The Importance of Earnings from Recycling

Maya A. Norman*

August 16, 2025

Abstract

Recycling—collecting discarded items and redeeming their scrap value—channels income to marginalized populations. This paper evaluates the importance and prevalence of these earnings throughout the US. Specifically, this paper shows that positive shocks in recycling earnings increase household food expenditures and improve birth outcomes. Bottle bills—a policy requiring refundable deposits on beverage containers—boost food expenditures by approximately 10.2%, increasing recycling earnings by roughly $1400/year (in 2015 dollars)—an amount achievable by collecting 38 containers per day during the sample period. The effect is concentrated among low-income, able bodied households—a population likely to recycle for cash. Moreover, recycling earnings increases improve birth outcomes—a proxy for economic well-being in low-income populations. Back-of-the-envelope calculations imply that informal recycling markets are 50% as effective as food stamps at reducing the incidence of low birth weight. The striking comparability between recycling earnings and a targeted welfare program can be explained by both the size of recycling-generated income and the distinct populations each income source reaches.

JEL Codes: D12, H23, I38, Q52, Q53, I12

---

*School of International and Public Affairs, Columbia University, email: man2185@columbia.edu

"Mr. Deane, a thin, bearded man who is 52 years old and unemployed,…headed for a recycling center that would pay him 21 cents a pound for…aluminum cans. [He] picks up 40 or 50 bucks a week, more on the weekends if the weather holds and there are lots of ball games. Last month, [Mr. Deane's] recycling center paid out more than $20,000, largely to the unemployed who brought boxes and carts filled with aluminum cans. Some make more than $300/week. One man hired three helpers and delivers his cans daily by the truckload. [Professional recyclers] go through the parking lot at Yankee Stadium after a game, social clubs, and delicatessens…'places where everybody drinks Bud'. [The recycling center manager suspects that] the new [bottle bill will] bring changes, although [professional recyclers] will still have plenty to collect.

—New York Times, the Bronx, 1983

Introduction

Anecdotal evidence suggests that informal recycling helps fill gaps in the U.S. social safety net. When faced with financial hardship, some individuals collect discarded beverage containers and redeem them for cash. Survey evidence from New York City indicate that recyclers predominantly use their earnings for food and to support family members. In addition to recycling income, many also rely on formal assistance such as Medicaid, Medicare and/or food stamps (Harvey, Hegel and Hartmann 2023).[1] Reliance on recycling income is neither new nor unique to New York City.[2] Despite its apparent reach, little non-anecdotal evidence documents the prevalence or importance of recycling income.

This paper systematically shows that recycling income fills gaps in the U.S. social safety net. Studying this relationship is challenging due to limited data on both informal income and recycling activity. To overcome this, I use two sources of variation in recycling income to identify its impact on economic well-being. First, I use policy-induced increases in container scrap value from the roll-out of bottle bills. Second, I use seasonal changes in beverage consumption, which affect the number of containers available for redemption. I examine how these changes in recycling earnings influence two measures of economic well-being: household food spending and infant health, measured by birth weight.

I begin by identifying the impact of bottle bills on recycling earnings. Bottle bills require consumers to pay a deposit when purchasing a containerized beverage. Anyone, not just the original consumer, can recover this deposit by returning the empty container. In effect, bottle bills increased the container scrap value eightfold, raising aggregate scrap value to levels comparable to total food stamp benefits (Sanger 1983).[3] This first

---

[1] In New York City, a 2023 survey found that recyclers earned roughly $120 per week from container redemption, using this income primarily to buy food while supporting an average of 2.4 dependents. 50% were enrolled in Medicaid or Medicare, and 32% received food stamps. The survey covered 257 recyclers at 38 redemption sites. Reported averages likely understate the income of full-time recyclers—some reportedly earn up to $60,000 annually. See (Harvey et al. 2023).

[2] For example, in 1983, a New York redemption center paid out about $20,000 per month to professional recyclers, with some individuals earning $300 per week (Sanger 1983). In 1982, a Vermont household earned 40% of its income—$165 per month—through container redemption (The Associated Press 1982). A 2002 survey of 660 patrons at eight redemption sites in Santa Barbara found that 12% of households earning less than $10,000 annually ($13,845 in 2015 dollars) derived 7% of their income from recycling, with Spanish-speaking households earning 9% (Ashenmiller 2009, Ashenmiller 2011).

[3] Aggregate scrap value is equivalent to total food stamp benefits in Michigan, where the deposit is 10 cents, and between 30–90% of food stamp benefits in other bottle bill states (see Figure 1b). Aggregate value is estimated by multiplying national per capita container sales, from the Container Recycling Institute (CRI), by census state population estimates. This approach produces conservative estimates. Porter (1978) reports Michiganders consumed 3.95 billion containers in 1974, whereas CRI suggests 2.6 billion. Similarly, Bingham and Mulligan (1972) estimated 101 billion containerized beverages would be sold in 1976, whereas CRI suggests 65 billion.

analysis tests whether this substantial increase in scrap value translates into an increase in earnings. For identification, I leverage the roll-out of bottle bills and food expenditure data. These implementations were highly idiosyncratic both in their timing and location due to delays and strong political opposition (Davis 1982, Franchot 1978, Peterson 1976, Ross 1982, White 2018, Costle, Kreps, Andrus, Marshall, Blumenthal, Warren, Cutler, Cornell and Alm 1978). I use household-level data on food spending from the Panel Study of Income Dynamics (PSID), treating food expenditures as a proxy for income. To conduct the analysis, I use a difference-in-difference quasi-experimental design in combination with the Generic Machine Learning Framework developed by Chernozhukov, Demirer, Duflo and Fernández-Val (2018), herafter CDDF. CDDF identifies subgroups who are and are not affected by bottle bills. This allows me to identify households who plausibly recycle without directly observing which households do and don't recycle. In addition, I also conduct a traditional event-study analysis to verify the results from the DiD CDDF analysis.

Bottle bill introductions increase food spending 10.2% among a subset of households who plausibly recycle for cash. Specifically, effects are concentrated among low-income, able-bodied households. CDDF formally tests if all households respond similarly to bottle bill introductions. The analysis rejects this possibility. It identifies a clear effect among likely recyclers—the aforementioned subset—and no effect for others. These results are robust to a traditional event-study analysis. In addition, the event study analysis shows that the effect among recyclers begins right after bottle bill implementations and there is no pre-trend. Back of the envelope calculations suggest that this increase in food expenditures is indicative of a $1400 (in 2015 dollars) increase in recycling earnings—an amount achievable by collecting 38 containers per day at a 5 cent deposit level (Hoynes and Schanzenbach 2009, Harvey et al. 2023).

With this evidence that bottle bills increase recycling income, I next examine the importance of this income to families. To do this, I estimate the impact of variation in recycling income on birth outcomes. Birth outcomes—particularly birth weight—is a widely used indicator of infant and maternal well-being. Prior work shows that small cash transfers to low-income mothers can improve birth outcomes (Almond, Hoynes and Schanzenbach 2011, Hoynes, Miller and Simon 2015). If recycling provides meaningful income for low-income or marginalized families, then increases in recycling earnings should lead to healthier births in this population.[4]

To identify the effect of recycling income on birth outcomes, I use two sources of variation that isolate changes in income unrelated to local economic conditions.[5] First, I use the introduction of bottle bills. Second, I use temperature-driven changes in beverage consumption. Warmer weather leads people to drink more, which increases the number of containers available for redemption. Because temperature is correlated with economic activity, I cannot use temperature shocks alone to identify the impact of recycling income. Instead, I combine temperature variation with the introduction of bottle bills to isolate plausibly exogenous, time- and place-specific changes in the aggregate scrap value of beverage containers.

Specifically, I employ two distinct, but closely related, approaches to identify the effect of increases in recycling income on birth weight. The first strategy—a triple difference (DiDiD)—compares (i) the years pre and post-bottle bills, (ii) the ten states that implemented bottle bills to the 40 states that did not, and (iii) winter to non-winter months, when beverage consumption is generally higher due to

---

[4]Almond et al. (2011) and Hoynes and Schanzenbach (2009) find that food stamps raise food spending by 20.8% and reduce low birth weight by 7.8%. If the income-birthweight relationship is similar here, the observed increase in food spending from bottle bills implies an expected 3.8% reduction in low birth weight.

[5]The value of discarded materials is likely procyclical. Rising production and consumption expand the waste stream and drive up scrap prices. For example, scrap metal prices tend to rise with demand for both virgin and recycled materials (Savov 2011).

higher temperatures. The second strategy—a difference in difference in temperature (DiDiT)—uses a triple interaction of temperature and bottle bill introductions. The estimator compares the temperature-outcome relationship in (i) the years pre and post-bottle bills, and (ii) the ten states that implemented bottle bills to the forty that did not (Colmer and Doleac 2023). These approaches assume that the relationship between temperature (or seasons) and economic activity does not change with bottle bill implementation—except through informal recycling (Olden and Møen 2022, Colmer and Doleac 2023). Leveraging both sources of variation has two main advantages over using either one alone. First, the aforementioned identifying assumption is weaker than the assumptions required when using either source in isolation. Second, combining the two sources allows me to fully exploit the rich spatial and temporal resolution of the birth outcome data.

Leveraging these two approaches, I find that a 1% increase in recycling earnings reduces the incidence of low birth weight (ILBW) among mothers without a high school diploma by 0.04%. The effect is largest in areas where informal recycling markets are likely to be active—specifically, places with higher income inequality.[6] The effect is concentrated among marginal births—that is, those weighing between 2,000 and 2,500 grams. There is no observed effect on births below 2,000 grams. This pattern suggests that recycling income reduces modest nutritional deficits among low-income mothers—deficits that would otherwise lead to marginal births.[7] This conclusion is consistent with the food expenditure findings. Specifically, the expenditure analysis shows that recycling income primarily reaches low-income households with excess able-bodied capacity. These households are plausibly living just below the margin of "okay," facing modest but meaningful shortfalls in nutrition.

Several findings support a causal link between increased recycling income and reductions in low birth weight. **First**, the effect is limited to mothers without a high school diploma—a group likely to include many informal recyclers. **Second**, there is no effect on the overall number of births across any demographic group. This finding rules out that changes in the population giving birth explains the changes in birth weights. **Third**, the timing of the effect strengthens the case for causality. The reduction in marginal births among less-educated mothers begins immediately after bottle bill implementation, persists for over a decade, and shows no evidence of a pre-trend. No similar pattern is seen in other education groups or weight categories. **Fourth**, the findings are consistent for both the DiDiD and DiDiT estimators. The distinct approaches rely on different assumptions but point to the same conclusion. This consistency makes it unlikely that the results are driven by any one modeling choice.

The findings show that that informal recycling markets in bottle bill states are 50% as effective as food stamps in reducing the incidence of low birth weight (ILBW). At first glance, the comparability may seem surprising. Food stamps are a formal government program, while recycling is a labor-intensive, informal activity. But the similarity becomes more plausible when recognizing that total recycling earnings are comparable in magnitude to total food stamp benefits. Specifically, recycling income represents 10–32% of food stamp spending. Also helping to explain the comparability, recycling income may be especially effective at reducing the ILBW. It reaches households with modest nutritional deficits relative to food stamp recipients.[8] As a result, smaller income increases are more likely to resolve the shortfalls that lead to low-

---

[6]Places likely to contain both a large number of individuals who forgo redeeming their own containers and a smaller group who rely on collecting those containers for income.

[7]The most modifiable determinants of birth weight are maternal nutrition and smoking. If a mother does not face nutritional deficits, an income shock is unlikely to affect birth weight (Almond et al. 2011, Kramer 1987a, Kramer 1987b).

[8]Food stamp recipients with children are disproportionately single-mother households (Hoynes and Schanzenbach 2009), while households that recycle for cash are more likely to include multiple working-aged adults. As previously discussed, recycling

weight births among recycling households. Specifically, I estimate that recycling households are $3.5\times$ more likely than the average food stamp recipient to avoid a low-weight birth from the same size income gain. Because recycling income is also larger, on average, than food stamp benefits, it may be more likely to close these nutritional gaps as well.[9]

In summary, I estimate that about 1% of the population recycles for cash. For these households, earnings from recycling help resolve meaningful nutritional deficits. It's well established that the formal social safety net plays an important role in the U.S. (Currie 2006, Hoynes, Page and Stevens 2011, Hoynes et al. 2015). This paper shows that informal earnings also matter. Moreover, their importance underscores the insufficiency of the formal safety net during the study period.

1   Bottle Bills: Origins and Legislative Challenges

This paper leverages bottle bill implementations as positive shocks to recycling income. Bottle bills have the potential to significantly raise total recycling earnings—possibly to a level comparable to food stamp benefits. The next section empirically estimates their impact on earnings. This analysis assumes that the passage of bottle bills is plausibly random—at least in their timing, if not in both timing and location. This section presents historical and political context to support this assumption.

Bottle bills responded to the rapid shift from refillable to disposable beverage containers between 1950 and 1970.[10] This transition led to a sharp rise in litter and waste management costs.[11] Policymakers proposed deposit-refund systems to make these new costs beverage companies' responsibility. These laws did their job. They reduced litter and enjoyed strong public support. Figure 1e summarizes survey evidence showing 56–91% approval ratings and redemption rates between 88–95% in early adopter states.

Despite their popularity and effectiveness, bottle bills faced strong opposition from beverage producers and container manufacturers. Industry lobbying aimed to preserve the profitable single-use container model.[12] This lobbying made bottle bill adoption almost impossible. Between 1969 and 1976, state legislatures introduced over 1,000 container bills, yet fewer than 0.5% passed (Lederer 1976, Arensman 1981, Wagenbach 1985). This pattern persisted from 1975 to 2000. During this time, industry lobbied against $\sim 2,000$ bottle bills nationwide (Moore and Scott 1983, Godush 2001). They spent $\sim$ \$20 million/year to ensure their defeat, generally outspending bill proponents (Arensman 1981).[13]

This extensive lobbying lead to uneven and delayed adoption of bottle bills, supporting the claim that the timing of bottle bill implementation was plausibly exogenous. For example, in Michigan, the governor called for a bottle bill in 1971, but repeated legislative defeats postponed its approval until a citizen vote in 1976 (White 2018). In Massachusetts, legislative and ballot efforts were repeatedly blocked due to coordinated lobbying, despite overwhelming public support (Ross 1982). Figure 1d shows the staggered adoption timeline

---

households—those with more able-bodied capacity—plausibly face relatively modest nutritional deficits.

[9]During the study period, food stamp benefits per household were only 54% of average earnings from recycling.

[10]In 1960, refillables made up 50% of beer and 95% of soda containers. By 1985, non-returnables dominated the market, accounting for over 85% of soft drink and 90% of beer sales (Moore and Scott 1983, United States General Accounting Office 1990).

[11]Beverage containers soon comprised 40–60% of roadside litter and 6% of municipal waste volume (Moore 1976, United States General Accounting Office 1990).

[12]This system externalizes disposal costs and limits material reuse, allowing for more vertical integration (Fraundorf 1975, United States General Accounting Office 1990, Saltzman, Levy and Hilke 1999).

[13]Figure 1c shows that in nearly all ballot campaigns, industry outspent bill proponents by wide margins.

across states, underscoring the idiosyncratic nature of adoption. Oregon and Vermont led in 1973, followed by seven other states over the next decade, with California joining in 1987.

2   The Impact of Bottle Bills on Recycling Earnings

This section evaluates the impact of bottle bills on earnings. I cannot directly observe informal recycling earnings, so I estimate bottle bills' impact on food expenditures. To isolate the effect, I leverage the Generic Machine Learning Framework introduced by Chernozhukov et al. (2018), hereafter CDDF, in combination with a Differnece-in-Difference quasi-experimental design. This approach identifies the subsample of households likely to recycle and estimates the effect of bottle bills on their earnings. I find that low-income households disproportionately make up the households identified as recyclers. In addition, recyclers are likely to have excess capacity for labor, i.e., have multiple adults in the household. I estimate that bottle bills increase recyclers' annual earnings by $\sim \$1,400$ (in 2015 dollars)—an amount achievable by collecting 38 containers/day during the sample period.

2.1   Data

For the Section 2 analysis, I use household food expenditure data from the Panel Study of Income Dynamics (PSID) from 1969 to 1993. The PSID began in 1968, tracking $\sim$5,000 households and all members (+ descendants) in subsequent years. The PSID sample period 1969-1993 was chosen to cover the ten bottle bill implementations between 1969 and 2000, while also minimizing survey differences across years. The sample excludes 1973—food consumption variables were not included in this survey (Hoynes and Schanzenbach 2009). Finally, the estimation sample drops surveys in which a household's total annual food spending falls below \$500 (2015 USD)—a threshold well below the USDA Thrifty Food Plan's benchmark (United States Department of Agriculture 2016).

There is some ambiguity in the time frame that food expenditure metrics pertain to. The survey, which takes place in the spring, asks households about their "typical food consumption." Their response is then annualized and applied to the previous year. Following prior work using this dataset, I designate a state as treated if a bottle bill was in effect by January of a given year (Hoynes and Schanzenbach 2009).

In addition to food expenditures, I employ a rich set of demographic data to explore treatment effect heterogeneity and control for factors that may influence food expenditures. I use household demographic data on the household head's gender, retirement status and race. I also employ household demographic data on: food stamp usage, state of residence, and the proportion of income from the head's labor. Finally, I use yearly state level average income per capita from the Bureau of Economic Analysis (BEA), Regional Economic Information System. I use state level average income to control for macro-economic factors and policy changes that may influence food consumption.

2.2   Empirical Strategy

Ideally, I aim to answer the following question: how did bottle bills impact recycling earnings? The PSID does not include any question identifying which households recycle for cash or how much they earn recycling. To address this limitation, I employ a Difference-in-Differences (DiD) quasi-experimental design in combination

with the Generic Machine Learning Framework introduced by Chernozhukov et al. (2018), hereafter CDDF. A DiD estimator captures the average impact of bottle bills. This average effect is likely to be zero because most households do not recycle for cash. CDDF provides a systematic approach to examining variation in treatment effects across the many dimensions of household heterogeneity documented in the PSID. Specifically, I leverage CDDF to identify the pre-existing family characteristics that best distinguish households experiencing the most positive or negative impacts of bottle bills. In other words, CDDF identifies households who are plausibly recyclers. I validate CDDF's findings with a traditional DiD analysis on the subsamples of housholds identified by CDDF as recyclers or not. This validation functions as a robustness check on the identifying assumptions underlying the CDDF framework.

CDDF Overview   CDDF relies on repeated sample splitting to robustly estimate features of the conditional average treatment effect (CATE). The procedure's confidence intervals and p-values incorporate uncertainty from both estimation and data splitting. Specifically, CDDF:

1. Splits the sample into an "auxiliary" and "main" sample.

2. Trains two separate algorithms with the "auxiliary" sample: one on treated units to predict outcomes under treatment, and the other using control units to predict outcomes without treatment.

3. Predicts the outcome with and without treatment for each observation in the "main" sample using the algorithms trained with the "auxiliary" sample.

4. Estimates features of the CATE with these predicted outcomes.

5. Repeats steps (i)-(iv) to ensure estimates reflect both estimation and sample-splitting uncertainty.

Even though CDDF was designed for randomized control trials, the method is generalizable to the quasi-experiemental DiD setting assuming the parallel trends assumption holds both in the entire sample and in considered sub-samples (Rehill 2025, Varela 2023). To the extent possible, I test these identification assumptions with event-study plots, as documented later in this section.

Applying CDDF in this Context   If the PSID directly asked whether households recycled for cash, I could estimate the impact of bottle bills on households who recycle using a Difference-in-Differences (DiD) estimator. This approach would recover an estimate reflecting the following difference in sample means:

$$\frac{1}{N_g} \sum_g \left[\underbrace{\frac{1}{N_{BB_g}} \sum_{j \in BB_g} \underbrace{\left(\overline{Y}_{j,t \geq t_g} - \overline{Y}_{j,t < t_g} \right)}_{\Delta_g Y_j}}_{\overline{\Delta_g Y}_{BB}} - \underbrace{\frac{1}{N_{\neg BB_g}} \sum_{j \in \neg BB_g} \underbrace{\left(\overline{Y}_{j,t \geq t_g} - \overline{Y}_{j,t < t_g} \right)}_{\Delta_g Y_j}}_{\overline{\Delta_g Y}_{\neg BB}} \right]$$

$\overline{Y}_{j,t \geq t_g}$ reflects the average outcome Y for household j in the period post year t_g. $\overline{Y}_{j,t < t_g}$ reflects the average in the period pre-t_g. $\Delta_g Y_j$ reflects the difference between the post and pre-period average. $\overline{\Delta_g Y}_{BB}$ reflects the average difference across all households in bottle bill states that implemented bottle bills in t_g. $\overline{\Delta_g Y}_{\neg BB}$ across all households in never treated states. Ultimately, the DiD estimate recovers a weighted average of

the difference of the average differences across all treatment groups g (Goodman-Bacon 2021). The estimate reflects the effect of bottle bills on Y, assuming parallel trends in outcomes.

I leverage the DiD estimator in combination with CDDF, given the PSID does not directly ask which households recycle for cash. CDDF was designed to apply to cross-sectional data. To accommodate a DiD estimator within this framework I take a number of steps, following the approach taken by Miao, Deng, Wang, Liu and Tang (2023) when applying the causal forest method in a DiD setting. **First**, I group bottle bill states based on their bottle bill implementation year t_g. In other words, all bottle bill states that implemented bottle bills in the same year t_g are assigned the same treatment group g. **Second**, for each treatment group g, I restrict the sample to households observed in the four years before and after t_g, households who did not move during this period, and households who either live in a state assigned to g during the period or never live in a bottle bill state. **Third**, for each treatment group and household, I difference the average outcome in the five years after t_g and four years before t_g. This transformation collapses the panel data into a cross-sectional dataset. In the transformed dataset, each observation corresponds to a household within a treatment group. The outcome variable for household j is now defined as ΔY_j the difference between the post and pre-period average, $\overline{Y}_{j,t \geq t_g} - \overline{Y}_{j,t < t_g}$. Treated households are those that live in a bottle bill state. With the transformed data, CDDF Step 2 trains algorithms to predict the difference in outcomes rather than the outcome itself. Similarly, CDDF Step 3 predicts the average difference in outcomes if an household was treated ($\widehat{\Delta Y}_{j,BB}$) and not treated $\widehat{\Delta Y}_{j,\neg BB}$. $\widehat{\Delta Y}_{j,\neg BB}$ is analogous to the "ML proxy" for the baseline conditional average—a potentially biased and noisy predictor of the expected change in outcome Y for household j conditional on covariates. $\widehat{\Delta Y}_{j,BB} - \widehat{\Delta Y}_{j,\neg BB}$ is analogous to the "ML proxy" for the CATE—a potentially biased and noisy predictor of the expected effect of a bottle bill on Y for household j conditional on covariates.

To account for PSID's small sample size and the fact that there are relatively few treated households, when I split the data in CDDF Step 1 into the "auxiliary" and "main" samples, I create sub-samples according to the following three rules. First, for each treatment group, I assign 50% of treated observations to the "auxiliary" sample. Second, I then assign the same number of untreated units for this treatment group to the "auxiliary" sample.[14] Third, all remaining observations are assigned to the "main" sample. This sub-sampling approach ensures that models used to predict outcomes in the "main" sample are trained on a sufficient number of treated units. Implicitly, this method also upweights treatment groups with more treated units when optimizing prediction performance. In addition, to ensure estimate stability across seeds and fully account for uncertainty due to sample splitting, I complete 750 splits—exceeding the 100 recommended.

For estimation, I use the *GenericML* package in R. The propensity score learner is set to *lasso*, and the estimation procedure checks that estimated propensity scores fall between 0.01 and 0.99—i.e., are sufficiently bounded away from 0 and 1. I implement the CDDF procedure using three learners: *random forest*, *lasso*, and *neural network*, meaning that CDDF Step 2 is completed separately with each method. The *GenericML* package draws on learner implementations from the *mlr3* framework. I allow the procedure to select the best-performing learner according to the model selection criteria outlined in Chernozhukov et al. (2018).

---

[14]Theoretically, there is potential for some untreated household j to be assigned to both the "auxiliary" sample and the "main" sample if the control household appears in multiple treatment groups. Given the relatively large number of control households in each treatment group, the likelihood of this form of contamination is incredibly low, so I do not further tailer the sampling procedure or data construction to eliminate this possiblity entirely.

Household Characteristics for Heterogenous Treatment Effect Estimation   I allow for heterogeneity along household characteristics that survey evidence suggests predict households who recycle. Specifically, I allow for heterogeneity along the following dimensions: household head gender, household head retirement status, if the household used food stamps in the pre-period, household head race, and proportion of household income from transfers in the pre-period.[15]   I use proportion of household income from transfers as a proxy for a household's able-bodiedness.[16]

Given the small sample size of PSID, to reduce noise in the ML algorithm training step and improve prediction performance, I transform household characteristics into binary or categorical variables. Most characteristics listed above naturally lend themselves to discretization. For example, I discretize household head's gender by creating an indicator variable for households with a female head. Finally, the proportion of household income from transfers is the only truly continuous variable considered in estimating treatment effect heterogeneity. This metric serves as a proxy for a household's able-bodied status. With this in mind, I discretize the variable into an indicator equal to one for households that earn the majority of their income from labor (i.e., more than 50% of their income does not come from transfers). Finally, to improve performance, I pre-interact indicator variables. For example, I include indicator variables for non-white households who use food stamps.

Recovering Features of Conditional Average Treatment Effects   CDDF provides a framework for estimating *features* of conditional average treatment effects. Specifically, CDDF provides a method to estimate "sorted group average treatment effects" as well as "sorted group average characteristics." CDDF estimates these objects by (i) assigning observations in the "main" sample to groups based on their predicted treatment effect, i.e., the treatment effect predicted by the "ML proxies" computed in CDDF Step 3, and (ii) estimating the average treatment effect and characteristics for each group. With CDDF, I estimate average treatment effects and characteristics for quartiles of the predicted treatment effect distribution, i.e., the first group are "main" sample households with predicted treatment effects between the 0th and 25th percentile of the predicted treatment effect distribution. This estimation procedure is completed for each sample-split, yielding confidence intervals and estimates that capture uncertainty in estimation and sample-splitting.

I use the "sorted group average characteristic" CDDF methodology to estimate the proportion of different subsamples in the data that were positively, negatively, or not impacted by bottle bills. The "sorted group average characteristics" denote the average of a given characteristic for each "sorted group". Given characteristics used to estimate heterogeneity are indicator variables, the average of a given characteristic is the proportion of a "sorted group" composed of households with that characteristic. I employ this proportion to back out the proportion of the estimation sample with this characteristic in the group.[17]

Validating CDDF Findings with a Traditional DiD Analysis   In a final step, I conduct a traditional DiD analysis on subsamples that CDDF identifies as recyclers and non-recyclers. This both (i) tests the CDDF identifying assumptions to the extent possible and (ii) validates the CDDF findings with an alternate estimation approach.

---

[15]PSID documentation of medicaid and or medicare utilization is incomplete during the sample-period, so I do not include indicator variables for either in the analysis.

[16]This proxy assumes that households with more able-bodied members are likely to retrieve a lower proportion of their income from transfers.

[17]For example, I recover the proportion of the estimation sample with female heads in the *most positively* affected group.

For each "sorted group", CDDF identifies the average effect of bottle bills on food expenditures under the parallel trends assumption for each *"sorted group"*. While the parallel trends assumption is inherently untestable, examining differential pre-trends serves as a useful sanity check. CDDF does not provide a means to identify exactly which households fall into different "sorted groups," so I cannot directly examine pre-trends in each group. Instead, I can examine pre-trends in subsamples identified as likely to be recyclers or unlikely to be recyclers. Specifically, I estimate an event-study version of the following regression equation:

$$Y_{jit} = \beta \mathbb{1}(\text{BB})_{it} + \mathbf{X}_{jt}\gamma_1 + \mathbf{Z}_{it}\gamma_2 + \alpha_j + \lambda_t + \epsilon_{jit} \tag{1}$$

where Y_{jit} denotes the log of total food expenditures for household j living in state i in year t. $\mathbb{1}(\text{BB})_{it}$ is an indicator variable equal to one if household j lives in a state i with an active bottle bill in year t. β is the coefficient of interest, reflecting the effect of bottle bills on food expenditures.

$\mathbf{Z}_{it}$ are state by year specific controls including the state average income per capita. α_j is a household specific fixed effect, controlling for all factors common to each household and constant across time. While a state rather than household specific fixed effect is necessary to operationalize β as a DiD estimator, the household specific fixed effects increase estimator precision. Finally, λ_t is a year specific fixed effect controlling for all factors common across households in a given year.

I use qualitative insights from CDDF to isolate subsamples in the panel data in which CDDF predicts bottle bills affect a high proportion of households. As previously discussed, I apply the CDDF procedure to a transformed estimation sample. This transformation provides a 'clean' way to assign household characteristics based on pre-period attributes. However, this assignment approach does not work in the panel data format, as pre-period characteristics may vary across treatment groups for never-treated households. Consequently, my subsampling approach differs slightly when applying a standard DiD framework. I view this distinction as a strength: if the CDDF findings are valid, the qualitative insights should carry over to the standard DiD analysis, regardless of the exact subsampling details.

2.3 Results

Figure 2a summarizes the results from estimating heterogeneous effects of bottle bills on food expenditures (in logs) using the CDDF procedure. The analysis rejects the null hypothesis of no treatment effect heterogeneity at the 5% level (p = 0.042). In theory, bottle bills should have no average effect on food expenditures, since most households do not recycle for cash. However, if bottle bills change recycling earnings, food expenditures should change for the subset of households that do recycle. The rejection of the null is consistent with this prediction, providing evidence that households respond differentially to bottle bills.

The figure's four panels report average treatment effects (ATEs) and average group characteristics for quartiles of households sorted by their predicted treatment effects.[18] For example, the left-most panel presents the ATE and average characteristics for households with predicted effects in the 0–25th percentile of the distribution. Average treatment effects for each group are listed at the top of their respective panels. For households in the first three quartiles (0th–75th percentiles), the average effect is positive but never statistically different from zero, increasing only modestly from the first to second quartile (2.6% to 3.4%) and from the second to third (3.2% to 4.7%). In contrast, the average effect rises sharply from the third to

---

[18]The procedure for sorting households is discussed in Section 2.2.

fourth quartile (4.4% to 10.2%), with the effect in the fourth quartile statistically significant at the 1% level (p = 0.010). Quartiles are labeled from *least affected* to *most affected* to reflect the lack of response at the lower quartiles.

Consistent with the hypothesis that bottle bills should not affect most households' food expenditures, 75% of households have effectively no estimated response. In contrast, for the remaining 25%, bottle bills increase food expenditures by 10.2%. This pattern demonstrates that the rejection of the null hypothesis of no treatment effect heterogeneity—i.e., the statistical evidence that households respond differentially to bottle bills—is driven by the contrast between households who are affected and those who are not. While the theoretical impact of bottle bills on recycling earnings is ambiguous, these empirical results indicate that, in practice, the policy increased earnings for the subset of households who recycle for cash.

Among these households who recycle for cash—i.e., households who fall into the *most affected* group—bottle bills lead to a $700 to $7,000 (in 2015 $) increase in household earnings.[19] The large range is due to uncertainty around the marginal propensity to consume food out of recycling income. The lower bound assumes that all recycling income is spent on food, while the upper bound assumes that only 10% is spent on food—an estimate drawn from the literature on the marginal propensity to consume food out of cash income (Hoynes and Schanzenbach 2009). Based on survey evidence, the marginal propensity to consume food out of recycling income is likely much higher: three-quarters of professional recyclers surveyed in New York City report spending the majority of their recycling income on food (Harvey et al. 2023). Assuming a marginal propensity of 50% yields an implied increase in earnings from recycling of approximately $1,400 (in 2015 dollars) per household annually. This estimate is used as the preferred benchmark throughout the paper.

I interpret households affected by bottle bills—i.e., those in the *most affected* group—as households that recycle for cash. While CDDF does not directly identify which households fall into this group, I approximate subsamples with a high or low proportion of recyclers based on their likelihood of falling into the *most affected* group. Each panel of Figure 2a plots the proportion of a given subsample in its respective group. Marker color and shape indicate subsample identity, as specified in the plot legend. Labels—*female*, *retired*, *non-white*, and *male*—refer to the household head. For example, *female* refers to the subsample of households with female heads. Labels—*food stamps* and *labor income > 50%*—refer to household-level attributes, i.e., households that used food stamps in the pre-period or earned more than 50% of their income from non-transfer sources, respectively.[20] For example, the red circle in the first panel plots the proportion of households with female heads among the households *least affected* by bottle bills. Vertical lines denote 90% confidence intervals, capturing sample-splitting uncertainty—i.e., variation in the proportion of a subsample within a given group across "auxiliary" and "main" sample splits. Figure 2b provides a more detailed view of the portion of each subsample who recycles—i.e., the proportion of the subsample in the *most affected* group—illustrating the information in the far-right panel of Figure 2a in greater detail. Bar height represents the percentage of a given subsample in the *most affected* group, with error bars indicating 90% confidence intervals. White text

---

[19] The CDDF estimate for the "most affected" group (10.2%) multiplied by the group's average food expenditures—roughly $7,000 (in 2015 $)—is $700.

[20] More specifically, marker colors indicate the subsample hierarchy. Red represents subsamples defined by a single indicator variable. Green, gold, and turquoise represent subsamples constructed using two indicators—one for food stamp use and another household characteristic. Among these, green differentiates by household head gender, gold by household head race, and turquoise by labor status. Purple represents subsamples defined by three indicators, while pink denotes a subsample constructed with four. Marker shape differentiates subsamples within a given color.

labels each bar with its corresponding subsample. Bar colors are consistent with Figure 2a marker colors. In both figures, dashed lines at 25% indicate parity—if each subsample were equally distributed across groups, approximately 25% of the subsample would appear in each group. Subsamples that are overrepresented (underrepresented) in a group will have proportions above (below) 25%.

Consistent with survey evidence and the hypothesis that recycling earnings primarily benefit low-income, able-bodied households, bottle bills disproportionately benefit households with limited income and greater available labor capacity—measured in terms of number of adults or income sources. Food stamp recipients—indicative of lower economic status—disproportionately benefit from bottle bills (Figure 2a, red star marker, far-right panel or Figure 2b, bottom red bar). Specifically, 53% of households in the sample who use food stamps benefit. Among these households, a much larger share of those with male heads (green triangle, bottom green bar)—predominantly representing households with multiple adults—benefit compared to those with female heads (green circle, top green bar), who are more likely to have fewer adults, as female-heads are predominantly single.[21] Specifically, 65% of households with male heads who use food stamps recycle and only 43% with female heads recycle. The subsample with the second highest proportion of households in the *most affected* group (80%) consists of households with non-white male heads who use food stamps (blue square, bottom blue bar). When this subsample is further restricted to households earning at least 50% of their income through non-transfer sources, the share of households increases even further (pink star, pink bar), from 80% to 91%.

Table A2 shows that subsamples with a higher proportion of recyclers are, on average, poorer—as indicated by lower per capita annual food expenditures—and have more adults per household. Table A2, Panel A reports descriptive statistics for the entire estimation sample, while Panel B reports statistics for the subsample of households who used food stamps at some point in the sample period. The first row in each panel covers all households and the subsequent four rows split the sample by household-head demographics. The first column documents the number of households in each (sub)sample. Columns 2-4 report the portion of a row's sample with children, who used food stamps, and earn the majority of their income from labor. The final two columns report the mean per household member annual food expenditures and number of adults with standard errors in parentheses below. Table A2, Panel B, second row covers the subgroup that CDDF predicts as having the greatest prevalence of cash recycling: households headed by non-white men who have received food stamp benefits. Ninety-six percent of these households are classified as high-labor, meaning they derive most income from work rather than transfers.[22] On average they spend $2,341 per household member on food each year and contain 2.31 adults, compared with sample means of $2,904 and 2.02, respectively.

Figure 2c applies insights from CDDF to a traditional event-study analysis, providing a validation check on CDDF's identifying assumptions and results. The event-study findings indicate no differential trends in food expenditures between bottle bill and non-bottle bill states prior to implementation, in any of the subsamples CDDF identifies as containing a high proportion of households likely or unlikely to recycle for cash. This absence of pre-trends supports the identification assumptions underlying CDDF. Specifically, Figure 2c, top panel plots event-study estimates from a DiD style regression (as outlined in equation 1) of the impact of bottle bills on food expenditures (in logs). The x-axis denotes policy-relative time, with coefficients

---

[21]See Table A2 for more descriptive statistics comparing the two subsamples.

[22]CDDF predicts that high-labor households within this demographic are even more likely to recycle for cash than their lower-labor counterparts.

normalized to the bottle bill implementation year (policy-relative time -1). Grey markers denote estimates from the subsample CDDF identifies as having the lowest proportion of households who recycle—households with either female heads or who have never used food stamps. Blue and pink markers denote estimates from the two subsamples CDDF identifies as having the highest proportion of households who recycle for cash—households that use food stamps and have non-white male heads (blue markers) and those that also earn the majority of their income from non-transfer sources (pink markers). Finally, vertical lines indicate point-wise 90% confidence intervals, clustered at the state level—the level of treatment.

In the subsample with a low concentration of households who recycle, food expenditures in bottle bill states never deviate meaningfully from non-bottle bill states—grey markers before and after bottle bill implementation oscillate tightly around zero never approaching statistical significance. In subsamples with a high concentration of household who recycle, there is no clear trend in expenditures pre-implementation, post implementation, expenditures clearly increase. These patterns—(i) parallel trends in the pre-period across estimation samples and (ii) no effect for households unlikely to recycle but a significant response among households likely to recycle—validates CDDF's findings within an alternative empirical framework.

Coefficients in Figure 2c, top panel are estimated using a two-way fixed effects (TWFE) estimator. Goodman-Bacon (2021) highlights potential concerns with TWFE estimation in Difference-in-Differences settings. However, given that only ten states are treated while forty remain untreated, these concerns are minimal in this paper's setting, as problematic comparisons receive very little weight. Nonetheless, as a robustness check, I also estimate the Callaway and Sant'Anna (2021) estimator, which is plotted in the bottom panel (red) alongside the TWFE estimates (blue) for the subsample associated with pink markers in the top panel. Reassuringly, while a littlie noisier, the Callaway and Sant'Anna (2021) estimates track the TWFE estimates.

3 Establishing Seasonality in Recycling Earnings

I leverage variation in recycling earnings to identify their importance and prevalence. Specifically, I use two sources of variation in earnings: the introduction of bottle bills and seasonal fluctuations in beverage consumption. This section motivates the use of seasonal fluctuations in beverage consumption as a proxy for variation in earnings. It demonstrates that aggregate recycling earnings are highly seasonal, and that this seasonality is primarily driven by temperature-induced changes in beverage consumption.[23]

To document these relationships, I leverage (i) monthly data on refund values paid out at redemption centers in forty California counties, (ii) monthly grocery store scanner data from across the U.S., and (iii) daily weather data aggregated to the county-by-month level. I first show that seasonality in recycling earnings—measured via redemption center payouts—is predominantly explained by temperature variation. I then show that this pattern is driven by fluctuations in the size of the empties reservoir, which are a function of temperature-induced variation in beverage consumption.

---

[23]Identifying temperature-driven beverage consumption as the primary mechanism is essential for interpreting the triple-difference estimator as isolating the effect of an increase in recycling earnings, since this variation is exogenous to recycling labor supply choices. I further explore this point later in this section.

3.1 Data

To establish the seasonality of earnings, temperature's important role, and the contribution of temperature driven fluctuations in beverage consumption, I combine *weather* data with data on beverage container *refunds* and *deposits*. I obtained county level monthly refund data from California through public record request. Uniquely, CalRecycle, California's Department of Resources Recycling and Recovery, tracks refunds paid out by redemption centers and to curbside collection companies within the state, making county level data publicly available (Reloop Platform 2020).[24] I construct beverage deposit data, using the Homescan (HMS) Nielsen data between 2005 and 2022. The HMS data tracks 40,000-60,000 US households' retail purchases. To approximate the deposits paid by a household, I curated a list of containerized beverage products that likely require deposits upon sale in bottle bill states (see Table A6). I scale household deposits by HMS market weights to estimate the aggregate amount of deposits paid within a market for a given month. To construct a county by month panel of realized temperature and precipitation, I aggregate PRISM weather data with population weights.

3.2 Empirical Strategy

I examine seasonality in recycling earnings using monthly data on refund values paid out at redemption centers in forty California counties. These refunds are paid directly to individuals and predominantly to those who did not originally consume the beverage, so seasonality in these payouts reflects seasonality in aggregate recycling earnings (Ashenmiller 2009, Ashenmiller 2011). To distinguish between container-supply and labor-driven mechanisms, I also use data on deposits paid by households and refunds issued to curbside recycling companies. This allows me to assess the extent to which seasonal (and temperature-driven) variation in recycling earnings is explained by beverage consumption rather than changes in recycling labor supply.

To begin I estimate weather dose response functions for redemption center refunds, i.e., aggregate recycling earnings. I flexibly model refunds/deposits as a function of daily temperature and precipitation within a month, following the climate econometrics literature (Hsiang 2016). Specifically, I estimate the following equation:

$$\log\left(y_{m(t)c}\right) = \boldsymbol{\theta}\mathbf{P}_{m(t)c} + \boldsymbol{\beta}\mathbf{T}_{m(t)c} + \delta_{ct} + \gamma_{cs(m)} + \nu_{s(m(t))} + \epsilon_{cm(t)} \tag{2}$$

Let $y_{m(t)c}$ denote the refund value (\$) paid out at redemption centers in county c in month m during year t. $\mathbf{T}_{m(t)c}$ represents a fourth order polynomial in temperature, i.e., the kth element denotes $\sum_{j\in c}\sum_{d(t)\in m(t)} T_{d(t)j}^k$, where $T_{d(t)j}$ is the maximum temperature at grid cell j on day d in year t. Similarly, $\mathbf{P}_{m(t)c}$ represents a fourth order polynomial in precipitation. δ_{ct} is a county by year fixed effect, $\gamma_{cs(m)}$ county by season, and $\nu_{s(m(t))}$ season by year. The response is identified from plausibly random year-to-year variation in weather within a county and season.

In addition, I estimate overall seasonality in recycling earnings with the following equation to calculate overall seasonality:

$$\log\left(y_{m(t)c}\right) = \beta_s \mathbf{1}_{s(m)} + \delta_{ct} + \epsilon_{m(t)c} \tag{3}$$

---

[24]Curbside collection companies in California can claim refunds they collect through curbside residential recycling.

As before, $y_{m(t)c}$ reflects recycling earnings in month m during year t in county c. $\mathbf{1}_s$ is an indicator variable for season. Winter is dropped for normalization. δ_{ct} is a county by year fixed effect restricting identifying variation to within a county and year. The coefficient of interest β_s reflects the effect of season s on the outcome in percent, i.e., the outcome is $\beta_s \times 100\%$ higher in season s relative to winter.

In a final step, I compute seasonality as predicted by the weather to relate estimates of seasonality from equation (3) to equation (2). Specifically, I first compute the monthly refunds as predicted by the weather for each season as follows:

$$\widehat{\log(y)_s} = \frac{1}{N_s \times N_i} \left(\sum_i \sum_{m(t) \in s} \hat{\boldsymbol{\theta}} \mathbf{P}_{m(t)i} + \hat{\boldsymbol{\beta}} \mathbf{T}_{m(t)i} \right)$$

N_i and N_s denote the number of counties i and the number of months in season s, respectively. $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ are estimates from equation (2), representing the outcome's temperature and precipitation dose response, respectively. The seasonality predicted by the weather (relative to winter) for season s is then:

$$\%\Delta_s = 100 \times \left(\widehat{\log(y)_s} - \widehat{\log(y)_w} \right) \tag{4}$$

$\widehat{\log(y)_w}$ denotes the outcome predicted by winter weather. $\%\Delta_s$ reflects the predicted percent difference in monthly earnings in season s relative to winter in the absence of seasonal differences beyond weather.

To compute seasonality predicted by temperature or precipitation, individually, rather than weather generally, I calculate $\widehat{\log(y)_s}$ with only realized temperature or precipitation. For example, the outcome predicted by the temperature is:

$$\widehat{\log(y)_s}^T = \frac{1}{N_s \times N_i} \left(\sum_i \sum_{m(t) \in s} \hat{\boldsymbol{\beta}} \mathbf{T}_{m(t)i} \right)$$

To compute, the seasonality predicted by the temperature for season s, I simply replicate the procedure from equation (4) as follows:

$$\%\Delta_s^T = 100 \times \left(\widehat{\log(y)_s}^T - \widehat{\log(y)_w}^T \right) \tag{4a}$$

As in the case of weather generally, $\%\Delta_s^T$ reflects the predicted percent difference in season s relative to winter in the absence of seasonal differences beyond temperature.

Seasonality in recycling earnings could arise from variation in either the supply of recyclable beverage containers or the labor supply of individuals who collect them. This distinction is critical for interpreting seasonal variation in the impact of bottle bills. If the primary driver is container supply—i.e., increased beverage consumption in warmer months—then seasonal changes in recycling earnings reflect variation in economic opportunity and plausibly map to changes in total welfare via increased income. In contrast, if the primary channel is labor supply—i.e., individuals spending more time recycling in warmer months—then, in the extreme, observed earnings changes may not reflect improved economic opportunity. Instead, they may reflect selection into recycling based on amenity value, with individuals choosing to recycle only when conditions are more favorable than their outside option, even if both activities offer similar earnings. This

interpretation complicates the use of seasonality as an interaction term in a triple-difference design, since seasonal variation in recycling earnings may no longer cleanly map onto changes in household income.

Neither plausible mechanism can be directly tested, as no administrative data exist on recycler labor supply or the size of the empties reservoir. Instead I leverage proxies for the size/value of the empties reservoir to investigate if one mechanism dominates. I use two proxies for the size of the empties reservoir: (i) refunds paid to curbside collection companies and (ii) deposits paid by households. Both proxies shed light on the extent to which seasonality in redemption center refunds is driven by consumption rather than collection effort. Their distinct limitations make using both informative. **Curbside refunds** reflect non-redeemed containers collected through residential recycling and primarily capture at-home beverage consumption. To the extent that, at-home beverage consumption closely tracks away-from-home consumption curbside refunds provide a useful proxy for total consumption and the empties reservoir value. However, a key concern is that collection effort from professional recyclers may vary with weather. If so, curbside refunds reflect both consumption and weather-sensitive harvesting effort, introducing measurement error. For instance, increased harvesting in warm months could mask underlying consumption responses, leading to a flat curbside-weather relationship even when consumption is temperature-sensitive. **Deposits paid by households** avoid this issue. They are unaffected by collection effort and better reflect total consumption. If monthly purchases proxy well for monthly consumption, deposits provide a clean proxy for the empties reservoir's size.

With both proxies, I conduct two auxiliary analyses to compare the weather and seasonal sensitivity of redemption center refunds using estimation equations (2), (3), and (4). In the **first auxiliary analysis**, I use counties in the Los Angeles and San Francisco Nielsen Homescan markets, excluding Sacramento due to insufficient refund data. I aggregate redemption center refunds to the Nielsen market level and estimate equations (2, 3, 4) for both redemption center refunds and household deposits. Hence, comparing results isolates differences in the underlying relationships, not sample composition. In the **second auxiliary analysis**, I restrict the sample to the 30 California counties with data on both redemption center and curbside refunds, and estimate equation (2) separately for each. This second auxiliary analysis ensures the conclusions from the first auxiliary analysis are not conflated by a lag between the time when deposits are paid and containers are available for redemption. The first auxiliary analysis ensures that conclusions from the second are not confounded by seasonal variation in recycling effort at residential curbside bins. Comparable responses across measures would suggest that weather-driven seasonality in recycling earnings is primarily explained by changes in beverage consumption—that is, by variation in the the empties reservoir value.

3.3   Results

Figure 3 shows recyclers' earnings are seasonal increasing by 20% in non-winter relative to winter months. The majority of this seasonality is driven by temperature induced fluctuations in the number of containers available to be recycled.

Figure 3a, left-most panel plots estimates from estimating equation 3. Orange (green) markers plot the percent change in deposits (refunds) paid by households (to recyclers) in California in non-winter relative to winter months. Red markers plot the percent change in containerized beverage consumption **across all** Nielsen markets in the US, benchmarking the extent to which seasonality in California is representative of seasonality in beverage consumption across the US. The vertical lines associated with markers denote

the 95% confidence intervals. The middle (right-most) panel plot seasonality as predicted by temperature (precipitation), computed with equation (4). Figure 3a, left-most panel illustrates recycling earnings—i.e., refunds paid to individuals at redemption centers—increase by 20% (green markers) in non-winter relative to winter months. 75% of this seasonality in earnings is driven by beverage consumption: the value of the empties reservoir—i.e., deposits paid by households in California—increases by 15% (orange markers) in non-winter months relative to winter months. This relationship holds nationally: seasonality in beverage consumption in California (orange markers) closely tracks patterns observed across the U.S. (red markers), supporting the inference that *recycling earnings are highly seasonal nationwide*. Figure 3a also shows that this seasonal variation in recycling earnings is primarily driven by temperature. Comparing the green markers in the left-most and middle panels, 70% of the seasonality in earnings is attributable to temperature. Furthermore, most of this temperature-driven seasonality can be explained by temperature-induced changes in the size of the empties reservoir, as indicated by the similarity between green and orange markers in the middle panel. By contrast, precipitation appears to have no explanatory power—coefficients in the right-most panel are effectively zero.

Figure 3b plots the temperature response of three measures: refunds paid to individuals at redemption centers (left panel, green), refunds paid to curbside collection companies (middle panel, yellow), and deposits paid by households (right panel, orange). Coefficients reflect estimates from regression equation (2) and two associated auxiliary analyses. Specifically, each response reflects the effect of temperature on a given day—measured as the day's maximum temperature—relative to a day with a maximum temperature of 10°C. The redemption center refund response (left panel, green) captures the effect of temperature on recycling earnings, while the curbside collection and deposit responses (middle and right panels) serve as two distinct proxies for how temperature affects the value of the empties reservoir. Dotted lines in the middle and right panels show the recycling earnings response re-estimated using the same estimation sample as each panel's primary response. Shaded areas represent 95% confidence intervals. Supporting the conclusion from Figure 3a, the consistency in temperature responses across recycling earnings and both measures of the empties reservoir value suggests that temperature-driven seasonality in aggregate recycling earnings—i.e., refunds paid to individuals—is primarily driven by fluctuations in the size of the empties reservoir, or the number of containers available to be recycled. Moreover, Figure A3 shows this conclusion is robust to alternative model specifications, specifically that the positive relationship between recycling earnings and temperature is robust. Panel A compares the temperature response across different polynomial orders and a binned model. Panel B examines sensitivity to the fixed effects specification. Across all approaches, the conclusion holds: recycling earnings consistently increase on hotter days.

Figure A2 shows the dose response of recycling earnings and the two measures of empties reservoir value to precipitation. While the direction of the response is broadly consistent across outcomes, the patterns are noisy and non-monotonic. At low levels, precipitation appears to reduce earnings and reservoir value relative to dry days; however, at higher levels, the relationship flattens or even turns slightly positive. Given the noise in the estimates—and the fact that California is not particularly rainy—I caution against overinterpreting these results. The apparent lack of a role for precipitation in earnings seasonality likely reflects limitations of the California-based sample rather than a true absence of effect. I leave this question for future research.

In conclusion, Figure 3 shows that seasonality in recycling earnings is primarily driven by seasonality in the value of the empties reservoir, rather than recycler labor supply choices. Moreover, this seasonality

is itself driven by temperature. The similarity in temperature dose responses across recycling earnings and multiple independent measures of the empties reservoir supports the conclusion that observed earnings fluctuations are exogenous to labor supply decisions and instead reflect temperature-induced variation in consumer behavior. Given this, I interpret seasonal differences in the impact of bottle bills as the product of the policy-induced increase in scrap value and the temperature-driven increase in the size of the empties reservoir in non-winter relative to winter months.

4   Recycling Earnings: Prevalence and Importance

In this section, I evaluate the importance and prevalence of recycling earnings in the U.S. Specifically, I estimate their impact on birth outcomes—a proxy for economic well-being among low-income populations. Unlike the food expenditure data, U.S. natality records are cross-sectional and do not follow individual households over time. As a result, I cannot apply the identification strategy used in the food expenditure analysis—a DiD estimator with household fixed effects—to the birth weight analysis. Instead, I exploit within-place, within-year variation in recycling earnings to credibly identify the impact of income increases on birth outcomes. Specifically, identification relies on two sources of variation: (i) bottle bill implementations, which I show in Section 2 increase recycling earnings, and (ii) seasonal and temperature-driven fluctuations in beverage consumption, which I show in Section 3 affect earnings by altering the number of containers available for redemption.

I view the evidence presented in this section as the most convincing causal evidence that recycling income is an important income source for low-income households in the US. I present this analysis last because the results from the previous two sections substantially strengthen its credibility.

4.1   Data

To recover the importance of recycling earnings as well as their prevalence, I leverage birth outcomes as a proxy for economic well-being. Specifically, the analysis employs U.S. Vital Statistics Natality Data from 1969-2002, consisting of the full census of births from the National Center for Health Statistics (NCHS). Table A3 summarizes birth statistics in the estimation sample. In addition to summarizing outcomes in the full sample, Table A3 summarizes statistics by season and mother's education two dimensions of heterogeneity the analysis leverages to causally identify the importance and prevalence of recycling earnings.

To isolate the effect of recycling earnings, I estimate treatment effect heterogeneity by education. As previously discussed, lower wage earners are more likely than higher wage earners to earn income recycling and thus experience income benefits from bottle bills. I use education to proxy for income because the natality data do not include information on wage. While the analysis in Section 2 shows that race is an important determinant of which households recycle, the natality data does not consistently allow mothers to report Hispanic ethnicity. As Ashenmiller (2009, 2011) show, low-income Hispanic individuals are disproportionately likely to recycle, and many may identify as white in the absence of an ethnicity indicator. As a result, I do not explore treatment effect heterogeneity by race, since misclassification could obscure meaningful variation. I explore educational heterogeneity by assigning birth outcomes to education groups using mother's education. Specifically, I use three education groups: less than a high school (low) education, a high school (middle) education, and more than a high school (high) education.

To conduct the analysis, I aggregate the natality data into county-education-month-year cells, employing data on birth month, year and weight as well as mother's education. I exclude all observations in which the mother's education is missing. Birth outcomes are assigned treatment based on when the pregnancy's third trimester began, as the third trimester is the most important in determining birth weight (Almond et al. 2011). I use Bailey, Clay, Fishback, Haines, Kantor, Severnini and Wentz (2018)'s infant health dataset to match NCHS county codes across years.[25]

Table A4 and A5 examine if metrics of data quality and missingness in mother's education systematically vary with the treatment—bottle bill introductions and/or the interaction of implementations and non-winter months. Table A4 does not find that bottle bills are correlated with any measure of data integrity. Table A5 shows that missingness in mothers' education does not systematically vary with the treatment in the estimation sample.

While data quality and missingness are not systematically related to treatment, states with extensive missingness can exert disproportionate influence on estimates depending on the weighting scheme. To inform robustness checks—particularly when estimators do not adequately downweight observations with high missingness—Figure A8 summarizes the prevalence of missing maternal education data. Figure A8a plots the distribution of the percent missing as a box and whiskers plot for each state. A single point represents a given county and year. Figure A8b plots percent missing time series for states with excessive missingness.[26] The top (bottom) panel of Figure A8b plots bottle bill (non-bottle bill) states with high levels of missingness as colored dotted lines. All states without excessive missingness are plotted in both panels as grey solid lines. Missingness in Connecticut, Delaware and New York do not obviously covary with bottle bill implementations. In Connecticut and Delaware, the missingness rates drop below 20% well before each state implemented in 1980 and 1983, respectively. In New York, the missingness rate peaked to roughly 50% towards the end of 1980, well after New York implemented in 1983. On the other hand, the missingness rate in California drops right around the time the state imposes a bottle bill. California implemented in 1987 and the state's missingness rate dropped from roughly 100% to 0% right around that same time. Finally, changes in missingness in non-bottle bill states with excessive missingness do not appear to systematically vary with bottle bill implementations. For robustness, the analysis considers subsamples that drop states with excessive missingness.

Additionally, I include control variables covering income, government welfare spending, and weather. I use data from the Bureau of Economic Analysis (BEA), Regional Economic Information System for state level quarterly information on personal income, wages and salaries, farm wages and salaries, personal current transfer receipts, medicare benefits, state unemployment insurance compensation, and social security benefits. Throughout the analysis, I refer to these variables as *income* controls. The BEA only has sub-annual data during the sample period at the state level, given the analysis's focus on seasonality, I use this data despite its spatial coarseness. To assign BEA control variable values to birth outcomes, I find the variable averages during the third trimester. Weather controls were aggregated with population weights to the county by month level from ERA5-Land. Weather controls include cumulative exposure during the third trimester to either cooling degree days (CDDs) and heating degree days (HDDs) or maximum daily

---

[25]This dataset does not include Alaska, so Alaska is excluded from the analysis. Additionally, I employ the Horan-Hargis IDs in Bailey et al. (2018) to construct geographic units that do not change over the duration of the sample period. The Horan-Hargis ID groups counties that merged or split during the sample period into a single geographic unit (Horan and Hargis 1995). The final estimation sample is aggregated to the Horan-Hargis ID (HHID) level.

[26]Excessive missingness is defined as states with at least 20% of records missing mother's education in a given year.

temperature and precipitation. The threshold used to construct both CDDs and HDDs is 20 degrees Celsius.

4.2 Empirical Strategy for Identifying the Bottle Bills' Impact on Births

Leveraging bottle bills and temperature, I employ two approaches—a DiDiD and DiDiT estimator—to identify the effect of an increase in recycling earnings on birth weight.

DiDiD I use a triple difference (DiDiD) estimator to isolate the association between bottle bills (BBs) and birth outcomes. Specifically, the estimator compares (i) the years pre and post-BBs, (ii) the ten states that implemented BBs to the forty states that did not, and (iii) winter to non-winter months, as BB related earnings increases should be largest during non-winter months. I assign season—non-winter or winter—to birth outcomes based on when the majority of a pregnancy's third trimester occurred, as the third trimester is the most important in determining birth weight (Almond et al. 2011). Specifically, babies born in January, February, March and April are designated as winter births. Babies born in all other months are designated as non-winter births. I take a similar approach to assign aggregate temperature and precipitation exposure to births. Aggregate exposure in the third trimester is assigned using the birth month and the two preceding months. Second trimester exposure covers the three months before that (birth month minus 3 to minus 5), and first trimester exposure is based on the three months prior to that. Lastly, as previously discussed, I look at heterogeneity in this estimator by education, as BBs should only impact individuals participating in informal recycling, likely lower wage earners, due to data constraints education proxies for income.

This paper's main regression—a DiDiD estimator with heterogeneity by education—serves three purposes. First, it tests whether the increase in recycling earnings from bottle bills, documented in Section 2 using PSID food expenditures, generalizes to a broader population and can be detected through a distinct measure of economic well-being, birth outcomes. Second, it evaluates the prediction from Section 3 that recycling earnings are larger in non-winter relative to winter months. Third, it estimates the impact of increased recycling earnings on birth outcomes—an important outcome in its own-right for human capital formation (Almond, Currie and Duque 2018). The regression equation for the DiDiD estimator *without* heterogeneity is:

$$Y_{ctem} = \phi \left[\mathbb{1}(\text{BB})_{it} * \mathbb{1}(\text{not winter})_s \right] + \alpha_{c,s(m)} + \delta_{c,t} + \gamma_{s(m),t} + \epsilon_{ctem} \tag{5}$$

Y_{ctem} is a given birth metric in county c in year t for education group e in month m. $\mathbb{1}(\text{BB})_{it}$ denoting if county c in state i has an active BB. $\mathbb{1}(\text{not winter})_s$ equals one when season s of month m is not winter.[27] To operationalize the DiDiD estimator, I include three sets of two way fixed effects: $\delta_{c,t}$, $\gamma_{c,t}$, and $\alpha_{c,s}$. $\delta_{c,t}$ is a county by year fixed effect, accounting for all factors common to a county in a given year. $\gamma_{s,t}$ is a season by year fixed effect, controlling for common factors in a season during a specific year. $\alpha_{c,s}$ is a county by season fixed effect; this fixed effect allows for differences between counties for each season. Standard errors are clustered by state, the level of treatment. While this results in only 10 clusters for treated units, following the design-based perspective, 10 clusters is sufficient for producing conservative standard errors (Roth, Sant'Anna, Bilinski and Poe 2023). Lastly, the regression is weighted by the number of births in each

---

[27]Season is assigned to birth outcome based on when the majority of the third trimester occurred, following the literature on when birth weights are most sensitive to cash transfers (Almond et al. 2011).

education x month x county x year cell.

The DiDiD estimator is ϕ. ϕ is associated with the interaction: $\mathbb{1}(\text{BB})_{it} * \mathbb{1}(\text{not winter})_s$, equaling one for states with active BBs in non-winter months. ϕ captures variation in a birth outcome specific to BB states, relative to non-BB states, when a BB was active relative to inactive, and in non-winter relative to winter months. To identify ϕ the following assumption must hold: in the absence of BBs, birth outcomes would not have changed differentially in BB and non-BB states, in winter versus non-winter months, during the sample period (Olden and Møen 2022).

Goodman-Bacon (2021) highlights potential concerns with the interpretation of difference-in-differences estimator in settings with staggered treatment timing. In my case, these concerns are likely of little consequence: only ten states are treated, while forty are never treated, so problematic comparisons receive very little weight. Nevertheless, following the approach in Section 2, I test this prediction by estimating the main result using both the traditional two-way fixed effects (TWFE) estimator and an alternative estimator proposed by Callaway and Sant'Anna (2021). To implement this difference-in-difference (DiD) estimator in my difference-in-difference-in-differences (DiDiD) context, I transform the panel from a *month* × year × county × education structure into a year × county × education format, where the outcome variable is the difference between birth outcomes in non-winter and winter months. This transformation reflects the insight that a DiDiD can be viewed as a DiD with a transformed outcome (Roth et al. 2023). A key feature of the Callaway and Sant'Anna (2021) framework is the aggregation of group-time average treatment effects—i.e., treatment effects for all units first treated at the same time in a given year—into summary measures. They propose using weights based on treatment probabilities. When sample weights are present, the ***average*** weight for each unit across the whole sample informs the aggregation. In my setting, however, sample weights vary substantially over time due to missing data. As a result, observations that are heavily downweighted in TWFE estimation can receive substantial weight in the Callaway and Sant'Anna (2021) procedure—particularly in earlier periods when birth counts are effectively zero. Hence, to ensure a fair comparison between estimators and to prevent missing data from biasing the estimated treatment effects, I impose two restrictions on the estimation sample. First, I limit the sample to states with relatively complete data, excluding those with extensive missingness: Texas, California, Washington, New Mexico, Idaho, and Arkansas (see Figure A8). Second, following Almond et al. (2011), I exclude county-season-year-education observations with fewer than seven births and county-year-education observations with fewer than 25 births to avoid estimation problems caused by sparse outcome data.

BBs are expected to increase earnings for less educated mothers more than for more educated mothers. To identify heterogeneity in the DiDiD estimator by education, I interact each term in equation 5 with an education group specific dummy variable.[28] Point estimates from this regression are equivalent to estimating a DiDiD estimator for each education group; however, standard errors vary in that correlation in errors is permitted across education groups. The analysis also considers DiDiD estimators that rely on pooled estimation of fixed effects across education groups. My preferred DiDiD *with* heterogeneity estimator takes the following form:

$$Y_{ctem} = \phi_e \left[\mathbb{1}(\text{BB})_{it} * \mathbb{1}_e * \mathbb{1}(\text{not winter})_s\right] + \alpha_{c,s(m)} + \delta_{c,e,t} + \gamma_{s(m),e,t} + \epsilon_{ctem} \tag{6}$$

---

[28]Education groups include less than high school (low), high school (middle), and more than high school (high).

The specification is just short of a stacked DiDiD estimator in education, i.e., interacting $\alpha_{c,s}$ with education would make the estimator equivalent to a DiDiD estimator for each education group. Generally, DiDiD estimators are better than DiD estimators in that they relax identifying assumptions. However, the cost of this relaxation tends to be overfitting.[29] I employ a main specification slightly more relaxed than a DiDiD estimator for each education group to increase power and reduce the risk of overfitting, but asses the cost in terms of identifying assumptions every step of the way. Intuitively, this approach reckons that some information can be jointly estimated across education groups and does so to the extent that results do not qualitatively vary from the case with the most relaxed identifying assumptions.

DiDiT Building off the result that 70% of seasonality in recycling earnings is driven by temperature (see Section 3 for details), I estimate a model examining the extent to which BB implementations modify the relationship between birth outcomes and temperature in the final trimester. The DiDiD analysis examines the extent to which BBs induce policy relevant increases in income, relying on the fact that in most places in most years aggregate recycling earnings are larger in non-winter relative to winter months. The DiDiT analysis leverages temperature's role as a driver of recycling earnings' seasonality, more precisely modeling fluctuations in earnings within a given year and place relative to the crude seasonal difference employed in the DiDiD analysis. Namely, I estimate the following model:

$$Y_{ctem} = \beta_e \left(T_{ctem} * \mathbb{1}(\text{BB})_{it} * \mathbb{1}_e \right) + \alpha_{cet} + \delta_{ce} * T_{ctem} + \gamma_{e,t} * T_{ctem} + \epsilon_{ctem} \tag{7}$$

T_{ctem} denotes aggregate temperature exposure, measured with maximum daily temperature, in birth month m and the preceding two months.[30] $\delta_{ce} * T_{ctem}$ reflects a county by education group specific temperature response and $\gamma_{et} * T_{ctem}$ a year by education specific response. As additional controls, I include precipitation exposure in the first, second and third trimesters as well as temperature in the first and second trimester, allowing responses to vary by education group. The coefficient of interest β_e reflects the degree to which BBs modify the third trimester-temperature birth outcome relationship (Colmer and Doleac 2023). To identify β_e, the following assumption must hold: in the absence of bottle bills, the temperature-outcome relationship would not have changed differentially in bottle bill and non-bottle bill states for each maternal education group.

I scale the DiDiT estimates to conduct a robustness check of the DiDiD results and to recover the implied effect of a 1% increase in recycling earnings on birth outcomes. For the robustness check, I scale the DiDiT estimates using realized temperature differences to recover the DiDiD estimates under alternative identifying assumptions. Specifically, I multiply the DiDiT coefficient, β_e, by the average temperature in bottle bill states during winter months ($\overline{T}_w$), non-winter months ($\overline{T}_n$), and their difference ($\overline{T}_\Delta = \overline{T}_n - \overline{T}_w$). This yields $\beta_e \times \overline{T}_w$, the predicted impact of bottle bills on birth outcomes via the temperature channel in winter months,[31] and $\beta_{e,\Delta} = \beta_e \times \overline{T}_\Delta$, the predicted impact in non-winter relative to winter months. Given that temperature explains approximately 70% of the seasonal variation in refunds, it is plausible that the DiDiD estimator ϕ_e and the scaled DiDiT estimator $\beta_{e,\Delta}$ are similar.

To convert DiDiT estimates into the implied effect of a 1% increase in recycling earnings on birth

---

[29]Fully saturated models have higher variance, as they are subject to overfitting (Angrist and Pischke 2009).

[30]See Section 3 for details on computing aggregate temperature exposure.

[31]See equation (4) for details.

outcomes, I consider the following stylized representation of birth outcomes:

$$Y_{itw} = f_{i,t}(w) + d \times e(w) + \epsilon_{itw}$$

Y_{itw} denotes a birth outcome observed in location $i \in \{BB, \neg BB\}$, where BB refers to a location that eventually adopts a bottle bill and $\neg BB$ to one that does not, during period $t \in \{\text{pre}, \text{post}\}$, representing the time before or after bottle bill implementation. The variable w represents the temperature of an *additional exposure day* during the third trimester. The function $f_{i,t}(w)$ captures the effect of one additional third-trimester day at temperature w on Y through any channel *other than* the size of the empties reservoir, while $e(w)$ captures the effect of such a day *via changes in the size of the empties reservoir*, with d denoting the scrap value of an empty beverage container. In other words, $d \times e(w)$ denotes the increase in recycling earnings from an additional exposure day at temperature w.

With this stylized representation, the DiDiT estimator β_e can be defined as follows:

$$
\begin{aligned}
\beta_e &= \mathbb{E}[(Y_{BB,\text{post},25} - Y_{BB,\text{post},10}) - (Y_{BB,\text{pre},25} - Y_{BB,\text{pre},10})] \\
&\quad - \mathbb{E}[(Y_{\neg BB,\text{post},25} - Y_{\neg BB,\text{post},10}) - (Y_{\neg BB,\text{pre},25} - Y_{\neg BB,\text{pre},10})] \\
&= \mathbb{E}[\Delta f_{BB,\Delta} + (d_b\ e(25) - d_b\ e(10)) - (d\ e(25) - d\ e(10))] \\
&\quad - \mathbb{E}[\Delta f_{\neg BB,\Delta} + (d\ e(25) - d\ e(10)) - (d\ e(25) - d\ e(10))] \\
&\quad \text{s.t. } \Delta f_{i,\Delta} = (f_{i,\text{post}}(25) - f_{i,\text{post}}(10)) - (f_{i,\text{pre}}(25) - f_{i,\text{pre}}(10)) \\
&= \mathbb{E}[\Delta f_{BB,\Delta} + (d_b - d)\ (e(25) - e(10))] - \mathbb{E}[\Delta f_{\neg BB,\Delta}] \\
&= (d_b - d)\mathbb{E}[e(25) - e(10)] \tag{8}
\end{aligned}
$$

where the final line follows from the fact that the the DiDiT identifying assumption implies that $\mathbb{E}[\Delta f_{BB,\Delta}] = \mathbb{E}[\Delta f_{\neg BB,\Delta}]$. In words, β_e captures the birth outcome effect of the increase in aggregate scrap value associated with an additional day at $25°C$ relative to a day at $10°C$ in bottle bill states.[32] Figure 3 shows that aggregate scrap value rises by 0.3% with an additional day at $25°C$ relative to a day at $10°C$, so β_e reflects the impact of a 0.3% increase in aggregate earnings or $\beta_e \times \frac{1}{0.3}$ represents the effect of a 1% increase. Finally, under the assumption that marginal changes in temperature fully pass through to collectors—that is, that households' decisions to recycle their own containers and individuals' decisions to recycle for cash do not vary (or vary only minimally) with temperature shocks—$\beta_e \times \frac{1}{0.3}$ reflects the effect of a 1% increase in *recycling earnings* on birth outcomes.

4.3 Results

DiDiD Estimates   Table 1 shows (i) bottle bills increase average birth weight 1.75 to 4.02 grams among mothers with less than a high school education (low education), (ii) bottle bills have no clear effect on mothers with a high school education or more and (iii) the association between bottle bills and low education mothers' birth outcomes is unlikely to be driven by compositional changes in the underlying population. Table 1 outlines the effect of bottle bills on average birth weight (Panel A) and fertility (Panel B), estimated with

---

[32]These temperature values are illustrative and could be substituted with any other pair; the comparison is intended to reflect the general relationship between temperature and the size of the empties reservoir. Since d_b is approximately eight times larger than d, the contribution of d can be reasonably ignored.

the DiDiD estimator discussed in equation 5 and 6. Rows within each panel differ by education group. Row 1 (2) (3) depicts the estimated effect on low (middle) (high) education mothers. Columns differ in the fixed effect controls, increasing in stringency from left to right. The fifth column is equivalent to a DiDiD estimator fully stacked in education, i.e., estimating a separate DiDiD estimator for each education group. The fourth column, the preferred specification, arguably best balances the bias-variance tradeoff of DiDiD estimators. Table 1, Panel A, first row shows that bottle bills (BBs) increase low education average birth weights consistently across all specifications. The effect ranges from a statistically significant 4.02 g increase in average birth weight to a statistically insignificant 1.75 g. The effect's statistical significance dwindles with fixed effect stringency. Panel A, row 2-3 show that middle and high education mothers are not obviously affected by bottle bills, estimates flip signs and attenuate with fixed effect stringency, never reaching statistical significance. Table 1, Panel B, first row shows that bottle bills do not impact low education mother fertility in non-winter relative to winter months, an important finding for supporting that BB birth weight effects likely act through an income channel rather than reflect compositional changes. The impact of bottle bills on low education mother fertility is noisy, never statistically significant, close to zero, and flips signs multiple times. Importantly, bottle bills' effect on low education average birth weight is stable—always positive—across fixed effect specifications, when the effect on the number of births is positive and negative, suggesting any potential compositional changes do not drive the observed effect on average birth weight.

Figure 4a shows that the consistently positive effect of bottle bills in non-winter relative to winter months on low education mothers' average birth weight is driven by a .18 pp in birth incidence between 2,000 and 2,500g—marginal births or a 1.8% reduction in the incidence of low birth weight (ILBW) with no clear impact on marginal births to middle and high education mothers.[33] In all other birth weight bins, bottle bills have no detectable effect or do not produce statistically or even substantively different impacts across education groups.[34] The figure plots estimates from regression equations (5)—green markers—and (6)—blue, red, and turquoise markers—on the incidence of births across the birth weight distribution. Green markers show the pooled effect across all births, while blue, red, and turquoise markers show effects by maternal education group—low, middle, and high, respectively. The y-axis indicates the change in the incidence of births within each weight bin, labeled on the x-axis, in percentage points. For example, the left-most panel shows the effect of bottle bills on births weighing ≤1,500g overall and by education group. 95% confidence intervals are calculated using standard errors clustered at the state level, the level of treatment.

The estimated effects on the incidence of births less than 1,500 g and between 1,500-2,000 g (first and second panels from the left) are close to zero across all groups. In the third panel (2,000–2,500g), there is effectively no effect for middle and high education mothers but a large and negative effect for low education mothers. This decline is statistically different from the effects for both middle and high education mothers. The overall negative effect in this bin (green marker), while statistically significant at the 10% level, is clearly driven by the low education subgroup. The pooled effect (green markers) is also statistically different from

---

[33]To express the impact on the incidence of low birth weight (ILBW) in percentage terms, I normalize the impact on marginal births—approximately 0.18 percentage points—by the overall ILBW rate, 9.62%, among low-education mothers, as reported in Table A3.

[34]Theory, survey evidence, and the effects of bottle bills on food expenditures suggest that: (i) bottle bills should primarily affect birth outcomes among lower-wage households, a population with a higher proportion of recyclers; and (ii) on average, bottle bills should have no effect. In addition to examining the magnitude, sign, and statistical significance of the estimates, I argue that assessing whether effects differ systematically across education groups is an important criteria in extracting signal from noise when evaluating bottle bills' effects.

the low education effect and substantially smaller, reflecting attenuation due to the null response among middle- and high-education mothers. While some bins above the low birth weight cutoff (2,500g) show modest movement away from zero, these effects are not substantively different across education groups and, as discussed later, are also not robust to alternative specifications or event study analyses.

Figure A6 Panels D-F show that these conclusions/results are robust to the inclusion of weather and income controls, assuaging concerns that the recovered relationship suffers from omitted variable bias. Panel D replicates Figure 4a. Panel E replicates Panel D, but includes weather controls in estimation and Panel F includes both weather and income controls.[35] In the presence of measurement error, balance tests are preferable to comparing coefficients with and without controls (Pei, Pischke and Schwandt 2019). Table A1 formalizes the conclusions from the coefficient comparisons in Figure A6, Panels D–F, by reporting estimates from a balance test. As expected, point estimates are small, since the test regresses large covariates on a treatment indicator. The only statistically significant association is between bottle bill introductions in non-winter months and farm salaries and wages, suggesting that seasonal variation in payments to hired farm labor differed systematically across treated and untreated units. This potential violation of the identifying assumption highlights the importance of interpreting the findings with caution. At the same time, potential confounders are weak predictors of the treatment, and the poor model fit supports the view that bottle bill implementations were highly idiosyncratic—particularly with respect to seasonal variation in confounders—bolstering the causal interpretation of the results (Hoynes and Schanzenbach 2009).

Figure 4b assuages concerns that the reduction in low education mother marginal births documented in Figure 4a is driven by overfitting or stringent identifying assumptions. It also supports the decision to down weight statistically significant effects in other bins that do not vary by education level, as these effects appear highly sensitive to model specification and are likely spurious. Figure 4b plots the effect of bottle bills on births to low-education mothers in non-winter relative to winter months across the birth weight distribution. Markers, as labeled in the legend, indicate the fixed effects specification used in estimation: red markers denote the preferred specification (replicating blue markers in Figure 4a), while purple markers represent fully stacked DiDiD estimates. Figure 4b shows that bottle bills consistently reduce the incidence of low education mother births between 2,000 and 2,500 grams (marginal births), with effects ranging from .15-.2 pp. No other birth weight bins have similarly robust effects.

Figure A6, Panels A-C demonstrate that the impact on marginal births to low-education mothers is the only effect that is both (i) robust to fixed effects specification and (ii) distinctly concentrated in one education group. Figure A6, Panel A replicates Figure 4b as a point of comparison; Panels B and C replicate Panel A for middle- and high-education mothers, respectively. In contrast to the consistent and robust effect among low-education mother marginal births (Panel A, third column from the left), estimated impacts across the birth weight distribution for middle- and high-education mothers are generally sensitive to the fixed effects specification. For example, the middle-education group shows a possibly "robust" increase in births between 1,500–2,000 grams, but the effect weakens under stricter fixed effects, suggesting it may reflect model artifacts rather than a causal impact. In addition, the effect is not statistically different from the high education group effect for this bin (see brown marker). Both low- and middle-education groups show increases in births between 3,500–4,000 grams, but these effects are not statistically or substantively different across education groups and are sensitive to specification. As discussed previously, differential

---

[35]See Section 4.1 for definitions of the variables included in weather and income controls.

effects across education groups are key to interpreting causal mechanisms in this context. Given the lack of clear group differences and the sensitivity of estimates, I interpret these patterns as likely spurious. I revisit these conclusions using event study plots to further assess (i) the validity of the identification strategy and (ii) the use of estimate robustness as a tool to distinguish true effects from spurious correlations.

Figure A4a further supports the robustness and causal interpretation of bottle bills' effect on marginal births among low-education mothers. Figure A4a demonstrates that the conclusions from Figure 4a and 4b are not sensitive to the definition of non-winter months. As described in Section 4.1, births are assigned to seasons based on when the majority of the third trimester occurred. Markers in the figure plot the effect of bottle bills on low education mother marginal births in non-winter relative to winter months, across various definitions of "non-winter," labeled on the x-axis and distinguished by marker color. Solid markers represent estimates from the preferred fixed effects specification (see equation 6); vertical dashed lines show 95% confidence intervals. Hollow markers represent alternative fixed effect specifications, corresponding in order to those shown in Figure 4b. The left-most definition—March through October—corresponds to the main specification. Its solid light blue marker aligns with the blue and red markers in the third panels of Figures 4a and 4b, respectively. From left to right, the season definitions become more restrictive, excluding spring and fall months. As expected, the estimated effects attenuate as the weather contrast between "winter" and "non-winter" narrows. Nevertheless, all estimates remain statistically significant and negative, ranging from 0.1 to 0.2 percentage points.

Figure 4c provides additional support for a causal interpretation of bottle bills' impact on marginal births among low-education mothers. The figure plots estimates from an event-study version of equation (6), where the outcome is the incidence of births weighing between 2,000 and 2,500 grams. Policy-relative year 0 denotes the year of bottle bill implementation. Solid markers indicate estimates from the preferred fixed-effects specification, consistent with the blue and red markers in Figures 4a and 4b, respectivelyt. Hollow markers represent results from a fully stacked regression by education level, matching the purple markers in Figure 4a. To improve precision and aid visual interpretation, effects are grouped into two-year bins spanning ten years before and after implementation, with the three years prior normalized to zero. The results show that reductions in marginal birth incidence begin shortly after bottle bill implementation and become more stable and pronounced five years post-implementation. Crucially, there is no evidence of differential pre-trends: pre-treatment estimates hover tightly around zero, supporting the parallel trends assumption. Figure A5b confirms that these conclusions are robust to the binning strategy, replicating Figure 4c with annual policy relative time estimates. Figure A7 shows that no other combination of birth weight bin and education group exhibits a clear or consistent association between policy timing and birth incidence. The figure presents event study plots, following the format of Figure 4c, for all education groups and birth weight bins above 1,500g and below 4,000g. Rows vary by education group (low, middle, high from top to bottom), and columns by birth weight bin: 1,500–2,000g, 2,000–2,500g (i.e., marginal births), and so on, up to 3,500–4,000g. Across all combinations, only marginal births among low-education mothers show a clear and consistent pattern following bottle bill implementation. This finding reinforces the conclusion that the only robust causal relationship between increased recycling earnings and birth outcomes is for marginal births among low-education mothers. Finally, Figure A5c confirms that these event-study conclusions are robust to the use of the TWFE estimator. Red and blue markers plot event-study estimates from Callaway and Sant'Anna (2021) and the TWFE specification, respectively. As discussed in Section 4.2, the estimation

sample for these estimates was modified to accommodate the Callaway and Sant'Anna (2021) estimator and ensure that differences in estimates reflect differences in the estimators rather than in the sample. Following the approach used in Figures 4c and A5a, event time is binned into two-year intervals spanning from nine years before to nine years after policy implementation, with any event time beyond this range grouped into a single bin. Even after these changes—namely, the modified estimation sample and binning procedure—the TWFE estimates remain consistent with the main results. More importantly, the Callaway and Sant'Anna (2021) estimates closely mirror the TWFE estimates, reinforcing the robustness of the findings. This alignment supports the theoretical prediction that concerns about TWFE estimators in staggered settings are of limited relevance in my context.

Figure A4b suggests that the effect of bottle bills on marginal births is concentrated in counties with high inequality—where, as predicted by theory (see Section A.1), there are more individuals who benefit from redeeming containers discarded by others and more individuals who discard containers without redeeming them. The figure plots treatment effect heterogeneity by place-based characteristics, with each x-axis group representing a distinct combination of county-level traits. For example, the left-most group shows estimates for counties with below-median income and below-median eligibility for federal transfers. Solid markers reflect estimates from the preferred fixed-effects specification (as in Figures 4a), while hollow markers represent alternative specifications, plotted in corresponding order to Figure 4b estimates. Pink lines denote 95% confidence intervals for the overall effect. Bottle bills have no detectable effect on marginal births in counties with below-median transfer eligibility (left panel), consistent with the idea that fewer residents in these areas rely on recycling. In contrast, reductions in marginal births emerge in counties with above-median transfer eligibility (right panel), particularly where average income is also above the median. While somewhat noisy and specification-sensitive in low-income areas, the effect is robust and precise in higher-income, high-transfer counties—the most unequal settings. This pattern aligns with theoretical predictions about where the economic gains from recycling are most likely to materialize.

DiDiT Estimates Figure 5b presents estimates from an alternative estimation strategy—a DiDiT model rather than DiDiD—based on distinct identifying assumptions. While the identification differs, the estimates in their scaled form should be comparable to the DiDiD results. Reassuringly, the DiDiT analysis yields similar conclusions, reinforcing the main findings and suggesting that the observed effects as well as conclusions are not artifacts of any single identification strategy. Specifically, Figure 5b plots estimates from the DiDiT specification detailed in equation (7), scaled by average winter temperature (blue markers), non-winter temperature (orange), and their difference (green).[36] The green markers in Figure 5b are comparable *in spirit* to the purple markers in Figure 4b. Importantly, both estimates lead to qualitatively similar conclusions: increases in recycling earnings reduce the incidence of marginal births to low-education mothers, with no consistent effects observed in other birth weight bins. Specifically, the DiDiT analysis shows that bottle bills reduced the incidence of low education marginal births .13 pp in non-winter relative to winter months (Figure 5b, third column from the left, green markers) and the DiDiD analysis estimates a .15 pp reduction (Figure 4b, third column from the left, purple markers).

The qualitative and empirical comparability of the DiDiT and DiDiD estimates are not guaranteed, and exact empirical equivalence should not be expected. The scaled DiDiT estimates are only comparable in

---

[36]See equation (4) for details on how dose response estimates can be scaled to compare with seasonal difference estimates.

spirit to the DiDiD estimates for several reasons.

1. As shown in Figure 3a, 70% of seasonality in recycling earnings is driven by weather. The green markers in Figure 5b isolate the effect of bottle bills on birth outcomes through the temperature channel alone, whereas the DiDiD estimates capture the effect through the broader seasonality channel—reflecting both seasonal variation in weather and in beverage consumption patterns (e.g., increased barbecues, tailgates, etc.).

2. To reduce the computational burden of estimation, the DiDiT model is estimated assuming linear effects of temperature and precipitation. This may attenuate the estimates, as Figures 3b and A2 suggest some non-linearity in the recycling earnings response to weather.

3. Finally, when scaling DiDiT estimates by seasonal differences in weather, I do not weight by the number of births, which may introduce additional differences.

In conclusion, it is reassuring that the DiDiD and DiDiT approaches yield remarkably similar estimates, despite relying on distinct strategies to answer the same question. Moreover, the modestly smaller DiDiT estimates are consistent with findings from Section 3 that seasonality in the size of the empties reservoir is not entirely driven by temperature.

Up to this point, the discussion has focused on the impact of bottle bills in non-winter relative to winter months. However, if bottle bills affect marginal births through the recycling earnings channel, their effects should be apparent in both seasons—albeit more pronounced in non-winter. The DiDiT approach enables separate estimation of effects in non-winter and winter months, providing a means to test this prediction. Figure 5b confirms this expected pattern: bottle bills reduce the incidence of marginal births among low-education mothers in both seasons, with a larger decline in non-winter months. This supports the conclusion that the DiDiD estimates are not driven by an increase in winter birth incidence, but rather by a more substantial reduction in non-winter months.

One potential concern is that rents from recycling should dissipate due to entry by new recyclers (as illustrated in Figure A1c), and that the DiDiD estimates may instead reflect individuals switching into a healthier line of work—recycling—during non-winter months, relative to their winter outside options. However, the DiDiT results, and their consistency with the DiDiD estimates, suggest this explanation is unlikely. Reductions in marginal birth incidence are observed whether identified through seasonal differences in recycling earnings or short-run, temperature-driven shocks to the value of the empties reservoir. Given labor market frictions, it is implausible that individuals would enter or exit recycling in response to such short-term variation. Thus, the DiDiT estimates likely reflect genuine changes in recycling wages—i.e., income shocks—rather than shifts in occupation-related health exposure.

Figure 5a plots DiDiT estimates scaled to reflect the impact of a 1% increase in recycling earnings on the incidence of low education mother births throughout the weight distribution. In addition, Figure 5a shows robustness in these estimates across model specifications, with red markers corresponding to the model specification described in equation 7 and plotted with different scaling factors in Figure 5b. The third column from the left documents the estimates of interest—the impact of a 1% increase in recycling earnings on the incidence of low education marginal births.

Figure 5a documents that a 1% increase in recycling earnings leads to a .0035 - .0048 pp reduction in

27

marginal births or a 0.04% decrease in the incidence of low birth weight.[37] As a point of comparison, the DiDiD estimates can also be scaled to reflect the effect of a 1% increase in recycling earnings. Assuming the DiDiD estimate captures the impact of a 20% increase in earnings—based on the fact that earnings are 20% higher in non-winter months relative to winter months (see Figure 3a)—this implies that a 1% increase in earnings reduces the incidence of low birth weight by 0.09%, i.e., $\frac{1.8\%}{20\%} = 0.09\%$. The similarity in magnitude between the DiDiD and DiDiT elasticity estimates is reassuring. However, the identifying assumptions underlying the DiDiT estimates are less stringent. DiDiD estimates rely on a relatively large seasonal change in aggregate earnings, making the assumption that aggregate changes reflect individual earnings changes more tenuous. In contrast, the DiDiT design leverages more marginal, and thus more plausibly reflects variation in individual households' earnings from recycling. For this reason, I focus on the DiDiT elasticity estimates.

Finally, I apply the findings from the birth outcome analysis to complete the following sanity check: what share of the total aggregate scrap value is redeemed by households who recycle others' discarded containers for cash? To answer this, I combine the finding from Section 2 that recycling earnings average approximately $1400 (in 2015$) annually among households who recycle, with the result that recycling earnings shocks affect only marginal births. Specifically, since effects are concentrated among marginal births to low-education mothers, I approximate the share of households who recycle as the share of all births that are both to low-education mothers and of marginal birth weight—approximately 1% (see Table A3). This approximation aligns closely with external estimates. For example, Ashenmiller (2011) finds that roughly 1% of households in the greater Santa Barbara area earn a substantial portion of their income through recycling, and Medina (2008) reports that approximately 1% of people in low- and middle-income countries rely on recycling to make ends meet. Assuming that the recycling earnings estimates in Section 2 capture the recycling income of most such recyclers, this would imply that approximately 30% of the total scrap value plotted as orange bars in Figure 1b is recycled for cash.

4.4   Recycling Earnings vs. Food Stamp Benefits: Prevalence and Importance

As a final step, I benchmark the estimated birth weight effects of recycling income against those of food stamp benefits. The finding that a 1% increase in recycling earnings reduces the incidence of low birth weight (ILBW) by 0.04% implies that informal recycling markets reduce the ILBW by approximately 4%. In other words, moving from a world with no informal recycling to one with the existing level of recycling activity reduces the ILBW rate by 4%. This implies that informal recycling markets are roughly 50% as effective as food stamps at reducing ILBW (Almond et al. 2011).

In this section, I show that differential household responses to the same size income increase is a key driver of the comparability of food stamps and informal recycling markets in their effectiveness at reducing the ILBW. Specifically, I show that the same size income increase is substantially more likely to reduce the risk of a low-weight birth for recycling households than for the average food stamp recipient. To demonstrate this, I introduce a stylized model of low birth weight risk and its sensitivity to income changes. The model allows me to decompose the effects of food stamps and informal recycling markets into their underlying

---

[37]See equation 8 for details on the scaling procedure.

components. The model is as follows:

$$\Delta_{\text{LBW}} = \sum_{n \in N_z} \mathbb{1}[t + e_n > l_w] \times \mathbb{1}[e_n < l_w] + \epsilon_n$$

Δ_{LBW} denotes the change in the number of births with low birth weight due to an increase of income of size t from source z. N_z denotes the number of households that receive t, e_n household n's earnings net of t. Finally, l_w denotes an earnings threshold below which a household n has a low weight birth and above which they do not. Hence in expectation:

$$\mathbb{E}[\Delta_{\text{LBW}}] = N_z \times \underbrace{\mathbb{P}(e < l_w)}_{P_l^z} \times \underbrace{\mathbb{P}(e + t_z > l_w | e < l_w)}_{P_t^z}$$

P_l^z denotes the share of households receiving transfer t from source z whose income **before** the transfer falls below the low birth weight threshold. P_t^z denotes the share of those households whose income **after** receiving the transfer rises above the low birth weight threshold. With this notation in hand, consider the following quantity:

$$\mathbb{E}[\Delta_{\text{LBW}}^f] = N_f \times P_l^f \times P_t^f$$

In words, this equation represents the impact of any transfer or increase in income on the number of low weight births. For example, in the case of food stamps $t_f = \$750$. Almond et al. (2011) show that, on average, a \$750 (in 2015 \$) increase in food stamp benefits reduces the incidence of low birth weight (ILBW) by 7.8% among beneficiaries, i.e. $\frac{\mathbb{E}[\Delta_{\text{LBW}}]}{\#_{\text{LBW}}} = 7.8\%$, where $\#_{\text{LBW}}$ is the number of low weight births. Back-of-the-envelope calculations imply that the equivalent quantity for recycling earnings is 2.2%.[38]

Taken together, the 7.8% and 2.2% impact of a \$750 increase in food stamps and recycling earnings respecively implies:

$$N^f \cdot P_l^f \cdot P_t^f > N^r \cdot P_l^r \cdot P_t^r \quad \text{and} \quad \frac{N^r \cdot P_l^r \cdot P_t^r}{N^f \cdot P_l^f \cdot P_t^f} = 28\% \tag{9}$$

Equation 9 implies that if recycling households and food stamp recipients each received the same income increase—\$750—the resulting reduction in the incidence of low birth weight (ILBW) among recycling households would be 28% as large as that observed for food stamp recipients. This estimate is lower than the unadjusted comparison. With no adjustments, informal recycling markets are 50% as effective as food stamps. This lower adjusted efficiency reflects the fact that recycling income reaches far fewer people than food stamps.

Next I adjust for differences in the size of the population reached by food stamps and informal recycling markets, respectively. Specifically, I divide equation (9) through by the population shares reached by food stamps (13%; Almond et al. (2011)) and by recycling earnings (1%; see Section 4.3), yielding:

$$P_l^f \cdot P_t^f < P_l^r \cdot P_t^r \quad \text{and} \quad \frac{P_l^r \cdot P_t^r}{P_l^f \cdot P_t^f} = 360\% \tag{10}$$

Equation 10 shows that households who recycle are 3.5× more likely to avoid a low birth weight outcome

---

[38]Section 4.3 shows that a 1% increase in recycling earnings reduces the incidence of low birth weight 0.04% ILBW. If most recipients earn \$1400 in recycling income following bottle bill implementations (see Section 2 for details), then a \$750 increase in recycling earnings is equivalent to a 54% increase in earnings or a 2.2% decrease in the incidence of low birth weight.

from the same income increase as the average food stamp recipient.

This conclusion aligns with findings from the food expenditure analysis, which suggest that recycling households are relatively better off than the average food stamp recipient. Food stamp recipients—who are disproportionately single mothers—are plausibly further from meeting basic nutritional needs (Hoynes and Schanzenbach 2009). For some of these households, a \$750 increase may be insufficient to eliminate the nutritional deficits contributing to low weight births. In contrast, recycling households tend to include multiple working-aged adults (see Section 2) and may face less severe deprivation. For these households, a comparable income gain is plausibly more likely to push them above the "nutritional adequacy threshold," preventing marginal low-weight births.

This interpretation is also consistent with the empirical pattern that recycling income primarily reduces the incidence of marginal births—those weighing between 2,000 and 2,500 grams—while food stamp benefits have the greatest effect among births below 2,000 grams (Almond et al. 2011). This distinction reinforces the view that food stamp benefits help households with more severe nutritional deficits on average, whereas recycling income benefits those who are slightly better off.

5   Conclusion

In low- and middle-income countries, roughly 1% of people rely on recycling to make ends meet. In high-income countries, however, far less is known about this livelihood (Medina 2008). This paper examines the extent to which informal recycling constitutes an important and prevalent livelihood in the United States. Earlier work by Ashenmiller (2009, 2011) uses survey data to show that about 1% of households in the greater Santa Barbara region earn 7% of their income recycling. Ashenmiller (2010) further demonstrates that recycling income can have broader societal effects, finding that bottle bills reduce petty crime by 10%. Extending this literature, I study the effect of bottle bills on household food expenditures and birth outcomes—two proxies for economic well-being and material hardship.

This paper provides new evidence that informal recycling markets partially filled gaps in the U.S. social safety net between 1970 and 2000. The findings show that 30% of aggregate scrap value is recycled for cash—an amount equivalent to 10–32% of total food stamp benefits. These earnings accrue to roughly 1% of the population. Most impressively, informal recycling markets are found to be 50% as effective as food stamps at increasing food expenditures and reducing the incidence of low birth weight (ILBW). The efficiency with which recycling earnings improve these canonical measures of well-being reflects the implicit targeting built into this income source: earnings accrue to households who are both able-bodied enough to engage in informal recycling and in sufficient economic need to choose to do so.

To provide this evidence, I exploit the introduction of bottle bills alongside seasonal variation in earnings from recycling to establish their prevalence and importance. These two sources of variation in earnings allow me to study recycling income despite limited data on both the waste sector and informal labor markets.

First, I estimate average recycling earnings using the introduction of bottle bills and food expenditure data from the Panel Study of Income Dynamics (PSID). I show that bottle bills increase earnings from recycling. To estimate the effect on household earnings, I combine a difference-in-differences quasi-experimental design with the Generic Machine Learning Framework (CDDF) (Chernozhukov et al. 2018). I find that bottle bills increase food expenditures by 10% among households plausibly engaged in informal recycling. I validate

this result using a traditional event-study analysis. This effect implies that recycling households earn at least $1400 (in 2015 $) annually from recycling or collect an average of 38 containers/day. Additionally, the analysis shows that households who recycle tend to be more able-bodied than similarly situated households who do not. This pattern is consistent with the conjecture that, because informal recycling is labor-intensive, it disproportionately benefits those with the time and physical capacity to perform the work.

Second, using monthly redemption data from thirty California counties from 2005 to the present, I show that temperature drives fluctuations in beverage consumption, which in turn affects the number of containers available for redemption and, ultimately, recycling earnings. Specifically, aggregate recycling earnings are 20% higher in non-winter months relative to winter months, and 70% of this seasonality is explained by temperature variation. I leverage this relationship to relax the identifying assumptions needed to estimate the impact of recycling earnings on birth outcomes and to infer the effects of marginal changes in earnings.

Third, I estimate the prevalence and importance of recycling earnings. Specifically, I estimate the impact of variation in recycling earnings on birth outcomes. For estimation, I use changes in recycling earnings from bottle bill introductions and seasonal fluctuations in beverage consumption. Unlike the household-level food expenditure data, the natality data are cross-sectional and do not follow individuals over time, limiting the ability of covariate controls to fully address identification concerns in a DiD framework. To overcome this, I exploit within-place, within-year variation in recycling earnings to identify the effect of income changes on birth outcomes. Specifically, I employ two complementary estimators: a triple-difference (DiDiD) design and a difference-in-difference-in-temperature (DiDiT) strategy. Both approaches yield consistent results: increases in recycling earnings reduce the incidence of marginal low birth weight births—those weighing 2,000–2,500 grams—among mothers without a high school diploma. This reduction plausibly reflects modest improvements in maternal nutrition, consistent with earlier findings that recycling income primarily accrues to relatively better-off low-income households facing moderate nutritional deficits. I find no comparable effects in other birth weight bins or among mothers with a high school diploma. In addition, recycling earnings shocks do not impact fertility, ruling out the concern that compositional changes drive the effect. Event-study analyses further support the causal interpretation: the decline in marginal births among low-education mothers begins immediately after the earnings shocks, persists for over a decade, and shows no evidence of pre-trends. No similar pattern is observed in other demographic groups or weight bins.

References

Almond, Douglas, Hilary Hoynes, and Diane Whitmore Schanzenbach, "Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes," *The Review of Economics and Statistics*, May 2011, *93* (2), 387–403.

⎯⎯ , **Janet Currie, and Valentina Duque**, "Childhood Circumstances and Adult Outcomes: Act II," *Journal of Economic Literature*, December 2018, *56* (4), 1360–1446.

Angrist, Joshua D. and Jörn Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2009.

Arensman, Russ, "Time Has Come for a Deposit Bill," *The Rocky Mountain Collegian*, June 1981, *89* (117).

Ashenmiller, Bevin, "Cash Recycling, Waste Disposal Costs, and the Incomes of the Working Poor: Evidence from California," *Land Economics*, 2009, *85* (3), 539–551.

⎯⎯ , "Externalities from Recycling Laws: Evidence from Crime Rates," *American Law and Economics Review*, February 2010, *12* (1), 245–261.

⎯⎯ , "The Effect of Bottle Laws on Income: New Empirical Results," *American Economic Review*, May 2011, *101* (3), 60–64.

Bailey, Martha, Karen Clay, Price Fishback, Michael R. Haines, Shawn Kantor, Edson Severnini, and Anna Wentz, "U.S. County-Level Natality and Mortality Data, 1915-2007," 2018.

Bingham, Tayler and Paul Mulligan, *The Beverage Container Problem: Analysis and Recommendations*, Washington, D.C.: U.S. Environmental Protection Agency, 1972.

Callaway, Brantly and Pedro H.C. Sant'Anna, "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Iván Fernández-Val, "Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, with an Application to Immunization in India," Working Paper 24678, National Bureau of Economic Research June 2018.

Colmer, Jonathan and Jennifer L. Doleac, "Access to Guns in the Heat of the Moment: More Restrictive Gun Laws Mitigate the Effect of Temperature on Violence," *The Review of Economics and Statistics*, 11 2023, pp. 1–40.

Costle, Douglas, Juanita Kreps, Cecil Andrus, F. Ray Marshall, W. Michael Blumenthal, Charles Warren, Eliot Cutler, Nina Cornell, and Alvin Alm, "Committee Findings and Staff Papers on National Beverage Container Deposits of the Resource Conservation Committee," Technical Report SW-733, National Service Center for Environmental Publications 1978.

Currie, Janet, *The Invisible Safety Net: Protecting the Nation's Poor Children and Families*, Princeton: Princeton University Press, 2006.

Davis, Joel, "The Bottle Bill," *Environs*, 1982, *6* (2), 1,7–9.

Davison, Tamara, "A Guide To Extended Producer Responsibility (EPR) in the US," *CleanHub*, 2024.

Eichner, Thomas and Rüdiger Pethig, "Product Design and Efficient Management of Recycling and Waste Treatment," *Journal of Environmental Economics and Management*, 2001, *41* (1), 109–134.

Franchot, Peter, *Bottles and Cans: The Story of the Vermont Deposit Law*, The National Wildlife Federation, 1978.

Fraundorf, Kenneth C., "The Social Costs of Packaging Competition in the Beer and Soft Drink Industries," *Antitrust Bulletin*, Winter 1975, *20* (4), 803–831.

Fullerton, Don, "The Circular Economy," Working Paper 32419, National Bureau of Economic Research May 2024.

___ **and Ann Wolverton**, "Two Generalizations of a Deposit-Refund Systems," *American Economic Review*, May 2000, *90* (2), 238–242.

___ **and Thomas C. Kinnaman**, "Garbage, Recycling, and Illicit Burning or Dumping," *Journal of Environmental Economics and Management*, 1995, *29* (1), 78–91.

Godush, Brett, "The Hidden Value of a Dime: How a Federal Bottle Bill Can Benefit the Country," *Vermont Law Review*, 2001, *25*, 855. Accessed: February 24, 2025.

Goodman-Bacon, Andrew, "Difference-in-differences with variation in treatment timing," *Journal of Econometrics*, 2021, *225* (2), 254–277.

Harvey, Jenna, Christine Hegel, and Chris Hartmann, "Independent Recyclers in New York City: Sector Profile and Pathways to Inclusion," 2023.

Holaday, Carsen, "First-in-the-nation law forces certain stores to take old clothes back for free as officials take aim at 'fast fashion'," *The US Sun*, 2024.

Horan, Patrick M. and Peggy G. Hargis, "County Longitudinal Template, 1840-1990," 1995.

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

___ **, Doug Miller, and David Simon**, "Income, the Earned Income Tax Credit, and Infant Health," *American Economic Journal: Economic Policy*, February 2015, *7* (1), 172–211.

___ **, 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* (7), 813–827.

Hsiang, Solomon, "Climate Econometrics," *Annual Review of Resource Economics*, 2016, *8*, 43–75.

Kramer, Michael S., "Determinants of low birth weight: methodological assessment and meta-analysis," *Bulletin of the World Health Organization*, 1987, *65* (5), 663–737.

___ , "Intrauterine Growth and Gestational Duration Determinants," *Pediatrics*, 10 1987, *80* (4), 502–511.

Lederer, Bob, "Bottle Bills: State-By-State," *Beverage World*, September 1976, p. 25.

Medina, Martin, "Organizing Waste Pickers to Enhance Their Impact," 2008.

Miao, Wei, Yiting Deng, Wei Wang, Yongdong Liu, and Christopher S. Tang, "The effects of surge pricing on driver behavior in the ride-sharing market: Evidence from a quasi-experiment," *Journal of Operations Management*, 2023, *69* (5), 794–822.

Moore, Marvin M., "The Case for the Regulation of Nonreturnable Beverage Containers," *Kentucky Law Journal*, 1976, *64* (4), 767–784.

Moore, W. Kent and David L. Scott, "Beverage Container Deposit Laws: A Survey of the Issues and Results," *Journal of Consumer Affairs*, 1983, *17* (1), 57–80.

Nichols, Albert L. and Richard J. Zeckhauser, "Targeting Transfers through Restrictions on Recipients," *The American Economic Review*, 1982, *72* (2), 372–377.

Olden, Andreas and Jarle Møen, "The triple difference estimator," *The Econometrics Journal*, March 2022, *25* (3), 531–553.

Palmer, Karen and Margaret Walls, "Optimal policies for solid waste disposal Taxes, subsidies, and standards," *Journal of Public Economics*, 1997, *65* (2), 193–205.

Pei, Zhuan, Jörn Steffen Pischke, and Hannes Schwandt, "Poorly Measured Confounders are More Useful on the Left than on the Right," *Journal of Business & Economic Statistics*, 2019, *37* (2), 205–216.

Peterson, Charles, "Price Comparison Survey of Beer and Soft Drinks in Refillable and Nonrefillable Containers," Technical Report SW-531, National Service Center for Environmental Publications 1976.

Porter, Richard C., "A social benefit-cost analysis of mandatory deposits on beverage containers," *Journal of Environmental Economics and Management*, 1978, *5* (4), 351–375.

Rehill, Patrick, "How Do Applied Researchers Use the Causal Forest? A Methodological Review," *International Statistical Review*, 2025.

Reloop Platform, "Global Deposit Book 2020: An Overview of Deposit Systems for One-Way Beverage Containers," 2020.

Ross, David Michael, "The Massachusetts bottle bill, 1967-1979: a study of policy failure from the perspective of interest-group liberalism," Master's thesis, McGill University 1982.

Roth, Jonathan, Pedro H.C. Sant'Anna, Alyssa Bilinski, and John Poe, "What's trending in difference-in-differences? A synthesis of the recent econometrics literature," *Journal of Econometrics*, 2023, *235* (2), 2218–2244.

Saltzman, Harold, Roy Levy, and John C. Hilke, "Transformation and Continuity: The U.S. Carbonated Soft Drink Bottling Industry and Antitrust Policy Since 1980," Technical Report, Federal Trade Commission 1999.

Sanger, David, "Resurgence of Aluminum Recycling Brings Cash to the Poor," *The New York Times*, August 1983.

Savov, Alexi, "Asset Pricing with Garbage," *The Journal of Finance*, 2011, *66* (1), 177–201.

Talbott, Taylor Cass, "Oregon's Bottle Bill: Opportunities and Challenges for Inclusive Waste Management," 2021.

The Associated Press, "Cottage Industry Thrives On Returnable Bottle Law," *The Associated Press*, October 1982.

United States Department of Agriculture, "Official USDA Food Plans: Cost of Food at Home at Four Levels, U.S. Average, December 2015," 2016.

United States General Accounting Office, "Solid Waste: Trade-offs Involved in Beverage Container Deposit Legislation," Technical Report RCED-91-25, U.S. General Accounting Office, Washington, D.C. November 1990. Prepared at the request of Senators Mark O. Hatfield, James M. Jeffords, and Representative Paul B. Henry.

Varela, Ana Varela, "Surge of Inequality: How Different Neighborhoods React to Flooding," *SSRN Electronic Journal*, 2023.

Wagenbach, Jeffrey B., "The Bottle Bill: Progress and Prospects," *Syracuse Law Review*, 1985, *36*, 759.

White, Russ, "Michigan's Bottle Deposit Law Celebrates 40 Years of Keeping Michigan Clean," *WKAR*, December 2018.

Table 1: The Effect of Bottle Bills on Birth Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Average Birth Weight						
$\mathbb{1}(\text{BB}) \times \mathbb{1}(\text{not winter}) \times \mathbb{1}(<\text{HS})$	3.30*	4.02**	2.31*	3.39*	1.75	2.38
	(1.72)	(1.65)	(1.19)	(1.79)	(1.51)	(1.77)
$\mathbb{1}(\text{BB}) \times \mathbb{1}(\text{not winter}) \times \mathbb{1}(\text{HS})$	0.86	0.80	0.42	0.12	-0.10	0.01
	(1.37)	(1.39)	(1.43)	(1.26)	(1.35)	(1.63)
$\mathbb{1}(\text{BB}) \times \mathbb{1}(\text{not winter}) \times \mathbb{1}(>\text{HS})$	-1.45	-1.57	0.23	-0.99	0.73	
	(1.43)	(1.41)	(1.60)	(1.32)	(1.87)	
within R^2	0.210816	0.000746	0.000469	0.000007	0.000001	0.000001
Panel B: Number of Births						
$\mathbb{1}(\text{BB}) \times \mathbb{1}(\text{not winter}) \times \mathbb{1}(<\text{HS})$	-0.44	-0.49	0.17	-0.30	0.04	-0.15
	(0.79)	(0.84)	(0.97)	(1.15)	(1.14)	(1.30)
$\mathbb{1}(\text{BB}) \times \mathbb{1}(\text{not winter}) \times \mathbb{1}(\text{HS})$	-0.18	0.06	-0.25	-0.36	-1.15*	-1.10
	(0.53)	(0.60)	(0.55)	(0.58)	(0.68)	(0.71)
$\mathbb{1}(\text{BB}) \times \mathbb{1}(\text{not winter}) \times \mathbb{1}(>\text{HS})$	-0.20	-0.57	-0.75	-0.20	0.06	
	(0.42)	(0.52)	(0.72)	(0.46)	(0.47)	
within R^2	0.049415	0.002953	0.002858	0.000002	0.000007	0.000006
Observations	2439765	2439765	2439765	2439765	2439765	2439765
County x Ed.		x				
County x Season	x	x		x		
County x Season x Ed			x		x	x
County x Season x Year						x
Season x Year	x	x	x			
County x Year	x	x	x			
Season x Ed x Year				x	x	x
County x Ed x Year				x	x	x

Notes: Table reports bottle bills' estimated effect on the average birth weight (Panel A) and number of births (Panel B) in non-winter relative to winter months. Specifically, table regresses the birth count on the following interaction $\mathbb{1}(\text{BB}) \times \mathbb{1}(\text{not winter})$, where $\mathbb{1}(\text{BB})$ is an indicator for counties with active bottle bills and $\mathbb{1}(\text{not winter})$ is an indicator for non-winter months. Estimates are broken out by education group, where $\mathbb{1}(<\text{HS})$ is an indicator for low education mothers, $\mathbb{1}(\text{HS})$ middle education mothers, and $\mathbb{1}(>\text{HS})$ high education mothers. Columns differ in the fixed effect included in estimation. The first column includes county x season, county x year, and season x year fixed effects, operationalizing the estimators with estimates reported in panel A, row 1 and panel B, row 1-3 as DiDiD estimators. Columns 2-4 include additional fixed effects, controlling for education specific characteristics in addition to those necessary to operationalize a DiDiD estimator. Column 5 includes fixed effects equivalent to those of column 1, but allowing for fixed effects to vary by education group, i.e., column 5 reports estimates from a fully stacked DiDiD estimator. Finally, column 6 reports estimates from a DiDiDiD estimator, where the fourth difference comes from comparing differences in the effect of bottle bills on low and middle education mothers in non-winter relative to winter months to that of high education mothers. Standard errors are clustered by state, the level of treatment.

Figure 1: Bottle Bills: Implementation, Economic Significance, and Political Opposition
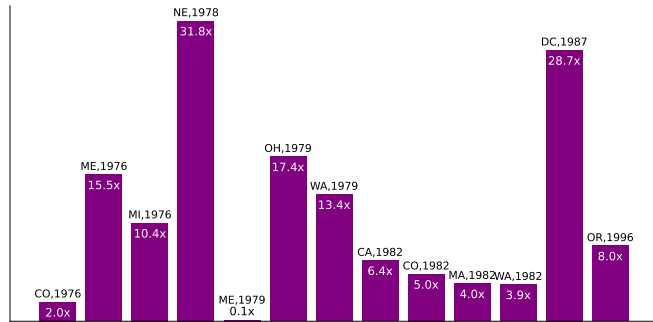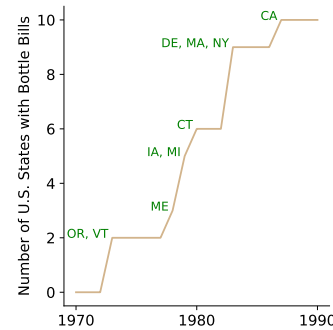
a How Bottle Bills Work
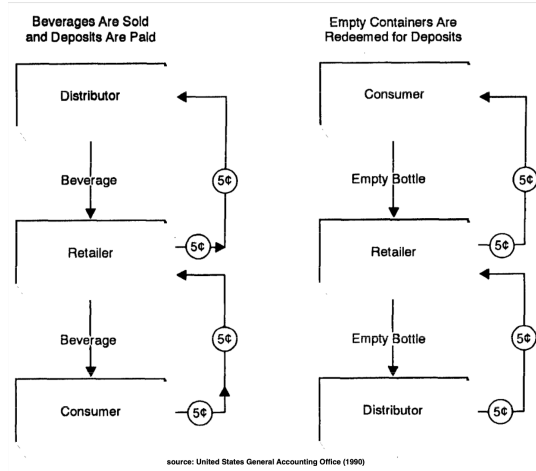
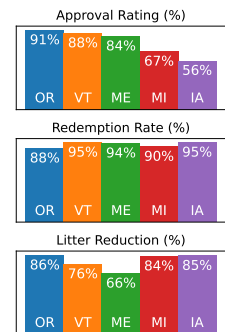**b** Total Container Scrap Value Relative to Food Stamp Benefits
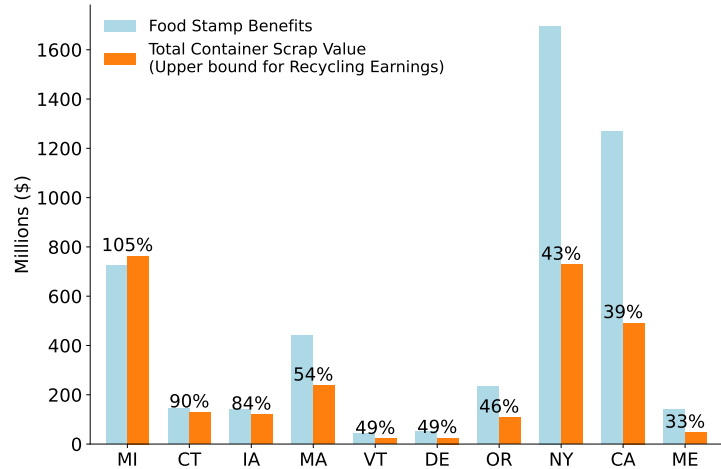
c Ratio of Opponent to Proponent Spending

**d** Bottle Bill Introductions

e Initial Responses

Notes: Figure summarizes bottle bill implementations, their economic significance, and political opposition to the policy. **Panel a** is an illustrations from United States General Accounting Office (1990) on the intended inter-workings of bottle bills from the consumer, retailer and beverage distributor perspective. **Panel b** compares the total scrap value of beverage containers (orange bars) by bottle bill state (denoted on the x-axis) to aggregate food stamps benefits (blue bars) in millions of dollars. **Panel c** reports campaign spending by bottle bill opponents relative to proponents on bottle bill ballot measures as reported by bottlebill.org. Each bar reflects how many dollars opponents spent for every $1 proponents spent on a given ballot measure denoted with the state and year annotation above each bar. **Panel d** plots when each bottle bill state implemented bottle bills between 1969 and 2002. The yellow line documents the total number of U.S. states with an active bottle bill in each year. The green annotations document which states implemented the program in a given year. For example, in 1980, Connecticut implemented a bottle bill, making it the sixth U.S state to introduce a deposit refund program for containers. **Panel e** documents initial responses to bottle bills as reported by Moore and Scott (1983) in the first five states to implement. The top panel documents the approval rating of bottle bills post implementation by state, the middle the redemption rate—of purchased containers how many were returned for a refund, and the bottom the extent to which bottle bills reduced beverage container litter.

Figure 2: Effect of Bottle Bills on Recycling Earnings—Evidence from Food Expenditures

a Group Average Treatment Effects + Characteristics
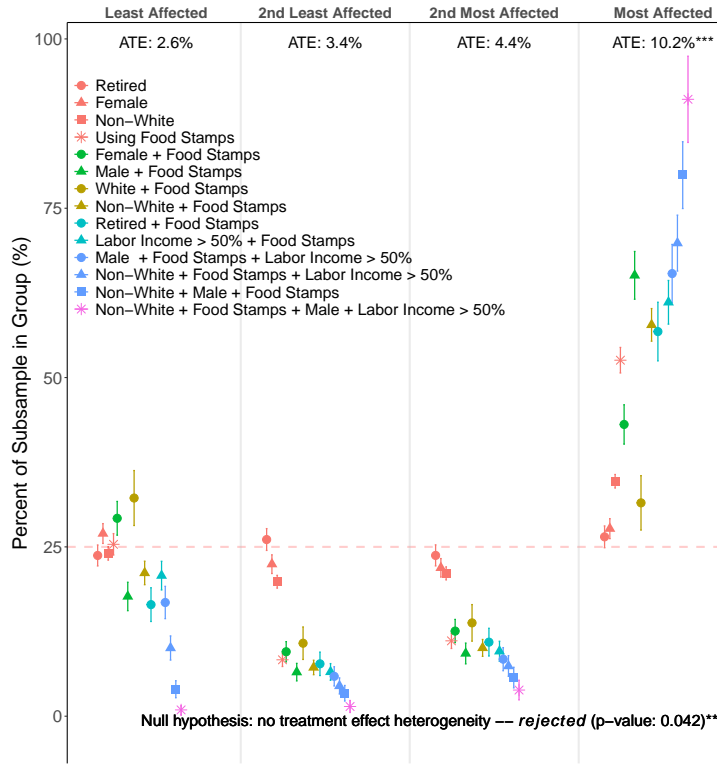
b Proportion of Households in Each Subsample Who Recycle
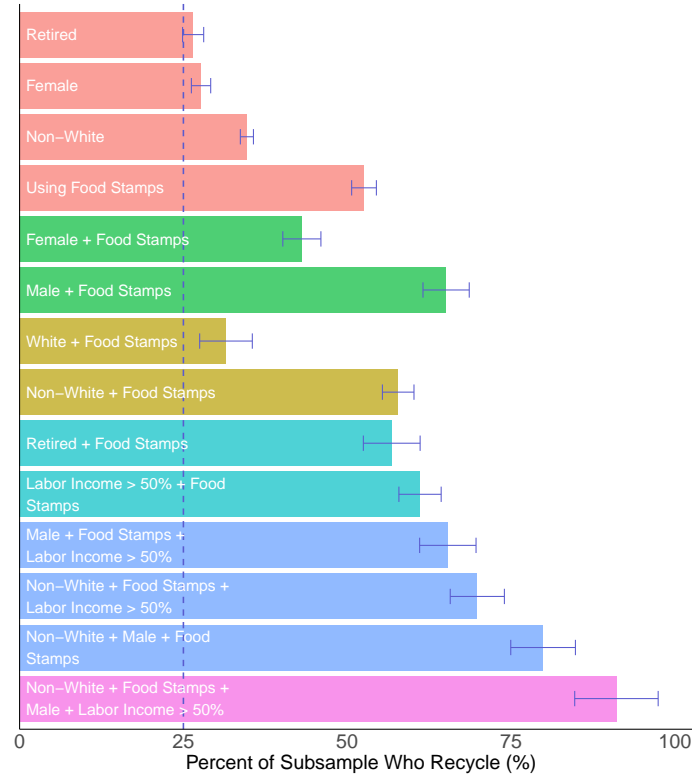
c Event Study: All + Recycling Households
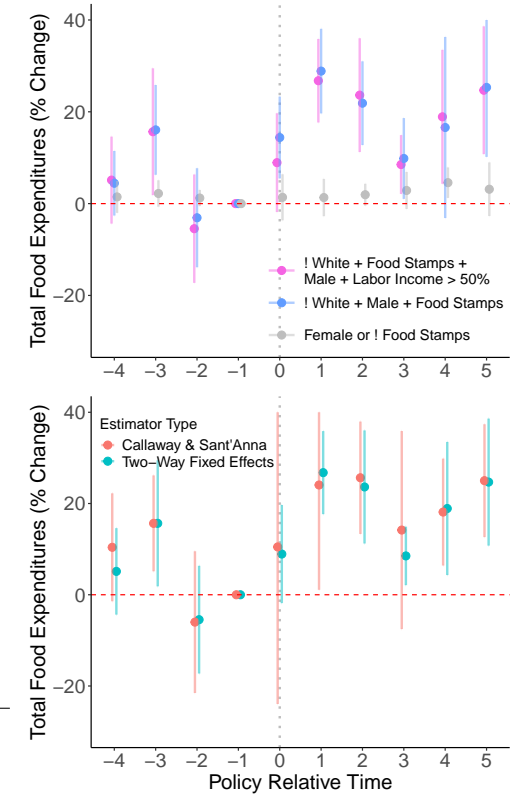
Notes: Figure presents estimates of bottle bills' effect on household food expenditures estimated with (i) the Generic ML Framework (referred to as CDDF) applied to a Difference-in-Differences (DiD) quasi-experimental design and (ii) a traditional DiD approach, with subsample analysis informed by findings from the Generic ML Framework (Chernozhukov et al. 2018). **Panel A** summarizes the results of estimating heterogeneous effects of bottle bills on food expenditures (in logs) with CDDF. The panels' four sections detail the proportion of different household subsamples in each group and average treatment effects (ATEs) for the *least affected* (left-most section) to *most affected* (right-most section) households. ATEs are written at the top of each groups section with statistical significance at the 10% level denoted by * and 5% with **. Each section plots the proportion of a given subsample in each group. Point color and shape indicate subsample identity as specified in the plot legend. For example, red circles represent the average proportion of the subsample of households with female heads in each group. All demographic related subsample identifiers pertain the household head, i.e., female, retired, non-white, and male refer to subsamples with heads who take on the aforementioned characteristics. Food stamps and labor income > 50% refer to subsamples with households who used food stamps in the pre-period and who earn more than 50% of their income through means other than transfers, respectively. Vertical lines denote 90% confidence intervals, which reflect sample-splitting uncertainty—i.e., variation in the proportion of a subsample within a given group across "auxiliary" and "main" sample splits. **Panel B** replicates the information plotted in Panel A's right-most section regarding the *Most Affected* group in greater detail. Assuming most households assigned to this group recycle for cash, the proportion of each subsample in this group provide an estimate for the proportion of each subsample who recycle for cash, as emphasized in the panel's title and y-axis label. **Panel C**, top figure plots event-study estimates from equation 1, with markers colored by subsample classification from CDDF. Gray markers denote estimates from a subsample with a low proportion of households who recycle for cash, while pink and purple markers denote estimates from subsamples with high proportions of recyclers. All coefficients are normalized to the year before bottle bill implementation (policy relative time = −1). Bottom figure replicates the analysis for the subsample associated with the pink markers in the top panel using the Callaway and Sant'Anna (2021) estimator. TWFE estimates are shown in blue (reproducing the pink series above); red markers plot the Callaway and Sant'Anna (2021) estimates. In both the top and bottom plot, vertical lines denote point-wise 90% confidence intervals clustered at the state level.

**Figure 3:** Seasonality in Recycling Earnings

a
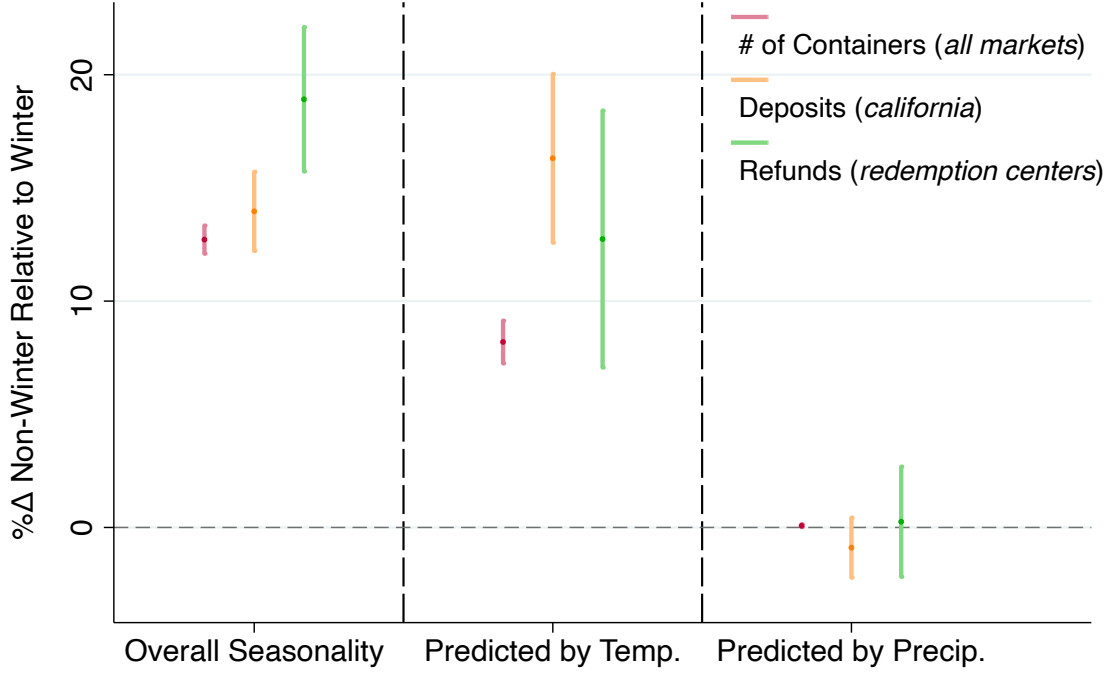
b
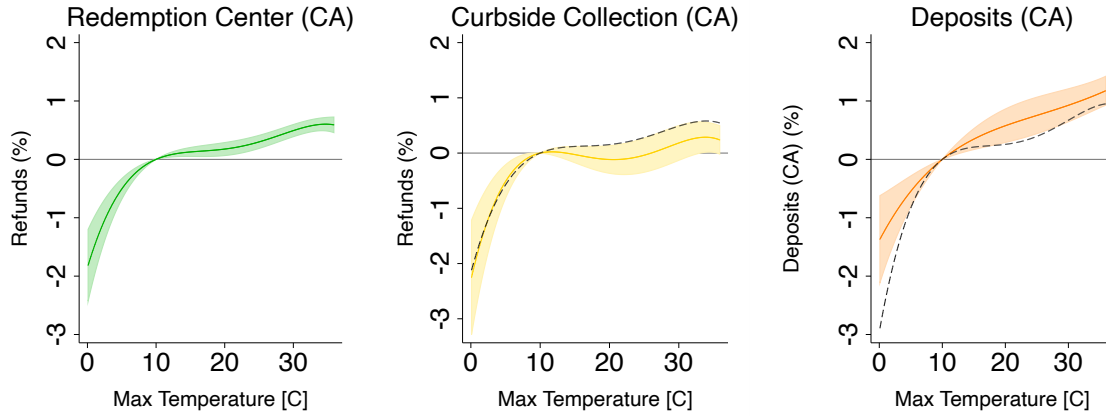
Notes: Figure plots seasonality in the empties reservoir value (deposits) and recycling earnings (refunds paid out at redemption centers). **Panel A**, column 1 plots the overall increase in non-winter relative to winter months in empties reservoir value (orange) and recycling earnings (green) in California (see estimation equation 3). For example, a coefficient of 25 implies that on average monthly earnings/scrap value are 25% higher in non-winter relative to winter. Dots mark point estimates and shading depict 95% confidence intervals. Panel A, column 2(3) plots the relative increase as predicted by temperature (precipitation), where the temperature (precipitation) prediction is computed from the estimated refund/deposit dose response (see estimation equation 4). **Panel B** plots the temperature response of recycling earnings or empties reservoir value (in logs) on f(Weather) estimated with regression equation (2). Specifically, Panel B illustrates the response in percent terms to maximum daily temperature, relative to 10°C. Solid lines represent point estimates and shading depicts 95% confidence intervals. Redemption center refunds (green, column 1) proxy for recycling earnings, as the majority of refunds paid out at redemption centers go to individuals other than the original consumer (Ashenmiller 2009, Ashenmiller 2011). Refunds paid to curbside collection companies (yellow, column 2) and deposits paid by households (orange, column 3) proxy for the empties reservoir value. Columns differ in the estimation data sample. Column 1 uses the full data sample, Column 2 the sample with refund data on both curbside collection and redemption centers, and Column 3 the sample with refund and deposit data. Dashed lines in Columns 2 and 3 depict recycling earnings responses estimated with each columns' respective sample.
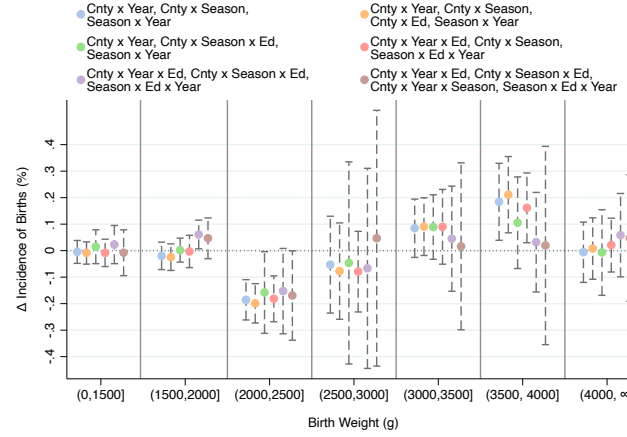
Figure 4: Bottle Bills' Effect on Birth Outcomes in Non-winter relative to Winter Months
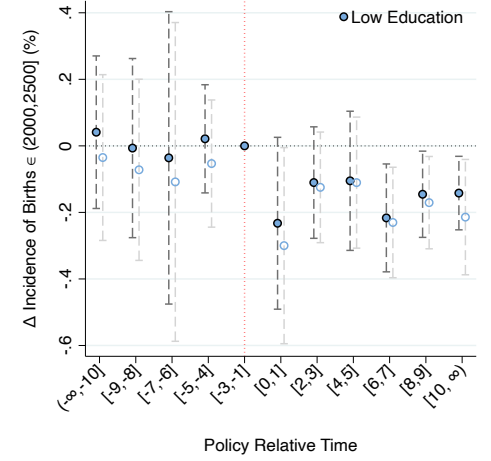
a Effect by Education
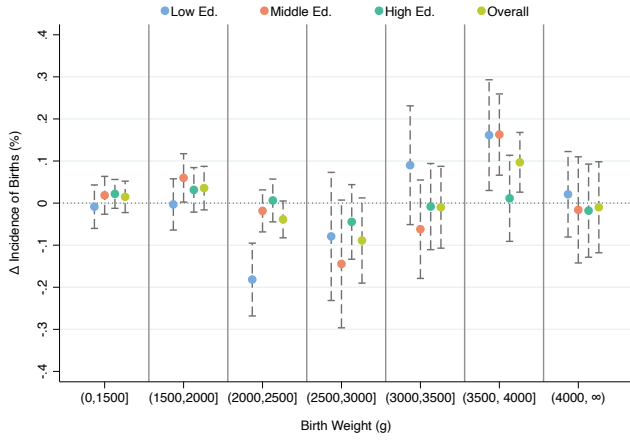
**b** Low Ed. Effect's Sensitivity to Fixed Effects Specification

c Low Ed. Marginal Birth Event Study

Notes: Figure plots bottle bills' estimated effect on birth outcomes in non-winter relative to winter months, plotting estimates from variations of regression equations (5) and (6). **Panel A and B** plot the effect throughout the birth weight distribution. The first column of all panels, denoted by (0, 1500], plots the effect on the incidence of birth incidence less than 1,500 grams. Columns 2-7 plot birth incidence effects in bins sequentially increasing in birth weight, where the x-axis denotes the bin specification. **Panel A** plots estimates for each education group as well as the whole sample. Marker colors denote the (sub)sample associated with the estimated effect: blue markers (low—less than high school—education mothers), red (middle—high school—ed.), and turquoise (high—more than high school). Green markers denote the average effect on all mothers. **Panel B** plots coefficients for *low education* mothers estimated with different fixed effect specifications. Marker colors denotes the fixed effect specification used to estimate each coefficient, increasing in stringency from left to right within each column. Red markers denote the preferred specification (detailed in eq. 6), replicating blue markers in Panel A. Purple markers are equivalent to separately estimating DiDiD estimators for each education group, and brown markers denote a DiDiDiD estimator where the fourth difference is between low education mothers and high education mothers. **Panel C** plots the effect in policy relative time on low education mother marginal births—the incidence of births between 2,000 and 2,500g to mothers without a high school diploma. Year zero denotes the bottle bill implementation year. Specifically, the effect is plotted for every two years after 10 years before implementation and before 10 years after implementation. The effect in the three years before bottle bill implementation is normalized to zero. Solid markers denote estimates from a event study style regression of equation (6) (red markers in Panel B). Hollow markers denote estimates from a fully stacked estimator (purple markers in Panel B). Across all panels, vertical dashed lines show 95% confidence intervals with standard errors clustered at the state level.

Figure 5: Recycling Earnings' Impact on Low Education Mother Birth Outcomes

a Effect of a 1% Increase in Recycling Earnings
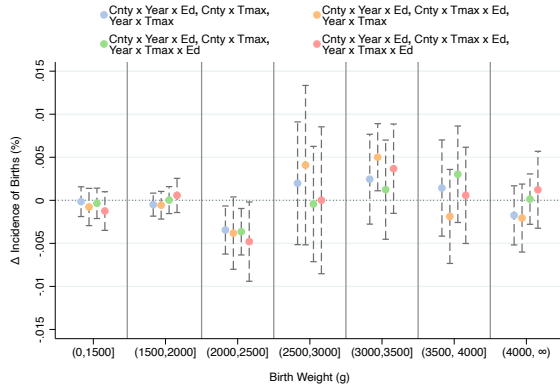
b Seasonally Scaled Effect of Recycling Earnings
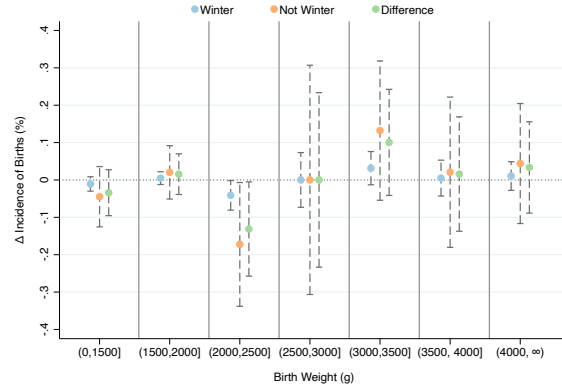
Notes: Figure plots the impact of recycling earnings on birth outcomes among mothers without a high school diploma estimated with the DiDiT quasi-experimental design described in equation 7. Replicating Figure 4a and 4b, Figure plots impacts throughout the birth weight distribution. Panels vary in how DiDiT estimates are scaled. **Panel A** plots estimates scaled to reflect the impact of a 1% increase in recycling earnings on birth outcomes. Marker colors denote different fixed effect specifications. Red markers correspond to the fixed effect specification described in equation 7. **Panel B** plots estimates scaled by average maximum temperature in bottle bill states in winter months (blue markers), non-winter months (orange) as well as the difference between non-winter and winter months (green). The estimates scaled correspond to red markers in Panel A. In both panels, standard errors are clustered by state and dashed vertical gray lines denote 95% confidence intervals.

A Supplemental Appendix

List of Tables

List of Figures

A.1 Theoretical Framework

This section develops a simple model of how recycling markets function and identifies the conditions under which bottle bills increase recycling income. This section formalizes the intuition discussed throughout the main text.

While bottle bills may substantially raise earnings—potentially to levels comparable to aggregate food stamp benefits—their theoretical impact is ambiguous. Bottle bills substantially raise containers' scrap value. This shocks' impact depends on the responses of consumers and recyclers. The model clarifies when higher scrap values, translate into higher informal earnings. The paper's core results either explicitly test or implicitly rely on the assumption that bottle bills increase recycling income. Hence, the model provides a useful framework for interpreting those findings.

Bottle bills are an extended producer responsibility policy (EPR)—a broad policy class requiring producers to be responsible for the societal costs of consuming their product. Bottle bills require producers to pay a refund for and recycle properly returned containers. Numerous other EPR policies have recently gained traction in the US, including bills mandating firms finance the management of packaging waste and requiring clothing brands to fund the collection and recycling of discarded clothes (Davison 2024, Holaday 2024). EPR policy often aims to redirect scrap from the waste stream, which may in turn reduce recycling earnings. If this "redirection" excludes recyclers—for example, by incentivizing all consumers to recycle their own containers—recyclers' incomes could fall to zero. On the other hand, if policy explicitly or implicitly leverage professional recycler labor, recyclers' incomes may increase (Cass Talbott 2021).

To the best of my knowledge, no existing economic frameworks evaluate the optimal design of EPR with respect to earnings from recycling. Fullerton (2024) acknowledges the importance of considering marginalized populations in EPR design but does not provide a formal framework for doing so. Palmer and Walls (1997), Fullerton and Kinnaman (1995), Fullerton and Wolverton (2000), and Eichner and Pethig (2001) outline frameworks for evaluating the optimal design of EPR. However, their frameworks overlook societal benefits of waste—earnings from recycling. Ashenmiller (2009, 2010, 2011) show this income is critical for many households. In addition, Ashenmiller (2009, 2010, 2011) discuss the mechanisms generating this income. Logic I formalize in this section.

A Stylized Model of Recycling Markets Consider an economy in which individuals' outside wage option w_i is uniformly distributed between 0 and some upper bound w_h. Their outside wage option is the highest wage they can earn in a profession other than beverage container collection. All individuals work T hours and consume $\overline{n}$ beverages. Figure A1a plots the income schedule for individuals with an outside wage option of w_i. In the absence of any form of recycling, individuals earn an income of $T\,w_i$ (the solid black line in figure A1a). If we allow individuals to recycle their own empties, they can now choose to earn (i) $T\,w_i$ **or** (ii) $(T-r)\,w_i + \overline{n}\,d$ (dashed black line), where r is the fixed cost of recycling, i.e., the amount of time it takes to visit a redemption center, and d is the scrap value of an empty container. An individual will choose to recycle their own empties if their income when doing so exceeds their income when they do not, i.e.,:

$$T\,w_i < (T-r)\,w_i + \overline{n}\,d \implies w_i < \frac{\overline{n}d}{r}$$

In other words, an individual will choose to recycle their own empties if the opportunity cost of not recycling $\overline{n}\,d$ normalized by the opportunity cost of recycling r exceeds their marginal value of time w_i. Finally, if we allow individuals to recycle their **own** empties **and other people's**, they can decide to earn at a third income option (iii) $(T-r)\overline{w}_r + \overline{n}\,d$ (dotted black line). They forgo their outside option wage w_i for the recycling wage $\overline{w}_r$. Ultimately, individuals will choose the option—(i), (ii) or (iii)—that maximizes their income (the solid blue line).

The recycling wage is defined as follows:

$$\overline{w}_r = \frac{d\,\overline{n}\,N}{P} = \frac{d\,\overline{n}\,\sum_i \mathbb{1}\left\{w_i > \frac{\overline{n}\,d}{r}\right\}}{\sum_i \mathbb{1}\left\{w_i \leq \overline{w}_r\right\}}$$

where N is the number of people who optimally choose not to recycle their own empties (option (i)), and P is the number of people who optimally choose to professionally recycle (option (iii)). The recycling wage $\overline{w}_r$ is an equilibrium outcome—reached when no individual can be made better off by switching to (or from) the third type (professional recycler). Figure A1b depicts how the non-equilibrium recycling wage (w_r) evolves outside of equilibrium and when the recycling wage is in equilibrium. The solid black line plots the non-equilibrium recycling wage w_r defined in terms of w_j the professional recycler outside option wage cutoff—all individuals with an outside option wage below w_j professionally recycle and all those with a wage above do not, i.e.,

$$w_r = \frac{d \, \overline{n} \, N}{\sum_j \mathbb{1} \left\{ w_i \leq w_j \right\}}$$

The dotted black line plots the outside option wage w_j of the marginal recycler. The equilibrium recycling wage $\overline{w}_r$ is reached when $w_j = w_r$—at this point no individual can be made better off by switching to or from professional recycling.

Assuming a uniform distribution, the teal rectangle in the bottom right of Figure A1a depicts the numerator $d \, \overline{n} \, N$ —the value of the empties reservoir. Professional recyclers recoup this value in the form of earnings from work, depicted by the two teal triangles on the left. The upper triangle depicts the rents from waste stream redistribution:

$$\sum_i \mathbb{1} \left\{ w_i \leq \overline{w}_r \right\} \, (\overline{w}_r - w_i)$$

The rents denote the earnings in excess of what an individual would have earned in the absence of redistribution via the waste stream.

Figure A1c shows how relaxing the uniform wage distribution assumption affects the recycling wage $\overline{w}_r$. Holding all else constant, as the number of individuals with an outside wage at $w_i = 0$ increases (plotted on the x-axis) the recycling wage (plotted on the y-axis) decreases. In other words, beyond the scrap value d and the fixed cost of recycling r, the wage distribution plays a meaningful role in determining the recycling wage and the rents from professional recycling.

Theoretical Impact of Bottle Bills on Recycling Earnings Bottle bills altered key aspects of the recycling economy (as modeled in the previous section), effectively increasing the scrap value of beverage containers eightfold—bringing total scrap value to a level comparable to aggregate food stamp benefits (see Figure 1b) (Sanger 1983). Despite this economically meaningful change, the implications for recycling earnings remain theoretically ambiguous.

How Bottle Bills Work Figure 1a depicts the policy's intended operation—consumers pay deposits, and they are refunded upon proper return of containers. Figure 1a, left panel, illustrates how, under this system, retailers pay a deposit to distributors for each containerized beverage they purchase, and consumers in turn pay that deposit at the point of sale. As shown in the right panel, consumers can recover the deposit by returning empty containers to the retailer, who then passes the containers back to the distributor and is reimbursed for the deposit. In some states, distributors also pay retailers a handling fee to compensate them for managing container returns (United States General Accounting Office 1990).

In practice, however, bottle bills reward *whoever* returns empty containers, effectively increasing the scrap value of discarded beverage packaging eightfold. While this increase in scrap value is economically significant, it represents only an upper bound on total recycling earnings. In theory, the policy's effect on individual earnings could have been negative, depending on the underlying structure of recycling markets.

Theoretical Implications for Recycling Earnings The recycling markets model provides structure to understand the extent to which the bottle bill increase in total scrap value translates into a change in recycling earnings. While bottle bills massively increased the total scrap value of containers, their impact on total earnings from recycling, i.e., the value of the empties reservoir $d \, \overline{n} \, N$, is theoretically ambiguous. Bottle bills increase d to d_b; this *direct price effect* increases the aggregate scrap value $d \, \overline{n} \, N$ by $\frac{d_b}{d}$. However, an *indirect behavioral effect* will reduce the value by reducing N, leading to an ambiguous net effect of bottle

bills. The *indirect behavioral effect* reduces the number of people who choose *not* to recycle their *own* empty containers, by increasing the own-recycling threshold $\frac{d\,\overline{n}}{r}$ by $\frac{d_b}{d}$. Additionally, bottle bills strive to reduce the fixed cost of recycling r as well, which would also reduce N, by decreasing the opportunity cost of recycling one's own empties also shifting the own-recycling threshold right.

The extent to which a bottle bill–induced change in the empties reservoir's value translates into a change in the recycling wage is not one-to-one. *If* bottle bills increase aggregate earnings from recycling $\%\Delta_b$, holding all else constant, the new recycling wage $w_r^b = \overline{w}_r \times \%\Delta_b$, the pre-bottle bill equilibrium recycling wage times bottle bill's net effect on aggregate scrap value. More explicitly:

$$w_r^b = \frac{\%\Delta_b \times d\,\overline{n}\,N}{P} = \frac{\%\Delta_b \times d\,\overline{n}\,N}{\sum_i \mathbb{1}\{w_i \leq \overline{w}_r\}}$$

However, this new recycling wage would not be an equilibrium recycling wage, i.e., there exists some individuals with an outside option wage $w_i \in (\overline{w}_r, w_r^b]$ who could be made better off by electing to professionally recycle. Hence, the new equilibrium recycling wage would take the following form:

$$\overline{w}_r^b = \frac{\%\Delta_b \times d\,\overline{n}\,N}{P + P_b} = \frac{\%\Delta_b \times d\,\overline{n}\,N}{\sum_i \mathbb{1}\{w_i \leq \overline{w}_r\} + \sum_i \mathbb{1}\{\overline{w}_r < w_i \leq \overline{w}_r^b\}}$$

In words the extent to which a bottle bill induced change in aggregate scrap value $\%\Delta_b$ changes the recycling wage is mediated by the extensive margin labor supply of recyclers—i.e., changes in the number of recyclers P—attenuating the extent to which a percentage change in aggregate scrap value translates into a corresponding change in recycling wages. This attenuation is unlikely to fully offset the aggregate effect. A complete offset could occur only if labor market frictions are minimal and a sufficiently large number of individuals have outside option wages at $\overline{w}_r$. Given this, if the net impact on aggregate recycling earnings is positive, incumbent and entrant recyclers would experience a wage increase. By assumption, entrant recyclers must receive a wage bump; otherwise, they would not choose to enter professional recycling. The existent of entrants implies incumbent recyclers receive a wage bump as well, otherwise there would be no entrants.

In conclusion, if data on recycling earnings existed, a Difference-in-Differences (DiD) analysis could empirically test the theoretically ambiguous effect of bottle bills on earnings. Specifically, one could estimate the impact on total household earnings—including recycling income—for individuals who recycle for cash. Such estimates would reveal the policy's welfare impact on both incumbent and entrant recyclers. Although data on recycling earnings are unavailable, in Section 2, I outline an approach to estimate the earnings impact of bottle bills and empirically assess the theoretical ambiguity.

Framework Assumptions   This model makes strong assumptions about why individuals choose to recycle both their own empty containers and others'. In reality, individuals whose outside option wage exceeds the recycling wage may still recycle others' empties. They may be constrained in the amount they can work outside of recycling, have access to large quantities of empties through their job, or simply value time spent collecting recyclables differently than time spent in formal employment. Moreover, individuals may choose to recycle their own containers for reasons beyond economic necessity. The decision may reflect a belief that the value of the rebate exceeds their marginal value of time, or be driven by intrinsic utility derived from recycling—independent of the rebate (Ashenmiller 2009). Conversely, if they derive negative utility from recycling, they may forgo recycling even when the recycling wage exceeds their outside option wage. In reality, survey evidence from redemption centers in California suggests that while a diverse group of people recycle for cash, outside option wage is a primary driver of individuals' choice to redeem (Ashenmiller 2009, Ashenmiller 2011).

Finally, ignoring individuals' utility (or disutility) from recycling obscures important features of the waste stream as a redistribution mechanism. When the disutility of rummaging through others' trash exceeds the monetary gains, recycling for cash may exhibit desirable characteristics from a target efficiency perspective. Among those able to work, the stigma and effort associated with scavenging may help preserve earnings for

individuals who truly need the additional income (Nichols and Zeckhauser 1982). This reasoning further supports the argument that a net increase in the empties reservoir will lead to an increase in the recycling wage and not be fully offset by an influx of professional recyclers. Even if enough individuals exist at the margin to attenuate the effect of increased aggregate earnings on the recycling wage, only those who need the extra income most will opt in.

Table A1: Balance Test

	Weighted $\mathbb{1}(\text{BB}) * \mathbb{1}(\text{not winter})$	Unweighted $\mathbb{1}(\text{BB}) * \mathbb{1}(\text{not winter})$
Weather (Cumulative $°C$)		
Temperature (Daily Max)	-0.000000215825	-0.000000672573
	(0.000000890027)	(0.000000937826)
Precipitation	0.024631763125***	0.010587384656
	(0.008737927909)	(0.006630330344)
Income and Government Spending (Millions $2015)		
Personal Income	0.000011169888	0.000007733938
	(0.000011019788)	(0.000013043285)
Personal Current Transfer Receipts	0.000006106729	-0.000018627930
	(0.000007345663)	(0.000033700601)
State Unemployment Insurance Benefits	-0.000069761188	-0.000013062573
	(0.000082941577)	(0.000127615907)
Medicare Benefits	-0.000112371001	0.000009646940
	(0.000085916483)	(0.000101300186)
Social Security Benefits	0.000057990191	0.000088344187
	(0.000067804562)	(0.000062135182)
Farm Wages and Salaries	0.000882474462***	-0.000409620799
	(0.000324064051)	(0.001141428684)
Wages and Salaries	-0.000027881425	-0.000022025025
	(0.000021250907)	(0.000031478762)
Observations	2364748	2364748
within R-squared	.0036	.0012

Notes: Table displays the association between potential confounders and active bottle bills in non-winter months. Each row provides the association between a different confounder and the causal variable of interest. The first column displays estimates weighted by the number of births, while the second column displays estiamtes from an unweighted regression. Point estimates are small, as the test regresses large covariates on the interaction of two dummy variables. Standard errors are clustered by state, the level of treatment. Specifically, Table reports estimates from the following regression equation:

$$\mathbb{1}(\text{BB})_{it} * \mathbb{1}(\text{not winter})_s = \sum_v \beta_v * v + \alpha_{i,s} + \delta_{i,t} + \gamma_{s,t} + \epsilon_{i,s,t}$$

$\mathbb{1}(\text{BB})_{it} * \mathbb{1}(\text{not winter})_s$ is a dummy for states with active bottle bills in non-winter months. v is a potential confounder, i.e., weather or income control—see Section 4.1 for definitions of the variables included in weather and income controls. $\delta_{i,t}$ is a state by year fixed effect. $\gamma_{s,t}$ is a season by year fixed effect. $\alpha_{i,s}$ is a state by season fixed effect.

Table A2: Food Expenditure Descriptive Statistics by Household Demographic

	# of Households	Households with (%)			Mean (SD)	
		Children	Food Stamps	High Labor	Food per Person ($2015)	# of Adults
Panel A: Full Sample						
All Households	4,887	61%	33%	88%	2,904 (1,805)	2.02 (0.85)
By Household Head's Sex and Race						
male, non-white	1,170	68%	43%	98%	2,562 (1,798)	2.27 (1.00)
male, white	2,290	62%	15%	99%	3,141 (1,823)	2.16 (0.66)
female, non-white	860	66%	78%	54%	2,396 (1,547)	1.59 (0.92)
female, white	567	25%	25%	62%	3,317 (1,795)	1.33 (0.63)
Panel B: Food Stamp Sub-Sample						
All Households	1,602	71%	100%	73%	2,368 (1,524)	1.94 (0.97)
By Household Head's Sex and Race						
male, non-white	494	70%	100%	96%	2,341 (1,697)	2.31 (1.05)
male, white	332	76%	100%	98%	2,536 (1,397)	2.16 (0.67)
female, non-white	652	74%	100%	45%	2,240 (1,463)	1.60 (0.94)
female, white	124	50%	100%	41%	2,589 (1,378)	1.39 (0.69)

Notes: Table summarizes demographic and economic characteristics of Panel Study of Income Dynamics (PSID) households observed from 1969 to 1993. The sample contains households that (i) appear in every survey wave during the ten-year window surrounding a bottle-bill introduction, (ii) remain in the same state during that window, and (iii) have a non-retired household head. **Panel A** reports statistics for the full sample; **Panel B** repeats the tabulations for the subset of households that received food-stamp benefits at any point during the sample period. Within each panel, the first row covers all households, and the next four rows split the sample by household-head demographics. Column 1 reports the number of households in each row's sample. Columns 2–4 report the share of households with children (column 2), that ever used food stamps (column 3), and who earn more of their income from labor than transfers (column 4). Columns 5–6 report the mean per household member annual food expenditures and mean number of adults; standard errors appear in parentheses below each mean.

Table A3: Birthweight Descriptive Statistics

	Mean		Incidence of Births (%) by Weight Bin (kg)						
	# of Births/Year	Birthweight (g)	(0,1.5]	(1.5,2]	(2,2.5]	(2.5,3]	(3,3.5]	(3.5,4]	(4,∞)
Panel A: All Mothers									
	2,682,730	3,323.65	1.27	1.41	4.62	16.84	37.23	28.52	10.10
By Season									
winter	849,315	3,330.32	1.25	1.37	4.54	16.67	37.11	28.74	10.34
non-winter	1,833,415	3,320.57	1.28	1.43	4.66	16.92	37.29	28.42	10.00
Panel B: Mothers with Less than a High School Education									
	636,519	3,223.93	1.61	1.87	6.14	20.68	38.11	24.13	7.46
By Season									
winter	201,882	3,234.25	1.55	1.77	5.96	20.42	38.13	24.47	7.70
non-winter	434,636	3,219.14	1.64	1.91	6.22	20.80	38.10	23.97	7.35
Panel C: Mothers with a High School Education									
	1,047,760	3,322.18	1.27	1.39	4.61	16.97	37.35	28.35	10.06
By Season									
winter	330,550	3,330.31	1.24	1.35	4.52	16.75	37.16	28.63	10.36
non-winter	717,210	3,318.43	1.29	1.40	4.65	17.08	37.43	28.23	9.92
Panel D: Mothers with more than a High School Education									
	998,451	3,388.77	1.05	1.15	3.67	14.26	36.55	31.49	11.83
By Season									
winter	316,882	3,391.53	1.05	1.14	3.66	14.19	36.39	31.57	12.00
non-winter	681,569	3,387.49	1.05	1.15	3.67	14.29	36.63	31.46	11.76

Notes: Table summarizes birth outcomes as reported in the U.S. Vital Statistics Natality Data from 1969-2002, consisting of the full census of birth from the National Center for Helath Statitics (NCHS). Panels differ the sample of births summarized. **Panel A** reports statistics for the full sample; **Panels B-D** repeats the tabulations for the subset of mothers with less than a high school education (Panel B), a high school education (Panel C) and more than a high school education (Panel D). Within each panel, the first row covers all births, and the next two rows summarize births in winter and non-winter months. Babies born in January, February, March and April are designated as winter birhts and babies born in all other months are designated as non-winter births. Column 1 (2) reports the average number of births per year (average birthweight) in each row's sample. Columns (3-10) report the incidence of births throughout the birthweight distribution. Birth weight bins are defined at the top of each column in kilograms. Incidence statistics are defined in %, i.e., the average number of births per year with a weight $\leq 1,500g$ among all mothers in all months is $2,682,730 \times .0127 \approx 34,071$ births.

Table A4: The Effect of Bottle Bills on Data Quality

	# Records (%Δ)	Records Dropped (%)				ILBW (%Δ)			
$\mathbb{1}$(BB)	-0.04	0.05	0.05	0.02	0.02	-0.01	-0.01	-0.01	-0.01
	(0.04)	(0.07)	(0.07)	(0.06)	(0.06)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	67657	67657	67067	66397	65777	67078	66500	65835	65215
within R^2	0.0009	0.0003	0.0003	0.0001	0.0001	0.0005	0.0003	0.0006	0.0006
County FE	x	x	x	x	x	x	x	x	x
Year FE	x	x	x	x	x	x	x	x	x
No Missing Birth Records			x	x	x		x	x	x
Max Dropped Record Rate < p99				x	x			x	x
Max %Δ ILBW < p99					x				x

Notes: Table regresses various measures of data integrity on bottle bill implementations denoted as $\mathbb{1}$(BB). Column 1 reports the effect of implementations on %Δ in birth counts between Bailey et al. (2018) and the estimation sample. Columns 2-5 document bottle bills' effect on the percent of observations dropped (record dropping rate) due to missing birth weight data or the mother's age falling below 15. Columns differ in the estimation sample. Column 2 includes the entire sample, column 3 restricts the sample to counties with no missing birth records, column 4 additionally restricts the sample to counties with a maximum record dropping rate below the 99th percentile, and column 5 further restricts the sample to counties with a maximum %Δ ILBW below the 99th percentile. Columns 6-9 document the effect of bottle bills on the %Δ in ILBW between Bailey et al. (2018) and the estimation sample. As with Columns 2-5, columns 6-9 differ in the estimation sample. Column 6 employing the same sample as column 2, column 7 as column 3 and so forth. Standard errors are clustered at the state level, or the level of treatment.

Table A5: The Effect of Bottle Bills on (%) Birth Records Missing Mother's Education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\mathbb{1}$(BB)	-18.92	4.24					
	(23.25)	(12.95)					
$\mathbb{1}$(BB) $\times$ $\mathbb{1}$(not winter)			2.66	2.73	0.36	0.50	0.47
			(1.63)	(1.68)	(0.70)	(0.66)	(0.66)
Observations	796450	796450	796403	796403	667230	667230	651425
within R^2	0.034174	0.212887	0.0026	0.0103	0.0001	0.0042	0.0041
Weather Controls		x		x		x	x
Income Controls		x		x		x	x
County FE	x	x					
Year FE	x	x					
County x Season FE			x	x	x	x	x
County x Year FE			x	x	x	x	x
Season x Year FE			x	x	x	x	x
Sample Excludes TX, CA, WA, NM, ID, AR					x	x	x
No Missing Birth Records							x
Max Dropped Record Rate < p99							x
Max %Δ ILBW < p99							x

Notes: Table regresses the percent of birth records missing mother's education on bottle bill implementations $\mathbb{1}$(BB) and bottle bill implementations in non-winter months $\mathbb{1}$(BB) $\times$ $\mathbb{1}$(not winter). Column 1-2 report the effect of implementations on missingness. Columns differ in the controls included. Column 1 includes no controls beyond those needed to operationalize the DiD estimator, county and year fixed effects. Column 2 additionally includes weather and income controls. Columns 3-7 report the effect of bottle bill implementations in non-winter relative to winter months on missingness. Columns differ in the controls included and the estimation sample. All columns include county x season, county x year, and season x year specific fixed effects to operationalize the DiDiD estimator. Columns 3-4 report estimates from the full sample. Columns 5-6 exclude states with median missingness rates exceeding the mean. Column 7 further restricts the sample to counties without missing birth records (i.e., Bailey et al. (2018) birth counts match the estimation sample), counties with maximum record dropping rates below the 99th percentile, and counties with maximum %Δ ILBW below the 99th percentile. Additionally, columns 4, 6 and 7 include weather and income controls. Standard errors are clustered at the state level, or the level of treatment.

## Table A6: Nielsen Product Modules Used to Assign Deposits

Product Module Code	Product Module Description	Product Group Description
452	REFERENCE CARD FOUNTAIN BEVERAGE	CARBONATED BEVERAGES
1030	FRUIT DRINKS & JUICES-CRANBERRY	JUICE, DRINKS - CANNED, BOTTLED
1031	CIDER	JUICE, DRINKS - CANNED, BOTTLED
1032	FRUIT JUICE - GRAPEFRUIT - OTHER CONTAINERS	JUICE, DRINKS - CANNED, BOTTLED
1033	FRUIT JUICE - APPLE	JUICE, DRINKS - CANNED, BOTTLED
1034	FRUIT JUICE - GRAPE	JUICE, DRINKS - CANNED, BOTTLED
1035	FRUIT JUICE-GRAPEFRUIT-CANNED	JUICE, DRINKS - CANNED, BOTTLED
1036	FRUIT JUICE - LEMON/LIME	JUICE, DRINKS - CANNED, BOTTLED
1037	FRUIT JUICE-ORANGE-CANNED	JUICE, DRINKS - CANNED, BOTTLED
1038	FRUIT JUICE - PINEAPPLE	JUICE, DRINKS - CANNED, BOTTLED
1039	FRUIT JUICE-PRUNE	JUICE, DRINKS - CANNED, BOTTLED
1040	FRUIT JUICE - ORANGE - OTHER CONTAINER	JUICE, DRINKS - CANNED, BOTTLED
1041	FRUIT DRINKS-CANNED	JUICE, DRINKS - CANNED, BOTTLED
1042	FRUIT DRINKS-OTHER CONTAINER	JUICE, DRINKS - CANNED, BOTTLED
1044	FRUIT JUICE-REMAINING	JUICE, DRINKS - CANNED, BOTTLED
1045	FRUIT JUICE-NECTARS	JUICE, DRINKS - CANNED, BOTTLED
1049	REMAINING DRINKS & SHAKES-NON REFRIGERATED	SOFT DRINKS-NON-CARBONATED
1054	VEGETABLE JUICE - TOMATO	JUICE, DRINKS - CANNED, BOTTLED
1055	VEGETABLE JUICE AND DRINK REMAINING	JUICE, DRINKS - CANNED, BOTTLED
1189	COOKING WINE & SHERRY	CONDIMENTS, GRAVIES, AND SAUCES
1484	SOFT DRINKS - CARBONATED	CARBONATED BEVERAGES
1487	WATER-BOTTLED	SOFT DRINKS-NON-CARBONATED
1553	SOFT DRINKS - LOW CALORIE	CARBONATED BEVERAGES
1635	PERISHABLE JUICE AND DRINK	JUICE, DRINKS - CANNED, BOTTLED
1636	PERISHABLE LIQUID COFFEE AND COFFEE DRINK	COFFEE
1637	PERISHABLE LIQUID TEA AND TEA DRINK	TEA
1638	PERISHABLE SOFT DRINK AND WATER	SOFT DRINKS-NON-CARBONATED
5000	BEER	BEER
5001	NEAR BEER/MALT BEVERAGE	BEER
5010	LIGHT BEER (LOW CALORIE/ALCOHOL)	BEER
5020	MALT LIQUOR	BEER
5040	ALCOHOLIC COCKTAILS	LIQUOR
5041	WINE-VERMOUTH	WINE
5049	WINE-APERITIFS	WINE
5050	WINE-DOMESTIC DRY TABLE	WINE
5052	WINE-IMPORTED DRY TABLE	WINE
5053	WINE-FLAVORED/REFRESHMENT	WINE
5054	WINE-KOSHER TABLE	WINE
5055	WINE-SAKE	WINE
5056	WINE-SANGRIA	WINE
5057	WINE-SPARKLING	WINE
5058	WINE-SWEET DESSERT-DOMESTIC	WINE
5059	WINE-SWEET DESSERT-IMPORTED	WINE
5060	WINE - NON ALCOHOLIC	WINE
6074	RBC BEER	BEER
6079	RBC JUICE AND DRINK	JUICE, DRINKS - CANNED, BOTTLED
6086	RBC SOFT DRINK AND WATER	SOFT DRINKS-NON-CARBONATED
6088	RBC WINE	WINE

Notes: Table lists product modules with products for which consumers likely paid deposits. The first column gives the Nielsen product module code, the second the product module description, and the third an example product with the product module. Deposits are assigned based on the UPC (i.e., product) size and the California legislation in the purchase year. During the sample period California changed deposit size in 2007 from 4 cents to 5 cents for containers less than 24 ounces and 8 cents to 10 cents for containers more than 24 ounces.[39] Container purchases prior to 2007 are assigned a deposit of 4 (8) cents for containers smaller (larger) than 24 ounces. After 2007, containers smaller (larger) than 24 ounces are assigned 5 (10) cent deposits.
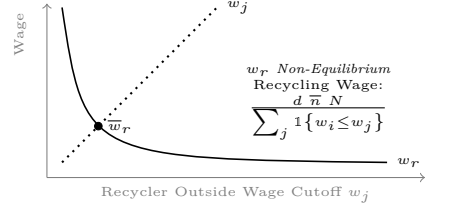
---

[39]https://www.bottlebill.org/index.php/current-and-proposed-laws/usa/california

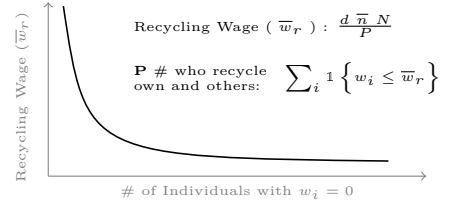**Figure A1:** A Stylized Depiction of Waste Stream Redistribution

a The Income Schedule with Recycling
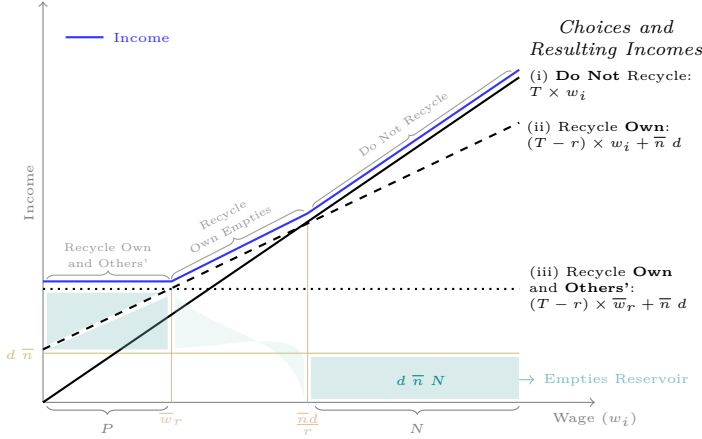
b Equilibrium Recycling Wage

c Recycling Wage + Wage Distribution

Notes: Figure depicts how recycling for cash can redistribute money from high to low wage earners. **Panel A** illustrates the relationship between income and individuals' outside option wage w_i depending on whether they: (i) do **not** recycle (solid black line), (ii) recycle their **own** beverage containers (dashed black line), and (iii) recycle both their **own** and **others'** beverage containers (dotted black line). The solid blue line plots the income individuals optimally earn by choosing the option—(i), (ii), or (iii)—that maximizes their income. Assuming a uniform wage distribution, P denotes the number of individuals who professionally recycle and N denotes the number of individuals who do not recycle, generating the empties reservoir, the value of which is depicted in the teal rectangle in the lower right corner of the graph. Professional recyclers recover this value for a wage of $\overline{w}_r = \frac{d\,\overline{n}\,N}{P}$, generating the rents $\sum_i \mathbb{1}\{w_i \leq \overline{w}_r\}(\overline{w}_r - w_i)$ depicted in the teal triangle between the dotted line (professional recyclers' income) and the dashed line (own recyclers' income), professional recyclers' next best income level. **Panel B** illustrates the equilibrium recycling wage $\overline{w}_r$ in terms of w_r—the non-equilibrium recycling wage—and w_j—the professional recycler outside option wage cutoff. All individuals with an outside option wage below w_j professionally recycle and all those with a wage above do not. When $w_j < w_r$ individuals with an outside option wage $\in (w_j, w_r]$ can be made better off by choosing to professionally recycle, hence $w_r \neq \overline{w}_r$. In equilibrium, w_j is equivalent to w_r, i.e., w_j defines the outside option wage at which an individual is indifferent between the outside option and professionally recycling. **Panel C** illustrates how assumptions around the wage distribution impact the recycling wage and subsequently the rents derived from professionally recycling. Holding all else constant, as the number of individuals with an outside option wage at $w_i = 0$ increases the recycling wage decreases driving rents towards zero.

**Figure A2:** Refund and Deposit Sensitivity to Precipitation
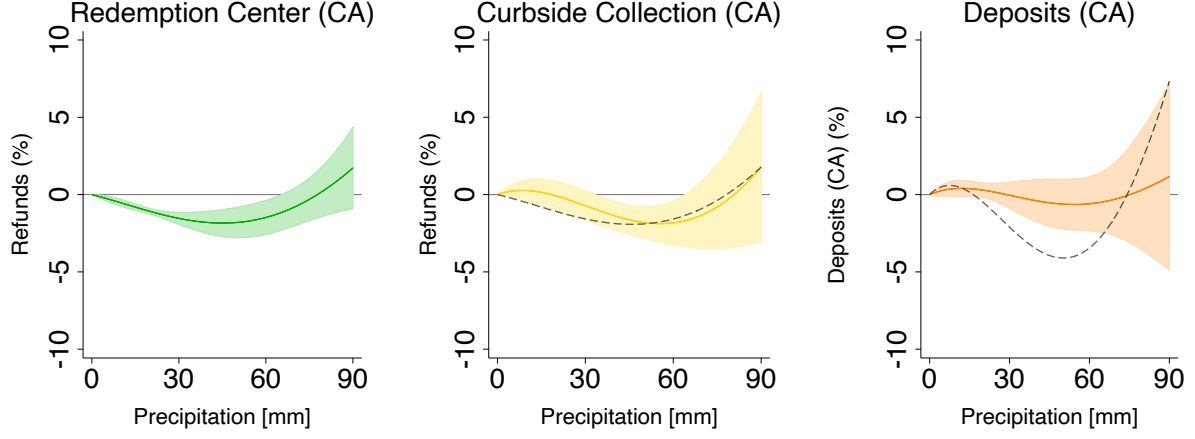
Notes: Figure plots the precipitation responses from estimating equation (2), illustrating additional refunds/deposits (%) in response to cumulative daily precipitation, relative to no precipitation. Solid lines represent point estimates and shading depicts 95% confidence intervals. Panels differ in the outcome variable. The left-most panel plots the response of refunds paid out at redemption centers, the middle of refunds paid to curbside collection companies, and the right-most of deposits paid at grocery stores. Panels differ in the estimation data sample. The left panel uses the full data sample, middle the sample with refund data on both curbside collection and redemption centers, and right the sample with refund and deposit data. Dashed lines in the middle and right panels depict redemption center refund responses estimated using the each columns' respective sample, i.e., the dashed line in the middle panel plots the redemption center refund response estimated using counties with redemption center and curbside refund data. The left and middle panel cluster errors at the county level. The right panel includes heteroskedasticity robust standard errors. This specification does not cluster as the regression is at the market level and there are only two markets.

Figure A3: Recycling Earnings Temperature Response Robustness

a Functional Form
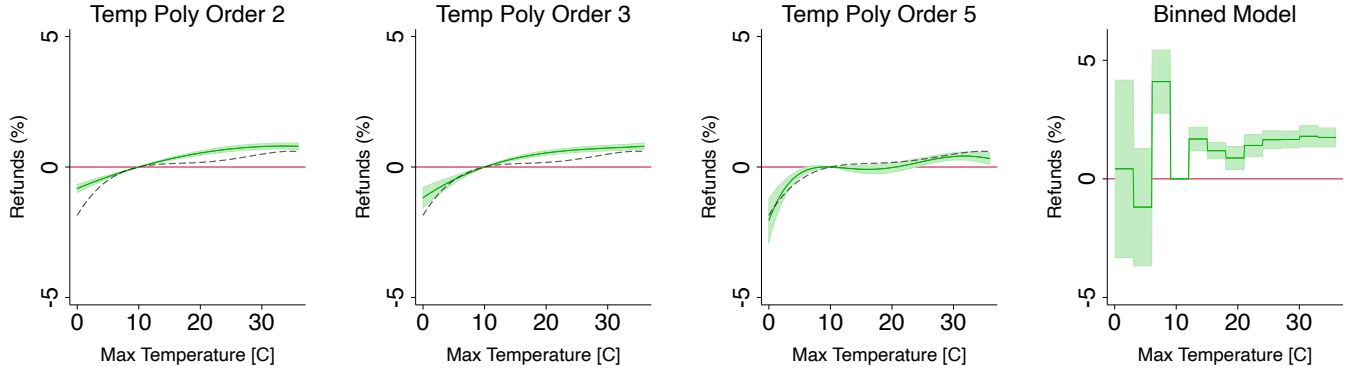
b Fixed Effect Specification
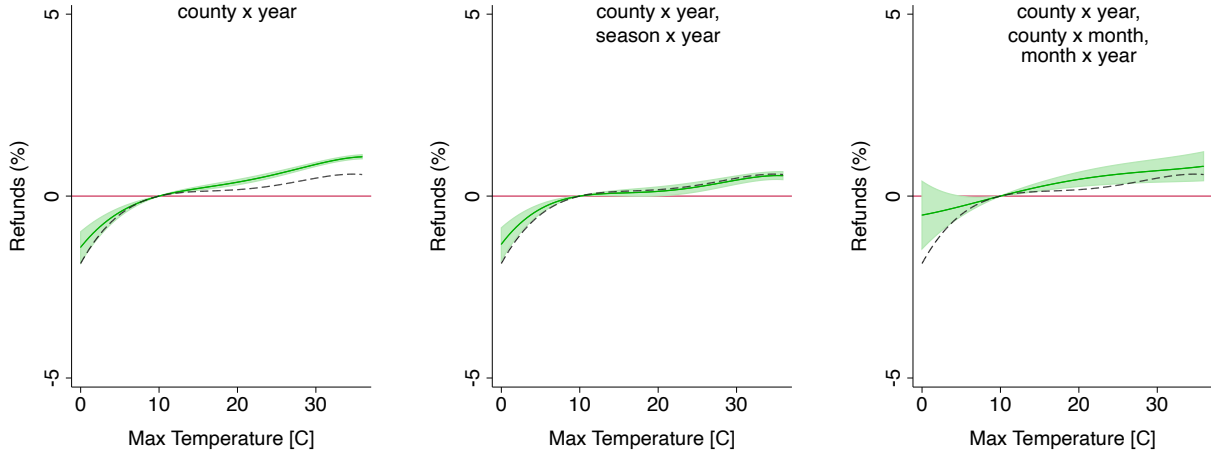
Notes: Figure plots model specification robustness in the recycling earnings temperature response. In all plots, solid lines represent point estimates of the recycling earnings response to a day's maximum temperature, measured relative to $10°C$ (or to the 9–$12°C$ range in the binned model). Shaded regions denote 95% confidence intervals. Where applicable, dashed lines replicate the main specification response from Figure 3b, Column 1, estimated using regression equation (2). **Panel A** plots the response estimated with different functional forms in temperature. Columns 1-3 replicate Figure 3b, Column 1 re-estimating regression equation (2) using second-, third-, and fifth-order polynomials in temperature, respectively, instead of the fourth-order polynomial used in the original specification. The final column presents estimates from a binned temperature model, offering a functional form that is closer to non-parametric. The response is estimated using equation (2), where temperature is represented by a vector of indicators for maximum daily temperature in 3°C bins. The first bin includes all days below 0°C, while the last bin includes all days above 33°C. The bin covering 9–12°C is omitted as the reference category. **Panel B** displays the sensitivity of the main specification to alternative fixed effects. Column 1 includes only county-by-year fixed effects. Column 2 adds season-by-year fixed effects. Column 3 includes county-by-year, county-by-month, and month-by-year fixed effects.

Figure A4: Heterogeneity in Bottle Bills' Effect on Low Ed. Marginal Births

**a** Effect Sensitivity to Season Definition
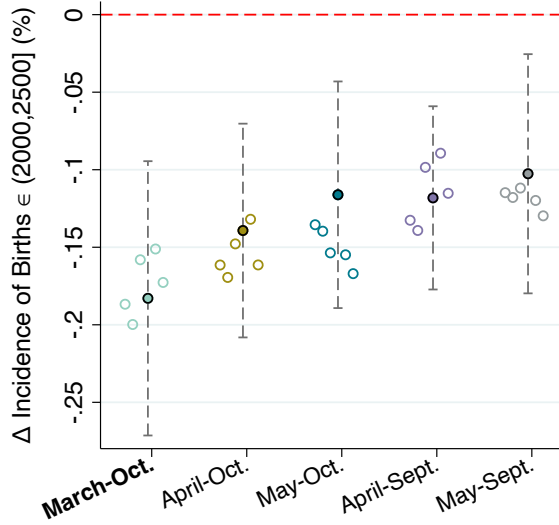
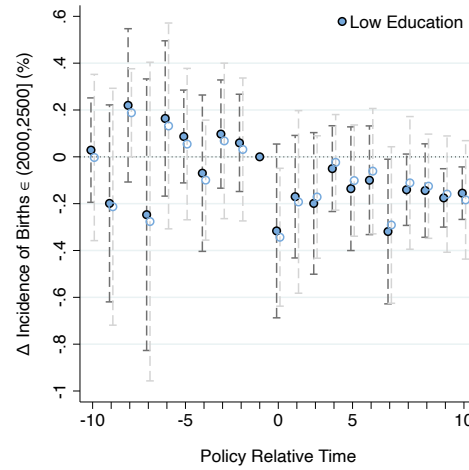**b** Heterogeneity by Place Based Characteristics
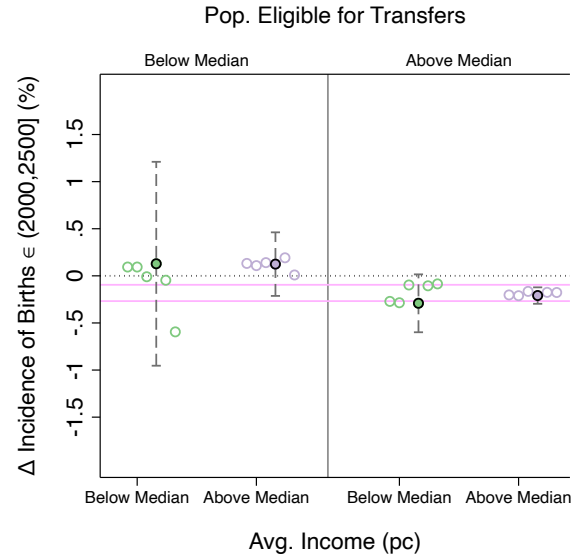
Notes: Figure plots the effect on low education mother marginal births—the incidence of births between 2,000 and 2,500g to mothers without a high school diploma. **Panel A** plots the sensitivity of the effect to the definition of non-winter months. The x-axis labels the definition of "non-winter" used in estimation. The left-most definition—March-October—corresponds to the main specification. Its solid light blue marker aligns with the blue and red markers in the third from left columns of Figure 4a and 4b, respectively. Hollow markers represent alternative fixed effect specifications, corresponding in order to those shown in Figure 4b. **Panel B** plots heterogeneity in the effect by location characteristics. Solid markers denote the preferred fixed effect specification equivalent to red markers in Figure 4b. As in Panel A, Hollow markers denote all other fixed effect specifications plotted in Figure 4b in a corresponding order. Across all panels, vertical dashed lines show 95% confidence intervals with standard errors clustered at the state level.

Figure A5: Bottle Bills' Effect on Low Ed. Marginal Births–Event-Study Robustness
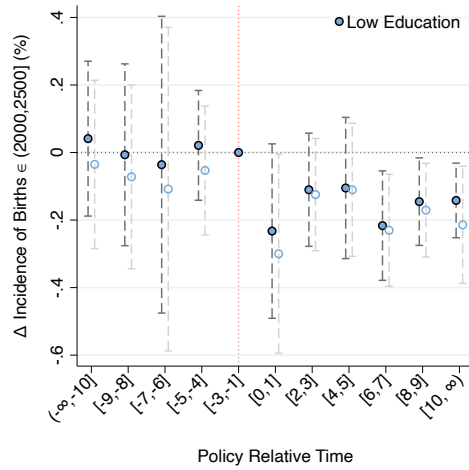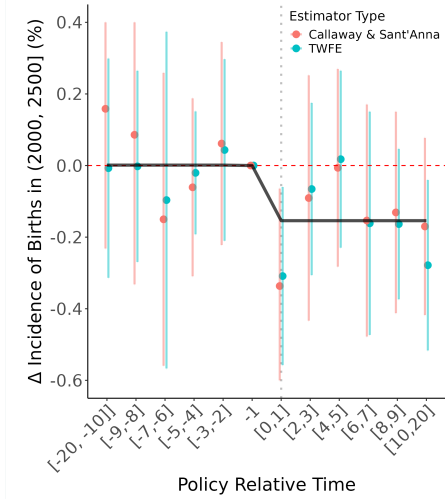
**a** Main Specification **b** Single Year Event Study **c** Estimator Robustness

Notes: Figure plots sensitivity analyses for the event-study estimates of marginal births to mothers without a high school diploma, as originally shown in Figure 4c. All panels plot the effect in policy relative time on low education mother marginal births—the incidence of births between 2,000 and 2,500g. Year zero denotes the bottle bill implementation year. **Panel A** replicates Figure 4c as a point of comparison. **Panel B** replicates Panel A without binning estimates, i.e., the effect is plotted for every year after 10 years before implementation and before 10 years after implementation. In both panels all years more than 10 years before or after policy implementation are grouped. Solid markers denote estimates from an event study style regression of equation (6) (red markers in Figure 4b). Hollow markers denote estimates from a fully stacked estimator (purple markers in Figure 4b). **Panel C** plots event-study estimates from two-way fixed effect (TWFE) and Callaway and Sant'Anna (2021) estimators. Red plot event-study estimates from Callaway and Sant'Anna (2021) and the TWFE specification, respectively. As discussed in Section 4.2, the estimation sample was modified to accommodate the Callaway and Sant'Anna (2021) estimator. To ensure that differences in estimates reflect differences in the estimators rather than in the sample, TWFE estimates using the same modified sample are plotted with blue markers. As in Panel A, event time is binned into two-year intervals spanning from nine years before to nine years after policy implementation, with any event time beyond this range grouped into a single bin. Vertical dashed lines show confidence intervals with standard errors clustered at the state level. Panels A and B plot 95% confidence intervals while Pane C plots 90% confidence intervals.

Figure A6: Bottle Bills' Effect on Birth Outcomes in Non-winter relative to Winter Months—Robustness
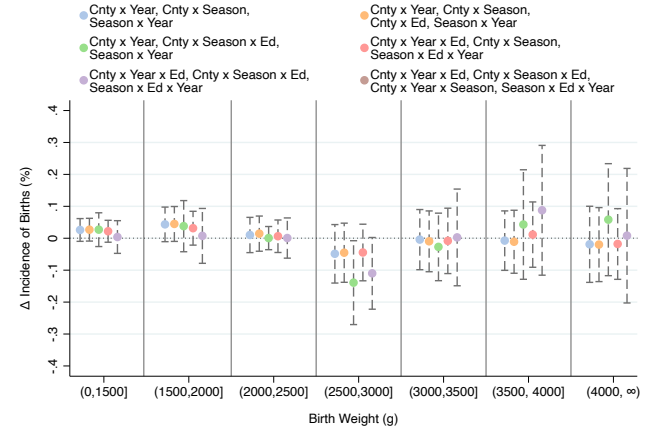
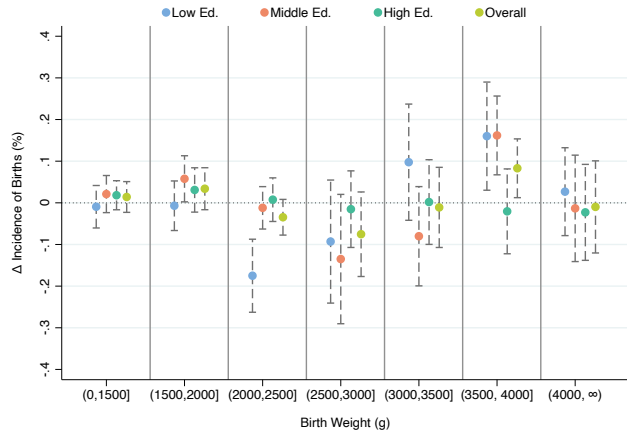**a** Low Education, FE Spec. Sensitivity
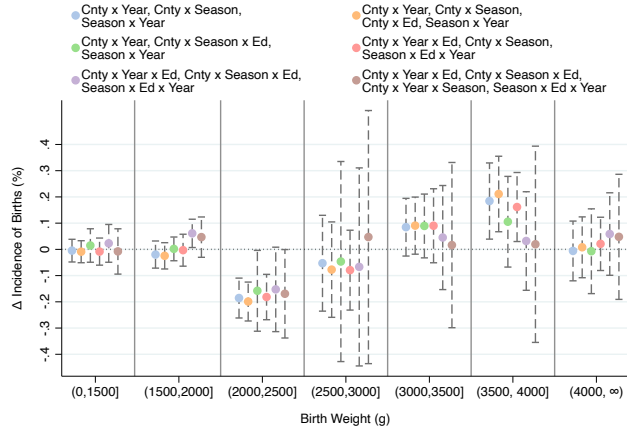
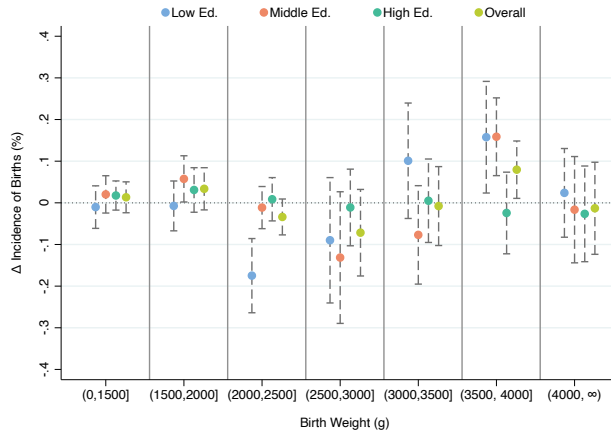**b** Middle Education, FE Spec. Sensitivity
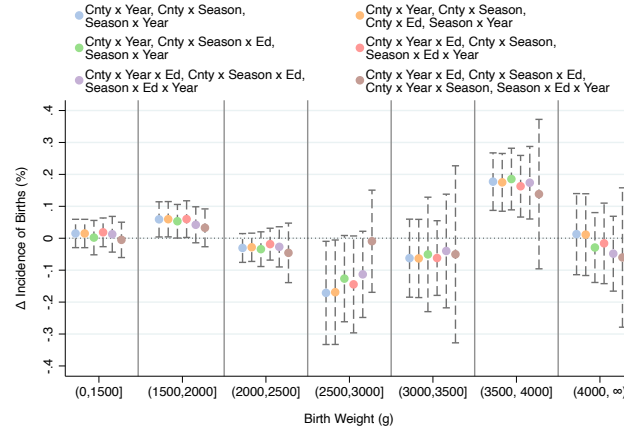
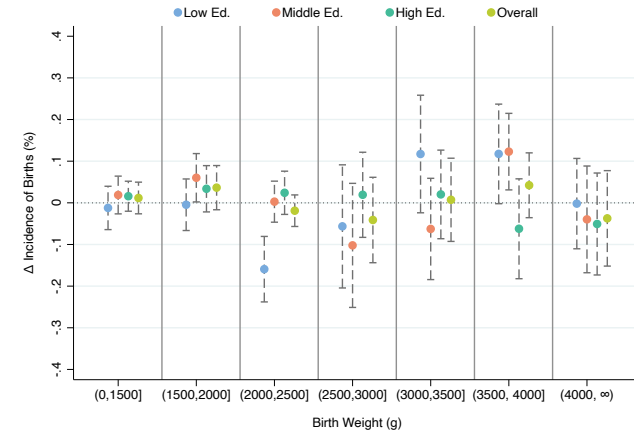**c** High Education, FE Spec. Sensitivity
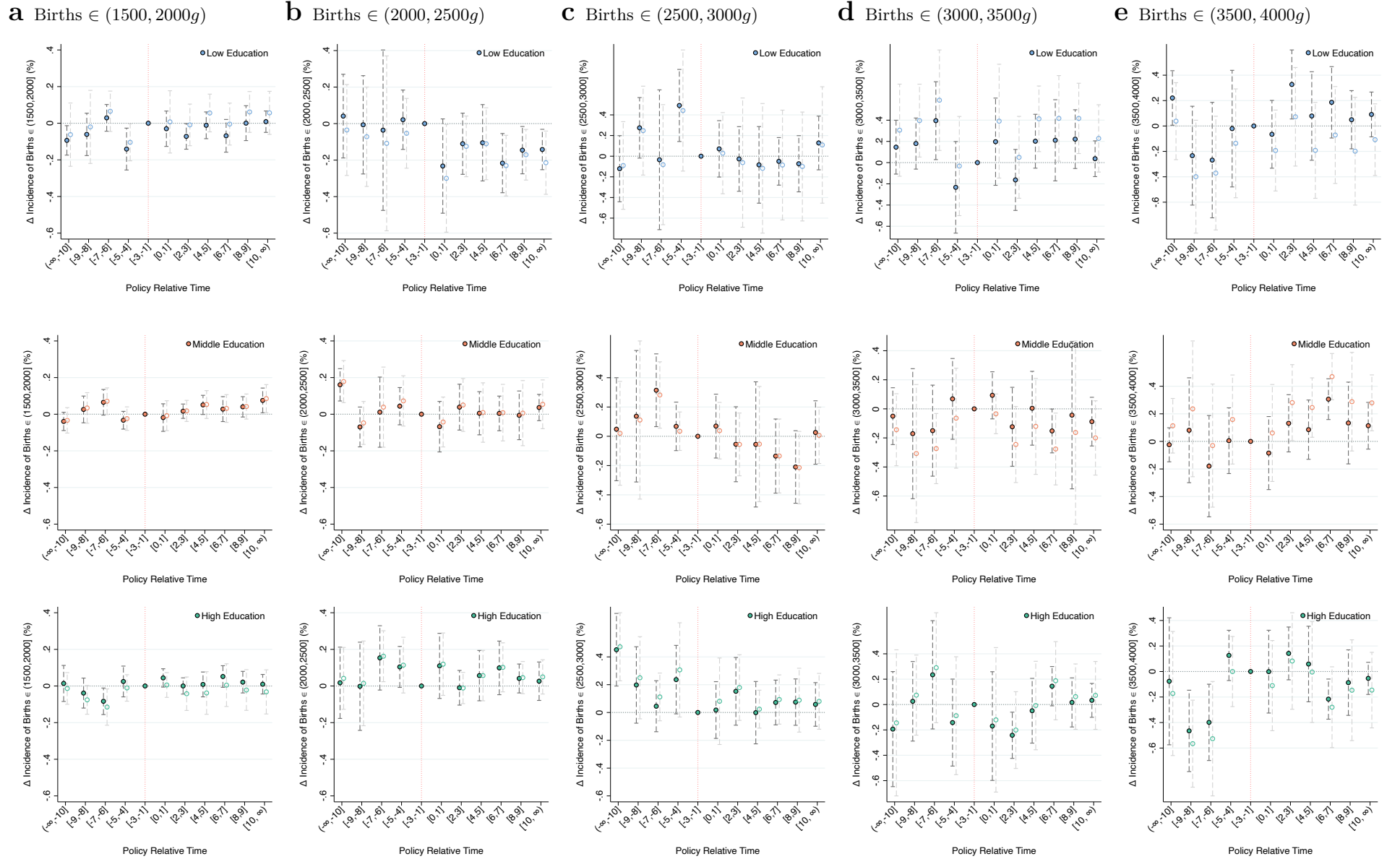
d No Controls

**e** Weather Controls

f Weather + Income Controls

Notes: Figure plots model specification robustness in the birth outcome response to bottle bill introductions in non-winter relative to winter months. **Panels A-C** plot estimates for low, middle and high education mothers, respectively, replicating Figure 4b for each education group. Marker color denotes the fixed effect specification used to estimate each coefficient, increasing in stringency from left to right within each column. **Panels D–F** replicate Figure 4a under three specifications: without controls (i.e., a perfect replication), with weather controls, and with both weather and income controls. See Section 4.1 for a description of the included controls.

Figure A7: Bottle Bills' Effect Throughout Birth Weight Distribution By Education Group–Event Studies

a Births $\in (1500, 2000g)$ **b** Births $\in (2000, 2500g)$ **c** Births $\in (2500, 3000g)$ **d** Births $\in (3000, 3500g)$ **e** Births $\in (3500, 4000g)$

xviii

Notes: Figure replicates Figure 4c for (i) birth weight bins between 1,500-4,000g and (ii) maternal education group. Columns differ by birth weight bin and rows differ by education group. All figures plot the effect of bottle bill implementations in non-winter relative to winter months on birth incidence in policy relative time. Year zero denotes the bottle bill implementation year. Specifically, the effect is plotted for every two years after 10 years before implementation and before 10 years after implementation. The effect in the three years before bottle bill implementation is normalized to zero. Solid markers denote estimates from a event study style regression of equation (6) (red markers in Figure 4b). Hollow markers denote estimates from a fully stacked estimator (purple markers in Figure 4b). Across all panels, vertical dashed lines show 95% confidence intervals with standard errors clustered at the state level.

Figure A8: Birth Records Missing Mother's Education

a

b
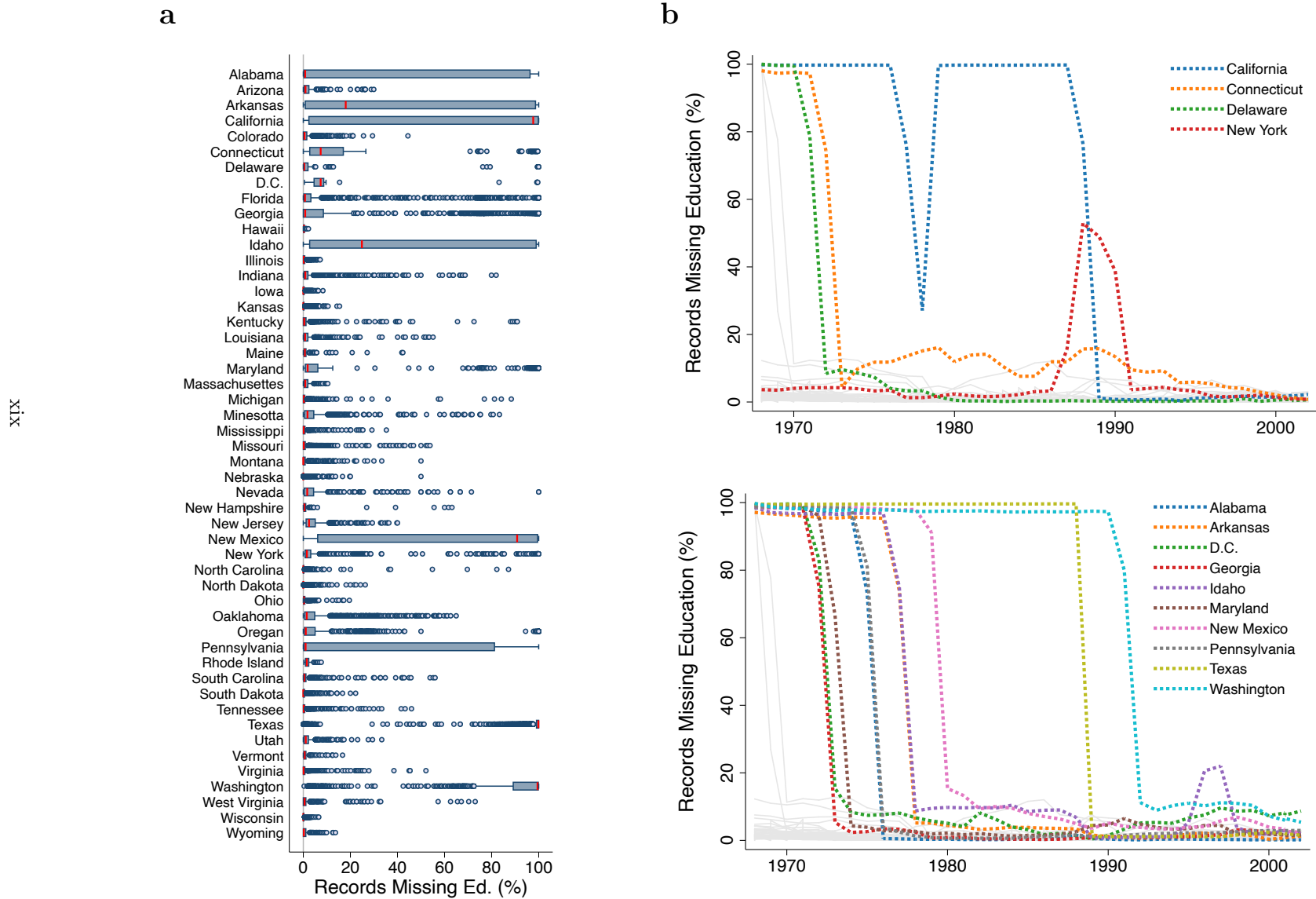
Notes: Figure documents variation in the percent of observations missing mother's education across states, counties and time. **Panel a** plots the distribution of missingness by state. One point in each box plot denotes the rate for a given county and year. **Panel b** plots the percent of observations missing mother's education for states with missingness levels exceeding 15%.