Delta, 2 \theta_0^{UB} + \Delta].
\]

A proof of this proposition is provided in the appendix.

If Assumptions ME1 and ME2 hold under arbitrary errors but the researcher does not know the value of \( P^* \), sharp bounds on the ATE are given as follows:\(^{20}\)

\[
ATE^{LB} = P(H = 1, W = 1) - P(H = 1, W = 0) - P(W = 1) - \min \{Q_a, \theta_i^{UB} + \theta_0^{UB} \}
\]

\[
ATE^{UB} = P(H = 1, W = 1) - P(H = 1, W = 0) + P(W = 0) + \min \{Q_a, \theta_0^{UB} + \theta_i^{UB} \}.
\]

Under no false positives, these bounds narrow to:

\[
ATE^{LB} = P(H = 1, W = 1) - P(H = 1, W = 0) - P(W = 1) - \theta_0^{UB}
\]

\[
ATE^{UB} = P(H = 1, W = 1) - P(H = 1, W = 0) + P(W = 0) + \theta_i^{UB}.
\]

Adding the MTR and MIV assumptions to these models is straightforward. To add MTS, recall that the upper bound on the ATE in this case is the difference in the conditional means:

\[
ATE(1,0) \leq P(H = 1|W^* = 1) - P(H = 1|W^* = 0).
\]

---

\(^{20}\) The ATE is given by \( P[H(1) = 1] - P[H(0) = 1] = P[H(1) = 1|W^* = 1]P(W^* = 1) + P[H(1) = 1|W^* = 0]P(W^* = 0) - P[H(0) = 1|W^* = 1]P(W^* = 1) - P[H(0) = 1|W^* = 0]P(W^* = 0) \). Letting \( P[H(1) = 1|W^* = 0] \) and \( P[H(0) = 1|W^* = 1] \) vary within \([0,1]\), the ATE is no smaller than \( P[H = 1,W^* = 1] - P(W^* = 1) - P(H = 1,W^* = 0) = [P(H = 1,W = 1) + \theta_i^* - \theta_i^*] - [P(W = 1) + \theta_0^* + \theta_0^* - \theta_i^* - \theta_i^*] - [P(H = 1,W = 0) + \theta_i^* - \theta_i^*] \). The lower bound is obtained by setting \( \theta_0^* = \theta_i^* = 0 \) and recognizing that the sum of errors cannot exceed \( Q_a \). The upper bound is derived analogously.
This upper bound is the health outcome gap $\beta$ evaluated in Section 4, the unobserved difference in mean birth outcomes among recipients and nonrecipients. Proposition 1 provides bounds on this difference.

6.2. Results for Measurement Error and Endogenous Selection

Figures 2A and 2B trace out the Proposition 2 worst case ATE selection bounds across conjectured values of $P^*$ for normal birth weight and normal gestation length with no monotonicity restrictions. Except when $P^*$ is close to 0 or 1, these endogenous selection bounds are much wider than the exogenous selection bounds in Figures 1A and 1B.21 Under no measurement error ($P^* = P = 0.687$ with verification), recall that the exogenous selection bounds are point-identified as 0.019 for normal birth weight and -0.022 for normal gestation length. Under endogenous selection, Manski’s (1995) classic worst-case bounds apply. As seen in Figures 2A and 2B, the ATE bounds expand to $[-0.361, 0.639]$ for birth weight and to $[-0.417, 0.583]$ for gestation. By construction, Manski’s (1995) bounds always include 0 with width of 1. Once we allow for potential measurement error at $P^* = P$ (no net misreporting), these bounds expand further to $[-0.462, 0.837]$ with width 1.30 and $[-0.565, 0.938]$ with width 1.50.

Identification decays yet further as $P^*$ rises to the estimated administrative rate $P^o = 0.74$. Even under the stronger verification, the ATE bounds have a width of 1.054 for birth weight and 1.06 for gestation. Compared with exogenous selection, the verification assumptions have little power in this setting. Without further restrictions on the selection process, we cannot come close to identifying the signs of the average treatment effects.

Much more can be learned about the average treatment effects in Figures 3A (normal birth weight) and Figure 3B (normal gestation age) once we consider the monotonicity restrictions. We

---

21 The bounds converge to $[-P(H = 1), P(H = 0)]$ and $[-P(H = 0), P(H = 1)]$ as $P^*$ approaches 0 and 1, respectively.
rescale these figures to focus on lower bounds over the range $P^* \in [0.60, 0.80]$. Both lower and upper bounds, along with confidence intervals, are reported in the tables beneath the figures. Starting towards the bottom and moving up, the figures trace out lower bounds layering successively stronger sets of assumptions: (i) MTS + income-MIV, (ii) MTS + income-MIV + ineligibles-MIV, and (iii) MTS + income-MIV + ineligibles-MIV + MTR. As before, the solid curves trace out the bounds under arbitrary errors. The dashed and dotted curves trace out the bounds under the no false positives and stronger verification assumptions, respectively. To save space, the tables present estimates only for cases (ii) and (iii). Income-MIV and ineligibles-MIV are abbreviated as “inc-MIV” and “ineligibles.”

We begin with Figure 3A for the ATE on normal birth weight. Starting at $P^* = P$, recall there is no net misreporting of WIC participation under arbitrary errors, and no misreporting at all under either verification assumption. Under MTS + inc-MIV (case i) with arbitrary errors (no net misreporting at $P^* = P$), the lower bound on the ATE is $-0.197$, a considerable improvement compared with the lower bound of $-0.462$ in Figure 2A. More dramatically, the MTS + inc-MIV assumptions combined with no misreporting signs the ATE as strictly positive at $0.065$ and the confidence interval lies strictly above 0. This implies that WIC reduces the prevalence of unhealthy birth weight by at least 33%.22

Lower bounds for the ATE on normal gestation length are traced out in Figure 3B. Under MTS + inc-MIV at $P^* = P$ with no net measurement error, the ATE exceeds $-0.343$ (arbitrary errors, not shown), a 22 point improvement in the lower bound of $-0.565$ when no monotonicity assumptions are imposed (Figure 2B). The ATE is identified to exceed $0.0059$ when classification errors are ruled out completely, corresponding to at least a 2.3 percent reduction in the prevalence of unfavorable gestation lengths. The confidence interval in this case, however, includes 0.

---

22 The percentage reduction in unfavorable outcomes under the program is $-\{P[H(1) = 1] - P[H(0) = 1]\} / P[H(0) = 1]$. 

For both outcomes, identification of a positive ATE under no measurement error deteriorates rapidly once classification errors are allowed. We have already seen that the ATEs can be substantially negative at $P^* = P$ if the no errors assumption is replaced with the no net errors assumption. We also see in Figures 3A and 3B that inferences are sensitive to small departures of $P^*$ from $P$ under the verification assumptions. For example, we found that prenatal WIC increases the probability of normal birth weight by at least 6.5 percent points under accurately reported program participation (Figure 3A). Yet if even 2 percent of eligible women may fail to acknowledge the receipt of WIC benefits, we can no longer identify a positive ATE. At our preferred estimate of $P^* = 0.74$, the lower bound falls to $-0.148$ under the no false positives assumption and to $-0.040$ under stronger verification.

Notice, however, that the estimated lower bound on the ATE begins to rise for sufficiently large values of $P^*$. Intuitively, the bounds can change direction as particular constraints in Equation (7a) and (7b) become binding. Thus, perhaps counterintuitively, a researcher’s ability to make inferences about the efficacy of a program does not necessarily deteriorate with the degree of misreporting.

Layering on the ineligibles-MIV assumption with accurately measured ineligibility ($q = 1$), we can identify the ATE for normal birth weight as strictly positive under the stronger verification assumption, with a confidence interval lying above 0, even at the estimated true participation rate $P^* = 0.74$. In this case, we estimate that WIC improves the healthy birth weight outcome by at least 2.54 percentage points, corresponding to at least a 14 percent decline in the probability of an unhealthy birth weight. Results are sensitive, however, to the value of $q$ in Equation (11) or (12). If

---

23 As $P^*$ rises for the MTS lower bound, for example, the $\Delta$ constraint in Equation (5) remains satisfied by continuously raising the value of $\theta_v = P(H = 0, W = 0, Z = 0)$, the unobserved fraction of households that experienced an unfavorable health outcome and failed to report WIC benefits actually received. At about $P^* = 0.72$, however, $\theta_v$ becomes maxed out in Equation (7a) under the stronger verification: its value cannot exceed the observed value $P(H = 0, W = 0, V = 0)$, the fraction of households that experienced an unfavorable health outcome, reported no WIC benefits, and true participation status is unverified. Because $\theta_v$ is capped, it now becomes $\theta_v = P(H = 1, W = 0, Z = 0)$ instead of $\theta_v$ that rises with $P^*$ to satisfy the $\Delta$ constraint. As these false-negative-reporting households experienced healthy instead of unhealthy outcomes, the lower bound on the ATE reverses direction and begins to rise with $P^*$. 

27
adjunctive eligibility may be arbitrarily misclassified by more than 2 percent (3 percent under no false positives), we can no longer sign the ATE without further assumptions.

If we additionally impose monotone treatment response (MTR), the healthy birth weight outcome improves by at least 3.66 percentage points, or 20 percent. In fact, the average treatment effect is strictly positive, and statistically significant, even for large degrees of arbitrary WIC misreporting (no verification) under this joint MTS + inc-MIV + ineligibles-MIV + MTR model. In this case, results are robust to large values of \( q \). This model also identifies that WIC decreases the probability of an unfavorable gestation length by at least 0.0148 points, or 5.8 percent (Figure 3B). The confidence interval includes 0 for this outcome, but the lower bound is actually higher at \( P^r \) than under an assumption of perfectly measured data.

### 6.3. Point Estimates Using Covariance Restrictions

The preceding nonparametric approach partially identifies the causal effects of prenatal WIC participation under a variety of different assumptions to address the classification and selection problems. To complement this analysis, we apply Lewbel’s (2012) linear simultaneous equations model that relies on second moment restrictions to point identify the ATE. In Lewbel’s model, identification is achieved through the presence of covariates related to the conditional variance of the first-stage errors, but not the conditional covariance between first- and second-stage errors. This model circumvents the need for traditional instruments by generating a pseudo-instrument to identify the structural parameters.\(^{24}\)

To implement this approach, we use the set of covariates listed in Table 1 including mother’s age, father’s age, an indicator for married parents, an indicator for whether the mother has a high school (HS) degree or less, a corresponding indicator for the father, an indicator for whether the child is male, an indicator for whether the child is Black or Hispanic, an indicator for living in a rural area,

\(^{24}\)The literature has struggled to find credible exclusion restrictions, and this is even more problematic in light of nonclassical measurement errors where true WIC participation is negatively correlated with the error. In this setting, any plausible instrument for true WIC participation is likely to be correlated with the measurement error as well.
three regional dummies, and four SES quintile dummies. Specification (1) includes the basic set of covariates. Specification (2) includes mother’s age squared, an interaction term between the mother’s age and whether she has a HS degree or less, an interaction term between whether the child is black or Hispanic and whether the mother has at most a HS degree, and an indicator of whether the child is black or Hispanic with married parents. Specification (3) includes the mother’s age cubed and an interaction term between squared mother’s age and whether the mother has at most a HS degree.

The estimates and standard errors reported in Table 3 suggest that WIC has a modest positive effect on the prevalence of normal birth weight. First, consider the estimates associated with the probability of normal birth weight (Panel III). The Lewbel model estimate of 0.031 implies a small but beneficial effect of prenatal WIC receipt on birth weight. Interestingly, the estimate of 0.031 lies outside the estimated bound on the ATE, [0.065, 0.639] under the joint MIV-MTS model with no classification error. With modest errors, however, the bounds on the ATE include this IV estimate. A similar result holds for the probability of birth weight of at least 2500 grams and the probability of birth weight above 1500 grams.

In contrast, the estimated effects of WIC on gestation outcomes are negative. The negative estimate of prenatal WIC on the probability of full term pregnancy lies outside of estimated bounds when there is no classification error and is inconsistent with the MTR assumption. However, when at least two percent of the eligible women misreport participation, the bounds on the ATE under the MTS-MIV models include zero as well as the IV estimates. In sum, there is less evidence of a beneficial effect of WIC on gestation age (see Figlio et al., 2009).

---

25 Estimating this model involves a three step process. We begin by estimating the first stage reduced form equation by OLS and then use the Breusch-Pagan test for heteroskedasticity on the residuals of the reduced form. Using this test, we selected instruments that are significantly associated with the conditional variance, namely mother’s age and four indicators of SES quintile. These instruments are assumed to be unrelated to the conditional covariance between first- and second-stage errors. Then, following Lewbel, we estimate the structural equation using two-stage least squares estimation.
7. CONCLUSION

Social scientists are often confronted with the dual identification problems of unknown counterfactual outcomes and imperfectly measured data. Using a framework that can be applied to a wide range of topics, this paper studied what can be learned about the efficacy of the prenatal WIC program in improving birth outcomes when program participation is underreported. Classical measurement error assumptions do not apply. Errors are mean-reverting, systematic in one direction, and may be correlated with the outcomes of interest. Since WIC participation is self-selected (not randomly assigned), even with perfectly measured data we could not infer how infants would have fared had the mother made a different participation decision.

While our framework does not allow us to point-identify average treatment effects, we derived sharp bounds under arbitrary misclassification of the treatment indicator. Extending KPGJ (2012), we considered the identifying power of “partially verified” treatments using a model that combines information from ECLS-B survey data and auxiliary administrative data from the USDA to derive logical restrictions on WIC misreporting patterns. We studied inference under both exogenous and endogenous WIC participation settings. After some initial deterioration, lower bounds on the average treatment effects rise, not fall, with the degree of misclassification. Modifying a regression discontinuity framework, we also introduced the notion of a partially observed monotone instrument. Specifically, households with income above the statutory eligibility threshold (185 percent of the poverty guidelines) are adjunctively eligible for WIC through the receipt of TANF, SNAP, or Medicaid benefits. Like WIC, participation in these programs may be mismeasured.

Under relatively weak assumptions that allow for some measurement error, we estimate that WIC reduces the prevalence of unhealthy birth weights by at least a few percentage points, or around 20 percent. We also identify a positive impact on normal gestation length of at least 6 percent. This latter effect is not statistically significantly different from zero, but the lower bound is higher under moderate underreporting than under an assumption of perfectly measured data. For both birth weights and gestation length, our bounds on average treatment effects are consistent with much larger beneficial impacts. Larger effects if they exist, however, cannot be identified without stronger assumptions on the selection and classification error processes.
References


APPENDIX

Proof of Proposition 1:

We can write the average treatment effect as

\[ ATE = \frac{P(H = 1, W = 1) - P(W = 1)P(H = 1) + \theta^{-}_{1} - \theta^{+}_{1}}{P(W = 1)P(W = 0)}. \]  \hspace{1cm} (A1)

Thus, subject to restrictions on the unknown classification error rates in Equations (5)-(7) and the laws of probability,\(^{26}\) the upper bound is found by maximizing \(\theta^{-}_{1}\) and minimizing \(\theta^{+}_{1}\). Likewise, the lower bound is found by minimizing \(\theta^{-}_{1}\) and maximizing \(\theta^{+}_{1}\). Let \(\Psi = \theta^{-}_{1} - \theta^{+}_{1}\).

If \(\Delta \geq 0\) (net false negative reporting): For the upper bound, first consider the case that \(\theta^{-}_{1} - \Delta < \theta^{+}_{0}\). Here, \(\theta^{-}_{1}\) cannot exceed \(\theta^{-}_{1} - \Delta\) and \(\theta^{+}_{1}\) cannot fall below \(\theta^{+}_{0}\) so that \(\Psi \leq \theta^{-}_{1} - \Delta\). At this upper bound, Equation (5) implies that \(\theta^{-}_{0} = \Delta - \theta^{-}_{1} - \Delta\) and \(\theta^{+}_{0} = 0\). If \(\theta^{-}_{1} - \Delta \geq \Delta\), \(\theta^{-}_{1}\) cannot exceed \(\theta^{-}_{1} - \Delta\) and \(\theta^{+}_{1}\) cannot fall below \(\theta^{+}_{1} - \Delta\) so that \(\Psi \leq \min\{\theta^{-}_{1} - \Delta, \theta^{+}_{1} - \Delta\}\). At this upper bound, Equation (5) implies \(\theta^{-}_{0} = \theta^{+}_{1} - \Delta\) and \(\theta^{+}_{0} = \min\{\theta^{-}_{1} - \Delta, \theta^{+}_{1} - \Delta\}\). For the lower bound, first consider the case that \(\theta^{+}_{0} < \Delta\). Here, \(\theta^{+}_{0}\) cannot exceed \(\theta^{+}_{0} + \Delta\) so from Equation (5) we know that \(\theta_{1}\) must be no less than \(\Delta - \theta^{+}_{0} - \Delta\). From Equation (7), we know that \(\theta^{+}_{j}\) can exceed \(\theta^{+}_{0}\) but any conjectured increase in the false positive error rate must be offset by an equivalent increase in the false negative error rate. So, in this case, the lower bound would be unchanged by increasing \(\theta^{+}_{j}\) above \(\theta^{+}_{0}\). Thus, we have \(\Psi \geq \Delta - \theta^{+}_{0}\). If \(\theta^{+}_{0} \geq \Delta\), \(\Psi\) is minimized when \(\theta^{-}_{0} = \min\{\theta^{-}_{1} - \Delta, \theta^{+}_{1} - \Delta\}, \theta^{+}_{0} = 0\), \(\theta^{+}_{1} = \min\{\theta^{-}_{1} - \Delta, \theta^{+}_{1} - \Delta\}, \text{ and } \theta^{-}_{0} = 0\) so that \(\Psi \geq -\min\{\theta^{-}_{1} - \Delta, \theta^{+}_{1} - \Delta\}\).

If \(\Delta \leq 0\) (net false positive reporting): For the upper bound, first consider the case that \(\theta^{+}_{0} - \Delta < \theta^{+}_{1}\). Here, \(\theta^{+}_{0}\) cannot exceed \(\theta^{+}_{0} + \Delta\) so from Equation (5) we know that \(\theta_{1}\) must be no less than \(\Delta - \theta^{+}_{0}\). From Equation (7), we know that \(\theta^{-}_{j}\) can exceed \(\theta^{+}_{0}\) but any conjectured increase in the false negative error rate must be offset by an equivalent increase in the false positive error rate. So,

\(^{26}\) In particular, \(P(H = 1|W = 1) = P(H = 1|W = 0)\) lie within \([0, 1]\). A more detailed proof showing these restrictions is available from the authors.
in this case, the upper bound would be unchanged by increasing \( \theta_j \) above 0. Thus, we have

\[ \Psi \leq \Delta - \theta_0^{UB} \]. If \( \theta_0^{UB} \geq \Delta \), \( \Psi \) is maximized when \( \theta_0^* = \min \{ \theta_0^{UB}, \theta_0^{LB} - \Delta \} \), \( \theta_1^* = 0 \),

\[ \theta_1^- = \min \{ \theta_1^{UB}, \theta_0^{UB} + \Delta \} \], and \( \theta_0^- = 0 \) so that \( \Psi \leq \min \{ \theta_1^{LB}, \theta_0^{LB} + \Delta \} \). For the lower bound, first consider the case that \( \theta_1^{UB} < -\Delta \). Here, \( \theta_1^+ \) cannot exceed \( \theta_1^{UB} \) and \( \theta_1^- \) cannot fall below 0 so that \( \Psi \geq -\theta_1^{UB} \). At this lower bound, Equation (5) implies that \( \theta_0^* = \Delta - \theta_1^{UB} \) and \( \theta_0^- = 0 \). If \( \theta_1^{UB} \geq -\Delta \), \( \theta_1^+ \) cannot exceed \( \min \{ \theta_1^{UB}, \theta_0^{LB} - \Delta \} \) and \( \theta_1^- \) cannot fall below 0 so that \( \Psi \geq -\min \{ \theta_1^{UB}, \theta_0^{LB} - \Delta \} \).

At this lower bound, Equation (5) implies \( \theta_0^* = 0 \) and \( \theta_0^- = \min \{ \theta_0^{LB}, \theta_0^{UB} + \Delta \} \). Combining these results, it follows that \( -\min \{ \theta_1^{UB}, \theta_0^{LB} - \Delta \} \leq \Psi \leq \min \{ \theta_1^{UB}, \theta_0^{UB} + \Delta \} \). \[ \square \]

**Proof of Corollary 1**

Under the arbitrary errors model, \( Q_n = 1 \). From Equation (7), we know that

\[ \theta_0^{UB} + \Delta = \min \left\{ P(W^* = 1) - P(H = 1,W = 1), P(W = 0), \frac{1}{2} [P(W^* = 1) + P(W = 0)] \right\} \]

and

\[ \theta_1^{UB} = \min \left\{ P(H = 1,W = 0), P(W^* = 1), \frac{1}{2} [P(W^* = 1) + P(W = 0)] \right\} \]. Thus, it follows that

\[ \min \{ \theta_1^{UB}, \theta_0^{UB} + \Delta \} = \min \{ P(W^* = 1) - P(H = 1,W = 1), P(H = 1,W = 0) \} \]. \[ \text{A2} \]

Likewise, from Equation (7) we know that

\[ \theta_0^{LB} - \Delta = \min \left\{ P(W^* = 0) - P(H = 1,W = 0), P(W = 1), \frac{1}{2} [P(W^* = 0) + P(W = 1)] \right\} \]

and

\[ \theta_1^{LB} = \min \left\{ P(H = 1,W = 1), P(W^* = 0), \frac{1}{2} [P(W^* = 0) + P(W = 1)] \right\} \]. Thus, it follows that

\[ \min \{ \theta_1^{LB}, \theta_0^{LB} - \Delta \} = \min \{ P(W^* = 0) - P(H = 1,W = 0), P(H = 1,W = 1) \} \]. \[ \text{A3} \]

Substituting (A2) into Proposition 1 obtains an upper bound of \( \frac{P(H = 0)}{1 - P^*} \) if \( P^* \leq P(H = 1) \) and

\[ \frac{P(H = 1)}{P^*} \] if \( P^* \geq P(H = 1) \), which reduces to \( \min \left\{ \frac{P(H = 0)}{1 - P^*}, \frac{P(H = 1)}{P^*} \right\} \). Similarly, the lower bound is obtained using Equation (A3). \[ \square \]
Proof of Proposition 2

This proof follows the structure of the proof of Proposition 1 in KPGJ but allows for verification and an upper bound error rate. These restrictions are embedded in Equation (7). Subject to the restrictions in Equations (5) and (7), the upper bound is found by maximizing \( (\theta_i^- + \theta_0^-) \) and minimizing \( (\theta_0^+ + \theta_i^+) \), and vice versa for the lower bound.

If \( \Delta \geq 0 \): For the upper bound, first consider the case that \( \theta_i^{-UB} \geq \Delta \). Then \( (\theta_0^- + \theta_i^-) \) is minimized at 0 and Equation (5) simplifies to \( \theta_i^- = \Delta + \theta_0^- \). It follows that \( \theta_0^- \) cannot exceed \( \min \{ \theta_0^{-UB}, \theta_i^{-UB} - \Delta \} \) and \( \theta_i^- \) cannot exceed \( \min \{ \Delta + \theta_0^{-UB}, \theta_i^{-UB} \} \). The upper bound follows directly.

Second, consider the case that \( \theta_i^{-UB} < \Delta \). We know that \( \theta_i^- \) cannot exceed \( \theta_i^{-UB} \) and, to satisfy the restriction in Equation (5), \( \theta_0^- \) must be no less than \( \Delta - \theta_i^{-UB} \). As before, \( \theta_i^- \) is minimized at 0. While \( \theta_0^- \) can exceed 0, any conjectured increase in the false positive error rate must be offset by an equivalent increase in the false negative error rate. So, in this case, the upper bound would be unchanged by increasing \( \theta_0^- \) above 0. Thus, we have the upper bound on \( \Theta \) of \( 2\theta_i^{-UB} - \Delta \) which can be shown to be no greater than \( 2\theta_0^{-UB} + \Delta \).

For the lower bound, first consider the case that \( \theta_0^{-UB} \geq \Delta \). Then \( (\theta_0^- + \theta_i^-) \) is minimized at 0 and Equation (5) simplifies to \( \theta_0^- = \Delta + \theta_i^- \). It follows that \( \theta_i^- \) cannot exceed \( \min \{ \theta_i^{-UB}, \theta_0^{-UB} - \Delta \} \) and \( \theta_0^- \) cannot exceed \( \min \{ \Delta + \theta_i^{-UB}, \theta_0^{-UB} \} \) so that max \( \{-2\theta_i^{-UB} - \Delta, -2\theta_0^{-UB} + \Delta\} \) provides the lower bound on \( \Theta \). Second, consider the case that \( \theta_i^{-UB} < \Delta \). We know that \( \theta_0^- \) cannot exceed \( \theta_i^{-UB} \) and, to satisfy the restriction in Equation (5), \( \theta_i^- \) must be no less than \( \Delta - \theta_i^{-UB} \). As before, \( \theta_0^- \) is minimized at 0. While \( \theta_i^- \) can exceed 0, any conjectured increase in the false positive error rate must be offset by an equivalent increase in the false negative error rate. So, in this case, the lower bound would be unchanged by increasing \( \theta_0^- \) above 0. Thus, we have the lower bound on \( \Theta \) of \(-2\theta_i^{-UB} + \Delta \) which can be shown to be no smaller than \(-2\theta_0^{-UB} - \Delta \).

If \( \Delta < 0 \): For the upper bound, first consider the case that \( \theta_0^{-UB} \geq -\Delta \). Then \( (\theta_0^+ + \theta_i^+) \) is minimized at 0 and Equation (5) simplifies to \( \theta_0^+ = -\Delta + \theta_i^+ \). We know that \( \theta_i^+ \) cannot exceed \( \min \{ \theta_i^{-UB}, \theta_0^{-UB} + \Delta \} \) and \( \theta_0^+ \) cannot exceed \( \min \{ \theta_0^{-UB}, -\Delta + \theta_i^{-UB} \} \). The upper bound follows directly.
Second, consider the case that \( \theta_0^{UB} < -\Delta \). We know that \( \theta_0^+ \) cannot exceed \( \theta_0^{UB} \) and, to satisfy the restriction in Equation (5), \( \theta_1^+ \) must be no less than \( -\Delta - \theta_0^{UB} \). As before, \( \theta_0^- \) is minimized at 0. While \( \theta_1^- \) can exceed 0, any conjectured increase in the false negative error rate must be offset by an equivalent increase in the false positive error rate. So, in this case, the upper bound would be unchanged by increasing \( \theta_1^- \) above 0. Thus, we have the upper bound on \( \Theta \) of \( 2\theta_0^{UB} + \Delta \) which can be shown to be no greater than \( 2\theta_1^{UB} - \Delta \).

For the lower bound, first consider the case that \( \theta_i^{UB} \geq -\Delta \). Then \( (\theta_i^- + \theta_i^+) \) is minimized at 0 and Equation (5) simplifies to \( \theta_1^+ = -\Delta + \theta_0^- \). We know that \( \theta_0^- \) cannot exceed \( \min\{\theta_0^{UB}, \theta_1^{UB} + \Delta\} \) and \( \theta_1^+ \) cannot exceed \( \min\{\theta_1^{UB}, -\Delta + \theta_0^{UB}\} \) so that \( \max\{-2\theta_1^{UB} - \Delta, -2\theta_0^{UB} + \Delta\} \) provides the lower bound on \( \Theta \). Second, consider the case that \( \theta_i^{UB} < -\Delta \). We know that \( \theta_1^+ \) cannot exceed \( \theta_i^{UB} \) and, to satisfy the restriction in Equation (5), \( \theta_0^+ \) must be no less than \( -\Delta - \theta_i^{UB} \). As before, \( \theta_1^- \) is minimized at 0. While \( \theta_0^- \) can exceed 0, any conjectured increase in the false negative error rate must be offset by an equivalent increase in the false positive error rate. So, in this case, the lower bound would be unchanged by increasing \( \theta_0^- \) above 0. Thus, we have the lower bound of \( -2\theta_1^{UB} - \Delta \) which can be shown to be no smaller than \( -2\theta_0^{UB} + \Delta \). \( \square \)
Table 1. Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>WIC – no WIC</th>
<th>Mean</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prenatal WIC Receipt (1 = Yes)</td>
<td>0.69</td>
<td>0.46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Child’s Health Indicators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth Weight (in grams)</td>
<td>3276.23</td>
<td>574.04</td>
<td>-35.74</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Birth Weight ≥ 2500 grams, Not Low Birth Weight (1 = yes)</td>
<td>0.93</td>
<td>0.26</td>
<td>0.00</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>Birth Weight ≥ 1500 grams, Not Very Low (1 = yes)</td>
<td>0.99</td>
<td>0.11</td>
<td>0.00</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>Birth Weight ≤ 4000 grams, Not Macrosomic (1 = yes)</td>
<td>0.92</td>
<td>0.26</td>
<td>0.02</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Normal Birth Weight: 2500-4000 grams (1=yes)</td>
<td>0.85</td>
<td>0.36</td>
<td>0.02</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Gestation Age (in weeks)</td>
<td>38.78</td>
<td>2.57</td>
<td>0.00</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>Gestation Age: 38-42 weeks, Normal Gestation Age (1 = yes)</td>
<td>0.75</td>
<td>0.43</td>
<td>-0.02</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>Gestation Age ≥ 33 weeks, Not Very Premature (1 = yes)</td>
<td>0.97</td>
<td>0.18</td>
<td>0.00</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>Gestation Age ≥ 37 weeks, Not Premature (1 = yes)</td>
<td>0.88</td>
<td>0.33</td>
<td>-0.02</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td><strong>Covariates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (in months)</td>
<td>10.53</td>
<td>2.07</td>
<td>-0.09</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Gender (1 = male)</td>
<td>0.50</td>
<td>0.50</td>
<td>0.00</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>White (1 = yes)</td>
<td>0.35</td>
<td>0.48</td>
<td>-0.12</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Black (1 = yes)</td>
<td>0.21</td>
<td>0.40</td>
<td>0.07</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Hispanic (1 = yes)</td>
<td>0.37</td>
<td>0.48</td>
<td>0.07</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Asian (1 = yes)</td>
<td>0.02</td>
<td>0.14</td>
<td>-0.02</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Urban Area (1 = yes)</td>
<td>0.69</td>
<td>0.46</td>
<td>-0.04</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Urban Cluster (1 = yes)</td>
<td>0.14</td>
<td>0.35</td>
<td>0.03</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Northeast (1 = yes)</td>
<td>0.14</td>
<td>0.35</td>
<td>-0.04</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Midwest (1 = yes)</td>
<td>0.19</td>
<td>0.40</td>
<td>0.00</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>South (1 = yes)</td>
<td>0.40</td>
<td>0.49</td>
<td>0.01</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>West (1 = yes)</td>
<td>0.26</td>
<td>0.44</td>
<td>0.04</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Household Socioeconomics Status (SES)</td>
<td>-0.67</td>
<td>0.53</td>
<td>-0.29</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Household SES Quintile 1 (1 = yes)</td>
<td>0.41</td>
<td>0.49</td>
<td>0.17</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Household SES Quintile 2 (1 = yes)</td>
<td>0.33</td>
<td>0.47</td>
<td>0.05</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Household SES Quintile 3 (1 = yes)</td>
<td>0.18</td>
<td>0.39</td>
<td>-0.11</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Household SES Quintile 4 (1 = yes)</td>
<td>0.07</td>
<td>0.26</td>
<td>-0.08</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Household SES Quintile 5 (1 = yes)</td>
<td>0.02</td>
<td>0.12</td>
<td>-0.04</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Household Size</td>
<td>4.65</td>
<td>1.69</td>
<td>-0.17</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Parents Married (1 = yes)</td>
<td>0.47</td>
<td>0.50</td>
<td>-0.15</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Mother’s Age</td>
<td>25.88</td>
<td>5.88</td>
<td>-2.23</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Kessner Index of Prenatal Care (1 = Adequate)</td>
<td>0.63</td>
<td>0.48</td>
<td>0.03</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Kessner Index of Prenatal Care (1 = Intermediate)</td>
<td>0.23</td>
<td>0.42</td>
<td>0.03</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Kessner Index of Prenatal Care (1 = Inadequate)</td>
<td>0.08</td>
<td>0.28</td>
<td>-0.04</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Kessner Index of Prenatal Care (1 = Unknown)</td>
<td>0.06</td>
<td>0.23</td>
<td>-0.01</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>Father’s Age</td>
<td>30.36</td>
<td>5.72</td>
<td>-0.92</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Mother’s Weight (in kilograms)</td>
<td>72.69</td>
<td>17.83</td>
<td>2.03</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Mother’s Education - High School (1=yes)</td>
<td>0.38</td>
<td>0.49</td>
<td>0.05</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Mother’s Education - Some College (1 = yes)</td>
<td>0.21</td>
<td>0.41</td>
<td>-0.09</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Mother’s Education - Bachelor’s Degrees (1 = yes)</td>
<td>0.03</td>
<td>0.18</td>
<td>-0.05</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Mother’s Education - Advanced College Degree (1=yes)</td>
<td>0.01</td>
<td>0.11</td>
<td>-0.02</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Father’s Education - High School (1=yes)</td>
<td>0.36</td>
<td>0.48</td>
<td>-0.01</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>Father’s Education - Some College (1 = yes)</td>
<td>0.18</td>
<td>0.38</td>
<td>-0.06</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Father’s Education - Bachelor’s Degrees (1 = yes)</td>
<td>0.04</td>
<td>0.19</td>
<td>-0.05</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Father’s Education - Advanced College Degree (1=yes)</td>
<td>0.02</td>
<td>0.13</td>
<td>-0.04</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

Note: The sample includes 9-month old children from households with income at or below 185% of the Federal Poverty Guidelines. All analyses are weighted using Wave 1 specific sample weights. Urban Cluster is defined as less densely populated than an Urbanized Area. The omitted category for race is ‘other,’ area type is ‘rural,’ mother’s education is ‘less than high school,’ and father’s education is ‘less than high school.’
Table 2. Sharp Bounds on the ATE of WIC Under No Classification Error

<table>
<thead>
<tr>
<th>Model</th>
<th>Birth Weight</th>
<th>Gestation Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal (2500-4000 g.)</td>
<td>Not Low (≥ 2500 g.)</td>
</tr>
<tr>
<td>(i) Worst Case</td>
<td>p.e.(^a) ([-0.361, 0.639])</td>
<td>CI ([-0.340, 0.660])</td>
</tr>
<tr>
<td></td>
<td>p.e. ([-0.373, 0.651])</td>
<td>CI ([-0.351, 0.671])</td>
</tr>
<tr>
<td>(ii) MTS(^c)</td>
<td>p.e. [0.019, 0.639]]</td>
<td>CI [-0.003, 0.651]\</td>
</tr>
<tr>
<td></td>
<td>p.e. [-0.003, 0.651]]</td>
<td>CI [-0.010, 0.671]]</td>
</tr>
<tr>
<td>(iii) MTS+MIV(^d)</td>
<td>p.e. [0.065, 0.639]]</td>
<td>CI [0.003, 0.651]]</td>
</tr>
<tr>
<td></td>
<td>p.e. [-0.008, 0.672]]</td>
<td>CI [0.001, 0.696]]</td>
</tr>
<tr>
<td>(iv) MTS+MIV +ineligibles(^e)</td>
<td>p.e. [0.065, 0.639]]</td>
<td>CI [0.028, 0.651]]</td>
</tr>
<tr>
<td></td>
<td>p.e. [-0.008, 0.672]]</td>
<td>CI [0.001, 0.696]]</td>
</tr>
<tr>
<td>(v) MTS-MIV+MTR(^f) +ineligibles</td>
<td>p.e. [0.065, 0.639]]</td>
<td>CI [0.028, 0.651]]</td>
</tr>
<tr>
<td></td>
<td>p.e. [-0.008, 0.672]]</td>
<td>CI [0.001, 0.696]]</td>
</tr>
</tbody>
</table>

Notes:

a. Bias-corrected point estimates of the bounds
b. 90% Imbens-Manski confidence internals (CI) using 1000 pseudosamples
c. MTS denotes Monotone Treatment Selection
d. MIV denotes the income monotone instrument
e. ‘ineligibles’ denotes the ineligibles monotone instrument
f. MTR denotes Monotone Treatment Response
Figure 1A. Sharp Bounds on the ATE for “Normal Birth Weight” (2500-4000 grams) as a Function of $P^*$, the Unobserved True WIC Participation Rate: **Exogenous Selection**

Exogenous selection

<table>
<thead>
<tr>
<th>(a) Arbitrary errors</th>
<th>p.e. $\uparrow$</th>
<th>LB</th>
<th>UB</th>
<th>width</th>
<th>LB</th>
<th>UB</th>
<th>width</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CI $\downarrow$</td>
<td>-0.217, 0.478</td>
<td>0.695</td>
<td>[-0.202, 0.575]</td>
<td>0.777</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) No false positives</td>
<td>p.e.</td>
<td>[0.019, 0.019]</td>
<td>0.000</td>
<td>[-0.202, 0.062]</td>
<td>0.264</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CI</td>
<td>[-0.003, 0.041]</td>
<td>[0.002, 0.062]</td>
<td>0.139</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) Verified if reported any gov’t benefits</td>
<td>p.e.</td>
<td>[0.019, 0.019]</td>
<td>0.000</td>
<td>[-0.077, 0.062]</td>
<td>0.139</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CI</td>
<td>[-0.003, 0.041]</td>
<td>[-0.093, 0.082]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^\dagger$ Point estimates of the population bounds

$^\ddagger$ Imbens-Manski (2004) 5th and 95th percentile bounds (1,000 pseudosamples)
Figure 1B. Sharp Bounds on the ATE for “Normal Gestation” (38-42 weeks) as a Function of $P^*$, the Unobserved True WIC Participation Rate: Exogenous Selection

### Exogenous selection

<table>
<thead>
<tr>
<th>Type</th>
<th>Self-reported participation rate: $P^* = P = 0.687$</th>
<th>Administrative participation rate: $P^* = P^o = 0.74$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Arbitrary errors</td>
<td>p.e. [-0.367, 0.803] 1.169</td>
<td>LB [-0.340, 0.968] UB [0.308]</td>
</tr>
<tr>
<td></td>
<td>CI [-0.384, 0.848]</td>
<td>CI [-0.356, 1.000]</td>
</tr>
<tr>
<td>(b) No false positives</td>
<td>p.e. [-0.022, -0.022] 0.000</td>
<td>LB [-0.232, 0.046] UB [0.278]</td>
</tr>
<tr>
<td></td>
<td>CI [-0.051, 0.008]</td>
<td>CI [-0.258, 0.072]</td>
</tr>
<tr>
<td>(c) Verified if reported any gov’t benefits</td>
<td>p.e. [-0.022, -0.022] 0.000</td>
<td>LB [-0.120, 0.046] UB [0.165]</td>
</tr>
<tr>
<td></td>
<td>CI [-0.051, 0.008]</td>
<td>CI [-0.142, 0.072]</td>
</tr>
</tbody>
</table>
Figure 2A. Sharp Bounds on the ATE for “Normal Birth Weight” (2500-4000 grams) as a Function of $P^*$, the Unobserved True WIC Participation Rate: Worst Case Selection Bounds

**ATE**

<table>
<thead>
<tr>
<th>Endogenous selection</th>
<th>Self-reported participation rate: $P^* = P = 0.687$</th>
<th>Administrative participation rate: $P^* = P^o = 0.74$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Arbitrary errors</td>
<td>p.e. [-0.462, 0.837] 1.299</td>
<td>p.e. [-0.409, 0.899] 1.299</td>
</tr>
<tr>
<td></td>
<td>CI [-0.447, 0.851]</td>
<td>CI [-0.424, 0.903]</td>
</tr>
<tr>
<td>(b) No false positives</td>
<td>p.e. [-0.361, 0.639] 1.000</td>
<td>p.e. [-0.409, 0.692] 1.102</td>
</tr>
<tr>
<td></td>
<td>CI [-0.373, 0.651]</td>
<td>CI [-0.423, 0.704]</td>
</tr>
<tr>
<td>(c) Verified if reported any gov’t benefits</td>
<td>p.e. [-0.361, 0.639] 1.000</td>
<td>p.e. [-0.361, 0.692] 1.054</td>
</tr>
<tr>
<td></td>
<td>CI [-0.373, 0.651]</td>
<td>CI [-0.375, 0.704]</td>
</tr>
</tbody>
</table>
Figure 2B. Sharp Bounds on the ATE for “Normal Gestation” (38-42 weeks) as a Function of $P^*$, the Unobserved True WIC Participation Rate: Worst Case Selection Bounds

**ATE**

![Graph showing the ATE as a function of $P^*$]

<table>
<thead>
<tr>
<th>Endogenous selection</th>
<th>Self-reported participation rate: $P^* = P = 0.687$</th>
<th>Administrative participation rate: $P^* = P^o = 0.74$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Arbitrary errors</td>
<td>p.e. [-0.565, 0.938] 1.503</td>
<td>p.e. [-0.512, 0.992] 1.503</td>
</tr>
<tr>
<td></td>
<td>CI [-0.581, 0.954]</td>
<td>CI [-0.527, 1.000]</td>
</tr>
<tr>
<td>(b) No false positives</td>
<td>p.e. [-0.417, 0.583] 1.000</td>
<td>p.e. [-0.470, 0.637] 1.107</td>
</tr>
<tr>
<td></td>
<td>CI [-0.430, 0.596]</td>
<td>CI [-0.483, 0.650]</td>
</tr>
<tr>
<td>(c) Verified if reported any gov’t benefits</td>
<td>p.e. [-0.417, 0.583] 1.000</td>
<td>p.e. [-0.427, 0.637] 1.064</td>
</tr>
<tr>
<td></td>
<td>CI [-0.430, 0.596]</td>
<td>CI [-0.441, 0.650]</td>
</tr>
</tbody>
</table>
Figure 3A. Sharp Lower Bounds on the ATE for “Normal Birth weight” (2500-4000 grams) as a Function of $P^*$, the Unobserved True WIC Participation Rate: MTS, Income MIV, and Ineligibles MIV

<table>
<thead>
<tr>
<th>MTS+ inc MIV + ineligibles:</th>
<th>width</th>
<th>width</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Arbitrary errors</td>
<td>p.e.</td>
<td>[-0.017, 0.837] 0.853</td>
</tr>
<tr>
<td></td>
<td>CI</td>
<td>[-0.047, 0.851]</td>
</tr>
<tr>
<td></td>
<td>bias*</td>
<td>+0.013 -0.001</td>
</tr>
<tr>
<td>(b) No false positives</td>
<td>p.e.</td>
<td>[0.065, 0.639] 0.574</td>
</tr>
<tr>
<td></td>
<td>CI</td>
<td>[0.028, 0.651]</td>
</tr>
<tr>
<td></td>
<td>bias</td>
<td>+0.019 -0.005</td>
</tr>
<tr>
<td>(c) Verified if reported any gov’t benefits</td>
<td>p.e.</td>
<td>[0.065, 0.639] 0.574</td>
</tr>
<tr>
<td></td>
<td>CI</td>
<td>[0.028, 0.651]</td>
</tr>
<tr>
<td></td>
<td>bias</td>
<td>+0.019 -0.012</td>
</tr>
</tbody>
</table>

MTS+ inc MIV +ineligibles+MTR:

<table>
<thead>
<tr>
<th>MTS+ inc MIV+ineligibles+MTR:</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Arbitrary errors</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>(b) No false positives</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>(c) Verified if reported any gov’t benefits</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

* Corrected finite sample bias
Figure 3B. Sharp Bounds on the ATE for “Normal Gestation” (38-42 weeks) as a Function of $P^*$, the Unobserved True WIC Participation Rate: MTS, Income MIV, and Ineligibles MIV

$P^* = P = 0.687$

MTS+ inc MIV + ineligibles:

(a) Arbitrary errors
- p.e. $[-0.138, 0.908]$ 1.046
- CI $[-0.179, 0.948]$ $[-0.117, 0.928]$ 1.045
- bias $+0.017$ $-0.008$ $+0.014$ $-0.007$

(b) No false positives
- p.e. $[0.0059, 0.576]$ 0.570
- CI $[-0.0444, 0.596]$ $[-0.077, 0.630]$ 0.707
- bias $+0.027$ $-0.005$ $+0.027$ $-0.005$

(c) Verified if reported any gov’t benefits
- p.e. $[0.0059, 0.576]$ 0.570
- CI $[-0.0444, 0.596]$ $[-0.092, 0.650]$ 0.670
- bias $+0.027$ $-0.005$ $+0.017$ $-0.005$

MTS+ inc MIV+MTR+ineligibles:

(a) Arbitrary errors
- p.e. $[0.0012, 0.908]$ 0.907
- CI $[0.0000, 0.948]$ $[0.0012, 0.928]$ 0.927
- bias $+0.017$ $-0.008$ $+0.014$ $-0.007$

(b) No false positives
- p.e. $[0.010, 0.576]$ 0.566
- CI $[0.000, 0.596]$ $[0.0012, 0.630]$ 0.629
- bias $+0.027$ $-0.005$ $+0.027$ $-0.005$

(c) Verified if reported any gov’t benefits
- p.e. $[0.010, 0.576]$ 0.566
- CI $[0.000, 0.596]$ $[0.015, 0.630]$ 0.616
- bias $+0.027$ $-0.005$ $+0.017$ $-0.005$
### Table 3. IV Estimates of the Effect of Prenatal WIC Participation on Infant Birth Outcomes

<table>
<thead>
<tr>
<th></th>
<th>Birth Weight</th>
<th></th>
<th>Gestation Age</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>Not Low</td>
<td>Not Very</td>
<td>Normal</td>
</tr>
<tr>
<td></td>
<td>(2500-4000 g.)</td>
<td>(≥ 2500 g.)</td>
<td>Low (≥ 1500 g.)</td>
<td>(≤ 4000 g.)</td>
</tr>
<tr>
<td><strong>Specification 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>OLS</strong></td>
<td>0.021</td>
<td>0.008</td>
<td>0.004**</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.007)</td>
<td>(0.002)</td>
<td>(0.013)</td>
</tr>
<tr>
<td></td>
<td><strong>IV</strong></td>
<td>0.031*</td>
<td>0.009</td>
<td>0.005***</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.008)</td>
<td>(0.002)</td>
<td>(0.014)</td>
</tr>
<tr>
<td><strong>Specification 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>OLS</strong></td>
<td>0.022</td>
<td>0.008</td>
<td>0.004**</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.008)</td>
<td>(0.002)</td>
<td>(0.013)</td>
</tr>
<tr>
<td></td>
<td><strong>IV</strong></td>
<td>0.031*</td>
<td>0.010</td>
<td>0.005***</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.008)</td>
<td>(0.002)</td>
<td>(0.014)</td>
</tr>
<tr>
<td><strong>Specification 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>OLS</strong></td>
<td>0.022</td>
<td>0.008</td>
<td>0.004**</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.007)</td>
<td>(0.002)</td>
<td>(0.013)</td>
</tr>
<tr>
<td></td>
<td><strong>IV</strong></td>
<td>0.030*</td>
<td>0.010</td>
<td>0.005***</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.008)</td>
<td>(0.002)</td>
<td>(0.014)</td>
</tr>
</tbody>
</table>

Notes:  * p<0.10, ** p<0.05, *** p<0.01. Robust standard errors are in parentheses. All analyses are weighted using Wave 1 specific sample weights. Specification (1) includes the basic set of covariates. Specification (2) includes mother’s age squared, an interaction term between the mother’s age and whether she has a HS degree or less, an interaction term between whether the child is black or Hispanic and whether the mother has at most a HS degree, and an indicator of whether the child is black or Hispanic with married parents. Specification (3) includes the mother’s age cubed and an interaction term between squared mother’s age and whether the mother has at most a HS degree.