The Causal Effect of WIC on Breastfeeding: Partial Identification under Misclassification of Treatment Assignment [‡]

Peter Anderson[§]

October 13, 2025— latest version

Abstract

Breastfeeding promotion and support is a core service of the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), which serves approximately half of all U.S. infants; however, program benefits include vouchers for infant formula, encouraging substitution away from breastfeeding among participating mothers. To estimate WIC's effect on breastfeeding, I formalize a partial identification approach to address a common challenge in regression discontinuity designs: misclassification of treatment assignment due to measurement error in the running variable. Leveraging the program's income eligibility cutoff, I apply this approach to bound the local average treatment effect (LATE) on the initiation and duration of breastfeeding. Using data from the Pregnancy Risk Assessment Monitoring System (PRAMS), I find that participation in WIC reduces the probability of initiating breastfeeding by at least 3% and the duration of breastfeeding by at least 1 week. My findings are stable across alternative specifications and robust to more conservative identifying assumptions.

Keywords: WIC, Breastfeeding, Partial-Identification, PRAMS
JEL Codes: C14, I18, I38

[†]This manuscript was supported by the Research Innovation and Development Grants in Economics (RIDGE) Partnership, with funding from the U.S. Department of Agriculture (USDA) Economic Research Service and Food and Nutrition Service. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the USDA.

[‡]I am especially grateful to my committee members John Pepper, Leora Friedberg, Lee Lockwood, and Derek Wu; to Carolina Caetano, Désiré Kédagni, Adam Rosen, Yuya Sasaki, and George Zuo for valuable conversations; and to participants in the University of Virginia Applied Micro/Public Workshop for helpful comments. Earlier versions were presented at the 2025 Midwest Economics Association Annual Meeting and the 2024 Association for Public Policy Analysis and Management (APPAM) Fall Research Conference. All errors are my own.

[§]University of Virginia; pa7kd@virginia.edu.

1 Introduction

The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)—the third largest federally funded nutrition program—serves approximately half of all infants in the United States and 6.3 million participants monthly. While the program improves birth outcomes and maternal health during pregnancy (Carlson and Senauer, 2003; Bitler and Currie, 2005; Kreider et al., 2020; Figlio et al., 2009; Hoynes et al., 2011), its effects on infant feeding practices remain unclear. Mandated to promote breastfeeding and support childhood nutrition, WIC provides both breastfeeding support services and quantity vouchers for infant formula to participating mothers. The existing academic literature provides no consensus on the program's net effect on breastfeeding: studies have found no effect (Jiang et al., 2010; Topolyan and Xu, 2017), substantial decreases (Ryan and Zhou, 2006; Bullinger and Gurley-Calvez, 2016), and even increases in breastfeeding (Joyce et al., 2008; Chatterji and Brooks-Gunn, 2004). Identifying WIC's causal effect poses a methodological challenge due to the non-random nature of program participation and limitations in the available survey data.

I examine how WIC participation affects breastfeeding initiation and duration among new mothers. Understanding these effects is critical due to the well-documented health benefits associated with breastfeeding for both infants and mothers, including reduced risks of childhood obesity, Type II diabetes, and certain cancers (Dieterich et al., 2013; Belfield and Kelly, 2012).[1] Among income-eligible mothers, WIC participants are less likely to initiate breastfeeding (88% vs. 92%) and, conditional on initiation, breastfeed for shorter durations (11.8 vs. 13.4 weeks) than non-participants. With WIC's broad reach, even modest changes in breastfeeding behavior may have substantial public health implications. In 2018, subsidies for infant formula accounted for 49.1% (18.2%) of the program's \$4.5 (\$2.8) billion of spending on food assistance before (after) manufacturer rebates (Kline et al., 2020). The potential for WIC's formula subsidies to discourage breastfeeding—through reducing the relative price of formula—may undermine the program's nutritional objectives (Rose et al., 2006; Joyce et al., 2008).

Two problems have confounded previous attempts to identify WIC's causal effects on breastfeeding. First, participation in WIC is endogenous: unobserved infant feeding preferences affect a mother's decision to enroll in WIC. Mothers with stronger preferences for formula feeding may be more likely to enroll in the program precisely because it provides formula subsidies, and as a result, simple comparisons of observed differences between participants and non-participants offer a biased measure of the program's effect. Second, few datasets contain detailed measures of infant feeding outcomes, and the existing datasets are either small, retrospective, or suffer from

[1] While numerous observational studies document positive associations between breastfeeding and later health outcomes, the causal interpretation of these relationships is debated. See Oster (2013) for a popular discussion.

measurement error in key variables. My primary data source, the Pregnancy Risk Assessment and Monitoring System (PRAMS), is large and collected contemporaneously with infant feeding decisions but suffers from one important limitation common to survey data: household income is reported in rounded brackets, creating uncertainty regarding eligibility near WIC's income cutoff, set nationally at 185% of the federal poverty level (FPL).[2] This bracketed reporting complicates empirical designs, such as regression discontinuity designs (RDD), that rely on precise measurement of income and treatment assignment to leverage policy discontinuities as a source of local exogeneity, motivating the need for an alternative identification strategy.[3]

I extend the regression discontinuity framework by developing a novel partial-identification approach that accommodates mismeasurement in the running variable and misclassification in treatment assignment, enabling estimation of the local average treatment effect (LATE) of WIC participation on breastfeeding outcomes. My approach mirrors traditional regression discontinuity designs (RDD), using WIC's income eligibility threshold as a discontinuity in program participation. In contrast to conventional designs, my method requires neither precise knowledge of the assignment variable, continuity of potential outcomes at the threshold, nor locally exogenous treatment assignment. Instead, I impose minimal and empirically supported shape restrictions of two flexible forms: (1) sign restrictions and (2) magnitude restrictions on the slope of the conditional expectation function. My insight is that observations far from the threshold—where treatment status is known with certainty despite bracketed income—can be used to bound treatment effects near the threshold, where classification is uncertain. By imposing restrictions on the slope of conditional expectation functions across income, I extrapolate from regions with certain assignment to the local neighborhood around the cutoff. The resulting estimates are transparent, robust, and credible, even under non-classical measurement error. This framework is well suited to applications where eligibility is measured with error or only partially observed, making it broadly applicable to the evaluation of means-tested programs using survey data.

To validate and contextualize my findings, I use the midpoint of each household's reported income bracket as a proxy for true income and implement two conventional RDD estimators as benchmarks. First, I estimate a traditional fuzzy regression discontinuity design, which uses the 185% FPL eligibility cutoff as an instrument for WIC participation. Second, I estimate a donut RDD, which excludes observations near the threshold in order to mitigate the influence of income misreporting. Both designs rely on precise measurement of the running variable, continuity of

[2] Two commonly used alternatives—the National Health and Nutrition Examination Survey (NHANES) and the WIC Infant and Toddler Feeding Practices Study (ITFPS)—are similarly limited. NHANES is small and retrospective, sampling only 5,000 individuals nationwide, and collects infant feeding data only for children under age five. ITFPS, while more detailed, samples only infant WIC participants, precluding comparison across program participation

[3] Even when income is reported as a continuous value, misreporting or recall error can lead to misclassification of treatment assignment near program eligibility thresholds.

counterfactual outcomes, and locally exogenous treatment assignment—assumptions that may be violated in this setting due to the bracketed nature of income reporting.

My shape-restricted partial-identification estimates provide evidence that WIC participation reduces both the initiation and duration of breastfeeding. Under weak and empirically grounded assumptions, I find that breastfeeding initiation declines by at least 3 percentage points, and duration falls by at least 1 week. The estimated effects represent 60% and 44% of the observed difference between WIC participants and income-eligible non-participants for breastfeeding initiation and duration, respectively. These bounds are robust to a range of slope restrictions and remain negative even under conservative assumptions about the rate of change in outcomes across income levels. To assess the sensitivity of my findings, I explicitly model household income uncertainty under a uniform distribution within reported income brackets. The resulting simulation-based estimates fall within my partial-identification bounds, providing additional support for the credibility of my approach.

In contrast, the conventional RDD estimates rely upon stronger identifying assumptions and exhibit considerable sensitivity to modeling choices. These designs serve as useful benchmarks, but my partial identification approach offers clear improvements over both. Leveraging minimal shape restrictions, I credibly identify the sign and magnitude of the treatment effect in the presence of income mismeasurement and treatment assignment misclassification. Despite stronger identifying assumptions, the conventional RDD estimates do not consistently identify the sign of LATE—particularly at narrower bandwidths, where local exogeneity is most plausible but precision is lowest.

My partial-identification approach contributes to the literature on inference under imperfect data, particularly regression discontinuity designs (RDD) affected by misclassified treatment assignment and measurement error in the running variable. While recent advances address various measurement issues in regression discontinuity designs—including heaping bias, discrete running variables, and rounding error—the specific problem of misclassified treatment assignment remains distinct and underexplored (Dong, 2015; Barreca et al., 2016; Kolesár and Rothe, 2018). A key limitation of existing methods is their reliance on correctly observed treatment status, which does not hold in many applied contexts with survey data. Standard polynomial selection methods are ill-suited for settings with error-prone running variables (Gelman and Imbens, 2019; Imbens and Wager, 2019; Pei et al., 2022), and existing solutions rely upon dropping misclassified observations, modeling manipulation patterns, or leveraging auxiliary datasets (Dong, 2015; Dong and Kolesár, 2023; Gerard et al., 2020; Davezies and Le Barbanchon, 2017). In my setting, none of these solutions are viable: income is reported in broad intervals; mismeasurement is bidirectional; and there are no comparable auxiliary data with which to validate assignment and treatment.

Building on the framework of Manski and Pepper (2018), who introduced bounded-variation assumptions for difference-in-differences designs, I develop a shape-restricted partial-identification method tailored to RDDs with interval-reported running variables and misclassified treatment assignment. This approach offers a middle ground between the strong continuity and local randomization assumptions of traditional RDD and the conservatism of worst-case nonparametric bounds. My method is specifically designed to address the challenges of fuzzy designs with mismeasurement of treatment assignment, offering applied researchers a transparent and tractable framework for bounding treatment effects when point identification is not credible.

This paper is organized as follows. Section 2 describes the institutional details of WIC eligibility and benefit structure, and introduces the PRAMS data. Section 3 formalizes the identification challenges arising from interval-reported income and presents my shape-restricted partial-identification approach, including the theoretical assumptions and supporting evidence. Section 4 details the implementation of my empirical strategy, beginning with the partial-identification method and comparing it to two conventional benchmarks: a traditional RDD and a donut RDD. Section 5 presents my main findings, with a particular focus on my partial-identification bounds under varying slope restrictions and a simulation-based sensitivity analysis. Section 6 concludes the paper.

2 Institutional Setting & Data

2.1 Institutional Setting

The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) is a federal assistance program that aims to improve the health and nutrition of low-income pregnant and postpartum women, infants, and children under age five. In 2021, WIC served an average of 6.3 million participants per month, including approximately half of all infants born in the United States.[4] Eligibility requires: (1) household income below 185% of the Federal Poverty Level (FPL), (2) categorical eligibility as a pregnant or postpartum woman, infant, or child under five, (3) documented nutritional risk, and (4) state residency.[5] Participants in SNAP, Medicaid, or TANF are considered adjunctively income-eligible and are exempted from WIC's standard income requirement, enabling participation at higher incomes. Although cross-program participation is substantial—in 2020, 74% of WIC participants were enrolled in Medicaid and 28% in SNAP—only 3.5% reported household incomes above the program's standard income eligibility threshold of 185% of the Federal Poverty Level (United States Department of Agriculture, Food and Nutrition Service, 2022b).

[4] WIC Program Overview (https://www.ers.usda.gov/topics/food-nutrition-assistance/wic-program/)

[5] WIC Eligibility (https://www.fns.usda.gov/wic/about-wic)

WIC provides a combination of nutrition education, referrals to public assistance programs and healthcare providers, and supplemental food packages. Food benefits include quantity vouchers—restricted to specific brands, package sizes, and other characteristics that vary by state—for designated food items, as well as cash-value vouchers for fresh produce (Table 1). For infants aged 0 to 6 months, food packages consist exclusively of infant formula, tailored to fully meet caloric needs. Between 6 and 12 months of age, packages introduce age-appropriate solid foods alongside formula. Throughout the first year of life, the quantity of infant formula provided is determined by the mother's self-reported feeding practice: *fully formula feeding*, *partially breastfeeding*, or *fully breastfeeding*. Although food packages may be adjusted if feeding practices change, such transitions are biologically constrained, and the cessation of breastfeeding is often irreversible due to declining milk production.[6]

WIC is federally mandated to promote and support breastfeeding as the preferred method of infant feeding. In line with this mandate, WIC provides educational programs and offers "anticipatory guidance, counseling, and breastfeeding educational materials" to participating pregnant mothers.[7] Postpartum, additional support includes: (1) education and counseling through group classes or individual sessions; (2) breastfeeding supplies, such as breast pumps and storage containers; and (3) peer counseling from experienced mothers (National WIC Association, 2019). The 2009 Food Package Revision—occurring before the period of my analysis—was designed, in part, to strengthen pecuniary incentives for breastfeeding. Following the reform, fully breastfeeding mothers became eligible for more comprehensive food benefits and extended postpartum participation—up to one year, compared to six months for mothers who do not breastfeed.

Despite evidence that the 2009 reforms increased both the initiation and exclusive duration of breastfeeding among WIC participants (Whaley et al., 2012), overall breastfeeding rates remain low (Table 2). Only 22% of WIC infants are still breastfeeding at 6 months, well below the CDC's 12-month recommendation. The high cost of formula feeding—estimated at $1600 to $2100 during the first year—makes infant formula the most financially valuable component of the WIC food packages (Centers for Disease Control and Prevention, 2023; United States Department of Agriculture, Food and Nutrition Service, 2022a).[8] By providing infant formula in-kind, the program reduces its relative price, thereby creating price incentives that discourage breastfeeding initiation and encourage earlier cessation. Across all levels of household income, WIC participants are less likely to initiate breastfeeding and breastfeed for shorter durations than income-eligible non-participants (Panel B of Figure 2).

[6]Regular milk expression is required to maintain supply and extended interruptions may result in permanent supply reduction.

[7]Breastfeeding Priority Programs (https://www.fns.usda.gov/wic/breastfeeding-priority-program)

[8]U.S. Surgeon General's U.S. 2011 Call to Action to Support Breastfeeding, adjusted to 2022 dollars (https://www.hhs.gov/surgeongeneral/reports-and-publications/breastfeeding/factsheet/index.html)

The documented association between early introduction of infant formula and adverse health outcomes—including childhood obesity, maternal obesity, breast cancer, and Type II diabetes—raises concerns about the health benefits and public value of WIC's in-kind provision of infant formula (Dieterich et al., 2013; Belfield and Kelly, 2012). While WIC participation has been shown to improve health outcomes among pregnant women (Carlson and Senauer, 2003; Bitler and Currie, 2005; Kreider et al., 2020; Figlio et al., 2009; Hoynes et al., 2011), there is limited evidence of comparable health benefits for infants. More than 30% of WIC infants and children exceed the 90th percentile for standardized growth measurements—weight-for-length (ages 0–2) and BMI (ages 2–4)—and the maternal obesity rate among participants has continued to rise (Figure 1). If these outcomes are causally linked to infant formula, the in-kind transfer may diminish the public value of program participation.

2.2 Data

This study uses data from the Pregnancy Risk Assessment and Monitoring System (PRAMS), a surveillance project conducted jointly by the Centers for Disease Control and Prevention (CDC) and participating state health departments. PRAMS collects state-specific, population-based data on maternal attitudes, behaviors, and experiences before, during, and shortly after pregnancy.[9] Mothers are surveyed 2–6 months postpartum, with annual state-level sample sizes ranging from approximately 1,000 to 3,000 respondents. The dataset links birth certificate records with survey responses collected through personalized mailing packets, use of incentives and rewards, and telephone follow-up for mail non-respondents (Shulman et al., 2018). To ensure adequate representation of key subpopulations, participating states employ stratified sampling that oversamples "mothers of infants with low-birthweight, those living in high-risk geographic areas, and racial/ethnic minority groups" (Shulman et al., 2018).

My analysis focuses on PRAMS Phases 6–8 (2009–2022), a period selected for two key reasons: the consistent structure of the WIC infant food packages following its October 2009 revision and the systematic collection of WIC participation data. The PRAMS questionnaire includes core questions asked across all participating states (55–60% of content), standardized CDC-developed questions, and state-specific additions (Shulman et al., 2018). From an initial sample of 508,387 mothers with live births, I construct the analytical sample by excluding observations with missing values for any of four key variables: breastfeeding initiation, breastfeeding duration, WIC participation during pregnancy, and household income relative to the Federal Poverty Level (FPL). This yields a final sample of 422,537 observations.[10] I next describe sample characteristics and the

[9] PRAMS Overview (https://www.cdc.gov/prams/index.htm). State participation in PRAMS is voluntary, and as of 2023, California, Idaho, North Carolina, and Ohio are the only states not participating.

[10] Sample sizes with complete data for individual variables are: breastfeeding initiation (484,964), breastfeeding

construction of the key variables, including measures of household income and WIC eligibility, breastfeeding initiation and duration, and program participation status.

Sample Characteristics Table 2 summarizes the characteristics of the analytical sample by WIC participation status among income-eligible households. In addition to differences in infant feeding practices, WIC participants differ systematically from eligible non-participants along several socioeconomic dimensions. WIC mothers are more likely to have low educational attainment (24% have less than a high school degree, compared to 13% of non-participants), less likely to be married (37% versus 58%), and report lower household incomes on average ($16,672 versus $24,748). These differences highlight the challenge of selection bias in identifying causal effects and suggest that simple mean comparisons may overstate WIC's negative impact on breastfeeding, as participants disproportionately possess characteristics associated with lower breastfeeding rates.

Income PRAMS reports income in rounded brackets that vary in both width and number across survey phases. Each bracket defines upper and lower bounds for household income, with the highest income bracket left open-ended (e.g., income > $100,000). I combine these brackets with household size to calculate upper (u_i) and lower (ℓ_i) bounds on each respondent's household income as a percentage of the Federal Poverty Level (FPL), adjusted for state- and year-specific thresholds.

For my analysis, I use the midpoint of these bounds as a proxy for true FPL: $X_i = \frac{u_i+\ell_i}{2}$. Using this measure, the estimated density of household income displays substantial and abrupt variation, reflecting the coarseness of the bracketed income reporting (Panel A of Figure 2). The average bracket width, as a percentage of FPL, is approximately 33 percentage points. Importantly, for 94% of observations with an income proxy within 10 percentage points of the eligibility threshold ($|X_i - 185| \in [0, 10]$), the upper and lower bounds straddle the 185% eligibility threshold, rendering treatment assignment uncertain (Figure 3). The distribution of incomes within these bounds is unobserved, and this measurement error is likely to attenuate the identifying variation in WIC participation near the income eligibility threshold.

Outcomes of Interest I examine two primary measures of infant feeding: the initiation and duration of breastfeeding. Initiation, a binary indicator of whether a mother attempts to breastfeed, is recorded by both birth certificate data and the PRAMS core questionnaire throughout my sample period. As initiation is documented at the time of birth, rather than solely through retrospective self-report, the risk of misclassification is minimal and unlikely to reflect recall or social desirabil-

duration (487,002), WIC participation (505,976), and household income (450,785). The analytical sample requires non-missing values for all four variables.

ity bias. For observations with missing initiation but non-missing duration, I impute initiation by coding mothers who breastfed for at least one week as having initiated breastfeeding.[11]

Breastfeeding duration, defined as the number of weeks a mother breastfeeds or pumps breast milk, is recorded in the PRAMS core questionnaire. While duration is self-reported, the PRAMS survey is administered shortly after delivery, typically within two to six months postpartum, limiting the risk of recall error relative to longer retrospective surveys. I additionally construct a measure of the exclusive duration of breastfeeding, defined as the number of weeks until the introduction of liquids other than breast milk or solid foods, using responses on the timing of these introductions.[12] Measurement of duration is right-censored by the age of the infant at the time of the survey, creating heterogeneous censoring across observations.

Table 2 and Panel B of Figure 2 reveal significant differences in infant feeding practices by WIC participation among income-eligible households (Income $\leq 185\%$ FPL). WIC participants initiate breastfeeding at lower rates (88% versus 92%), and conditional on initiation, breastfeed for shorter durations (11.84 weeks versus 13.38 weeks) than eligible non-participants. Exclusive breastfeeding duration follows a similar pattern, with WIC participants introducing non-breastmilk liquids earlier (11.52 versus 12.68 weeks). These differences may reflect both selection into WIC and potential treatment effects, motivating the need for causal identification.

The descriptive evidence reveals a clear positive relationship between household income and breastfeeding. On average, both the initiation and duration of breastfeeding increase with income (Figures 4 and 5). This pattern persists when excluding households with indeterminate treatment assignment, those with FPL bounds straddling the income eligibility threshold. Within eligibility groups, the income-breastfeeding relationship appears approximately linear, although slopes may differ between eligible and ineligible households, suggesting heterogeneity in the income gradient across eligibility status.

WIC Participation I define treatment as participation in WIC during pregnancy, which is consistently documented in both birth certificate data and PRAMS questionnaires across all sample years. While some mothers may only participate postpartum or enroll their infants directly, data on postnatal participation is recorded in only a limited set of states.[13] This measurement limitation has two implications. First, I underestimate total WIC exposure by missing postpartum enrollment among mothers and infants. Second, my treatment variable overstates the share of prenatal participants, the period when breastfeeding promotion is most intensive. Together, these issues may lead

[11] The one-week cutoff used for imputation is a consequence of questionnaire data, which records duration in weekly increments.

[12] As an example, for a mother who reports initiating breastfeeding and introducing a liquid other than breastmilk, e.g. infant formula, at 8 weeks and solid food at 21 weeks, I record exclusive breastfeeding for 8 weeks.

[13] Between 2009 and 2022, only Illinois (Phase 6), Maine (Phases 6–8), and New Hampshire (Phase 8) collected data on postpartum and infant WIC participation.

me to underestimate the negative effects of participation in WIC on breastfeeding outcomes.

My measure of prenatal WIC participation benefits from administrative recording at the time of birth. Participation is drawn from the birth certificate rather than retrospectively self-reported, reducing the risk of misclassification commonly associated with survey-based data. In contrast, self-reported WIC participation in the CPS-ASEC yields markedly lower participation rates at all levels of income (see Figure 6), suggesting that birth certificate-based recording in PRAMS provides a more reliable measure of take-up. While my estimates remain robust to potential under-reporting of program participation, the timing of data collection and use of birth certificate records suggest that my WIC participation variable is comparatively reliable.

Household income and WIC participation have a negative relationship: as household income increases, average program participation decreases. Figure 7 documents both this relationship and a distinct discontinuity at the 185% FPL eligibility threshold. The pattern and threshold discontinuity persist when varying bandwidth and excluding households with indeterminate treatment assignment—those whose income brackets straddle the eligibility threshold (Figures 7b, 7d, and 7f). Within eligibility groups, the relationship appears approximately linear, although the slope differs between eligible and ineligible households. External validation using the Current Population Survey Annual Social and Economic Supplement (CPS-ASEC) confirms this approximately linear negative gradient between household income and WIC participation (Flood et al., 2022), supporting my findings (Figure 6).

3 Identification

This section formalizes my partial-identification framework and outlines the threats to identification posed by measurement error in the running variable, an issue increasingly examined in the economic literature (Barreca et al., 2011, 2016; Dong and Kolesár, 2023). I impose theoretically and empirically grounded assumptions to identify expected treatment take-up and mean outcomes. Under these assumptions, the local average treatment effect (LATE) is partially identified.

3.1 Identification Conundrum

I am interested in identifying the causal effect of WIC participation ($T \in \{0, 1\}$) on breastfeeding outcomes (Y). Following the potential outcomes framework, I observe treatment status $T = T_Z$, where $Z \in \{0, 1\}$ denotes treatment assignment. The observed outcome is given by

$$Y = Y_0 + T_Z(Y_1 - Y_0)$$

where Y_0 and Y_1 are the potential outcomes for the treated and untreated, respectively. Much of the existing academic literature focuses on estimating the average treatment effect (ATE), defined as ATE $= \mathbb{E}[Y_1 - Y_0|W]$, conditional on covariates W. However, even after conditioning on observables, unobserved heterogeneity in infant feeding preferences may influence enrollment in WIC, introducing selection bias and rendering ATE estimates non-causal.

The income eligibility threshold for WIC, set nationally at 185% of FPL, provides a natural source of exogenous variation in program eligibility, offering a solution to the problem of selection endogeneity. Let $X^\star$ denote a household's true income relative to the threshold, centered so that the cutoff occurs at zero. Eligibility for WIC is then defined as

$$Z = \mathbb{1}[X^\star \leq 0]$$

Following Dong and Kolesár (2023), I define the conditional probability of treatment under assignment $z \in \{0, 1\}$ as $p^\star_z(x) := \mathbb{E}[T_z|X^\star = x]$ and the observed treatment probability as

$$p^\star(x) := \mathbb{E}[T|X^\star = x] = \mathbb{1}[x > 0]p^\star_0(x) + \mathbb{1}[x \leq 0]p^\star_1(x)$$

Similarly, let $g^\star_t(x) := \mathbb{E}[Y_t|X^\star = x]$ denote the conditional mean of the potential outcomes, where $t \in \{0, 1\}$ denotes treatment, and define the conditional mean of the observed outcomes as

$$g^\star(x) := \mathbb{E}[Y|X^\star = x] = g^\star_0(x) + p^\star(x)[g^\star_1(x) - g^\star_0(x)]$$

Under the canonical assumptions of the fuzzy regression discontinuity design,

(i) Monotonicity: $\mathbb{P}(T_1 \geq T_0|X^\star = 0) = 1$, with $\mathbb{P}(T_1 > T_0|X^\star = 0) > 0$

(ii) Continuity: $p^\star_z(x)$ and $g^\star_t(x)$ are continuous at $x = 0$ for $z, t \in \{0, 1\}$

the local average treatment effect (LATE) is identified as:

$$\tau = \frac{\lim_{x\uparrow 0} g^\star(x) - \lim_{x\downarrow 0} g^\star(x)}{\lim_{x\uparrow 0} p^\star(x) - \lim_{x\downarrow 0} p^\star(x)}$$

PRAMS reports household income in categorical brackets rather than exact dollar amounts, precluding precise calculation of each household's income as a percentage of the Federal Poverty Level (FPL). This interval-based reporting induces measurement error in the running variable, as true household income is only known to fall within the reported bracket. As a result, I observe a noisy proxy $X = X^\star - e$, where e denotes measurement error. In my context, I define this proxy as the midpoint of each household's reported FPL interval. Regardless of the particular estimate

$X_i \in [\ell_i, u_i]$, the width of these brackets creates uncertainty about treatment assignment—that is, true eligibility status—for 32,928 observations near the eligibility threshold. Unlike the setting analyzed by Dong and Kolesár (2023), measurement error in this case not only affects the precision with which I observe the running variable but also misclassifies treatment assignment.

When using the noisy measure X in place of true income $X^\star$, I observe assignment

$$\tilde{Z} = \mathbb{1}[X \leq 0]$$

rather than true assignment

$$Z = \mathbb{1}[X^\star \leq 0]$$

for observations where X and $X^\star$ lie on opposite sides of the threshold. Under the potential outcomes framework, the observed conditional means based on X are given by:

$$\tilde{p}(x) := \mathbb{E}[T \mid X = x] = \int_{-\infty}^{\infty} p^\star(x^\star) f_{X^\star|X}(x^\star|x) dx^\star$$
$$\tilde{g}(x) := \mathbb{E}[Y \mid X = x] = \int_{-\infty}^{\infty} g^\star(x^\star) f_{X^\star|X}(x^\star|x) dx^\star$$

The corresponding estimator using the noisy running variable is:

$$\tilde{\tau} = \frac{\lim_{x\uparrow 0} \tilde{g}(x) - \lim_{x\downarrow 0} \tilde{g}(x)}{\lim_{x\uparrow 0} \tilde{p}(x) - \lim_{x\downarrow 0} \tilde{p}(x)}$$

However, in this context, the conventional regression discontinuity estimator $\tilde{\tau}$ does not generally recover the true LATE. The discontinuities in $\tilde{g}(x)$ and $\tilde{p}(x)$ at $X = 0$ reflect an unobserved, weighted average of income-eligible and ineligible households. The resulting bias depends both on the degree of misclassification and on differences in potential outcomes at the threshold. Without knowledge of these misclassification probabilities, which depend on the distribution of measurement error e near the threshold, neither the direction nor magnitude of bias is identified. Thus, standard RDD estimators do not yield valid causal estimates under these conditions. This identification challenge motivates my partial-identification approach, which establishes bounds on the treatment effect without requiring assumptions on e.

3.2 Partial-Identification

Given measurement error in treatment assignment, I establish bounds on the treatment effect rather than pursuing a point estimate. My approach leverages observations with unambiguous treatment assignment—households whose reported income intervals, as a percentage of the Federal Poverty Level (FPL), lie entirely above or below the eligibility threshold—to bound effects at the threshold, where assignment is uncertain.

I focus on bounding the local average treatment effect (LATE), defined in terms of the potential outcomes Y_0 and Y_1. As described above, let $g^\star(x) := \mathbb{E}[Y \mid X^\star = x]$ denote the conditional mean of outcome Y, and let $p^\star(x) := \mathbb{E}[T \mid X^\star = x]$ denote the observed treatment probability. My partial-identification strategy imposes three assumptions:

Assumption 1 (Positive First Stage)**.** Income eligibility increases treatment take-up at the threshold:

$$\tau_T := \lim_{x \uparrow 0} p^\star(x) - \lim_{x \downarrow 0} p^\star(x) > 0$$

Assumption 2 (Monotonicity in Income)**.** For any $x^-, x^+ \in \mathbb{R}$ with $x^- \leq x^+$ such that $0 \notin [x^-, x^+]$:

$$\begin{aligned} &\text{(i)} \quad p^\star(x^-) \geq p^\star(x^+) \\ &\text{(ii)} \quad g^\star(x^-) \leq g^\star(x^+) \end{aligned}$$

Assumption 3 (Bounded Rate of Change)**.** Let $\mathcal{R}_1 = [-\delta_1, 0]$ and $\mathcal{R}_0 = (0, \delta_0]$ denote neighborhoods on the eligible and ineligible sides of the threshold, respectively. For $z \in \{0, 1\}$ and all $x \in \mathcal{R}_z$:

$$\begin{aligned} &\text{(i)} \quad \left| \frac{dp^\star(x)}{dx} \right| \leq \lambda_z \\ &\text{(ii)} \quad \left| \frac{dg^\star(x)}{dx} \right| \leq \kappa_z \end{aligned}$$

where $\kappa_z, \lambda_z > 0$ are assignment-specific bounds, and $\delta_z > 0$ defines the size of the neighborhood.

These assumptions are grounded in economic theory and supported by empirical regularities in the data. Assumption 1, the first-stage assumption, reflects that removing a constraint on program participation, in this case moving from income ineligibility to eligibility, should weakly increase participation rates. Empirically, Figure 7 shows a discontinuous increase in WIC participation at the eligibility threshold, with consistently higher participation rates observed among income-eligible households. This pattern is consistent with administrative data and prior analyses: WIC Participant Characteristics Reports and survey-based analyses document that program take-up is concentrated below the 185% FPL threshold, with participation rates declining steadily as income rises (United States Department of Agriculture, Food and Nutrition Service, 2022b; Hoynes and Schanzenbach, 2015).

Assumption 2 imposes monotonicity in income for both treatment take-up and outcomes. As income rises, program benefits represent a decreasing share of household resources, decreasing the value of participation. Figure 7 shows decreasing rates of participation as income increases and a discontinuous decline at the eligibility threshold. With respect to infant feeding practices, mothers in higher-income households typically have greater access to paid leave, flexible work arrangements, professional lactation support, and familial resources that serve to facilitate breastfeeding. Existing research documents strong socioeconomic gradients in breastfeeding behaviors, with higher-income and higher-education mothers initiating and sustaining breastfeeding at higher rates (American College of Obstetricians and Gynecologists, 2021; Diaz et al., 2023). Figures 4 and 5 show clear upward trends in both initiation and duration of breastfeeding as household income increases, both across the income distribution and within assignment groups.

Assumption 3 places limits on the rate at which these conditional expectations may change with household income. This reflects that changes in participation and breastfeeding behaviors are driven by structural and socioeconomic forces—program eligibility, workplace policies, and access to information—that evolve smoothly with income. These determinants, such as paid leave, flexible work arrangements, health literacy, and social support, are closely tied to income and accumulate gradually across the income distribution (American College of Obstetricians and Gynecologists, 2021; Diaz et al., 2023). Figures 7, 4, and 5 display smooth relationships without abrupt changes or irregularities, and within-group patterns appear approximately linear. Figure 8 provides further empirical support, showing that race, access to paid leave, return-to-work timing, and maternal education—-key determinants of both WIC participation and breastfeeding behaviors—evolve smoothly with household income. The gradual evolution of these characteristics across income levels provides empirical support for the bounded-slope restriction in local neighborhoods.

Construction of Bounds

Under the maintained assumptions of a positive first stage, monotonicity in income, and bounded rates of change in the conditional expectation functions, I derive bounds on the discontinuities in treatment take-up (τ_T) and outcomes (τ_Y) at the eligibility threshold. These bounds rely on observations with unambiguous treatment assignment and partially identify the local average treatment effect (LATE). Define the left and right limits at the threshold as:

$$g^\star(0^-) := \lim_{x\uparrow 0} g^\star(x), \quad g^\star(0^+) := \lim_{x\downarrow 0} g^\star(x)$$
$$p^\star(0^-) := \lim_{x\uparrow 0} p^\star(x), \quad p^\star(0^+) := \lim_{x\downarrow 0} p^\star(x)$$

so that $\tau_Y = g^\star(0^-) - g^\star(0^+)$ and $\tau_T = p^\star(0^-) - p^\star(0^+)$.

Household income is observed in intervals $[\ell, u]$. For households whose brackets lie entirely below the threshold ($u \leq 0$), treatment assignment is known with certainty: $Z = 1$. Likewise, for households with brackets entirely above the threshold ($\ell > 0$), treatment assignment is likewise certain: $Z = 0$. A household's true income $X^\star$ is unobserved but bounded: $X^\star \in [\ell, u]$.

Let x^- denote any point in an eligible bracket (where $u < 0$), and x^+ denote any point in an ineligible bracket (where $\ell > 0$). Under my shape restrictions, observations at x^- provide bounds for the left limits $g^\star(0^-)$ and $p^\star(0^-)$, and observations at x^+ provide bounds for the right limits $g^\star(0^+)$ and $p^\star(0^+)$.

For certainly eligible households ($u < 0$), Assumptions 2 and 3 imply:

$$
\begin{aligned}
p^\star(0^-) &\in [p^\star(x^-) - \lambda_1 \cdot |x^-|, \quad p^\star(x^-)] \\
g^\star(0^-) &\in [g^\star(x^-), \quad g^\star(x^-) + \kappa_1 \cdot |x^-|]
\end{aligned}
$$

Similarly, for certainly ineligible households ($\ell > 0$):

$$
\begin{aligned}
p^\star(0^+) &\in [p^\star(x^+), \quad p^\star(x^+) + \lambda_0 \cdot x^+] \\
g^\star(0^+) &\in [g^\star(x^+) - \kappa_0 \cdot x^+, \quad g^\star(x^+)]
\end{aligned}
$$

I obtain bounds on discontinuities τ_T and τ_Y by considering the extreme cases:

$$
\begin{aligned}
\tau_T &\in \Big[p^\star(x^-) - p^\star(x^+) - \lambda_0 \cdot x^+ - \lambda_1 \cdot |x^-|, \quad p^\star(x^-) - p^\star(x^+)\Big] \\
\tau_Y &\in \Big[g^\star(x^-) - g^\star(x^+), \quad g^\star(x^-) - g^\star(x^+) + \kappa_1 \cdot |x^-| + \kappa_0 \cdot x^+\Big]
\end{aligned}
$$

These bounds on the treatment and outcome discontinuities partially identify the local average treatment effect (LATE). For the first stage, I obtain bounds $[\tau_T^{LB}, \tau_T^{UB}]$ on the discontinuity in treatment take-up, which is constrained to be positive by Assumption 1. For the outcome, I similarly obtain bounds $[\tau_Y^{LB}, \tau_Y^{UB}]$ on the discontinuity in the outcome—either the initiation or duration of breastfeeding. The ratio of these bounds yields an interval for LATE:

$$
\tau = \begin{cases}
\left[\dfrac{\tau_Y^{LB}}{\tau_T^{UB}}, \dfrac{\tau_Y^{UB}}{\tau_T^{LB}}\right] & \text{if } \tau_Y^{LB} \geq 0 \\[2em]
\left[\dfrac{\tau_Y^{LB}}{\tau_T^{LB}}, \dfrac{\tau_Y^{UB}}{\tau_T^{UB}}\right] & \text{if } \tau_Y^{UB} < 0 \\[2em]
\left[\dfrac{\tau_Y^{LB}}{\tau_T^{LB}}, \dfrac{\tau_Y^{UB}}{\tau_T^{LB}}\right] & \text{if } \tau_Y^{LB} < 0 < \tau_Y^{UB}
\end{cases}
$$

The applicable case depends on the sign of the outcome bounds. When $\tau_Y^{LB} \geq 0$, both numerator bounds are positive, and dividing the lower bound by the largest denominator (τ_T^{UB}) and the upper bound by the smallest denominator (τ_T^{LB}) yields the narrowest valid interval. When $\tau_Y^{UB} < 0$, both numerator bounds are negative, and the relationship is reversed. When the outcome bounds straddle zero, the treatment effect is unsigned, and I conservatively divide by the smallest denominator (τ_T^{LB}) in both expressions to ensure a valid identification region.

My approach departs from the traditional regression discontinuity design (RDD) assumption of exogeneity, which requires that households cannot precisely manipulate their income to fall just below the eligibility threshold. Instead, I exclude observations near the threshold, where manipulation is most likely, and restrict attention to households with unambiguous treatment assignment. Households capable of sorting to just below the threshold in response to incentives are thus omitted from my identification strategy. By imposing shape restrictions on the conditional expectation functions (CEF), I extrapolate from manipulation-free regions to bound treatment effects at the threshold. In so doing, I trade the local randomization assumption of the standard RDD for smoothness and monotonicity restrictions on the CEFs. As long as the first stage is positive, the local average treatment effect (LATE) is well-defined, and the resulting bounds remain valid—even in the presence of sorting around the eligibility threshold. Unlike conventional RDD approaches, my method does not require dense support or local randomization near the eligibility threshold. Identification is achieved through shape restrictions rather than localized comparisons, and my bounding procedure avoids reliance on narrow bandwidths and remains valid even with sparsely populated neighborhoods around the cutoff.

4 Empirical Strategy

This section outlines my empirical strategy for estimating the causal effect of WIC participation on breastfeeding outcomes. I begin with my primary partial-identification approach, which bounds the local average treatment effect (LATE) under flexible, empirically supported assumptions. To benchmark these results, I implement two conventional designs: (1) a traditional fuzzy regression discontinuity design (RDD) and (2) a donut RDD that excludes observations near the eligibility threshold. Both conventional approaches rely on stronger identification assumptions, including local exogeneity and precise classification of treatment assignment. All approaches use a common income proxy, $X_i = (\ell_i + u_i)/2$, representing the midpoint of each household's reported FPL interval.

4.1 Partial-Identification RDD

To implement my partial-identification approach, I construct bounds using households with unambiguous treatment assignment—those whose income brackets lie entirely above or below the eligibility threshold. I exclude observations near the threshold, where measurement error creates uncertain eligibility, and restrict the estimation sample to households with known treatment assignment. I then extrapolate from these manipulation-free regions to bound treatment effects at the threshold.

4.1.1 Bin Construction and Reference Sets

My identification strategy relies on households with certain treatment assignment in order to bound conditional expectation functions $g^\star(x)$ and $p^\star(x)$ at the eligibility threshold. For each bandwidth h, I partition observations into bins of width h, defined over the observed income proxy X_i.

I then identify reference sets: the nearest bins on either side of the threshold for which treatment assignment is unambiguous for the average household. On the left (eligible) side of the threshold, I select the first bin for which the mean upper bound of reported income falls below $-h$:

$$\frac{1}{n_b}\sum_{i\in b} u_i < -h$$

On the right (ineligible) side, I select the first bin for which the mean lower bound of income exceeds h:

$$\frac{1}{n_b}\sum_{i\in b} \ell_i > h$$

These conditions ensure that the average income bracket in each reference bin lies outside the neighborhood $[-h, h]$ surrounding the threshold. I denote these reference bins as b_h^- and b_h^+, requiring each to contain at least 100 observations. Under Assumption 2, the mean outcomes in b_h^- and b_h^+ provide one-sided bounds on the conditional expectations over the regions $[-h, 0]$ and $(0, h]$ adjacent to the threshold, corresponding respectively to marginally income-eligible and -ineligible households.

4.1.2 Estimation Strategy

Having established one-sided bounds on the conditional expectation functions from the reference bins on either side of the threshold, I construct complementary bounds by combining observed bin means with constraints on the allowable rate of change in outcomes with respect to income. Specifically, I bound the functions $g^\star(x) = \mathbb{E}[Y \mid X^\star = x]$ and $p^\star(x) = \mathbb{E}[T \mid X^\star = x]$ separately for each side of the income eligibility threshold. Together with Assumption 2, these constraints yield upper and lower bounds on the conditional expectations over the regions $[-h, 0]$ and $(0, h]$ adjacent to the threshold.

For each outcome—WIC participation, breastfeeding initiation, and breastfeeding duration—I fit separate polynomial models for certainly eligible ($X_i^\star \leq c$) and certainly ineligible ($X_i^\star > c$) households. Let $x_i := X_i - c$ denote income centered at the cutoff. Using estimation windows with width equal to the sample mean FPL-bracket range, I estimate

$$Y_i \;=\; h_z(x_i) \;+\; \mathbf{W}_i'\boldsymbol{\beta}_z \;+\; \varepsilon_i, \qquad h_z(x) \;=\; \alpha_z \;+\; \sum_{k=1}^{p} \gamma_{z,k}\, x^k, \tag{1}$$

on the corresponding side of the threshold. Y_i is the outcome of interest, h_z is a polynomial function of order p, and $\mathbf{W}_i$ is a vector of maternal controls.[14] For participation and initiation, equation (1) is a linear probability model. For duration, I use the full sample for estimation and do not condition on initiation.

To operationalize Assumption 3, I separately estimate the rate of change in the conditional expectation functions with respect to income, $\frac{\partial h_z}{\partial X}$, over the intervals $[-\frac{1}{N}\sum_{i=1}^{N}(u_i - \ell_i), 0]$ for the income-eligible ($z = 1$) and $[0, \frac{1}{N}\sum_{i=1}^{N}(u_i - \ell_i)]$ for the ineligible ($z = 0$). I then scale these estimates by multipliers in $\{1, 1.5, 2.0\}$ to form plug-in estimators for λ_z and κ_z, which bound the rate of change in the participation and outcome functions. This approach assumes that average slope estimates across these regions provide a conservative upper bound on the rate of change between each reference bin and the threshold. I use these bounds to extrapolate from the observed means in each bin to the threshold, allowing deviations of up to $\lambda_z \cdot |x^{-/+}|$ for participation and $\kappa_z \cdot |x^{-/+}|$ for breastfeeding outcomes, where x^- and x^+ denote the midpoints of the reference bins for $z = 1$ and $z = 0$, respectively.

4.1.3 LATE Bounds and Implementation Details

I construct bounds on the discontinuities in treatment take-up (τ_T) and outcomes (τ_Y) by combining the upper and lower bounds on conditional expectations at the threshold, derived from the monotonicity and shape restrictions. Figure 9 provides a stylized visualization of this procedure.

Consistent with Assumption 1, I impose the requirement that $\tau_T^{LB} > 0$. When the data yield $\tau_T^{LB} \leq 0$, I do not report LATE bounds, as the first stage is too weak for credible identification.

[14] Appendix Section C.5 details a secondary estimation strategy using probit for the binary outcomes and tobit for duration. My main analysis relies on the polynomial model in (1).

For specifications that satisfy the first-stage bounds, I construct LATE bounds as:

$$\hat{\tau} = \begin{cases} \left[\dfrac{\hat{\tau}_Y^{LB}}{\hat{\tau}_T^{UB}}, \dfrac{\hat{\tau}_Y^{UB}}{\hat{\tau}_T^{LB}}\right] & \text{if } \hat{\tau}_Y^{LB} \geq 0 \\[2em] \left[\dfrac{\hat{\tau}_Y^{LB}}{\hat{\tau}_T^{LB}}, \dfrac{\hat{\tau}_Y^{UB}}{\hat{\tau}_T^{UB}}\right] & \text{if } \hat{\tau}_Y^{UB} < 0 \\[2em] \left[\dfrac{\hat{\tau}_Y^{LB}}{\hat{\tau}_T^{LB}}, \dfrac{\hat{\tau}_Y^{UB}}{\hat{\tau}_T^{LB}}\right] & \text{if } \hat{\tau}_Y^{LB} < 0 < \hat{\tau}_Y^{UB} \end{cases} \tag{2}$$

The appropriate structure is determined by the outcome bounds, which govern whether the LATE bounds span only positive values, only negative values, or both.

I implement this procedure across multiple specifications, varying both the bandwidth h and polynomial order p. For each specification, I re-estimate the reference bins, refit the participation and outcome models, and recompute the slope-based extrapolation limits using multipliers $\in \{1, 1.5, 2.0\}$. This allows me to examine the sensitivity of the bounds to different modeling and bandwidth choices, paralleling the specification search common in conventional regression discontinuity designs.

To quantify sampling uncertainty, I employ a bootstrap procedure with 1,000 replications. For each bootstrap sample, I re-run the complete analysis from bin selection through LATE bound construction. I report the 2.5th and 97.5th percentiles of the bootstrap distribution as 95% confidence intervals for the bounds. This approach accounts for the complex dependence structure in my multi-step estimation procedure.

4.2 Traditional and Donut Regression Discontinuity Designs

As a benchmark, I implement two conventional regression discontinuity designs (RDD): a standard fuzzy RDD and a donut RDD that excludes observations with ambiguous treatment assignment due to income bracketing. Both approaches exploit the discontinuity in WIC eligibility at 185% of the Federal Poverty Level (FPL) and use a common income proxy, $X_i = (\ell_i + u_i)/2$, the midpoint of each household's reported income bracket.

The standard fuzzy RDD estimates the local average treatment effect (LATE) using households near the income eligibility threshold, where program participation shifts discontinuously at the locally arbitrary cutoff. I instrument for endogenous participation in WIC T_i using income eligibility $Z_i = \mathbb{1}[X_i \leq c]$, where $c = 185$ denotes the eligibility threshold in FPL percentage points. Specifically, I estimate:

$$Y_i = \alpha + \tau T_i + f^+(X_i - c) + Z_i \cdot f^-(X_i - c) + \epsilon_i \tag{3}$$

where Y_i is the infant feeding outcome and f^+ and f^- are polynomial functions of order p defined on either side of the threshold, allowing the relationship between outcomes and income to vary flexibly for eligible and ineligible households.

Identification under this fuzzy RDD framework relies on two standard assumptions:

(i) Monotonicity: $\mathbb{P}(T_1 \geq T_0 | X = 0) = 1$ and $\mathbb{P}(T_1 > T_0 | X = 0) > 0$

(ii) Continuity: $p_z(x)$ and $g_t(x)$ are continuous at $x = 0$ for $z, t \in \{0, 1\}$

Monotonicity requires that income eligibility weakly increases participation for all households, with a strictly positive effect for some. Continuity requires that both the conditional probability of participation, $p_z(x) = \mathbb{E}[T_z \mid X = x]$, and conditional mean outcomes, $g_t(x) = \mathbb{E}[Y_t \mid X = x]$, are continuous at the threshold for $z, t \in \{0, 1\}$. Under these conditions, τ identifies the local average treatment effect (LATE) among compliers at the eligibility cutoff.

As a secondary benchmark, I implement a "donut" regression discontinuity design (RDD) that explicitly excludes observations with uncertain eligibility status. PRAMS reports household income in rounded brackets, making treatment assignment ambiguous for those households whose income brackets straddle the eligibility threshold. Let $\mu = \frac{1}{N}\sum_{i=1}^{N}(u_i - \ell_i)$ denote the average width of these brackets, where u_i and ℓ_i are the upper and lower bounds of household i's income bracket as a percentage of the Federal Poverty Level (FPL). In my sample, $\mu \approx 33$ percentage points.

For the donut approach, I exclude all observations within μ of the eligibility threshold on each side and estimate equation 3 using the subsample $|X_i - c| \in [\mu, \mu + h]$. While this approach reduces potential bias from misclassified treatment assignment, it strengthens the identifying continuity requirements. In contrast with the standard fuzzy RDD, which requires continuity of potential outcomes only at the threshold, the donut RDD assumes that conditional relationships estimated from observations outside the exclusion zone approximate the conditional expectation functions at the threshold. The benefit is eliminating bias from misclassified treatment assignment at the cost of relying on observations further from the discontinuity.

For both designs, I estimate equation 3 across a range of bandwidths h. For the standard RDD, I restrict the sample to observations with $|X_i - c| \leq h$, and for the donut RDD, I use observations $|X_i - c| \in [\mu, \mu + h]$. For each bandwidth-polynomial combination, I compute the Akaike Information Criterion (AIC) and select the specification that minimizes this criterion. While useful as benchmarks, these models rely on precise income measurement and local randomization, assumptions explicitly relaxed in my partial-identification framework.

5 Results

In this section, I present estimates of the local average treatment effect (LATE) of prenatal WIC participation on breastfeeding initiation and duration using a partial-identification framework designed for a rounded running variable and potential misclassification of treatment assignment. I first report bounds constructed from projections selected using AIC under empirically grounded restrictions, alongside alternative implementations of the slope constraint (Linear–Mean, Linear–Bin, and Nonlinear), then examine sensitivity to relaxed slope assumptions and bandwidth choice. I assess the credibility of my estimates with a simulation-based RDD that draws incomes uniformly within reported brackets and produces point-identified estimates under particular distributional assumptions. As a benchmark, I compare these results to conventional point estimates from traditional and donut RDDs estimated at comparable bandwidths.

5.1 Partial-Identification Regression Discontinuity

I begin with my partial identification approach, which directly addresses rounding in the running variable and misclassification of treatment assignment. Table 3 reports bounds on the local average treatment effect (LATE) for breastfeeding initiation and duration at 50%, 75%, 100%, 150%, and 200% of the MSE-optimal bandwidth (h_{MSE}), using the selector of Calonico et al. (2014). I impose Assumptions 1, 2, and 3 to extrapolate from households with unambiguous treatment assignment and construct bounds at the eligibility threshold. Estimates are derived from first- or second-order polynomial models fit to certainly eligible and certainly ineligible subsamples, with the polynomial order selected using the Akaike Information Criterion (AIC). For my main estimates, I implement the slope restriction (Assumption 3) using three approaches: (i) Linear–Mean Slope, (ii) Linear–Bin Slope, and (iii) a Nonlinear projection that allows curvature between the reference bin and the cutoff.[15] I report 95% empirical confidence intervals from 1,000 bootstrap draws in parentheses, shown only when the first-stage lower bound is positive in at least 50% of draws. Unlike conventional RDD methods, this approach does not require local continuity or randomization assumptions, thereby reducing reliance on narrow bandwidths near the threshold.

For breastfeeding initiation, the linearized implementations (Mean Slope and Bin Slope) yield uniformly negative, tightly estimated LATE bounds across the five MSE-selected bandwidths (Table 3, Panel A). Lower bounds lie between -0.29 and -0.14, and upper bounds between -0.07 and -0.04. Both the bounds and the associated 95% bootstrap confidence intervals are strictly

[15] Additional implementation details for each method are provided in Appendix C.3.

negative across the reported bandwidths. At these bandwidths, AIC selects a first-order (linear) projection used to implement Assumption 3, and as a result, the two linearized implementations coincide in the table. The Nonlinear projection widens the sets and is most conservative at the narrowest windows. Its 95% confidence interval includes zero at $h = 9.97$ and is only marginally negative at $h = 14.95$. At wider bandwidths, the nonlinear bounds remain negative and close to the linearized magnitudes.

For breastfeeding duration, the linearized implementations again produce strictly negative bounds and 95% bootstrap confidence intervals at each bandwidth (Table 3, Panel B). Lower bounds lie between -56.72 and -8.25, and upper bounds between -2.45 and -0.83. At wider bandwidths, the lower bound becomes markedly more negative while the upper bound remains near zero. The Nonlinear projection is more conservative and, at some bandwidths, yields 95% confidence intervals that straddle zero; however, the reported nonlinear bounds remain negative across bandwidths.

For both outcomes, the bounds are strictly negative across the reported bandwidths, and I find no evidence of positive effects on either the initiation or duration of breastfeeding. I next assess credibility with a set of sensitivity exercises. First, I construct a Simulation RDD that explicitly models bracket-level income uncertainty and provides a point-identified benchmark. Second, I relax the slope restriction using slope multipliers that scale the allowable rate of change in mean outcomes and treatment. Third, I conduct bandwidth sweeps from $0.5h_{\mathrm{MSE}}$ to $2.0h_{\mathrm{MSE}}$.

5.1.1 Simulation RDD

As a first sensitivity check, I implement a simulation-based regression discontinuity design (RDD) that explicitly models income uncertainty. For each household with reported income bracket $[\ell_i, u_i]$, expressed as a percentage of the Federal Poverty Level (FPL), I draw simulated values from a uniform distribution:

$$X^{\star}_{i,s} \sim \text{Uniform}(\ell_i, u_i),$$

where $s \in \{1, ..., 1000\}$ indexes the simulation draws. This procedure yields 1,000 datasets, each with point-identified income values drawn from the reported household-specific brackets.

For each simulated dataset s, I assign treatment eligibility $Z_{i,s} = \mathbb{1}[X^{\star}_{i,s} \leq c]$ and estimate equation 3 as a traditional fuzzy RDD. This yields 1,000 treatment effect estimates, $\{\hat{\tau}_s\}_{s=1}^{1000}$. I report the mean estimate, $\bar{\tau} = \frac{1}{1000}\sum_{s=1}^{1000}\hat{\tau}_s$, as the simulation-based LATE. The 95% confidence interval is defined by the 2.5th and 97.5th percentiles of the empirical distribution.

This approach retains all observations—including those nearest the threshold, which are most informative for identification—and explicitly accounts for uncertainty in treatment assignment. By drawing from the full support of each household's income bracket, it avoids the attenuation bias introduced by midpoint proxies. The primary limitation is the uniform distributional assumption. Although the true within-bracket distribution of income is unknown, the uniform distribution offers a transparent baseline that assigns equal probability to all values within the reported range. For each simulation, I select the optimal polynomial order using the Akaike Information Criterion to ensure consistency with my other designs.

Table 3 reports simulation-based local average treatment effects (LATE), 95% confidence intervals, and first-stage F-statistics for the initiation and duration of breastfeeding. For initiation, AIC-selected LATE estimates are uniformly negative and exhibit limited variation across bandwidths, with magnitudes ranging from -0.041 to -0.058. At narrower bandwidths, the 95% confidence intervals include zero, with statistically significant estimates at only the widest bandwidths. First-stage strength increases with h, and the mean first-stage F-statistic exceeds 10 across specifications, indicating sufficient instrument strength. Simulation-based estimates are contained within the partial-identification bounds across the majority of bandwidths examined, lending support to my main results under the assumption that income is uniformly distributed within each bracket.

For breastfeeding duration, simulation LATE estimates are again consistently negative and stable across bandwidths, with magnitudes ranging from -3.06 to -3.29 weeks. The 95% confidence intervals exclude zero at all reported bandwidths. First-stage strength is high across bandwidths, indicating adequate instrument strength. As with initiation, simulation-based point estimates fall within the partial-identification bounds, reinforcing the credibility of my main results.

5.1.2 Sensitivity to Bounded Rate of Change

I next examine the sensitivity of my main partial identification results to the bounded rate of change, Assumption 3. In my main specification, I impose that the absolute rate of change in the conditional expectation function (CEF) is bound above by the rate of change among observations with unambiguous treatment assignment. Approaching the income eligibility threshold from the left (right) reference bin, I restrict the maximum variation in mean outcomes using the estimated rate of change among the income eligible (ineligible). To relax this assumption, I scale the estimated slope of the CEF using multipliers of 1.5 and 2.0. I then estimate the local average treatment effect using my partial identification and bootstrap procedure with these new, more conservative restrictions on the rate of change. I implement this scaling procedure symmetrically, scaling both the rate of change in WIC participation and breastfeeding outcomes, and asymmetrically, scaling only outcomes.

For the symmetric procedure, I report bounds on the discontinuity in treatment and breastfeeding outcomes at the income eligibility threshold, respectively, the denominator and numerator of my LATE estimator (Table 4). Examining these discontinuities at the reported multiples of h_{MSE}, I find that scaling the slope disproportionately affects the bounds on WIC participation in comparison with those on breastfeeding initiation and duration. At a multiplier of 1.5, the bounds on the discontinuity in WIC participation contain zero, leaving the sign indeterminate for three of five bandwidths examined. For a multiplier of 2.0, I am unable to identify the sign at any bandwidth. In contrast, for both the initiation and duration of breastfeeding, bounds are uniformly negative at all bandwidths with a multiplier of 1.5 and remain negative at the majority of bandwidths even under a multiplier of 2.0.

For the asymmetric procedure, I report bounds on LATE and the associated 95% confidence intervals from my bootstrap procedure (Table 5). For both outcomes, I find that relaxing the restriction on the rate of change widens the partial identification bounds on LATE but does not preclude identification of the sign of the treatment effect. Even under a multiplier of 2.0, I find strictly negative bounds for all initiation bandwidths and for three of five duration bandwidths examined. At the highest multiplier, the 95% confidence intervals include zero at all initiation bandwidths and at all but two duration bandwidths.

5.1.3 Sensitivity to Bandwidth Choice

I next examine the sensitivity of my partial identification results to the bandwidth used for monotonicity, Assumption 2. In my main results, I construct reference bins with bandwidths corresponding to 50%, 75%, 100%, 150%, and 200% of the MSE-optimal bandwidth, h_{MSE}. The difference in mean WIC participation and mean breastfeeding outcomes between reference bins forms one side of the partial identification bounds. To complement this analysis, I extend the number of bandwidths examined, including additional bandwidths distributed between 50% and 200% of h_{MSE}. For completeness, I asymmetrically scale the slope using multipliers of 1.5 and 2.0, as detailed above.

Figure 10 presents bounds with 95% confidence intervals over these additional bandwidths. For baseline estimates, bounds are uniformly negative for both initiation and duration, and the 95% confidence intervals exclude zero over most of the sweep. Across this same range, my simulation-based regression discontinuity produces point estimates that generally fall within the partial identification bounds (Figure 11). Relaxing the restrictions on the bounded rate of change, the 95% intervals continue to exclude zero for a multiplier of 1.5. For a multiplier of 2.0, intervals frequently include zero and, in some windows, the bounds cross zero, precluding sign identification. Across these additional bandwidths, I find consistent evidence, aligning with my main results, that WIC participation reduces breastfeeding.

Taken together, linearized bounds are negative across bandwidths and modeling choices. Allowing curvature or relaxing slope multipliers widens sets—as predicted by the identification logic—without producing credible evidence of positive effects.

5.2 Traditional and Donut Regression Discontinuity

I next report benchmark results from two conventional regression discontinuity designs (RDDs): a traditional specification using all observations within a symmetric window around the income eligibility threshold, $|X_i| \in [0, h]$, and a donut specification that excludes observations in an inner band equal to the mean bracket width μ and estimates on the ring $|X_i| \in [\mu, \mu+h]$. For each design and outcome, I estimate zero-, first-, and second-order polynomials and select the polynomial order by the Akaike Information Criterion (AIC). Bandwidths are chosen using the MSE-optimal selector with regularization from Calonico et al. (2014). Tables 6 and 7 report local average treatment effect (LATE) estimates and first-stage F-statistics at 50%, 75%, 100%, 150%, and 200% of the MSE-optimal bandwidth. Complementary figures plot AIC-selected point estimates with 95% confidence intervals for $h \in [0.5, h_{\text{MSE}}]$ and the corresponding first-stage F-statistics across the same range (Figures 12 and 13). These results provide a transparent reference for the trade-offs inherent in conventional RDD with a rounded running variable.

For the traditional RDD, AIC favors zero-order polynomial fits, and precision across polynomial orders improves as the estimation window widens. For initiation (Table 6, Panel A), AIC selects a zero-order specification at all five MSE-selected bandwidths, yielding uniformly negative and precisely estimated effects from -0.090 to -0.054 (s.e. 0.021–0.029), with large first-stage F-statistics. Higher-order fits exhibit volatility and weaker instruments at several bandwidths. First-order estimates attenuate toward zero as the window widens and the associated first-stage strength fluctuates, increasing from $[12.0]$ to $[303.2]$. Second-order estimates are unstable at the narrowest bandwidth, with very low first-stage strength, and improve only at wider windows. For duration (Panel B), the AIC-selected order varies but estimates remain consistently negative, increasing in magnitude and precision with bandwidth. Zero-order estimates range -1.47 to -3.03 weeks (s.e. 0.70–0.33), with high first-stage strength. Higher-order specifications again reveal instability and weak instruments, with second-order F-statistics as low as $[1.3]$ at narrow windows. Figures 12a and 12b plot AIC-selected point estimates with 95% confidence intervals for $h \in [0.5, h_{\text{MSE}}]$. These visuals mirror the table patterns, showing bandwidth sensitivity at narrow windows and tightening intervals with larger bandwidths. Figure 13 plots the AIC-selected first-stage F-statistics for the traditional RDD over this range, showing weak-instrument concerns at the narrowest windows (F < 10) that dissipate as the bandwidth expands and the F-statistics stabilize above conventional thresholds. Overall, the traditional design delivers predominantly negative point estimates for both outcomes, with inference sensitive to polynomial order and bandwidth, and AIC typically preferring zero-order fits.

For the donut RDD, excluding observations within the average bracket width around the threshold yields more stable and precisely estimated effects across outcomes. For initiation (Table 7, Panel A), AIC selects a zero-order specification at four of the five MSE-selected bandwidths, delivering uniformly negative and tightly estimated effects ranging from -0.094 to -0.081 (s.e. 0.004–0.009), alongside large first-stage F-statistics. Higher-order fits display occasional large swings and weak instruments; however, at the widest window the second-order estimate remains negative. For duration (Panel B), AIC again favors a zero-order specification at most bandwidths, producing consistently negative estimates ranging from -4 to -5 weeks with standard errors between 0.6–0.9 and first-stage strength above conventional thresholds. First- and second-order polynomials are less stable and exhibit weak instruments at several bandwidths. Figures 12c and 12d plot AIC-selected point estimates with 95% confidence intervals for $h \in [0.5, h_{\mathrm{MSE}}]$, showing uniformly negative estimates with tightening intervals as bandwidth increases. Figure 13 plots the AIC-selected first-stage F-statistics for the donut RDD over this range, indicating strong first-stage relationships across bandwidths. The donut design delivers generally negative point estimates with stronger and more stable first stages than the traditional RDD, consistent with bracket-induced measurement error in the untrimmed sample.

Taken together, these benchmark estimates underscore the core tension in conventional regression discontinuity designs with a rounded or bracketed running variable. Diagnostic checks are mixed: histogram-based density tests show no discontinuity in the running-variable distribution across the bandwidths I examine (Table 8), but several predetermined covariates exhibit statistically detectable jumps at the cutoff—especially in education, marital status, and race/ethnicity—with moderate standardized differences (Table 9). Taken together, these patterns are consistent with bracket-induced measurement error and treatment misclassification near the threshold, underscoring the fragility of conventional RDD identification. Widening the window strengthens the first stage and tightens confidence intervals, but it does so by drawing inference from observations increasingly distant from the cutoff, where local exogeneity and continuity of potential outcomes are less credible and functional-form restrictions exert greater influence. Trimming an inner band in the donut design mitigates weak-instrument problems by discarding most ambiguous observations near the threshold—those most affected by bracket-induced measurement error and treatment misclassification—but this gain in first-stage strength comes at the cost of a larger effective bandwidth and heavier reliance on extrapolation away from the cutoff. The net result is a clear trade-off: traditional RDDs face attenuation and imprecision at narrow windows, whereas donut RDDs improve power but increase sensitivity to specification and lessen the plausibility of local identifying assumptions.

5.3 Discussion

My analysis uses partial identification to address bracket reporting (or rounding) in the running variable, a common challenge in regression discontinuity designs (RDD). As formalized in Section 3, bracketed reporting creates two distinct difficulties: (1) measurement error in the running variable and (2) misclassification of treatment assignment for observations whose brackets straddle the policy threshold. In such a setting, conventional RDDs struggle to identify the local average treatment effect. Including observations near the threshold introduces misclassification bias from households with an observed running variable, X, and true running variable, $X^{\star}$, on opposite sides of the cutoff. Excluding these observations relies on increasingly tenuous identification assumptions. My partial-identification framework sidesteps this dilemma by focusing on households with unambiguous treatment assignment and imposing theoretically grounded shape restrictions to bound effects at the threshold.

Table 10 presents the AIC-selected traditional and donut estimates side-by-side with my partial identification bounds. For breastfeeding initiation (Panel A), the traditional RDD lies inside my bounds at the narrower windows but drifts above the upper bound as the window widens. The donut design largely remains inside my bounds at these bandwidths and delivers tighter standard errors at many windows; however, at the widest bandwidth ($h = 29.90$) the donut estimate is near zero (-0.005) with weak first-stage strength $[9.39]$ and a large standard error. For breastfeeding duration (Panel B), both conventional designs deliver negative estimates with strong first-stage relationships. The donut estimates fall inside the bounds at all reported bandwidths, and the traditional estimates fall inside at three of five bandwidths and sit just outside at the remaining two.

Although the conventional designs deliver statistically significant estimates at these bandwidths, the width of the estimation window meaningfully affects the credibility of the underlying identifying assumptions of local exogeneity and continuity of potential outcomes. At the MSE-selected bandwidth of 29.51, I find statistically significant differences in demographics, education, employment, and birth outcomes across the income eligibility threshold (Table 9). Examining narrower bandwidths, the traditional RDD estimator is sensitive to bandwidth selection and fails to identify the sign of the treatment effect. The donut design, which excludes observations near the threshold, strengthens the first stage on average, but increases reliance on observations farther from the cutoff where the identifying assumptions are more tenuous. For narrower estimation windows, the effective bandwidth for the donut RDD is larger than any bandwidth examined with the traditional RDD. In contrast, the partial identification bounds provide a range of plausible effects across both outcomes under empirically supported assumptions.

Three key advantages distinguish my approach from conventional methods. First, my framework explicitly addresses misclassification when observed treatment assignment, $\tilde{Z} = \mathbb{1}[X \leq 0]$, differs from true assignment, $Z = \mathbb{1}[X^{\star} \leq 0]$. With such misclassification, the conventional estimator $\tilde{\tau}$ captures an unidentified weighted average of eligible and ineligible households, whereas my bounds rely only on certainly eligible households (where $u < 0$) and certainly ineligible households (where $\ell > 0$), eliminating misclassification bias. Second, I replace the strong continuity assumption required for point identification with weaker shape restrictions. Assumptions 2 and 3 impose only monotonicity and bounded rates of change, conditions that are both theoretically motivated and empirically supported in my setting. Third, my bounds avoid reliance on localized comparisons and local randomization near the eligibility threshold, where measurement error and potential manipulation pose the greatest threat.

The theoretical foundation of my approach—formalized in Section 3—provides transparent guidance on the strength of identification. My framework makes explicit that identification depends on the plausibility of the imposed shape and monotonicity restrictions, allowing researchers to assess the robustness of conclusions to alternative assumptions. Although bounds may be wide under more conservative assumptions, my approach is more robust to sorting than conventional RDD, which relies on local randomization exactly where manipulation incentives are strongest. For the broad class of empirical settings with rounding in the running variable, my partial identification approach provides valid inference in the presence of misclassified treatment assignment.

Finally, my findings speak to the broader WIC literature that has documented lower breastfeeding among participants in observational settings but has faced challenges in isolating causal effects. By focusing on compliers at the cutoff and accommodating bracket-induced uncertainty directly, the partial identification results show negative local effects on both initiation and duration without invoking strong local continuity or randomization assumptions. My results provide clear evidence under explicit assumptions that WIC reduces the initiation and duration of breastfeeding among participating mothers.

6 Conclusion

This paper develops a partial-identification regression discontinuity design (RDD) for settings with rounding in the running variable. My approach uses shape restrictions on conditional expectations to bound the local average treatment effect (LATE) without requiring local continuity, random assignment near the threshold, or precise measurement of the running variable. The framework's core assumptions—monotonicity and bounded rates of change—are derived from empirical patterns, allowing flexible application to a broad class of settings with data limitations or sorting behavior.

Applying this framework to WIC participation and infant feeding practices, I find evidence that program participation reduces both the initiation and duration of breastfeeding. These findings suggest that despite WIC's substantial investments in breastfeeding promotion through counseling and peer support programs, the in-kind transfer of infant formula creates stronger countervailing incentives. While my partial-identification approach cannot pinpoint precise effect magnitudes, it provides policymakers with credible bounds on the trade-offs inherent in WIC's current benefit structure.

References

American College of Obstetricians and Gynecologists, "Barriers to Breastfeeding: Supporting Initiation and Continuation of Breastfeeding. Committee Opinion No.821," *Obstetrics and Gynecology*, February 2021, *137* (2), e54–e62.

Barreca, Alan I., Jason M. Lindo, and Glen R. Waddell, "Heaping-Induced Bias in Regression-Discontinuity Designs," *Economic Inquiry*, 2016, *54* (1), 268–293.

_ , Melanie Guldi, Jason M. Lindo, and Glen R. Waddell, "Saving Babies? Revisiting the Effect of Very Low Birth Weight Classification," *Quarterly Journal of Economics*, 2011, *126* (4), 2117–2123.

Belfield, Clive R. and Inas Rashad Kelly, "The benefits of breast feeding across the early years of childhood," *Journal of Human Capital*, 2012, *6.*

Bitler, Marianne P. and Janet Currie, "Does WIC work? The effects of WIC on pregnancy and birth outcomes," *Journal of Policy Analysis and Management*, 2005, *24.*

Bullinger, Lindsey Rose and Tami Gurley-Calvez, "WIC Participation and Maternal Behavior: Breastfeeding and Work Leave," *Contemporary Economic Policy*, 2016, *34*, 158–172.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Carlson, Andrea and Ben Senauer, "The impact of the special supplemental nutrition program for women, infants, and children on child health," *American Journal of Agricultural Economics*, 2003, *85.*

Centers for Disease Control and Prevention, "Breastfeeding Recommendations and Benefits," 2023. Accessed on May 2, 2024.

Chatterji, Pinka and Jeanne Brooks-Gunn, "WIC participation, breastfeeding practices, and well-child care among unmarried, low-income mothers," *American Journal of Public Health*, 2004, *94.*

Davezies, Laurent and Thomas Le Barbanchon, "Regression discontinuity design with continuous measurement error in the running variable," *Journal of Econometrics*, 2017, *200* (2), 260–281.

Diaz, Laura E., Lynn M. Yee, and Joe Feinglass, "Rates of Breastfeeding Initiation and Duration in the United States: Data Insights from the 2016–2019 Pregnancy Risk Assessment Monitoring System," *Frontiers in Public Health*, 2023, *11*, 1256432.

Dieterich, C. M., J. P. Felice, E. O'Sullivan, and K. M. Rasmussen, "Breastfeeding and health outcomes for the mother-infant dyad," *Pediatric Clinics of North America*, February 2013, *60* (1), 31–48.

Dong, Yingying, "Regression Discontinuity Applications with Rounding Errors in the Running Variable," *Journal of Applied Econometrics*, 2015, *30* (3), 422–446.

_ **and Michal Kolesár**, "When Can We Ignore Measurement Error in the Running Variable?," *Journal of Applied Econometrics*, 2023, *38* (4), 504–518.

Figlio, David, Sarah Hamersma, and Jeffrey Roth, "Does prenatal WIC participation improve birth outcomes? New evidence from Florida," *Journal of Public Economics*, 2009, *93.*

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, and Michael Westberry, "Integrated Public Use Microdata Series, Current Population Survey: Version 10.0," 2022.

Gelman, Andrew and Guido Imbens, "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs," *Journal of Business & Economic Statistics*, 2019, *37* (3), 447–456.

Gerard, François, Miikka Rokkanen, and Christoph Rothe, "Bounds on Treatment Effects in Regression Discontinuity Designs with a Manipulated Running Variable," *Quantitative Economics*, 2020, *11* (3), 839–870.

Hoynes, Hilary, Marianne Page, and Ann Huff Stevens, "Can targeted transfers improve birth outcomes? Evidence from the introduction of the WIC program," *Journal of Public Economics*, 2011, *95.*

Hoynes, Hilary W. and Diane Whitmore Schanzenbach, "U.S. Food and Nutrition Programs," Working Paper 21057, National Bureau of Economic Research 2015.

Imbens, Guido and Stefan Wager, "Optimized Regression Discontinuity Designs," *The Review of Economics and Statistics*, May 2019, *101* (2), 264–278.

Jiang, Miao, E. Michael Foster, and Christina M. Gibson-Davis, "The effect of WIC on breastfeeding: A new look at an established relationship," *Children and Youth Services Review*, 2 2010, *32*, 264–273.

Joyce, Ted, Andrew Racine, and Cristina Yunzal-Butler, "Reassessing the WIC effect: Evidence from the pregnancy nutrition surveillance system," *Journal of Policy Analysis and Management*, 2008, *27.*

Kline, Nicholas, Karen Meyers Mathieu, and Jennifer Marr, "WIC Participant and Program Characteristics 2018: Food Packages and Costs Report," Technical Report, U.S. Department of Agriculture, Food and Nutrition Service, Alexandria, VA 2020. Project Officer: Grant Lovellette.

Kolesár, Michal and Christoph Rothe, "Inference in Regression Discontinuity Designs with a Discrete Running Variable," *American Economic Review*, 2018, *108* (8), 2277–2304.

Kreider, Brent, John V. Pepper, and Manan Roy, "Does the Women, Infants, and Children Program Improve Infant Health Outcomes?," *Economic Inquiry*, 2020, *58*, 1731–1756.

Manski, Charles F. and John V. Pepper, "Right-to-Carry and Coping with Ambiguity with Bounds," *Review of Economics and Statistics*, 2018, *100* (2), 232–244.

National WIC Association, "WIC's Promotion and Support of Breastfeeding," 2019. Accessed on May 2, 2024.

Oster, Emily, *Expecting Better: Why the Conventional Pregnancy Wisdom Is Wrong–and What You Really Need to Know*, New York: Penguin Press, 2013.

Pei, Zhuan, David S. Lee, David Card, and Andrea Weber, "Local Polynomial Order in Regression Discontinuity Designs," *Journal of Business & Economic Statistics*, 2022, *40* (3), 1259–1267.

Rose, Donald, J Nicholas Bodor, and Mariana Chilton, "Symposium: Food Assistance and the Well-Being of Low-Income Families Has the WIC Incentive to Formula-Feed Led to an Increase in Overweight Children?," 2006.

Ryan, Alan S. and Wenjun Zhou, "Lower breastfeeding rates persist among the special supplemental nutrition program for women, infants, and children participants, 1978-2003," *Pediatrics*, 2006, *117.*

Shulman, Holly B., Denise V. D'Angelo, Leslie Harrison, Ruben A. Smith, and Lee Warner, "The Pregnancy Risk Assessment Monitoring System (PRAMS): Overview of Design and Methodology," *American Journal of Public Health*, August 2018, *e1*, e9.

Topolyan, Iryna and Xu Xu, "Differential Effects of Mother's and Child's Postnatal WIC participation on breastfeeding," *Applied Economics*, 5 2017, *49*, 2216–2225.

United States Department of Agriculture, Food and Nutrition Service, "Appendix of the WIC 2020 Participant and Program Characteristics Report," Appendix, United States Department of Agriculture, Food and Nutrition Service 2022.

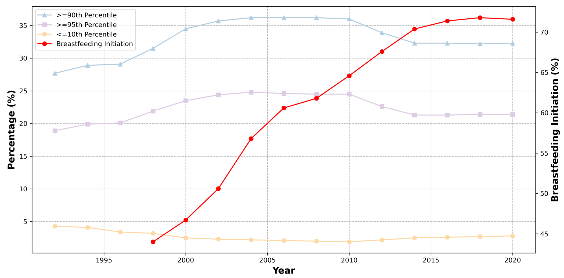
_ , "WIC 2020 Participant and Program Characteristics Report," Report, United States Department of Agriculture, Food and Nutrition Service 2022.

Whaley, Shannon E., Maria Koleilat, Mike Whaley, Judy Gomez, Karen Meehan, and Kiran Saluja, "Impact of policy changes on infant feeding decisions among low-income Women participating in the Special Supplemental Nutrition Program for Women, Infants, and Children," *American Journal of Public Health*, 12 2012, *102*, 2269–2273.

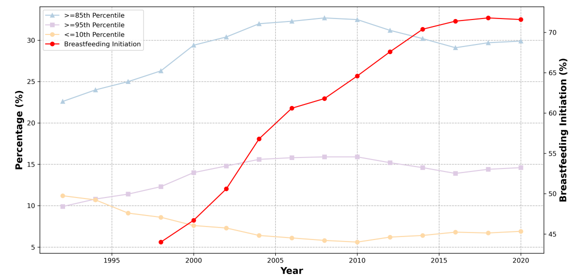
A Figures

Figure 1: Breastfeeding Initiation Rates and Participant Weight Distributions

Panel A: Children

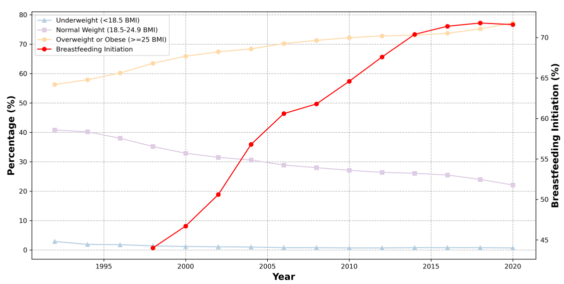


(a) 1-year-old Children

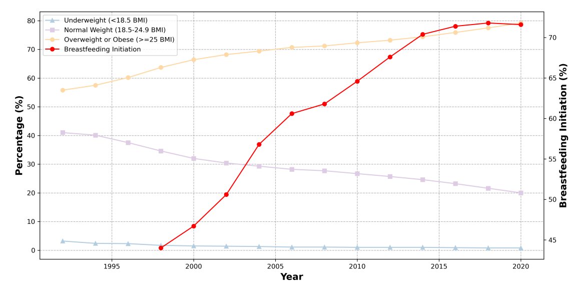


(b) 2- to 4-year-old Children

Panel B: Women



(c) Breastfeeding Women

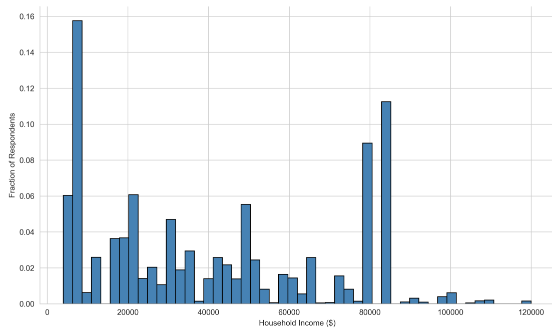


(d) Postpartum Women

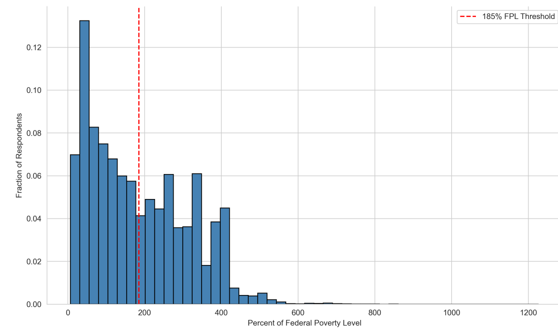
Notes: This figure presents breastfeeding initiation rates and the distribution of growth measurements for WIC participating women, infants, and children using data from the WIC Participant and Program Characteristics 2020 Appendix. Panel A shows anthropometric distributions for children: Subfigure (a) displays the distribution of 1-year-old children by standardized WHO Weight-for-Length percentiles; Subfigure (b) shows the distribution of 2- to 4-year-old children by standardized WHO BMI percentiles. Panel B presents maternal anthropometrics: Subfigure (c) displays the BMI distribution for breastfeeding women (those with fully or partially breastfeeding infants); Subfigure (d) shows BMI distribution for postpartum women without breastfeeding infants. All figures include breastfeeding initiation rates for participating 7- to 11-month-old infants. Percentiles above the 85th represent overweight and above the 95th represent obesity according to CDC standards.

Figure 2: Breastfeeding Behaviors by Household Income (%FPL)

Panel A: Distribution of Household Income

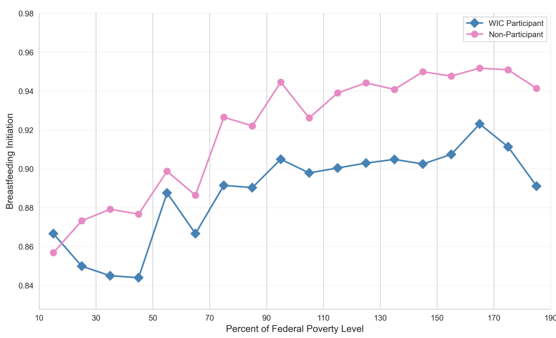


(a) Dollars ($)

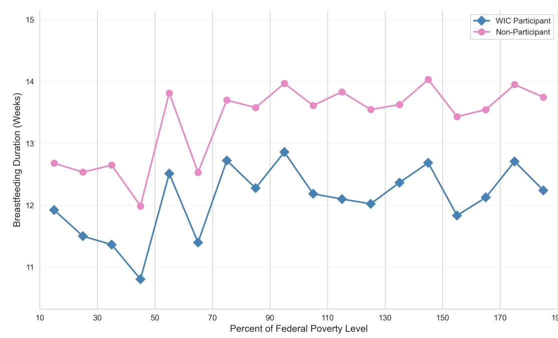


(b) Percentage of FPL (%FPL)

Panel B: Initiation and Duration of Breastfeeding



(c) Initiation

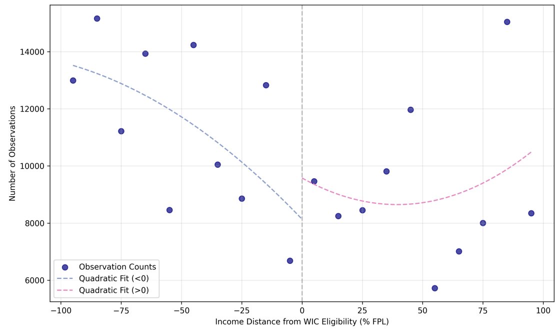


(d) Duration

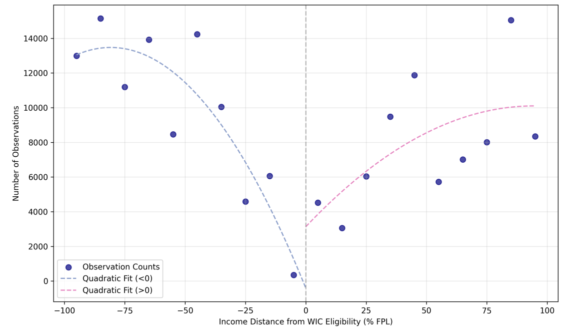
Notes: This figure presents the distribution income and infant feeding practice by income in the Pregnancy Risk Assessment Monitoring System (PRAMS). Panel A presents the distribution of household income—the midpoint of observed brackets as proxy—in dollars and as a percentage of the Federal Poverty Level (FPL). The vertical dashed line indicates the WIC eligibility threshold, set nationally at 185% FPL. Panel B displays breastfeeding initiation rates and duration by household income for WIC participants (blue diamonds) and income-eligible non-participants (pink circles). Duration is conditional on breastfeeding initiation.

Figure 3: Binned Scatter of Observation Count by FPL Bin

Panel A: 20 Bins (Width of 10 pct)

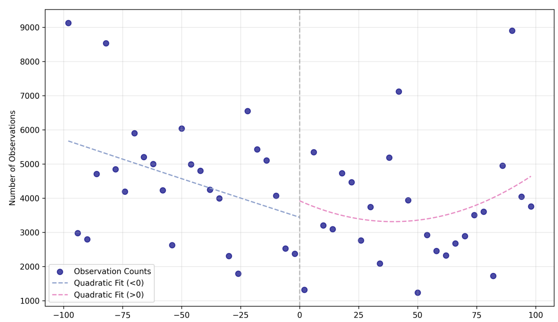


(a) Including indeterminate Assignment

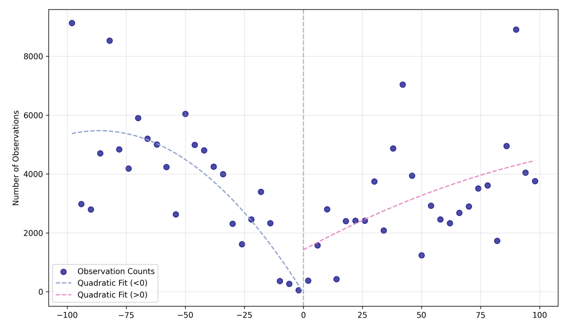


(b) Excluding indeterminate Assignment

Panel B: 50 Bins (Width of 4 pct)

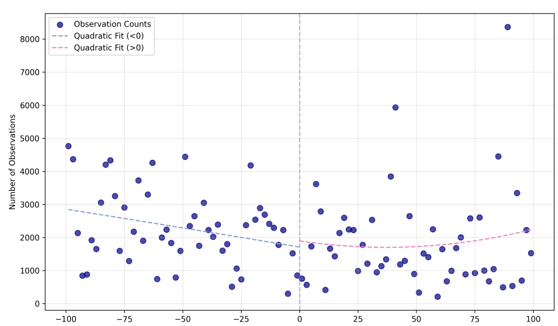


(c) Including indeterminate Assignment

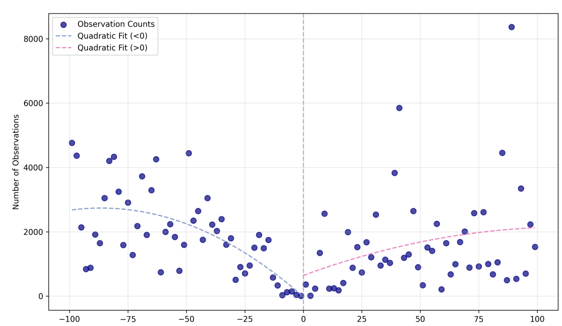


(d) Excluding indeterminate Assignment

Panel C: 100 Bins (Width of 2 pct)



(e) Including indeterminate Assignment

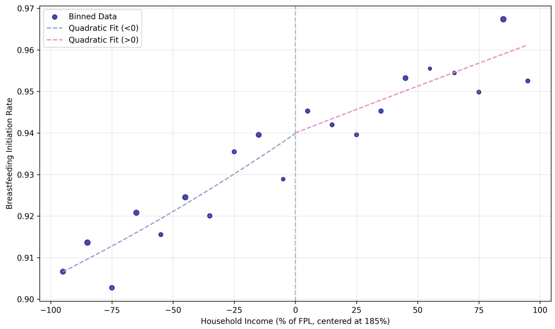


(f) Excluding indeterminate Assignment

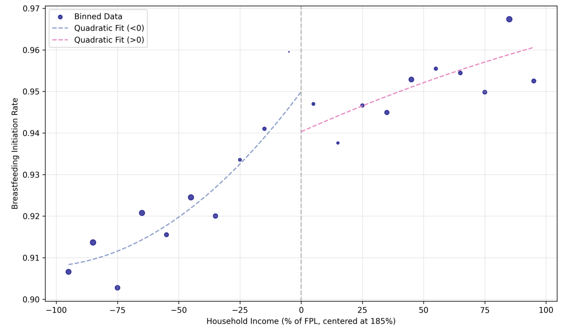
Notes: This figure displays the distribution of observations by household income, as a percentage of the Federal Poverty Level (FPL), in the Pregnancy Risk Assessment Monitoring System (PRAMS). Observations are assigned to particular FPL bins using the midpoint of their observed income bracket $(\text{FPL}_\ell + \text{FPL}_u)/2$. Household income, as a percentage of FPL, is centered at the WIC income eligibility threshold of 185%, which is marked vertically in red. Panels vary by bin width: (A) 10 percentage points, (B) 4 percentage points, and (C) 2 percentage points. Left figures include all observations; right figures exclude households with brackets straddling the threshold ($\text{FPL}_\ell \leq 185 < \text{FPL}_u$). Fitted curves show quadratic approximations.

Figure 4: Binned Scatter of Breastfeeding Initiation by FPL Bin

Panel A: 20 Bins (Width of 10 pct)

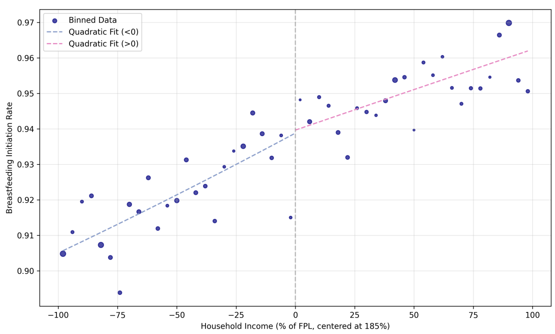


(a) Including indeterminate Assignment

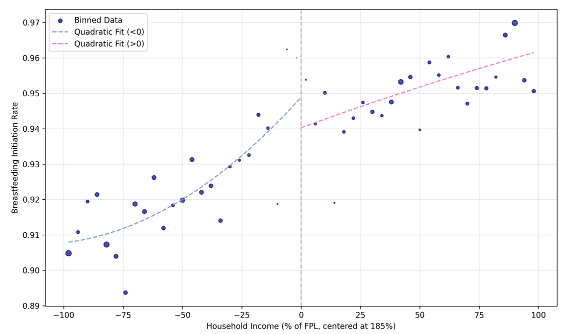


(b) Excluding indeterminate Assignment

Panel B: 50 Bins (Width of 4 pct)

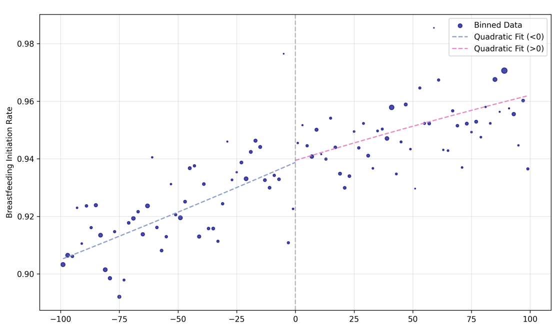


(c) Including indeterminate Assignment

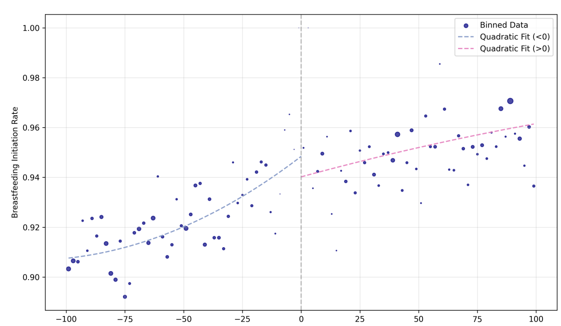


(d) Excluding indeterminate Assignment

Panel C: 100 Bins (Width of 2 pct)



(e) Including indeterminate Assignment

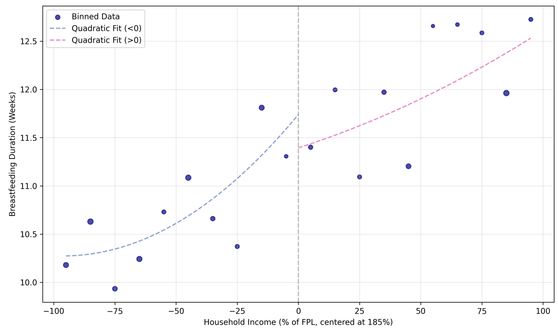


(f) Excluding indeterminate Assignment

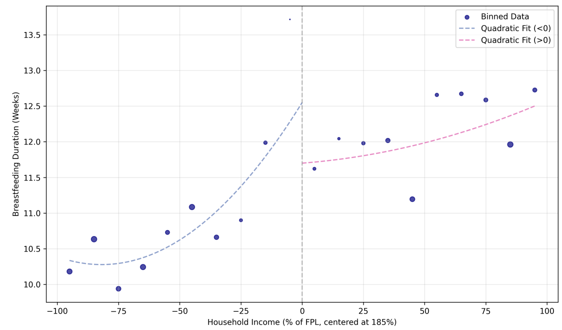
Notes: This figure displays the distribution of breastfeeding initiation by household income, as a percentage of the Federal Poverty Level (FPL), in the Pregnancy Risk Assessment Monitoring System (PRAMS). Observations are assigned to particular FPL bins using the midpoint of their observed income bracket $(\text{FPL}_\ell + \text{FPL}_u)/2$. Household income, as a percentage of FPL, is centered at the WIC income eligibility threshold of 185%, which is marked vertically in red. Panels vary by bin width: (A) 10 percentage points, (B) 4 percentage points, and (C) 2 percentage points. Left figures include all observations; right figures exclude households with brackets straddling the threshold $(\text{FPL}_\ell \leq 185 < \text{FPL}_u)$. Fitted curves show quadratic approximations. Observations are weighted using PRAMS survey weights.

Figure 5: Binned Scatter of Breastfeeding Duration by FPL Bin

Panel A: 20 Bins (Width of 10 pct)

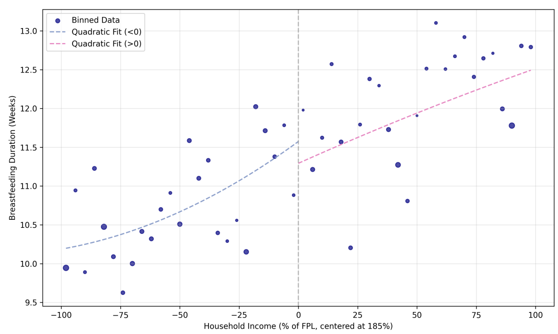


(a) Including indeterminate Assignment

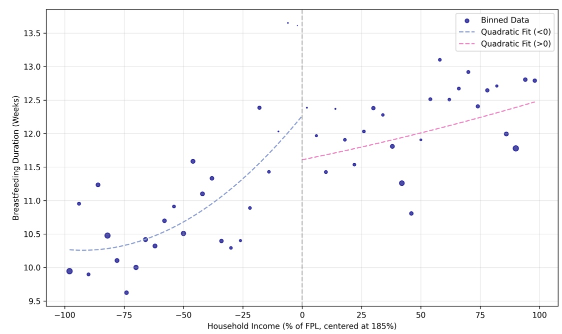


(b) Excluding indeterminate Assignment

Panel B: 50 Bins (Width of 4 pct)

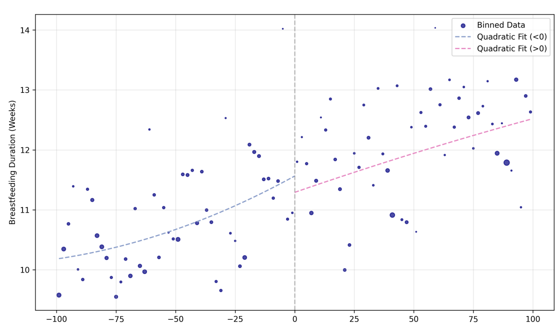


(c) Including indeterminate Assignment

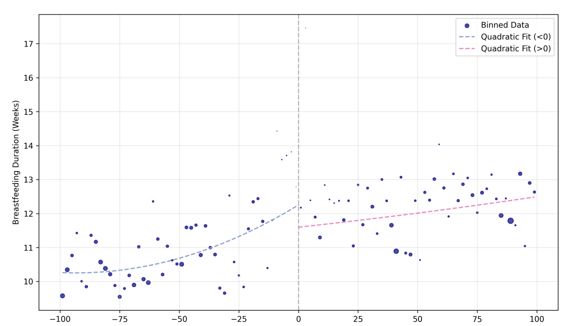


(d) Excluding indeterminate Assignment

Panel C: 100 Bins (Width of 2 pct)



(e) Including indeterminate Assignment

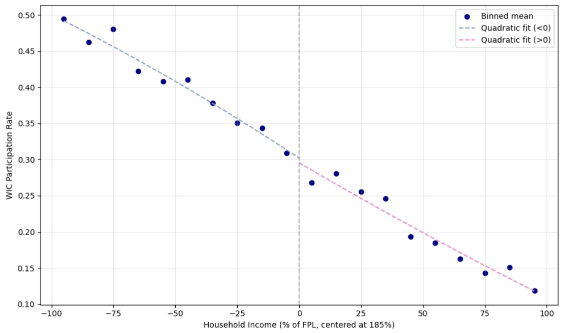


(f) Excluding indeterminate Assignment

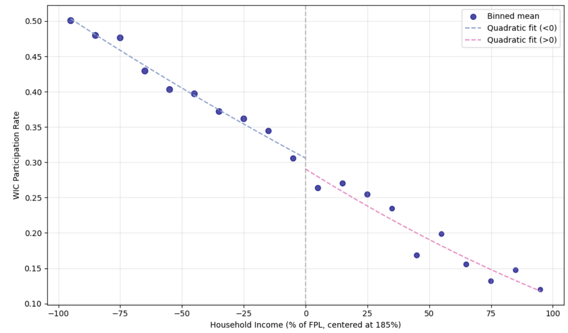
Notes: This figure displays the distribution of breastfeeding duration by household income, as a percentage of the Federal Poverty Level (FPL), in the Pregnancy Risk Assessment Monitoring System (PRAMS). Observations are assigned to particular FPL bins using the midpoint of their observed income bracket $(\text{FPL}_\ell + \text{FPL}_u)/2$. Household income, as a percentage of FPL, is centered at the WIC income eligibility threshold of 185%, which is marked vertically in red. Panels vary by bin width: (A) 10 percentage points, (B) 4 percentage points, and (C) 2 percentage points. Left figures include all observations; right figures exclude households with brackets straddling the threshold ($\text{FPL}_\ell \leq 185 < \text{FPL}_u$). Fitted curves show quadratic approximations. Observations are weighted using PRAMS survey weights.

Figure 6: Binned Scatter of WIC Participation by FPL Bin: CPS Data

Panel A: 20 Bins (Width of 10 pct)

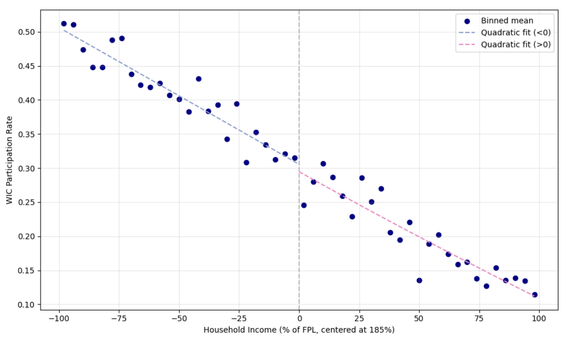


(a) Unweighted

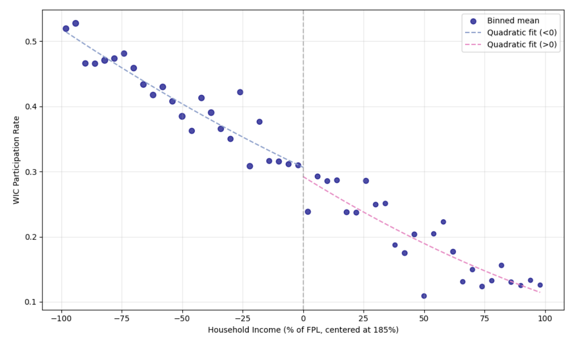


(b) Weighted

Panel B: 50 Bins (Width of 4 pct)

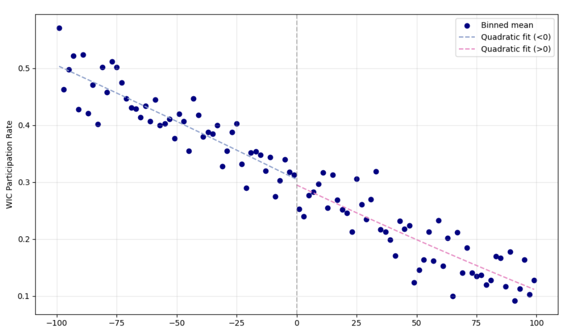


(c) Unweighted

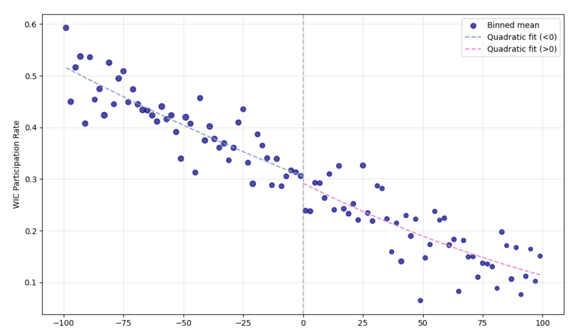


(d) Weighted

Panel C: 100 Bins (Width of 2 pct)



(e) Unweighted

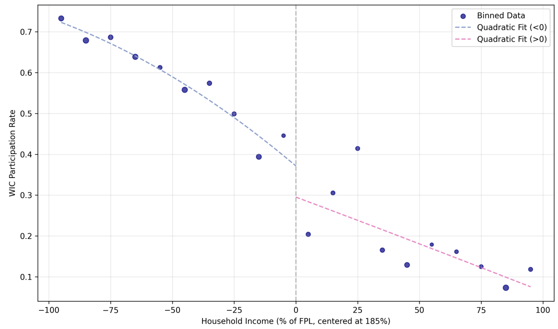


(f) Weighted

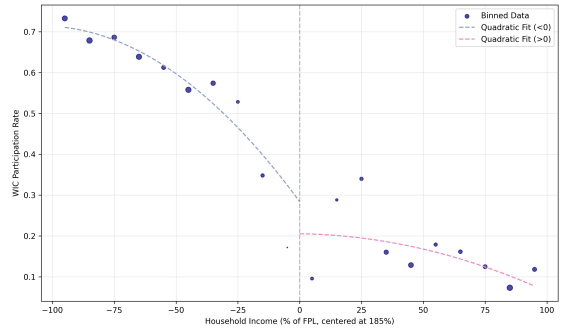
Notes: This figure displays the distribution of participation in WIC by household income, as a percentage of the Federal Poverty Level (FPL), in the Current Population Survey Annual Social and Economic Supplement (CPS-ASEC). Observations are assigned to particular FPL bins using point-identified household income. Household income, as a percentage of FPL, is centered at the WIC income eligibility threshold of 185%, which is marked vertically in red. Panels vary by bin width: (A) 10 percentage points, (B) 4 percentage points, and (C) 2 percentage points. Left figures show unweighted quadratic fits; right figures show quadratic fits weighted by bin size.

Figure 7: Binned Scatter of WIC Participation by FPL Bin

Panel A: 20 Bins (Width of 10 pct)

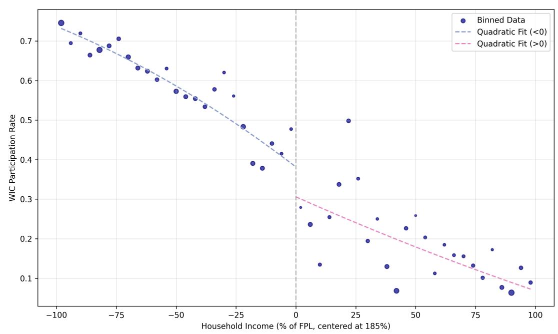


(a) Including indeterminate Assignment

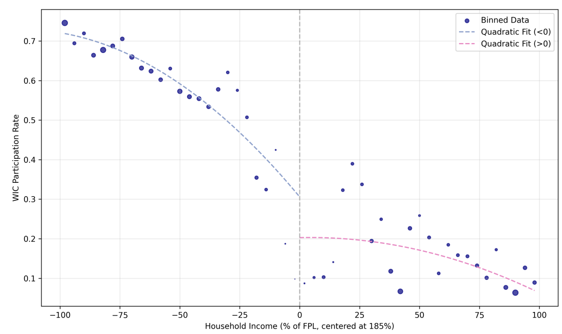


(b) Excluding indeterminate Assignment

Panel B: 50 Bins (Width of 4 pct)

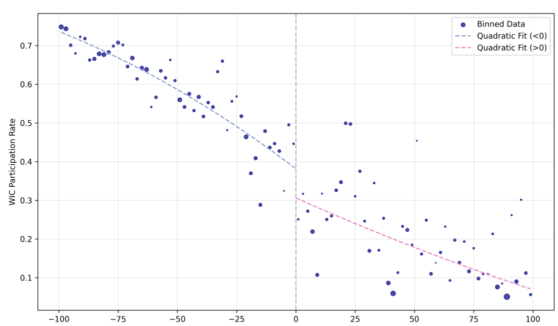


(c) Including indeterminate Assignment

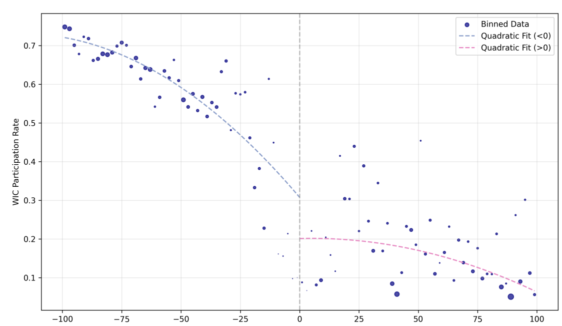


(d) Excluding indeterminate Assignment

Panel C: 100 Bins (Width of 2 pct)



(e) Including indeterminate Assignment

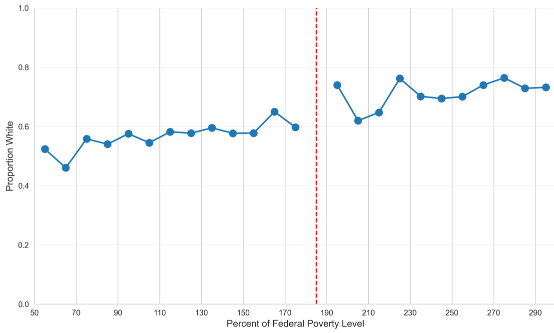


(f) Excluding indeterminate Assignment

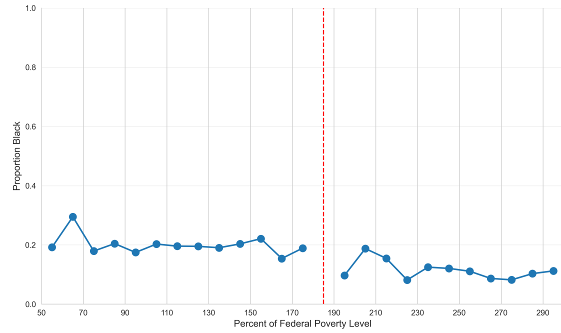
Notes: This figure displays the distribution of participation in WIC by household income, as a percentage of the Federal Poverty Level (FPL), in the Pregnancy Risk Assessment Monitoring System (PRAMS). Observations are assigned to particular FPL bins using the midpoint of their observed income bracket $(\text{FPL}_\ell + \text{FPL}_u)/2$. Household income, as a percentage of FPL, is centered at the WIC income eligibility threshold of 185%, which is marked vertically in red. Panels vary by bin width: (A) 10 percentage points, (B) 4 percentage points, and (C) 2 percentage points. Left figures include all observations; right figures exclude households with brackets straddling the threshold ($\text{FPL}_\ell \leq 185 < \text{FPL}_u$). Fitted curves show quadratic approximations. Observations are weighted using PRAMS survey weights.

Figure 8: Maternal Characteristics by Household Income

Panel A: Race/Ethnicity

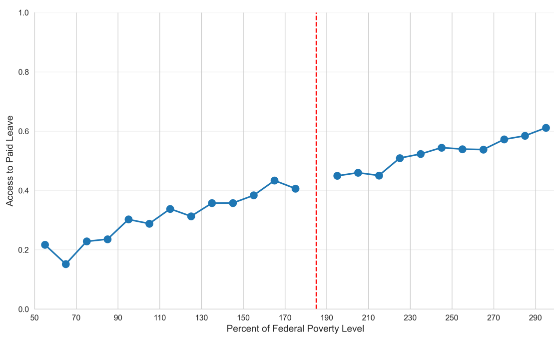


(a) White

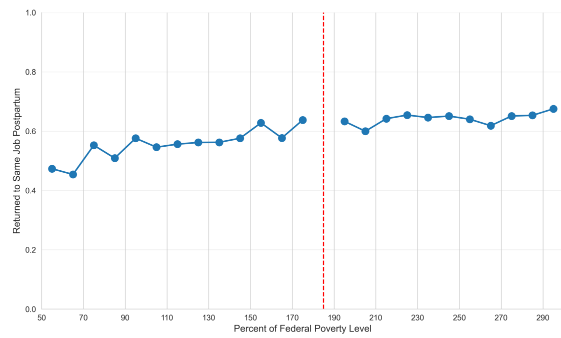


(b) Black

Panel B: Employment and Leave

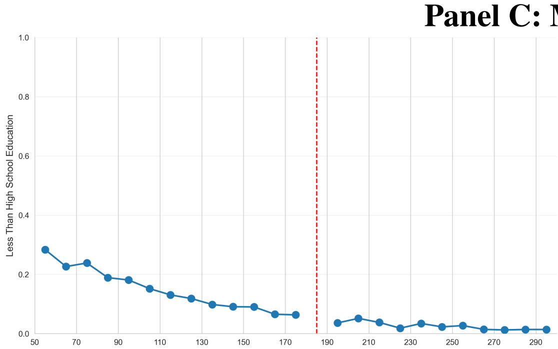


(c) Access to Paid Leave

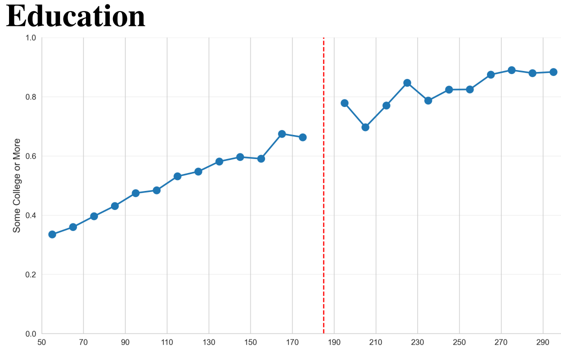


(d) Returned to Job Postpartum

Panel C: Maternal Education



(e) Less than High School

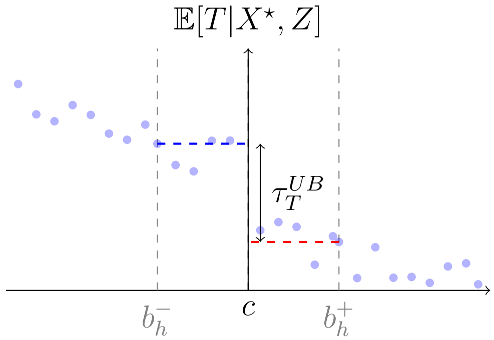


(f) Some College or More

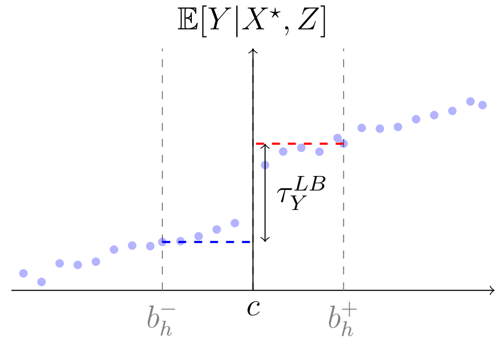
Notes: This figure shows how selected maternal characteristics evolve with household income, expressed as a percentage of the Federal Poverty Level (FPL). Each subplot presents binned proportions for a binary variable by FPL bracket using PRAMS data. The vertical dashed red line denotes the WIC income eligibility threshold at 185% FPL. The bin spanning 180–190% is excluded to avoid overlap with the threshold, and the plotted lines are segmented accordingly. All variables are defined as binary indicators and plotted as proportions within each income bin.

Figure 9: Partial Identification Bounds

Panel A: Monotonicity Assumptions

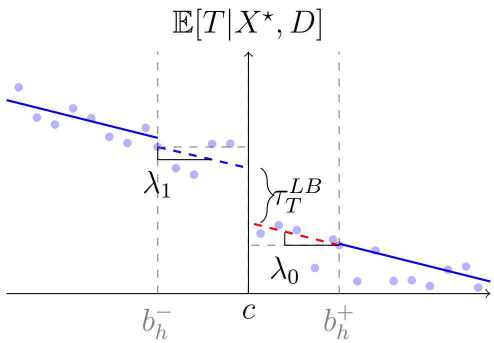


(a) Decreasing in Income

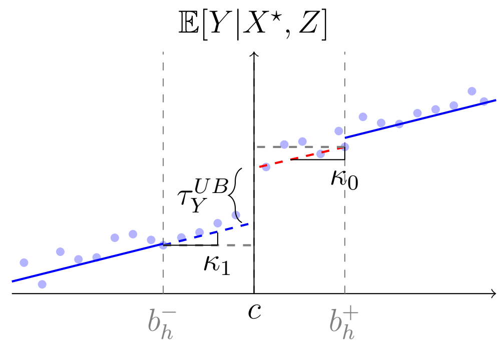


(b) Increasing in Income

Panel B: Bounded Slope Assumptions

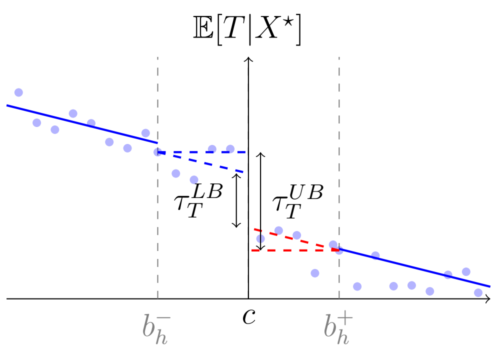


(c) Decreasing in Income

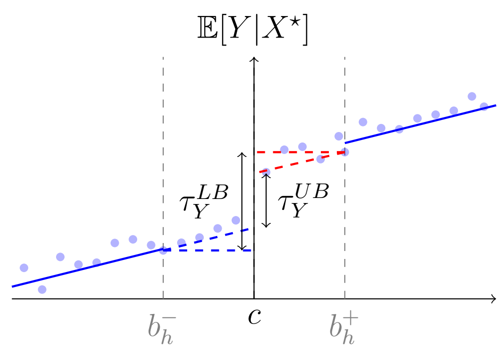


(d) Increasing in Income

Panel C: Combined



(e) Decreasing in Income

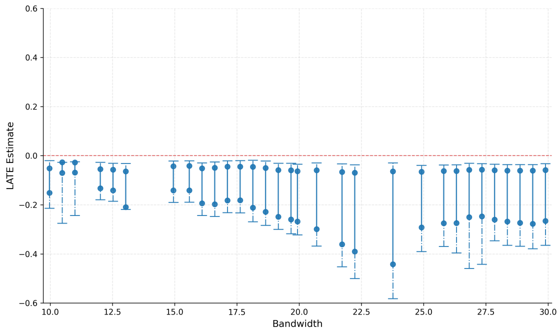


(f) Increasing in Income

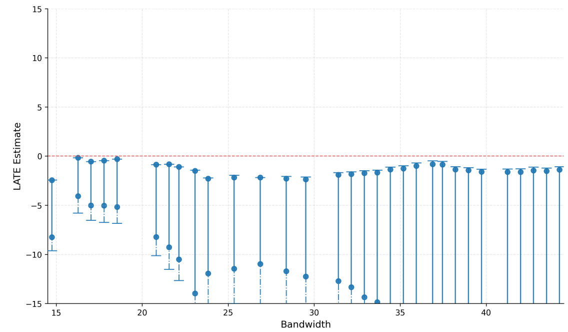
Notes: This stylized figure illustrates the bounds on discontinuities τ_T for WIC participation T and τ_Y for breastfeeding outcomes Y at the income eligibility threshold imposed by partial-identification Assumptions 2 and 3. The horizontal axis shows household income as a percentage of FPL. Panel A: Monotonicity bounds—upper bound (τ_T^{UB}) for WIC participation, lower bound (τ_Y^{LB}) for breastfeeding outcomes. Panel B: Bounded rate of change constraints—lower bound (τ_T^{LB}) for WIC participation, upper bound (τ_Y^{UB}) for breastfeeding outcomes. Panel C: Combined bounds, showing the cone of uncertainty as income approaches the threshold.

Figure 10: LATE Bounds from Partial Identification RDD: Mean Slope Projection

Panel A: Slope Multiplier $\times$ 1.0

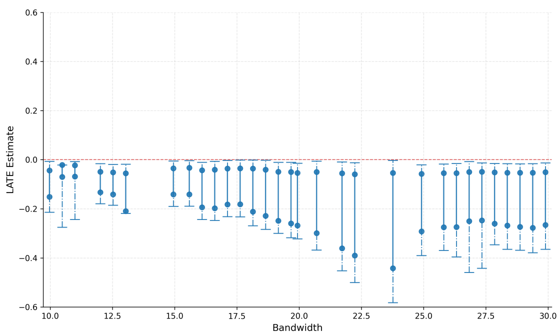


(a) Breastfeeding Initiation

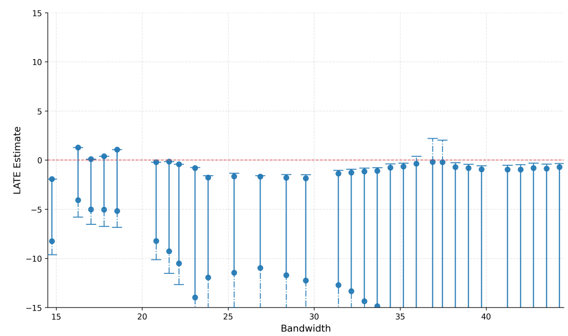


(b) Breastfeeding Duration (weeks)

Panel B: Slope $\times$ 1.5

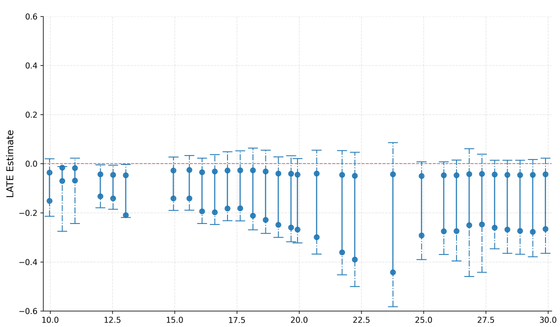


(c) Breastfeeding Initiation

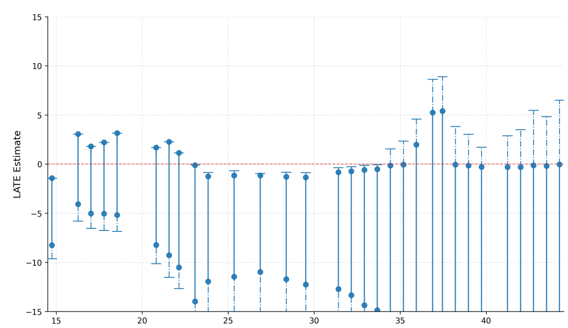


(d) Breastfeeding Duration (weeks)

Panel C: Slope Multiplier $\times$ 2.0



(e) Breastfeeding Initiation

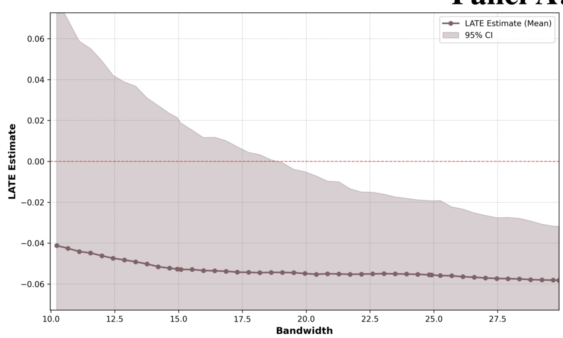


(f) Breastfeeding Duration (weeks)

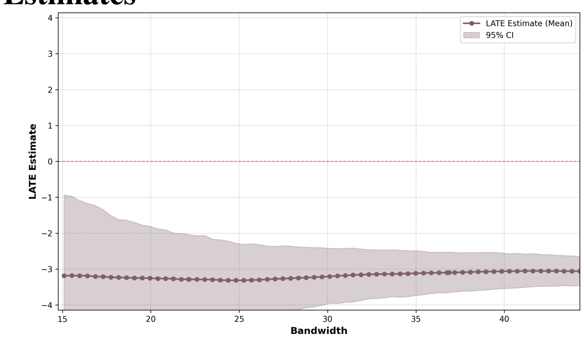
Notes: This figure reports bounds on the local average treatment effect (LATE) and the associated 95% confidence intervals for the effect of WIC participation on breastfeeding outcomes for a set of bandwidths $h \in [\frac{1}{2}h_{\text{MSE}}, 2h_{\text{MSE}}]$. Estimates are derived using the partial-identification regression discontinuity design (RDD) and Assumptions 1–3. The bounded rate of change (Assumption 3) is computed using a linearized slope based on the average derivative of the conditional expectation function across the estimation window. Panels A–C correspond to slope multipliers $\{1.0, 1.5, 2.0\}$. See Section 4 for estimation details.

Figure 11: Sensitivity Testing: Simulation RDD

Panel A: LATE Estimates

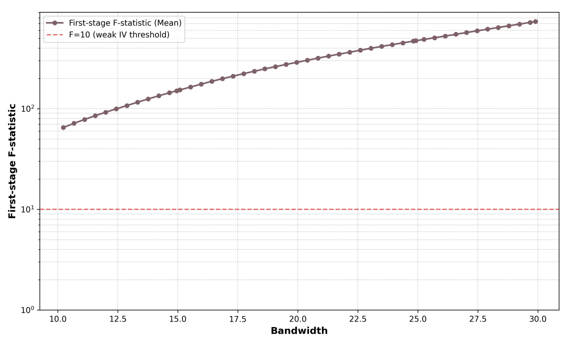


(a) Breastfeeding Initiation

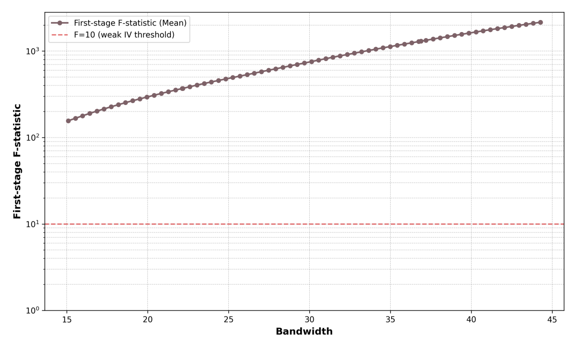


(b) Breastfeeding Duration (weeks)

Panel B: First-Stage F-Statistics



(c) Breastfeeding Initiation

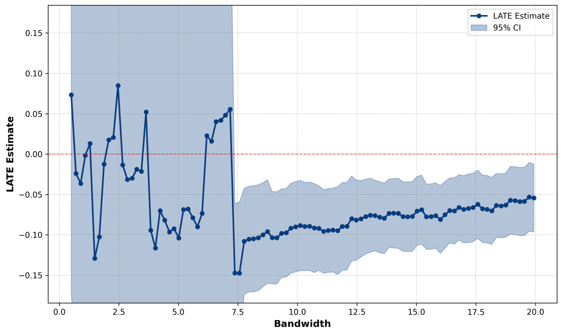


(d) Breastfeeding Duration (weeks)

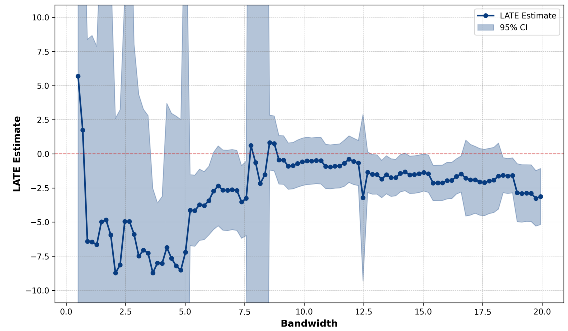
Notes: This figure presents local average treatment effect estimates (τ) from equation 3 and associated first-stage F-statistics from a simulation-based RDD for a set of bandwidths $h \in [\frac{1}{2}h_{\text{MSE}}, 2h_{\text{MSE}}]$. Household income is simulated using a uniform distribution: $X^{\star}_{i,s} \sim \text{Uniform}(\ell_i, u_i)$. The optimal polynomial order is selected using the Akaike Information Criterion (AIC). Panel A displays displays LATE estimates, with 95% confidence intervals indicated via shaded areas. Panel B presents first-stage F-statistics, with the conventional weak instrument threshold, F=10, indicated in red on the vertical axis. See Section 4 for estimation details.

Figure 12: LATE Point-Estimates by FPL Bandwidth

Panel A: Traditional Regression Discontinuity

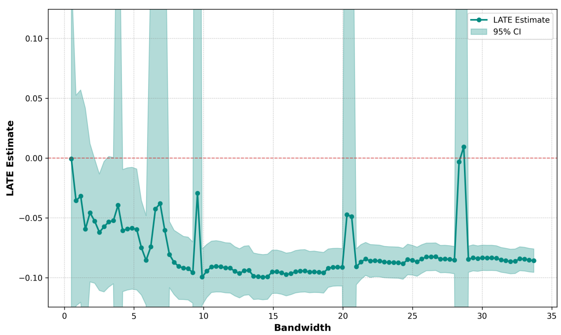


(a) Breastfeeding Initiation

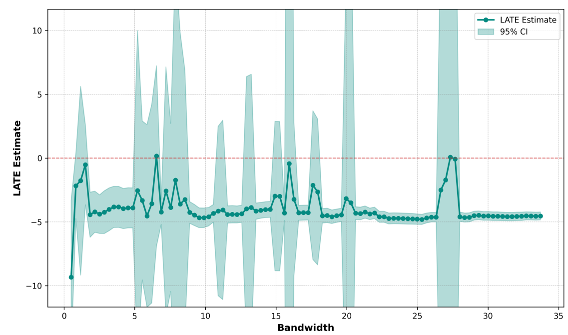


(b) Breastfeeding Duration (weeks)

Panel B: Donut Regression Discontinuity



(c) Breastfeeding Initiation

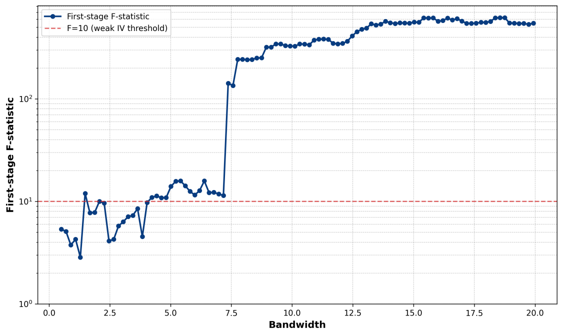


(d) Breastfeeding Duration (weeks)

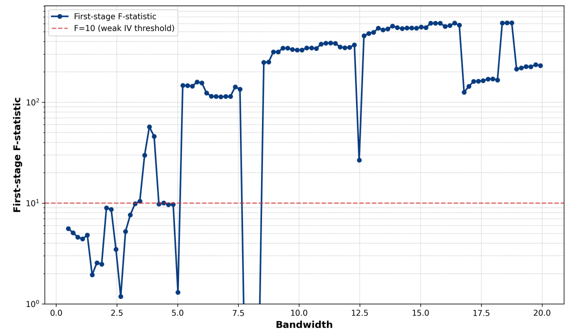
Notes: This figure presents local average treatment effect estimates (τ) from equation 3 using data from the Pregnancy Risk Assessment Monitoring System (PRAMS). For a narrow set of bandwidths $h \in [0.5, h_{\text{MSE}}]$ near the income eligibility threshold, the optimal polynomial order is selected using the Akaike Information Criterion (AIC). Panel A displays results from a traditional fuzzy regression discontinuity design (RDD); Panel B displays results from a donut RDD, excluding observations within $\mu = \frac{1}{N}\sum_{i=1}^{N}(u_i - \ell_i)$ of the income eligibility threshold. Left figures show effects on breastfeeding initiation (a coefficient estimate of -0.1 represents a 10 percentage point decline in probability); right figures show effects on duration (weeks). Shaded areas indicate 95% confidence intervals. See Section 4 for estimation details.

Figure 13: First Stage F-Stat by FPL Bandwidth

Panel A: Traditional Regression Discontinuity

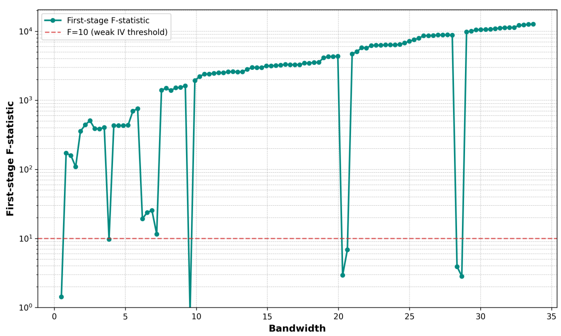


(a) Breastfeeding Initiation

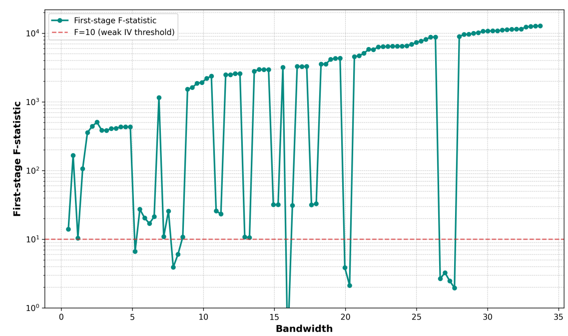


(b) Breastfeeding Duration (weeks)

Panel B: Donut Regression Discontinuity



(c) Breastfeeding Initiation



(d) Breastfeeding Duration (weeks)

Notes: This figure presents first-stage F-statistics for the instrumental variables regression of WIC participation on income eligibility using data from the Pregnancy Risk Assessment Monitoring System (PRAMS). For a narrow set of bandwidths $h \in [0.5, h_{\text{MSE}}]$ near the income eligibility threshold, the optimal polynomial order is selected using the Akaike Information Criterion (AIC). Panel A displays results from a traditional fuzzy regression discontinuity design (RDD); Panel B displays results from a donut RDD, excluding observations within $\mu = \frac{1}{N}\sum_{i=1}^{N}(u_i - \ell_i)$ of the income eligibility threshold. The conventional weak instrument threshold, F=10, is indicated in red on the vertical axis. See Section 4 for estimation details.

B Tables

Table 1: Maximum Monthly Allowances of Supplemental Foods for Infants, Children, and Women

Panel A: Infants

Foods	Fully Formula Fed Food Packages I and III A: 0–3 mo. B: 4–5 mo.	Fully Formula Fed Food Packages II and III 6–11 mo.	Partially Breastfed Food Packages I and III A: 0–3 mo. B: 4–5 mo.	Partially Breastfed Food Packages II and III 6–11 mo.	Fully Breastfed Food Package I 0–5 mo.	Fully Breastfed Food Package II 6–11 mo.
WIC Formula	A: Up to 806 fl. oz. B: Up to 884 fl. oz.	Up to 624 fl. oz.	A: Up to 364 fl. oz. B: Up to 442 fl. oz.	Up to 312 fl. oz.	N/A	N/A
Infant cereal	N/A	8 oz.	N/A	8 oz.	N/A	16 oz.
Baby food fruits and vegetables	N/A	128 oz.	N/A	128 oz.	N/A	128 oz.
Baby food meat	N/A	N/A	N/A	N/A	N/A	40 oz.

Panel B: Children and Women

Foods	Children Food Package IV A: 12–23 mo.	Children Food Package IV B: 2–4 yrs	Women Food Package V Pregnant / Mostly BF	Women Food Package VI Postpartum	Women Food Package VII Fully Breastfeeding	
Juice, single strength	64 fl. oz.	64 fl. oz.	64 fl. oz.	64 fl. oz.	64 fl. oz.	64 fl. oz.
Milk	A: 12 qt. B: 14 qt.	14 qt.	16 qt.	16 qt.	16 qt.	16 qt.
Breakfast cereal	36 oz.	36 oz.	36 oz.	36 oz.	36 oz.	36 oz.
Eggs	1 dozen	1 dozen	1 dozen	1 dozen	2 dozen	2 dozen
Fruit and vegetable CVB	$26.00	A: $47.00 B: $52.00	$47.00	$47.00	$52.00	$52.00
Whole wheat bread	24 oz.	48 oz.	48 oz.	48 oz.	48 oz.	48 oz.
Fish (canned)	6 oz.	A: 10 oz. B: 15 oz.	10 oz.	10 oz.	20 oz.	20 oz.
Legumes and/or Peanut Butter	1 lb. dry / 64 oz. canned or 18 oz. PB	1 lb. / 64 oz. canned and 18 oz. PB	1 lb. / 64 oz. canned or 18 oz. PB	1 lb. / 64 oz. canned or 18 oz. PB	1 lb. / 64 oz. canned and 18 oz. PB	1 lb. / 64 oz. canned and 18 oz. PB

Notes: This table summarizes the maximum monthly quantities of WIC-provided supplemental foods by age and participant category. Panel A presents allowances for infants based on age and breastfeeding status across Food Packages I, II, and III. Panel B presents allowances for children and women across Food Packages IV–VII. All values reflect monthly values under current federal guidelines.

Table 2: Summary Statistics by WIC Participation

Variable	WIC			Non-WIC			
	Mean	Obs	SD	Mean	Obs	SD	Diff (P-value)
Infant Feeding Practices							
Breastfeeding Initiation	0.88	163,328	0.33	0.92	68,753	0.27	-0.05 (0.00)
Breastfeeding Duration - Any	11.84	124,245	7.15	13.38	57,495	6.80	-1.54 (0.00)
Breastfeeding Duration - Exclusive	11.52	53,987	6.71	12.68	22,315	6.37	-1.17 (0.00)
Formula Duration - Any	44.95	88,110	7.64	43.35	32,709	7.91	1.60 (0.00)
Formula Duration - Exclusive	43.20	167,006	8.03	40.83	68,900	7.96	2.36 (0.00)
Infant and Maternal Health							
Maternal BMI	28.14	143,351	7.48	27.29	61,881	6.99	0.86 (0.00)
Maternal Diabetes	0.07	173,453	0.25	0.06	71,794	0.24	0.00 (0.01)
Preterm Labor	0.29	63,868	0.45	0.25	20,721	0.43	0.04 (0.00)
Infant Birth Weight - Grams	3015	173,804	767	3041	71,900	792	-26 (0.00)
Small for Gestational Age	0.17	167,740	0.38	0.15	69,662	0.36	0.02 (0.00)
Maternal Education and Employment							
Employment Before Pregnancy	0.61	38,123	0.49	0.62	22,667	0.49	-0.00 (0.26)
Employment During Pregnancy	0.56	47,025	0.50	0.61	18,649	0.49	-0.05 (0.00)
Returned to Job After Birth	0.49	21,440	0.50	0.55	9,420	0.50	-0.06 (0.00)
Less than High School	0.05	172,168	0.22	0.03	71,207	0.17	0.02 (0.00)
Some High School	0.19	172,168	0.39	0.10	71,207	0.31	0.08 (0.00)
High School Graduate	0.39	172,168	0.49	0.31	71,207	0.46	0.08 (0.00)
Some College	0.31	172,168	0.46	0.36	71,207	0.48	-0.06 (0.00)
College Graduate or Greater	0.07	172,168	0.26	0.19	71,207	0.40	-0.12 (0.00)
Household Characteristics							
Maternal Marital Status	0.37	173,696	0.48	0.58	71,890	0.49	-0.21 (0.00)
Household Size	3.42	174,079	1.74	3.79	72,054	1.73	-0.37 (0.00)
Household Income - Midpoint	16671	174,079	11404	24748	72,054	14631	-8076 (0.00)
Maternal Race - White	0.49	168,825	0.50	0.60	70,026	0.49	-0.11 (0.00)
Maternal Race - Black	0.26	168,825	0.44	0.17	70,026	0.38	0.09 (0.00)
Maternal Ethnicity - Hispanic	0.25	168,763	0.43	0.16	69,940	0.37	0.09 (0.00)

Notes: This table reports summary statistics for infant–mother pairs in PRAMS (2009–2022), restricted to income-eligible households ($\leq$ 185% FPL). Columns list variable-specific means, sample sizes (Obs), and standard deviations (SD). Differences in means are defined as WIC $-$ Non-WIC; t-test p-values are shown in parentheses. Breastfeeding duration is measured in weeks; exclusive duration is the number of weeks until any non-breastmilk liquid or solids are introduced. These are descriptive associations and should not be interpreted causally.

Table 3: Partial Identification Regression LATE Estimates

Panel A: Breastfeeding Initiation					
Bandwidth:	9.97	14.95	19.93	24.91	29.90
Ref. bins [L/R]	[-39,-29]/[29,39]	[-44,-29]/[29,44]	[-59,-39]/[39,59]	[-49,-24]/[49,74]	[-59,-29]/[59,89]
Specification:					
Mean Slope	-0.151, -0.051	-0.141, -0.043	-0.268, -0.063	-0.292, -0.066	-0.266, -0.058
	(-0.214, -0.020)	(-0.190, -0.022)	(-0.322, -0.035)	(-0.390, -0.039)	(-0.365, -0.033)
Bin Slope	-0.151, -0.051	-0.141, -0.043	-0.268, -0.063	-0.292, -0.066	-0.266, -0.058
	(-0.214, -0.020)	(-0.190, -0.022)	(-0.322, -0.035)	(-0.390, -0.039)	(-0.365, -0.033)
Nonlinear Projection	-0.191, -0.030	-0.198, -0.035	-0.432, -0.065	-0.402, -0.065	-0.495, -0.054
	(-0.362, 0.025)	(-0.364, -0.003)	(-0.999, -0.021)	(-1.235, -0.019)	(-2.839, -0.009)
Simulation	-0.041	-0.053	-0.055	-0.056	-0.058
	(0.065)	(0.041)	(0.029)	(0.023)	(0.018)
	[61.3]	[150.5]	[288.1]	[473.2]	[733.7]
Panel B: Breastfeeding Duration (Weeks)					
Bandwidth:	14.76	22.13	29.51	36.89	44.27
Ref. bins [L/R]	[-44,-29]/[29,44]	[-66,-44]/[44,66]	[-59,-29]/[59,88]	[-73,-36]/[73,110]	[-88,-44]/[88,132]
Specification:					
Mean Slope	-8.25, -2.45	-10.51, -1.10	-12.25, -2.36	-23.58, -0.83	-56.72, -1.38
	(-9.62, -2.43)	(-12.65, -1.13)	(-15.92, -2.13)	(-91.53, -0.49)	(-492.27, -1.07)
Bin Slope	-8.25, -2.45	-10.51, -1.10	-12.25, -2.36	-23.58, -0.83	-56.72, -1.38
	(-9.62, -2.43)	(-12.65, -1.13)	(-15.92, -2.13)	(-91.53, -0.49)	(-492.27, -1.07)
Nonlinear Projection	-10.74, -1.98	-16.50, -1.01	-22.43, -2.12	-31.33, -0.85	·
	(-16.90, -1.32)	(-56.67, -0.21)	(-127.10, -0.67)	(-661.69, 4.55)	(·)
Simulation	-3.186	-3.287	-3.219	-3.096	-3.060
	(1.304)	(0.824)	(0.591)	(0.438)	(0.342)
	[148.9]	[369.2]	[717.3]	[1302.8]	[2153.4]

Notes: This table reports bounds on the local average treatment effect (LATE) of WIC participation on breastfeeding outcomes using the partial identification regression discontinuity design (Assumptions 1–3). Bootstrap 95% confidence intervals (1,000 replications) appear in parentheses. Confidence intervals are reported only when the first-stage lower bound excludes zero in at least 500 replications. Panel A presents breastfeeding initiation; Panel B presents breastfeeding duration (weeks). Column headers indicate bandwidth h with corresponding reference bin ranges $[b_h^-]/[b_h^+]$ from which Assumption 2 is imposed. The bounded rate of change (Assumption 3) is estimated via three approaches using the slope of the conditional expectation function, $\partial \hat{h}_z(x)/\partial x$: (i) *Mean Slope* uses derivatives averaged over the estimation window; (ii) *Bin Slope* uses derivatives evaluated at midpoints of $[b_h^-]$ and $[b_h^+]$; (iii) *Nonlinear Projection* directly projects the estimated conditional expectation function. All implementations use a slope multiplier of one. The *Simulation* row presents simulation-based RDD estimates with 95% confidence intervals in parentheses and mean first-stage F-statistics in brackets. See Section 4 for details.

Table 4: Partial Identification: Discontinuities by Variable and Slope Multiple

Panel A: Breastfeeding Initiation					
Bandwidth:	9.97	14.95	19.93	24.91	29.90
Ref. bins [L/R]	[-39,-29]/[29,39]	[-44,-29]/[29,44]	[-59,-39]/[39,59]	[-49,-24]/[49,74]	[-59,-29]/[59,89]
WIC					
Slope × 1.5	0.052, 0.336	0.042, 0.348	-0.017, 0.380	-0.031, 0.369	-0.036, 0.419
Slope × 2.0	-0.042, 0.336	-0.060, 0.348	-0.149, 0.380	-0.164, 0.369	-0.188, 0.419
Initiation					
Slope × 1.5	-0.022, -0.015	-0.020, -0.012	-0.031, -0.021	-0.030, -0.021	-0.031, -0.021
Slope × 2.0	-0.022, -0.012	-0.020, -0.009	-0.031, -0.017	-0.030, -0.018	-0.031, -0.018
Panel B: Breastfeeding Duration (Weeks)					
Bandwidth:	14.76	22.13	29.51	36.89	44.27
Ref. bins [L/R]	[-44,-29]/[29,44]	[-66,-44]/[44,66]	[-59,-29]/[59,88]	[-73,-36]/[73,110]	[-88,-44]/[88,132]
WIC					
Slope × 1.5	0.046, 0.348	-0.057, 0.386	-0.036, 0.415	-0.147, 0.403	-0.206, 0.482
Slope × 2.0	-0.054, 0.348	-0.205, 0.386	-0.186, 0.415	-0.330, 0.403	-0.436, 0.482
Duration					
Slope × 1.5	-1.21, -0.67	-0.95, -0.16	-1.40, -0.77	-0.86, -0.07	-1.31, -0.34
Slope × 2.0	-1.21, -0.49	-0.95, 0.10	-1.40, -0.56	-0.86, 0.19	-1.31, -0.02

Notes: This table reports bounds on the partial treatment effects, i.e. discontinuities in treatment and outcomes, at the WIC income eligibility threshold using using my partial identification design (Assumptions 1–3). Panel A presents bounds on the discontinuities for WIC participation (τ_T) and breastfeeding initiation (τ_Y) at MSE-selected bandwidths for initiation; Panel B presents bounds on the discontinuities for WIC participation (τ_T) breastfeeding duration (τ_Y) at MSE-selected bandwidths using duration. Column headers indicate bandwidth h with corresponding reference bin ranges $[b_h^-]/[b_h^+]$ from which Assumption 2 is imposed. Rows indicate the slope multipliers used for Assumption 3.

Table 5: Partial Identification LATE Estimates: Robustness Across Slope Multiples

Panel A: Breastfeeding Initiation					
Bandwidth:	9.97	14.95	19.93	24.91	29.90
Ref. bins [L/R]	[-39,-29]/[29,39]	[-44,-29]/[29,44]	[-59,-39]/[39,59]	[-49,-24]/[49,74]	[-59,-29]/[59,89]
Slope Multiplier:					
Slope $\times$ 1.0	-0.151, -0.051	-0.141, -0.043	-0.268, -0.063	-0.292, -0.066	-0.266, -0.058
	(-0.214, -0.020)	(-0.190, -0.022)	(-0.322, -0.035)	(-0.390, -0.039)	(-0.365, -0.033)
Slope $\times$ 1.5	-0.151, -0.044	-0.141, -0.035	-0.268, -0.054	-0.292, -0.058	-0.266, -0.051
	(-0.214, -0.006)	(-0.190, -0.005)	(-0.322, -0.014)	(-0.390, -0.020)	(-0.365, -0.013)
Slope $\times$ 2.0	-0.151, -0.036	-0.141, -0.027	-0.268, -0.045	-0.292, -0.050	-0.266, -0.043
	(-0.214, 0.020)	(-0.190, 0.027)	(-0.322, 0.021)	(-0.390, 0.008)	(-0.365, 0.023)
Panel B: Breastfeeding Duration (Weeks)					
Bandwidth:	14.76	22.13	29.51	36.89	44.27
Ref. bins [L/R]	[-44,-29]/[29,44]	[-66,-44]/[44,66]	[-59,-29]/[59,88]	[-73,-36]/[73,110]	[-88,-44]/[88,132]
Slope Multiplier:					
Slope $\times$ 1.0	-8.25, -2.45	-10.51, -1.10	-12.25, -2.36	-23.58, -0.83	-56.72, -1.38
	(-9.62, -2.43)	(-12.65, -1.13)	(-15.92, -2.13)	(-91.53, -0.49)	(-492.27, -1.07)
Slope $\times$ 1.5	-8.25, -1.94	-10.51, -0.41	-12.25, -1.85	-23.58, -0.18	-56.72, -0.71
	(-9.62, -2.15)	(-12.65, -0.51)	(-15.92, -1.48)	(-91.53, 2.20)	(-492.27, -0.33)
Slope $\times$ 2.0	-8.25, -1.42	-10.51, 1.15	-12.25, -1.34	-23.58, 5.25	-56.72, -0.04
	(-9.62, -1.76)	(-12.65, 0.64)	(-15.92, -0.87)	(-91.53, 8.64)	(-492.27, 6.50)

Notes: This table reports bounds on the local average treatment effect (LATE) of WIC participation on breastfeeding outcomes using the partial identification regression discontinuity design (Assumptions 1–3). Bootstrap 95% confidence intervals (1,000 replications) appear in parentheses. Confidence intervals are reported only when the first-stage lower bound excludes zero in at least 500 replications. Panel A presents breastfeeding initiation; Panel B presents breastfeeding duration (weeks). Column headers indicate bandwidth h with corresponding reference bin ranges $[b_h^-]/[b_h^+]$ from which Assumption 2 is imposed. The bounded rate of change (Assumption 3) is estimated using the slope of the conditional expectation function, $\partial\hat{h}_z(x)/\partial x$, averaged over the estimation window and then scaled by slope multipliers $\in \{1.0, 1.5, 2.0\}$. For breastfeeding initiation and duration, the slope multiplier varies across rows. I hold the slope multiplier equal to 1.0 for WIC participation. See Section 4 for details.

Table 6: Traditional Regression Discontinuity Estimates

Panel A: Breastfeeding Initiation					
Bandwidth:	9.97	14.95	19.93	24.91	29.90
Polynomial Order:					
Zero	-0.090	-0.073	-0.054	-0.054	-0.054
	(0.029)	(0.022)	(0.021)	(0.024)	(0.021)
	[327.8]	[557.5]	[547.1]	[445.8]	[546.5]
One	-0.425	-0.208	-0.135	-0.103	-0.107
	(0.184)	(0.071)	(0.034)	(0.029)	(0.034)
	[12.0]	[57.0]	[230.4]	[303.2]	[228.4]
Two	7.691	-0.922	-0.514	-0.237	-0.144
	(78.319)	(0.828)	(0.337)	(0.078)	(0.039)
	[0.0]	[1.5]	[4.2]	[48.9]	[171.7]
Optimal Order	0	0	0	0	0
Observations	16,129	24,151	37,183	49,649	54,532

Panel B: Breastfeeding Duration (Weeks)					
Bandwidth:	14.76	22.13	29.51	36.89	44.27
Polynomial Order:					
Zero	-1.466	-0.653	-1.850	-3.274	-3.026
	(0.702)	(0.775)	(0.670)	(0.433)	(0.328)
	[541.8]	[424.7]	[555.2]	[1339.4]	[2368.1]
One	-0.292	-1.823	0.338	6.861	7.996
	(2.125)	(0.879)	(1.106)	(2.654)	(4.813)
	[57.6]	[326.6]	[213.5]	[48.0]	[15.7]
Two	-38.023	1.729	-3.562	-2.053	-0.091
	(34.819)	(8.586)	(1.159)	(0.954)	(1.039)
	[1.3]	[3.6]	[185.2]	[272.0]	[237.5]
Optimal Order	0	1	0	2	0
Observations	24,040	43,661	54,422	66,601	86,702

Notes: This table presents estimates of the treatment effect (τ) from WIC participation on the initiation and duration of breastfeeding using the traditional regression discontinuity design. Panel A shows results for breastfeeding initiation, and Panel B shows results for breastfeeding duration in weeks. Each column represents a different bandwidth (h), and each row represents a different polynomial order for the regression function. Standard errors are shown in parentheses. Observation counts reflect the number of observations i such that $X_i^\star \in [c-h, c+h]$ for bandwidth h. First-stage F-test values are shown in square brackets. The optimal polynomial order according to Akaike's Information Criterion is indicated for each bandwidth. Additional estimation details can be found in Section 4.

Table 7: Donut Regression Discontinuity LATE Estimates

Panel A: Breastfeeding Initiation					
Bandwidth:	16.86	25.28	33.71	42.14	50.57
Effective Bandwidth:	64.17	72.60	81.03	89.45	97.88
Polynomial Order:					
Zero	-0.094	-0.085	-0.086	-0.081	-0.082
	(0.009)	(0.006)	(0.005)	(0.004)	(0.004)
	[3277.7]	[7336.0]	[12688.7]	[17986.4]	[19342.0]
One	0.258	-0.692	0.049	-0.622	-0.133
	(0.314)	(0.895)	(0.231)	(0.411)	(0.046)
	[3.6]	[0.9]	[6.0]	[3.8]	[138.6]
Two	-0.181	-0.532	0.211	-0.299	-0.049
	(0.073)	(0.356)	(0.659)	(0.200)	(0.052)
	[56.6]	[4.2]	[0.9]	[9.7]	[112.9]
Optimal Order	0	0	0	0	2
Observations	27,970	45,924	62,964	90,399	104,579

Panel B: Breastfeeding Duration (Weeks)					
Bandwidth:	2.14	3.21	4.28	5.36	6.43
Effective Bandwidth:	49.46	50.53	51.60	52.67	53.74
Polynomial Order:					
Zero	-4.318	-4.015	-3.923	-4.148	-4.661
	(0.818)	(0.856)	(0.801)	(0.785)	(0.613)
	[446.3]	[383.8]	[433.0]	[453.4]	[770.2]
One	-13.439	-0.033	-5.101	4.343	-10.353
	(5.323)	(5.998)	(5.527)	(8.889)	(3.659)
	[15.9]	[7.6]	[8.8]	[4.1]	[27.7]
Two	-4.457	-9.556	-9.643	-42.640	-3.484
	(23.050)	(2.400)	(4.041)	(157.568)	(3.929)
	[0.5]	[54.8]	[19.6]	[0.1]	[17.4]
Optimal Order	0	0	0	0	2
Observations	5,628	6,780	8,318	8,653	9,637

Notes: This table presents estimates of the treatment effect (τ) from WIC participation on the initiation and duration of breastfeeding using the donut regression discontinuity design. Panel A shows results for breastfeeding initiation, and Panel B shows results for breastfeeding duration in weeks. Each column represents a different bandwidth (h), and each row represents a different polynomial order for the regression function. Effective bandwidth is the sum of the donut hole (μ) and the estimation bandwidth (h). Standard errors are shown in parentheses. Observation counts reflect the number of observations i such that $X_i^\star \in [c-\mu-h, c-\mu] \cup [c+\mu, c+\mu+h]$ for bandwidth h. First-stage F-test values are shown in square brackets. The optimal polynomial order according to Akaike's Information Criterion is indicated for each bandwidth. Additional estimation details can be found in Section 4.

Table 8: Histogram Density Test

Bandwidth	θ	Standard Error	P-Value	Density Ratio	Count Left	Count Right
14.76	0.158	0.710	0.826	1.171	12,356	11,816
22.13	0.196	0.567	0.732	1.216	23,693	19,949
29.51	0.022	0.507	0.966	1.022	28,375	26,157
36.89	-0.108	0.472	0.819	0.897	35,380	31,244
44.27	-0.111	0.434	0.799	0.895	43,227	43,096

Notes: I implement a McCrary-style density test for manipulation at the WIC eligibility cutoff. For each bin width b and window $h = m \cdot h_{\text{MSE}}$, where h_{MSE} is the duration-specific MSE-optimal bandwidth from the main RDD and $m \in \{0.5, 0.75, 1.0, 1.25, 1.5\}$), I construct histograms aligned at the cutoff. I use these histograms to fit a regression of $\log((\text{count} + \varepsilon)/b)$, with zero-count correction $\varepsilon = 0.5$, using side-specific pooled first-order polynomials in $(x - c)/h$. Estimation uses triangular kernel weights within h, and drops bins with zero weight. I report the estimated log-density jump $\hat{\delta}$, its standard error, the p-value for $H_0 : \delta = 0$, and the implied density ratio $\exp(\hat{\delta})$. Fits require minimum effective bins and counts on each side; and if the requested polynomial order fails, the regression falls back to order 0.

Table 9: Covariate Balance Table

	Control Mean	Coefficient	Standard Error	P-Value	SMD
Maternal Health Characteristics					
Maternal BMI	27.345	1.080	0.150	0.000	0.062
Maternal Diabetes	0.066	0.003	0.005	0.539	0.015
Preterm Labor	0.226	-0.010	0.015	0.514	0.011
Infant Outcomes					
Infant Birth Weight (grams)	3128.768	-79.086	15.115	0.000	-0.044
Small for Gestational Age	0.134	0.019	0.007	0.004	0.023
Employment Status					
Employment Before Pregnancy	0.645	0.015	0.018	0.408	-0.018
Employment During Pregnancy	0.742	0.115	0.020	0.000	-0.047
Returned to Job After Birth	0.612	0.021	0.028	0.448	-0.000
Educational Attainment					
Less than High School	0.008	0.005	0.002	0.012	0.049
Some High School	0.037	0.019	0.004	0.000	0.080
High School Graduate	0.213	0.085	0.008	0.000	0.116
Some College	0.381	0.096	0.010	0.000	0.062
College Graduate or More	0.362	-0.205	0.009	0.000	-0.220
Demographics					
Maternal Marital Status	0.697	-0.417	0.009	0.000	-0.152
Household Size	3.861	-1.599	0.033	0.000	-0.111
Maternal Race - White	0.670	-0.171	0.010	0.000	-0.099
Maternal Race - Black	0.148	0.112	0.007	0.000	0.069
Maternal Ethnicity - Hispanic	0.121	0.069	0.007	0.000	0.095

Notes: Covariate-balance regressions at the WIC income cutoff (185% of the FPL). Estimates use the primary symmetric bandwidth $h = 29.51$ FPL points. For each covariate, I regress the variable on an indicator for the eligible side ($\leq$ 185% FPL) and a linear function of the running variable with side-specific slopes. The reported coefficient is the estimated discontinuity at the cutoff, and the *Control Mean* is the mean just above the cutoff. Standard errors are heteroskedasticity-robust. *SMD* is the signed standardized mean difference across sides (difference in means divided by the pooled standard deviation).

Table 10: Comparison of Methods: Partial Identification vs. Conventional Benchmarks

Panel A: Breastfeeding Initiation					
Bandwidth:	9.97	14.95	19.93	24.91	29.90
Ref. bins [L/R]	[-39,-29]/[29,39]	[-44,-29]/[29,44]	[-59,-39]/[39,59]	[-49,-24]/[49,74]	[-59,-29]/[59,89]
Specification:					
Partial ID	-0.151, -0.051	-0.141, -0.043	-0.268, -0.063	-0.292, -0.066	-0.266, -0.058
	(-0.214, -0.020)	(-0.190, -0.022)	(-0.322, -0.035)	(-0.390, -0.039)	(-0.365, -0.033)
Traditional	-0.090	-0.073	-0.054	-0.054	-0.054
	(0.029)	(0.022)	(0.021)	(0.024)	(0.021)
	[327.8]	[557.5]	[547.1]	[445.8]	[546.5]
Doughnut	-0.096	-0.099	-0.063	-0.088	-0.005
	(0.014)	(0.011)	(0.133)	(0.076)	(0.182)
	[1511.9]	[2372.2]	[15.1]	[4998.1]	[9.39]
Panel B: Breastfeeding Duration (Weeks)					
Bandwidth:	14.76	22.13	29.51	36.89	44.27
Ref. bins [L/R]	[-44,-29]/[29,44]	[-66,-44]/[44,66]	[-59,-29]/[59,88]	[-73,-36]/[73,110]	[-88,-44]/[88,132]
Specification:					
Partial ID	-8.25, -2.45	-10.51, -1.10	-12.25, -2.36	-23.58, -0.83	-56.72, -1.38
	(-9.62, -2.43)	(-12.65, -1.13)	(-15.92, -2.13)	(-91.53, -0.49)	(-492.27, -1.07)
Traditional	-1.466	-1.823	-1.850	-2.053	-3.026
	(0.702)	(0.879)	(0.670)	(0.954)	(0.328)
	[541.8]	[326.6]	[555.2]	[272.0]	[2368.1]
Doughnut	-2.624	-2.764	-3.225	-3.451	-3.161
	(0.351)	(0.281)	(0.215)	(0.180)	(0.157)
	[2374.1]	[3737.9]	[6692.9]	[9878.9]	[13366.8]

Notes: This table reports AIC-selected bounds and point-estimates for the local average treatment effect (LATE) of WIC participation on breastfeeding outcomes using the novel partial identification regression discontinuity design (Assumptions 1–3) and conventional RDD approaches. For the partial identification results, bootstrap 95% confidence intervals (1,000 replications) appear in parentheses. The bounded rate of change (Assumption 3) is estimated by the slope of the conditional expectation function, $\partial\hat{h}_z(x)/\partial x$, averaged over the estimation window, with no subsequent scaling (i.e., a slope multiplier of 1.0). For the conventional (traditional and donut) RDD approaches, standard errors are shown in parentheses, and first-stage F-test values are shown in square brackets. Column headers indicate bandwidth h with corresponding reference bin ranges $[b_h^-]/[b_h^+]$ from which Assumption 2 is imposed. Panel A presents breastfeeding initiation; Panel B presents breastfeeding duration (weeks).

A Supplementary Tables

Table 11: Summary Statistics by FPL Interval

Variable	Bounds Overlap 185%			Bounds Do Not Overlap			Diff (P-value)
	Mean	Obs	SD	Mean	Obs	SD	
Infant Feeding Practices							
Breastfeeding Initiation	0.93	16,773	0.25	0.89	216,308	0.32	0.05 (0.000)
Breastfeeding Duration - Any	12.97	14,348	6.74	12.27	168,178	7.11	0.70 (0.000)
Breastfeeding Duration - Exclusive	12.36	5,987	6.47	11.81	70,515	6.65	0.55 (0.000)
Formula Duration - Any	43.78	8,985	7.69	44.57	112,145	7.74	-0.80 (0.000)
Formula Duration - Exclusive	41.00	16,899	7.68	42.62	220,047	8.10	-1.62 (0.000)
Infant and Maternal Health							
Maternal BMI	27.61	14,402	7.05	27.91	191,553	7.37	-0.30 (0.000)
Maternal Diabetes	0.06	17,376	0.24	0.07	228,983	0.25	-0.00 (0.319)
Preterm Labor	0.23	6,421	0.42	0.28	78,417	0.45	-0.06 (0.000)
Infant Birth Weight - Grams	3084	17,400	777.50	3018	229,425	774.49	66 (0.000)
Small for Gestational Age	0.15	16,761	0.35	0.17	221,718	0.37	-0.02 (0.000)
Maternal Education and Employment							
Employment Before Pregnancy	0.64	5,321	0.48	0.61	55,888	0.49	0.03 (0.000)
Employment During Pregnancy	0.74	4,398	0.44	0.56	61,506	0.50	0.17 (0.000)
Returned to Job After Birth	0.62	2,640	0.49	0.50	28,334	0.50	0.12 (0.000)
Less than High School	0.01	17,268	0.11	0.05	227,117	0.21	-0.03 (0.000)
Some High School	0.05	17,268	0.22	0.17	227,117	0.38	-0.12 (0.000)
High School Graduate	0.27	17,268	0.44	0.37	227,117	0.48	-0.10 (0.000)
Some College	0.42	17,268	0.49	0.31	227,117	0.46	0.10 (0.000)
College Graduate or Greater	0.25	17,268	0.43	0.10	227,117	0.30	0.15 (0.000)
Household Characteristics							
Maternal Marital Status	0.62	17,405	0.49	0.41	229,293	0.49	0.20 (0.000)
Household Size	3.41	17,435	1.45	3.54	229,822	1.77	-0.13 (0.000)
Household Income - Midpoint	37455	17,435	10750	17636	229,822	12017	19819 (0.000)
Maternal Race - White	0.62	16,781	0.49	0.51	223,113	0.50	0.10 (0.000)
Maternal Race - Black	0.17	16,781	0.38	0.24	223,113	0.43	-0.07 (0.000)
Maternal Ethnicity - Hispanic	0.15	16,734	0.36	0.23	223,028	0.42	-0.08 (0.000)

Notes: This table compares income-eligible PRAMS respondents (2009–2022) whose reported FPL interval overlaps the 185% FPL eligibility threshold—$[\ell_i, u_i]$ such that $\ell_i \leq 185\% \leq u_i$—to those entirely below or above 185% FPL. Columns list variable-specific means, sample sizes (Obs), and standard deviations (SD). Differences are defined as Overlap $-$ No overlap, and t-test p-values are shown in parentheses. Breastfeeding duration is measured in weeks, and exclusive duration is the number of weeks until any non-breastmilk liquid or solids are introduced. These are descriptive associations and should not be interpreted causally.

Table 12: Partial Identification Regression LATE Estimates: Mean Slope

Panel A: Breastfeeding Initiation					
Bandwidth:	9.97	14.95	19.93	24.91	29.90
Ref. bins [L/R]	[-39,-29]/[29,39]	[-44,-29]/[29,44]	[-59,-39]/[39,59]	[-49,-24]/[49,74]	[-59,-29]/[59,89]
Polynomial Order:					
One	-0.151, -0.051	-0.141, -0.043	-0.268, -0.063	-0.292, -0.066	-0.266, -0.058
	(-0.214, -0.020)	(-0.190, -0.022)	(-0.322, -0.035)	(-0.390, -0.039)	(-0.365, -0.033)
Two	-0.181, -0.010	-0.172, 0.000	-0.374, -0.011	-0.358, 0.031	-0.318, 0.074
	(-0.211, -0.025)	(-0.187, -0.024)	(-0.320, -0.037)	(-0.382, -0.039)	(-0.348, -0.033)
Panel B: Breastfeeding Duration (Weeks)					
Bandwidth:	14.76	22.13	29.51	36.89	44.27
Ref. bins [L/R]	[-44,-29]/[29,44]	[-66,-44]/[44,66]	[-59,-29]/[59,88]	[-73,-36]/[73,110]	[-88,-44]/[88,132]
Polynomial Order:					
One	-8.25, -2.45	-10.51, -1.10	-12.25, -2.36	-23.58, -0.83	-56.72, -1.38
	(-9.62, -2.43)	(-12.65, -1.13)	(-15.92, -2.13)	(-91.53, -0.49)	(-492.27, -1.07)
Two	-10.00, 0.81	-18.15, 18.58	-14.72, 3.58	-76.10, 113.77	·
	(-9.46, -2.27)	(-12.49, -0.82)	(-15.47, -1.87)	(-76.08, -0.23)	(·)

Notes: This table reports bounds on the local average treatment effect (LATE) of WIC participation on breastfeeding outcomes using the partial identification regression discontinuity design (Assumptions 1–3). Bootstrap 95% confidence intervals (1,000 replications) appear in parentheses. Confidence intervals are reported only when the first-stage lower bound excludes zero in at least 500 replications. Panel A presents breastfeeding initiation; Panel B presents breastfeeding duration (weeks). Column headers indicate bandwidth h with corresponding reference bin ranges $[b_h^-]/[b_h^+]$ from which Assumption 2 is imposed. The bounded rate of change (Assumption 3) is estimated using the slope of the conditional expectation function, $\partial\hat{h}_z(x)/\partial x$, averaged over the estimation window. All specifications use a slope multiplier of one. See Section 4 for details.

Table 13: Partial Identification Regression Discontinuity LATE Estimates: Reference Bin Slope

Panel A: Breastfeeding Initiation					
Bandwidth:	9.97	14.95	19.93	24.91	29.90
Ref. bins [L/R]	[-39,-29]/[29,39]	[-44,-29]/[29,44]	[-59,-39]/[39,59]	[-49,-24]/[49,74]	[-59,-29]/[59,89]
Polynomial Order:					
One	-0.151, -0.051	-0.141, -0.043	-0.268, -0.063	-0.292, -0.066	-0.266, -0.058
	(-0.214, -0.020)	(-0.190, -0.022)	(-0.322, -0.035)	(-0.390, -0.039)	(-0.365, -0.033)
Two	-0.134, -0.046	-0.116, -0.035	-0.152, -0.043	-1.081, -0.056	·
	(-0.341, 0.094)	(-0.405, 0.168)	(-6.576, 2.967)	(·)	(·)
Panel B: Breastfeeding Duration (Weeks)					
Bandwidth:	14.76	22.13	29.51	36.89	44.27
Ref. bins [L/R]	[-44,-29]/[29,44]	[-66,-44]/[44,66]	[-59,-29]/[59,88]	[-73,-36]/[73,110]	[-88,-44]/[88,132]
Polynomial Order:					
One	-8.25, -2.45	-10.51, -1.10	-12.25, -2.36	-23.58, -0.83	-56.72, -1.38
	(-9.62, -2.43)	(-12.65, -1.13)	(-15.92, -2.13)	(-91.53, -0.49)	(-492.27, -1.07)
Two	-6.75, -3.48	-4.96, -2.46	-245.32, -3.38	·	·
	(-19.42, -2.38)	(·)	(·)	(·)	(·)

Notes: This table reports bounds on the local average treatment effect (LATE) of WIC participation on breastfeeding outcomes using the partial identification regression discontinuity design (Assumptions 1–3). Bootstrap 95% confidence intervals (1,000 replications) appear in parentheses. Confidence intervals are reported only when the first-stage lower bound excludes zero in at least 500 replications. Panel A presents breastfeeding initiation; Panel B presents breastfeeding duration (weeks). Column headers indicate bandwidth h with corresponding reference bin ranges $[b_h^-]/[b_h^+]$ from which Assumption 2 is imposed. The bounded rate of change (Assumption 3) is estimated using the slope of the conditional expectation function, $\partial\hat{h}_z(x)/\partial x$, evaluated at midpoints of $[b_h^-]$ and $[b_h^+]$. All specifications use a slope multiplier of one. See Section 4 for details.

Table 14: Partial Identification Regression Discontinuity LATE Estimates: Nonlinear Projection

Panel A: Breastfeeding Initiation					
Bandwidth:	9.97	14.95	19.93	24.91	29.90
Ref. bins [L/R]	[-39,-29]/[29,39]	[-44,-29]/[29,44]	[-59,-39]/[39,59]	[-49,-24]/[49,74]	[-59,-29]/[59,89]
Polynomial Order:					
One	-0.191, -0.030	-0.198, -0.035	-0.432, -0.065	-0.402, -0.065	-0.495, -0.054
	(-0.362, 0.025)	(-0.364, -0.003)	(-0.999, -0.021)	(-1.235, -0.019)	(-2.839, -0.009)
Two	-0.307, 0.344	-0.334, 0.377	-0.718, -0.018	$\cdot$	$\cdot$
	(-0.315, 0.078)	(-0.335, 0.059)	(-1.162, -0.025)	$(\cdot)$	$(\cdot)$
Panel B: Breastfeeding Duration (Weeks)					
Bandwidth:	14.76	22.13	29.51	36.89	44.27
Ref. bins [L/R]	[-44,-29]/[29,44]	[-66,-44]/[44,66]	[-59,-29]/[59,88]	[-73,-36]/[73,110]	[-88,-44]/[88,132]
Polynomial Order:					
One	-10.74, -1.98	-16.50, -1.01	-22.43, -2.12	-31.33, -0.85	$\cdot$
	(-16.90, -1.32)	(-56.67, -0.21)	(-127.10, -0.67)	(-661.69, 4.55)	$(\cdot)$
Two	-17.04, 20.78	-19.66, -0.20	$\cdot$	$\cdot$	$\cdot$
	(-15.85, 1.11)	(-148.77, -0.77)	$(\cdot)$	$(\cdot)$	$(\cdot)$

Notes: This table reports bounds on the local average treatment effect (LATE) of WIC participation on breastfeeding outcomes using the partial identification regression discontinuity design (Assumptions 1–3). Bootstrap 95% confidence intervals (1,000 replications) appear in parentheses. Confidence intervals are reported only when the first-stage lower bound excludes zero in at least 500 replications. Panel A presents breastfeeding initiation; Panel B presents breastfeeding duration (weeks). Column headers indicate bandwidth h with corresponding reference bin ranges $[b_h^-]/[b_h^+]$ from which Assumption 2 is imposed. The bounded rate of change (Assumption 3) is estimated using the estimated conditional expectation function, which is directly projected to the threshold from the reference midpoints of $[b_h^-]$ and $[b_h^+]$. All specifications use a slope multiplier of one. See Section 4 for details.

Table 15: Simulation Regression Discontinuity Estimates

Panel A: Breastfeeding Initiation					
Bandwidth:	9.97	14.95	19.93	24.91	29.90
Polynomial Order:					
Zero	-0.041	-0.053	-0.055	-0.056	-0.058
	(0.065)	(0.041)	(0.029)	(0.023)	(0.018)
	[61.3]	[150.5]	[288.1]	[473.2]	[733.7]
One	0.496	-0.071	0.642	-0.028	-0.017
	(328.829)	(84.781)	(1474.011)	(0.292)	(0.171)
	[1.4]	[3.1]	[5.7]	[9.8]	[12.5]
Two	-0.008	0.010	0.128	0.083	0.038
	(547.613)	(111.384)	(193.148)	(383.705)	(47.082)
	[1.0]	[1.0]	[1.2]	[1.5]	[2.7]
Optimal Order	0	0	0	0	0
Observations	15,260	22,672	30,090	37,550	45,093

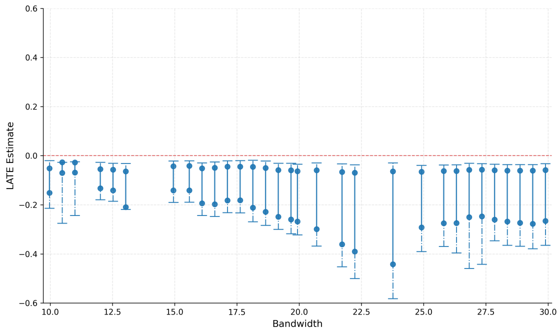
Panel B: Breastfeeding Duration (Weeks)					
Bandwidth:	14.76	22.13	29.51	36.89	44.27
Polynomial Order:					
Zero	-3.186	-3.287	-3.219	-3.096	-3.060
	(1.304)	(0.824)	(0.591)	(0.438)	(0.342)
	[148.9]	[369.2]	[717.3]	[1302.8]	[2153.4]
One	-18.877	-2.584	-3.385	-4.234	-4.104
	(7812.942)	(25.503)	(5.275)	(4.788)	(4.176)
	[2.9]	[7.5]	[12.9]	[14.5]	[17.5]
Two	28.428	69.343	5.799	1.196	-2.052
	(25209.891)	(107570.505)	(1616.146)	(1690.631)	(173.998)
	[1.0]	[1.2]	[2.5]	[5.1]	[6.5]
Optimal Order	0	0	0	0	0
Observations	22,492	33,552	44,741	56,375	68,421

Notes: This table presents estimates of the treatment effect (τ) from WIC participation on the initiation and duration of breastfeeding using the simulation regression discontinuity design. Each column represents a different bandwidth (h), and each row represents a different polynomial order for the regression function. Panel A shows results for breastfeeding initiation, and Panel B shows results for breastfeeding duration measured in weeks. Standard errors are shown in parentheses. Observation counts reflect the mean number of observations i across simulations j such that $X^{\star}_{i,j} \in [c-h, c] \cup [c, c+h]$ for bin-width h. The optimal polynomial order according to Akaike's Information Criterion is indicated for each bandwidth. Additional estimation details can be found in Section 4.

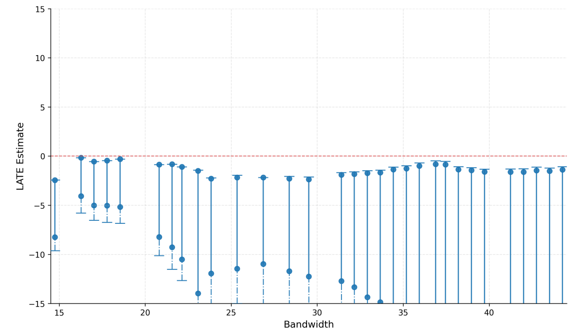
B Supplementary Figures

Figure 14: LATE Bounds from Partial Identification RDD: Reference Bin Slope

Panel A: Slope Multiplier - 1.0

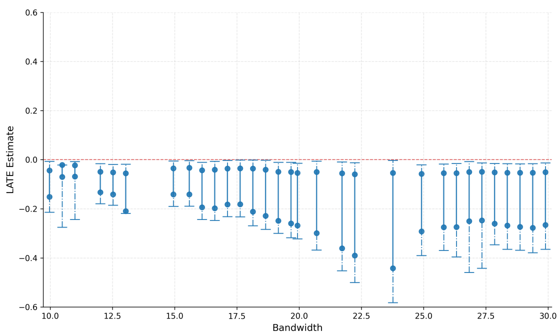


(a) Breastfeeding Initiation

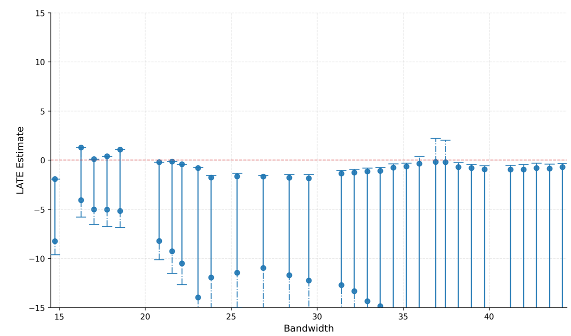


(b) Breastfeeding Duration (weeks)

Panel B: Slope Multiplier - 1.5

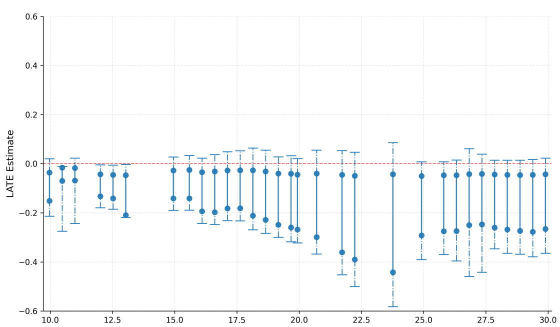


(c) Breastfeeding Initiation

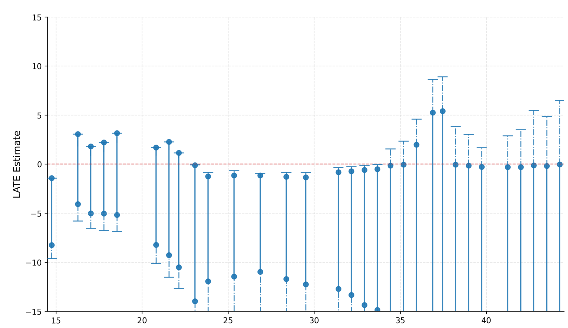


(d) Breastfeeding Duration (weeks)

Panel C: Slope Multiplier - 2.0



(e) Breastfeeding Initiation

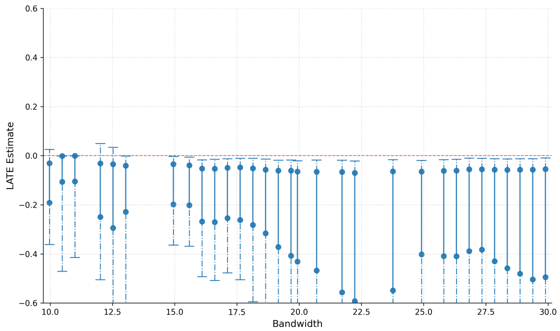


(f) Breastfeeding Duration (weeks)

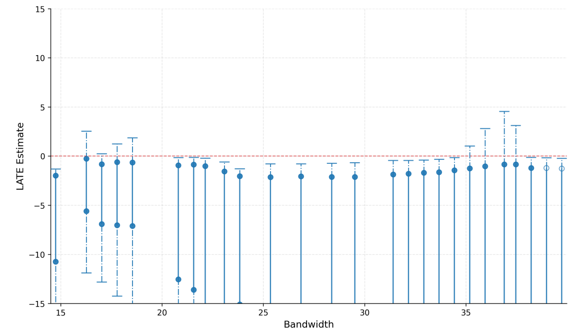
Notes: This figure reports bounds on the local average treatment effect (LATE) and the associated 95% confidence intervals for the effect of WIC participation on breastfeeding outcomes. Estimates are derived using the partial-identification regression discontinuity design (RDD) and Assumptions 1–3. The bounded rate of change (Assumption 3) is computed using a linearized slope estimate for the conditional expectation function at the midpoint of the reference bin. Panels A–C correspond to slope multipliers $\{1.0, 1.5, 2.0\}$. See Section 4 for estimation details.

Figure 15: LATE Bounds from Partial Identification RDD: Nonlinear Projection

Panel A: Slope Multiplier - 1.0

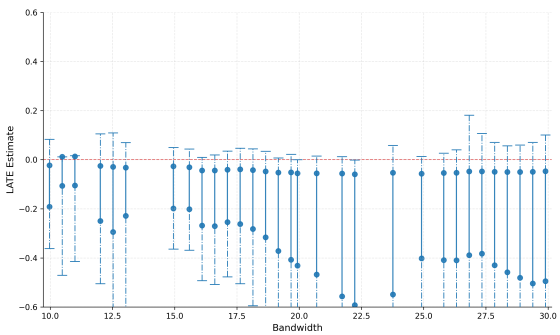


(a) Breastfeeding Initiation

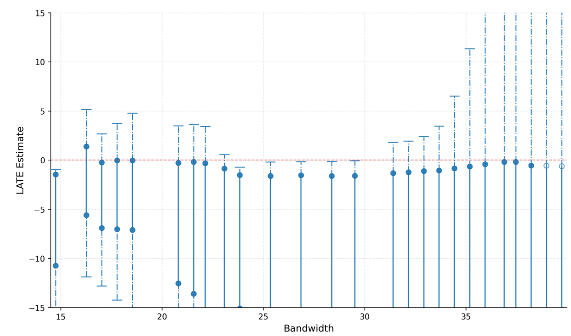


(b) Breastfeeding Duration (weeks)

Panel B: Slope Multiplier - 1.5

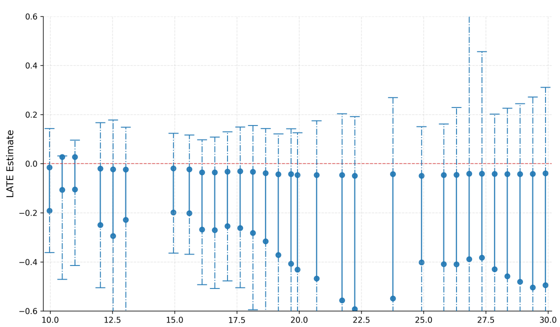


(c) Breastfeeding Initiation

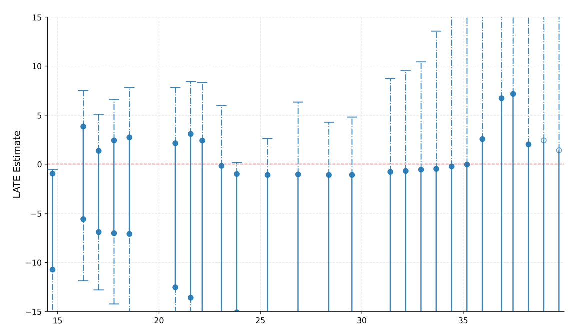


(d) Breastfeeding Duration (weeks)

Panel C: Slope Multiplier - 2.0



(e) Breastfeeding Initiation



(f) Breastfeeding Duration (weeks)

Notes: This figure reports bounds on the local average treatment effect (LATE) and the associated 95% confidence intervals for the effect of WIC participation on breastfeeding outcomes. Estimates are derived using the partial-identification regression discontinuity design (RDD) and Assumptions 1–3. The bounded rate of change (Assumption 3) is computed using a nonlinear projection of the estimated conditional expectation function. Panels A–C correspond to slope multipliers $\{1.0, 1.5, 2.0\}$. See Section 4 for estimation details.

C Empirical Appendix

C.1 Variable Construction (PRAMS)

C.1.1 Income

PRAMS reports income categorically. I harmonize the phase-specific income variables, e.g., 8/7/5-level codes, to lower and upper dollar bounds for each respondent, then merge U.S. Department of Health and Human Services (HHS) poverty guidelines by year and state group (contiguous United States, Alaska, Hawaii). This yields household income as a percentage of the Federal Poverty Level (FPL) and its reported interval $[\ell_i, u_i]$ for each observation. The midpoint, $(\ell_i + u_i)/2$, serves as the proxy for the running variable in the analysis.

C.1.2 Breastfeeding outcomes

Breastfeeding duration (weeks) is constructed from PRAMS items with special-case handling: non-breastfeeders and "never breastfed" codes are set to zero; "still breastfeeding" at interview is set to the infant's age in weeks; in all other cases, the reported weeks are used. I flag, but do not replace, durations that exceed the infant's age and those above one year. Initiation is coded as one if either PRAMS indicator records breastfeeding, or if a positive, non-missing duration is observed. Initiation is coded as zero only when both initiation indicators explicitly record "no."

C.1.3 WIC participation

Prenatal WIC participation is taken from the birth certificate linkage in PRAMS (not retrospective self-report), which offers better concordance than CPS-ASEC self-reports. This is the treatment indicator in the analysis window.

C.2 Bandwidth Selection (RDD)

I determine the appropriate estimation window around the WIC eligibility cutoff using a mean–squared–error (MSE) optimal bandwidth selector from Calonico et al. (2014). The selector targets the discontinuity parameter and chooses a window h_{MSE} that balances variance and squared bias in a local polynomial estimator at the cutoff. Intuitively, enlarging h reduces variance by increasing effective sample size but raises bias because observations farther from the cutoff are less comparable, while shrinking h has the opposite effect. The MSE criterion formalizes this trade-off: it uses side-specific estimates to approximate the bias term of a local polynomial of order p and then selects h_{MSE} that minimizes the estimated MSE of the discontinuity estimator.

Implementation follows standard choices: a triangular kernel concentrates weight near the cutoff, and the main fit uses a local linear specification ($p = 1$). The selector returns potentially different optimal windows on the left and right of the cutoff; I report and use a symmetric window formed by averaging these side-specific values to preserve comparability across outcomes and specifications. For transparency, I also conduct pre-specified sensitivity checks that scale the point-optimal window by fixed multiples $m \in \{0.5, 0.75, 1.0, 1.25, 1.5\}$. For donut designs that exclude an inner neighborhood of the cutoff, I define "optimal" with respect to the effective sample by recomputing the MSE-optimal window after trimming or, where noted, by applying the base h_{MSE} to the donut sample.

C.3 Implementation of Assumptions

C.3.1 Assumption 1: Positive First-Stage

I bound the treatment discontinuity $\tau_T := \lim_{x\uparrow 0} p^\star(x) - \lim_{x\downarrow 0} p^\star(x)$ using manipulation-free reference bins on each side of the eligibility threshold. For a given bandwidth h, I first partition observations into bins of width h using the centered income proxy. I then identify the nearest "certain assignment" bins b_h^- (eligible side) and b_h^+ (ineligible side) that lie strictly outside the $\pm h$ neighborhood: the mean upper bound in b_h^- must satisfy $\overline{u} < -h$ and the mean lower bound in b_h^+ must satisfy $\overline{\ell} > h$, with each reference bin required to contain at least 100 observations. These conditions ensure that average households in the reference bins have unambiguous eligibility and furnish one-sided information for the neighborhoods $[-h, 0]$ and $(0, h]$. This threshold applies only to the reference bins and is stricter than my general inclusion rule.

I fit separate models on windows with width equal to the sample mean FPL-bracket range on each side, producing estimates of $p^\star(\cdot)$ and $g^\star(\cdot)$ on each side. For all variables examined—WIC take-up, breastfeeding initiation, and breastfeeding duration—I use flexible polynomials of the running variable and maternal controls.

Given the slope restrictions described under Assumption 3, the reference-bin means and allowable rates of change yield lower and upper bounds on τ_T. I implement Assumption 1 by requiring $\tau_T^{\text{LB}} > 0$ when forming LATE, and I suppress confidence intervals when this condition fails in at least half of bootstrap draws. This operationalizes Assumption 1 via the bounded discontinuity mapping in the paper's identification result.

C.3.2 Assumption 2: Monotonicity in Income

Monotonicity governs both treatment and outcomes away from the threshold. I impose that $p^\star(x)$ is weakly decreasing in income and that outcome CEFs $g^\star(x)$ (initiation, duration) are weakly increasing in income on each side, excluding a neighborhood of the cutoff. In practice, monotonicity enters in two places.

First, it converts the reference-bin means into one-sided bounds on the adjacent neighborhoods. Under monotonicity, the mean in b_h^- bounds $[-h, 0]$ from *below* for $g^\star$ and from *above* for $p^\star$, and conversely for b_h^+ and $(0, h]$.

Second, when projecting model-based predictions from the reference bins to the cutoff, I enforce direction-preserving adjustments so that predicted values between x_{ref} and 0 cannot violate the assumed monotonicity. Concretely, I evaluate models with covariates fixed at region-specific means—“inner” bin means at x_{ref} and “middle” (interior-window) means at the cutoff—and apply a monotonicity check that caps or floors the projected value to preserve the implied slope sign on each side. This ensures that the bin-to-cutoff projection respects the global direction of change implied by Assumption 2. For my main estimates, these caps and floors do not bind.

C.3.3 Assumption 3: Bounded Rate of Change

I implement three approaches to the slope restriction: (i) *Linear—Mean Slope*, (ii) *Linear—Bin Slope*, and (iii) a *Nonlinear Projection* that allows curvature between the reference bin and the cutoff. Polynomial order is chosen independently on each side of the cutoff by AIC, allowing one side to select a first-order fit and the other a higher order. In the Linear implementations, the slope bound is computed either as an average rate of change over wide reference intervals (Mean Slope) or as the model-implied slope evaluated at the midpoint of each reference bin (Bin Slope). The Nonlinear procedure uses the full estimated polynomial to form two predictions—$\hat{g}(x_{\text{ref}}, \bar{W}_{\text{ref}})$ at the reference bin with local covariate means and $\hat{g}(0, \bar{W}_{\text{mid}})$ at the cutoff with interior-window covariate means—and then applies a vertical shift so the curve exactly matches the observed outcome at x_{ref}. This correction allows the slope constraint to reflect curvature between the bin and the cutoff while preserving fit at the reference point. When first-order fits are selected, the Linear-Mean and Linear-Bin implementations coincide by construction.

Operationally, the “Linear” methods treat the slope bound as a constant derivative over the short interval from the reference bin to the cutoff on each side, whereas the “Nonlinear” method computes the counterfactual path using the full estimated polynomial and then aligns levels at the bin before reading off the implied value at the cutoff. All three approaches are implemented separately on the eligible and ineligible sides using an estimation window with a width corresponding to the mean FPL bracket range and region-specific covariate means.

C.3.4 CEF bounds to LATE

Combining Assumptions 2 and 3 yields bounds on the discontinuities in the CEFs for treatment and outcomes, $[\tau_T^{\mathrm{LB}}, \tau_T^{\mathrm{UB}}]$ and $[\tau_Y^{\mathrm{LB}}, \tau_Y^{\mathrm{UB}}]$. The LATE identification region then follows from the ratio mapping in the paper, with cases determined by the sign of the outcome bounds and with the first-stage lower bound constrained to be positive. Inference uses a bootstrap, and I report confidence intervals only when at least half of the bootstrap draws satisfy $\tau_T^{\mathrm{LB}} > 0$, reflecting the operational status of Assumption 1 in finite samples.

C.4 Midpoint FPL as Proxy

Let X_i denote the midpoint proxy and $X_i^\star$ the true (unobserved) FPL. Define $e_i := X_i - X_i^\star$. Given interval bounds $X_i^\star \in [\ell_i,\ \ u_i]$, I have $e_i \in [\,X_i - u_i,\ \ X_i - \ell_i\,]$. For any points $\{a, b\}$, define the local set $A = [a - \frac{h}{2}, a + \frac{h}{2}]$ and $B = [b - \frac{h}{2}, b + \frac{h}{2}]$ with bandwidth h. For the proxy variable X to preserve ordering of conditional expectations:

$$\frac{1}{N_A}\sum_{i:X_i\in A} X_i \leq \frac{1}{N_B}\sum_{j:X_j\in B} X_j \Rightarrow \frac{1}{N_A}\sum_{i:X_i\in A} X_i^\star \leq \frac{1}{N_B}\sum_{j:X_j\in B} X_j^\star$$

Given measurement error bounds:

$$e_i \in [X_i - u_i,\ \ X_i - \ell_i] \quad \forall\, X_i$$

The above condition holds under the most extreme case – when measurement error is minimized for set A and maximized for set B – if:

$$\begin{gathered}
\frac{1}{N_A}\sum_{i:X_i\in A} X_i^\star \leq \frac{1}{N_B}\sum_{j:X_j\in B} X_j^\star \\
\frac{1}{N_A}\sum_{i:X_i\in A} [X_i - e_i] \leq \frac{1}{N_B}\sum_{j:X_j\in B} [X_j - \epsilon_j] \\
\frac{1}{N_A}\sum_{i:X_i\in A} [X_i - (X_i - u_i)] \leq \frac{1}{N_B}\sum_{j:X_j\in B} [X_j - (X_j - \ell_j)]
\end{gathered}$$

Simplifying yields the sufficient condition:

■

C.5 Alternative Estimation Approaches

In addition to the polynomial specifications reported in the main text, I also estimate probit and tobit models for completeness. The full results from these models are not presented in this paper but are available upon request. While the probit and tobit frameworks introduce additional distributional assumptions, they provide a convenient way to calibrate the slope restrictions that underpin the partial-identification strategy. The results are consistent with the polynomial specifications and do not alter the substantive conclusions of the analysis.

C.5.1 Probit Models for Binary Outcomes

For binary outcomes—WIC participation and breastfeeding initiation—I estimate probit models of the form

$$\Pr(Y_i = 1 \mid X_i, \mathbf{Z}_i) = \Phi(h_z(X_i - c) + \mathbf{Z}_i'\boldsymbol{\beta}_z)\,, \tag{4}$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function, $h_z(\cdot) = \sum_{k=1}^{p} \gamma_{z,k}(X_i - c)^k$ is a polynomial in income centered at the eligibility threshold, and $\mathbf{Z}_i$ is a vector of maternal controls. Estimation is conducted separately for income-eligible ($z = 1$) and ineligible ($z = 0$) households. The estimated slopes from these models are used to calibrate the allowable rate of change in the participation and outcome functions under Assumption 3.

C.5.2 Tobit Models for Breastfeeding Duration

For breastfeeding duration, which is right-censored at the infant's age at survey, I estimate Tobit models of the form

$$y_i^* = h_z(X_i - c) + \mathbf{Z}_i'\boldsymbol{\beta}_z + \epsilon_i, \tag{5}$$

$$y_i = \min(y_i^*, A_i), \tag{6}$$

$$\epsilon_i \sim N(0, \sigma^2), \tag{7}$$

where y_i^* is latent duration, y_i is observed duration, and A_i is the age of the infant at interview. Estimation is restricted to mothers who have initiated breastfeeding, isolating the intensive margin effect on duration. As with the probit models, the fitted slopes are used to calibrate the slope restrictions in Assumption 3.