When Food Stamps Became Free: Impacts and Implications of Removing the Purchase Requirement*

Amy Finkelstein Matthew Notowidigdo Charlie Rafkin

July 28, 2026

Abstract

We provide positive estimates of the impact of food stamps on infant health and normative analysis of this large-scale transfer program, using geographic variation in exposure to the 1979 elimination of the food stamp "purchase requirement." After this reform, eligible households received their food stamps for free, rather than paying 35 cents on the dollar. Making them free raised enrollment by 1.85 percentage points (standard error = 0.15), or about 25%, with new enrollees substantially poorer than existing participants. Using the free stamps reform as an instrument for enrollment, we find no evidence that food stamp enrollment improved infant health. For instance, 95% confidence intervals rule out that gains in birth weight exceed 2%; estimates that add the initial program rollout as an additional instrument also show similarly precise nulls. Under a revealed-preference assumption, the enrollment response to the reduced price of food stamps implies an average private value of 59 cents per dollar of free food stamps, with nonparametric bounds of 42–77 cents. While moving to free food stamps improved targeting on social value, it greatly reduced targeting on private value.

*Finkelstein: MIT and NBER, afink@mit.edu; Notowidigdo: Chicago Booth and NBER, noto@chicagobooth.edu; Rafkin: University of British Columbia, charlie.rafkin@ubc.ca. We thank seminar audiences, Eric Budish, Manasi Deshpande and Jesse Shapiro for helpful feedback. Maura Hogaboom, Ro Huang, Thomas Mazzuchi and Tracy Zhou provided excellent research assistance.

The food stamp program, known today as the Supplemental Nutrition Assistance Program (SNAP), is the second-largest means-tested transfer program in the United States (Hoynes and Schanzenbach, 2015; Carrington et al., 2013). This $100-billion-a-year program provides food assistance to 40 million Americans, about 35% of whom are children (USDA, 2025).

An important and perhaps underappreciated aspect of the program's history is that, for its first two decades, it provided only a partial subsidy on food bought through it. Under the original "purchase requirement," households had to pay the program in cash to get food stamps in return; before the purchase requirement was eliminated, the price of food stamps was about 35 cents per dollar. The current program, by contrast, is free: participants pay nothing, and food stamps subsidize 100% of program food up to the maximum benefit. Congress deemed the removal of the purchase requirement — which raised the subsidy on program food from about 65% to 100% — as consequential enough that, rather than amend the existing program, it repealed the original Food Stamp Act of 1964 and created the free Food Stamp Program in 1979 as part of the 1977 Food and Agriculture Act (Congressional Research Service, 1977; Richardson, 1979).

This major redesign of the food stamp program provides an opportunity to enhance our understanding of both its positive and normative aspects. On the positive side, we use the increase in the subsidy as an instrument for food stamp enrollment to estimate the causal impact of food stamps on infant health. Much of what we know about the program's impacts comes from a series of important empirical studies on the initial county-by-county rollout of the program in the 1960s and early 1970s, when the purchase requirement was still in place. This research — which has been very influential in both academic and policy circles (e.g., National Academies of Sciences, Engineering, and Medicine, 2019) — has found that food stamp enrollment during pregnancy increased infant birth weight (Almond et al., 2011), and that food stamp enrollment in childhood improved subsequent adult health, human capital, and economic outcomes (Hoynes et al., 2016; Bitler and Figinski, 2024; Bailey et al., 2024) and reduced adult criminal behavior (Bailey et al., 2024; Barr and Smith, 2023). Evidence on the causal impacts of food stamps from later periods is much more limited, reflecting the difficulty of finding quasi-experimental variation in this uniform, national program.[1] The free-stamps reform thus provides a valuable source of policy variation for learning about the impact of food stamps in a period when the program more closely resembled its current form.

On the normative side, we use the enrollment response to the elimination of the purchase requirement to estimate the private value of free food stamps to enrollees and to analyze the trade-offs involved in cost-sharing in means-tested transfer programs. Estimating the private value of goods when prices are not observed is a well-known challenge (Samuelson, 1954), and

[1]Two important exceptions are papers that use contemporary variation in local food prices which affect SNAP purchasing power (Bronchetti et al., 2019) and state-level changes in eligibility for documented immigrants (East, 2020), and which find improvements in child health.

one that pervades welfare analysis of in-kind transfer programs (e.g., Finkelstein et al., 2019a). A valuable feature of the free food stamps reform is that it provides variation in prices, enabling us to directly trace out a "demand curve" for food stamps.[2] This same feature also allows us to analyze empirically the welfare tradeoffs involved in setting the subsidy level. Conceptually, higher cost-sharing (i.e., a lower subsidy) can improve targeting on *private* value by screening out households with lower willingness to pay for the transfer, but can also reduce targeting on *social* value by screening out households with high social valuations of the transfer, such as cash-constrained households whom policymakers may want to reach for redistributive purposes.[3] Whether higher or lower subsidies are socially desirable is thus an empirical question. It is also a practically relevant one: there is considerable debate over the appropriate level of subsidy in a variety of current U.S. means-tested in-kind transfer programs, including public subsidies for health insurance (e.g., Finkelstein et al., 2019b; Tebaldi et al., 2023), higher education (e.g., Colas et al., 2021), and child-care (e.g., Montpetit et al., 2025).[4]

We study these positive and normative questions by exploiting geographic variation in exposure to the free food stamp reform. Prior to the reform, counties varied substantially in the share of the population that was eligible for — but not enrolled in — food stamps, creating differences in scope for increased enrollment in response to the increase in subsidy. Using this variation, we estimate that the move to free food stamps raised enrollment by 1.85 percentage points (standard error = 0.15), an increase of about 25 percent relative to the pre-reform enrollment rate. The new enrollees were substantially poorer — with about half the average household income of those who enrolled under the purchase requirement. We then use exposure to the free food stamp reform as an instrument for program enrollment to study the impact of food stamp enrollment on infant health, an important and well-measured contemporaneous outcome. For comparison — and to combine multiple natural experiments for greater precision — we also revisit the county-by-county rollout of the initial food stamps program, using this rollout as another instrument for enrollment.

We find no evidence of an impact of food stamp enrollment on infant health, whether we instrument for enrollment using the 1979 free food stamp reform, the initial rollout, or both. IV

[2] By contrast, inferring the private value of a transfer program from how enrollment responds to changes in benefit size (Anderson and Meyer, 1997; Pukelis, 2024; Naik, 2025; McQuillan and Moore, 2025) or to changes in ordeal costs (e.g., Deshpande and Li, 2019; Finkelstein and Notowidigdo, 2019) requires an additional step of needing to translate the ordeal or benefits into dollar values, precisely the object of interest.

[3] Evidence on cost-sharing for health products in developing countries illustrates that this tension between private and social value is relevant in practice: higher willingness to pay predicts higher use of the transferred product (Ashraf et al., 2010), but cost sharing can also screen out lower willingness to pay households for whom take up of the product is socially valuable due to externalities (Cohen and Dupas, 2010).

[4] There is also considerable variation across programs in whether they fully subsidize the in-kind transfer. Like the modern food stamp program, the Low Income Home Energy Assistance Program (LIHEAP) and the Women, Infants, and Children (WIC) nutrition program fully subsidize consumption of the in-kind good. Like the original food stamp program, the National School Lunch Program imposes copays on some households, and some state Medicaid programs have begun to charge premiums for some enrollees (e.g., Dague, 2014; Brooks et al., 2019, 2025); beginning in 2027, California will shift low-income undocumented adults on Medicaid from a zero-premium design to $100 monthly premiums (Associated Press, 2025).

estimates from the free food stamp reform indicate that food stamp enrollment reduces birth weight by a statistically insignificant 0.4%; our 95% confidence interval rules out an increase in birth weight of more than 2%. We find no statistically significant impact of food stamps on birth weight from the initial rollout either, and precision-weighted IV estimates that combine both reforms rule out that food stamps increase birth weight by more than 1.9%. We also do not find robust evidence of food stamps improving other measures of infant health such as rates of pre-term birth or infant mortality. Consistent with Almond et al. (2011), we do find some evidence that the initial rollout of food stamps led to statistically significant improvements in Black infant birth weight, but no such evidence from the free food stamp reform or from precision-weighted estimates that use both reforms.

Taken together, our findings suggest limited impacts of food stamps on infant health in its first few decades. This is particularly noteworthy in light of the substantial attention that the prior evidence on the impact of the food stamp rollout on infant health has received in both academic literature reviews (e.g., Almond et al., 2018; Hoynes and Schanzenbach, 2018; Aizer et al., 2022) and in policy discussions of safety net investments on children (e.g., National Academies of Sciences, Engineering, and Medicine, 2019; Carlson and Llobrera, 2022). A natural and important direction for further work will be to use the free food stamp reform to investigate impacts of increased food stamp enrollment in childhood on subsequent adult outcomes.

At the same time, these analyses do not speak directly to the normative value of the food stamp program or to the welfare consequences of changing its subsidy level. Means-tested transfer programs like food stamps serve a redistributive purpose. Two key forces affecting their redistributive value are recipients' private value of the transfer and recipients' income level or "need." Neither force depends on the causal effect of the transfer on infant health or indeed on other outcomes. We therefore use our empirical estimates of the free food stamp reform's impacts on *enrollment* to estimate recipients' private value of free food stamps, and to quantify the welfare tradeoffs involved in raising the subsidy to 100%.

The large enrollment response to the increase in the food stamp subsidy indicates that many enrollees act as if the value they receive from food stamps is less than dollar-for-dollar; if food stamps were fungible with cash and involved no participation costs ("ordeals"), enrollment would not have increased. We estimate that the enrollment response to the free-stamps reform, as well as to an earlier reform that raised the subsidy from 40 to 65%, suggests an average private value of free food stamps to enrollees of 59 cents per dollar, with nonparametric bounds of 42 to 77 cents. Our approach complements other estimates of food stamps' valuation coming from labor-supply responses to the food stamp rollout (Hendren and Sprung-Keyser, 2020) and survey evidence on willingness to pay for food stamps in the 1990s (Whitmore, 2002; Bailey et al., 2024), which have found private values of $0.62 and $0.80, respectively.

A per dollar private valuation of food stamps of 59 cents suggests that about 41 cents per dollar of government expenditure on food stamps is dissipated as deadweight loss. By contrast, the Congressional Budget Office's Supplemental Poverty Measure imputes food stamps' transfer value as a frictionless cash transfer valued at $1-per-dollar (Congressional Budget Office, 2025), as does influential research on distributional national accounts (Piketty et al., 2018; Blanchet et al., 2022; Auten and Splinter, 2024). While one should naturally exercise caution in extrapolating our specific estimates to today's program, the specter of dissipating potentially two-fifths of the private value of this major transfer program suggests that the U.S. transfer system may involve much more deadweight loss than assumed under current, standard accounting practices.[5]

The enrollment response to the free food stamp reform also sheds light on the welfare tradeoffs of cost-sharing in the program. Consistent with cost-sharing improving targeting on private value, we find that the compliers (those who enroll only when food stamps become free) have a private value of free food stamps that is only about one-quarter that of always-takers (those who would have also enrolled under the cost-sharing regime). However, our finding that compliers are substantially poorer than always-takers highlights that cost-sharing also reduced targeting on social value, as the free food stamp reform transferred resources to more economically disadvantaged households. Under standard redistributive preferences, we estimate that the magnitude of reduced targeting on private value appears to outweigh the magnitude of improved targeting on social value: cost-sharing produced more social welfare per dollar spent than free stamps.

Our welfare analysis of cost-sharing relates to the lively "self-targeting" literature examining whether non-monetary ordeal costs in means-tested transfer programs screen out recipients with low private value (Nichols and Zeckhauser, 1982) or, instead, exclude those with highest social value (often referred to as "need") (Mullainathan and Shafir, 2013; Herd and Moynihan, 2025).[6] For example, this tradeoff appears centrally in Shepard and Wagner (2025)'s study of ordeals in health insurance — where enrollees' social and private valuations may diverge due to adverse selection — and in Lieber and Lockwood (2019)'s study of targeting of in-kind transfers relative to cash. The literature on ordeals often takes as given that governments cannot allocate in-kind transfers through prices (e.g., Yang et al., 2025). It therefore studies the welfare effects of allocating resources with zero monetary price by dissipating private values via non-pecuniary ordeals. Our analysis instead allows governments to allocate resources through prices, but features the same welfare tradeoff between targeting on private and social value.

Importantly, our normative analyses require additional assumptions beyond what is needed to

[5]Conceptually, this deadweight cost may reflect either "ordeal costs" of enrollment (including stigma and hassle costs) and/or in-kind transfers distorting consumption relative to what the enrollee would consume with the equivalent cash. Estimates from both the 1970s and the current time period suggest that most food stamp enrollees are "inframarginal", i.e., they would have bought more food than their food stamps benefit even if those benefits had been paid in cash (e.g., Hoynes et al., 2015; Hastings and Shapiro, 2018; MacDonald, 1977; Moffitt, 1989). As a result, much of the 41 cents of deadweight loss per dollar of government expenditure on food stamps likely reflects ordeal costs.

[6]A partial list of recent empirical analyses on the self-targeting properties of ordeals includes Alatas et al. (2016), Deshpande and Li (2019), Finkelstein and Notowidigdo (2019), Wu and Meyer (2023) and Rafkin et al. (2025), among many others.

use the free food stamp reform for positive analysis of the impact of food stamps on infant health. In particular, they require that the 1979 reform did not change any aspects of the food stamp program other than the food subsidy. As we discuss below, this is not literally the case; however, it appears to be a reasonable approximation for our purposes. They also invoke an assumption of revealed preference, in the tradition of other work using revealed preference to estimate the private value of social insurance programs for health (Finkelstein et al., 2019b), unemployment (Landais and Spinnewijn, 2021), disability insurance (Haller and Staubli, 2023), as well as cash transfers (Naik, 2025). Relying on revealed preference provides a principled starting point for normative analysis, but involves strong assumptions about optimizing behavior which we investigate in sensitivity tests.

The rest of the paper proceeds as follows. Section 1 provides an overview of the food stamp program and the 1979 free food stamps reform. Section 2 describes the data. Section 3 presents reduced form evidence of the impact of the 1979 free food stamps reform on food stamp enrollment, characteristics of compliers, and birth weight. Section 4 presents IV analyses of the impact of food stamps on birth weight, using both the initial rollout of food stamps and the 1979 free food stamps reform as instruments for enrollment. Section 5 shows how empirical estimates of the enrollment response to changes in the food subsidy can inform both empirical estimates of the private value of free food stamps and welfare analysis of free food stamps compared to partially subsidized food stamps. The last section concludes.

1 Background: The 1979 Free Food Stamps Reform

The Food Stamp Act of 1964 initiated the staggered rollout of the Food Stamp Program across counties nationwide, completed a decade later in 1974.[7] However, unlike the current food stamp program (now called the Supplemental Nutrition Assistance Program, or SNAP) the original program required participating households to buy their stamps. This was known as the "purchase requirement." The amount of the purchase requirement increased with household income and initially averaged about 60 cents (i.e., a 40% subsidy).

Starting in the late 1960s, advocates from the "hunger lobby" began calling for the elimination of the purchase requirement, arguing that it presented a barrier to participation for the poorest eligible individuals (Citizens' Board of Inquiry, 1968). In response, in 1970, the average subsidy was raised to about 65% (Hoover, 1974; Richardson, 1979).[8] However, with the worsening economy, calls for further reductions or removal of the purchase requirement continued.

The ultimate result was the Food And Agriculture Act of 1977 (implemented in January

[7]The information in this section draws heavily on Richardson (1979), Congressional Budget Office (1977), and USDA Food and Nutrition Service (2025). Appendix A provides additional details on program evolution.

[8]This was known as the "Bonus Reform" and we will use it in the normative analysis in Section 5 to provide additional subsidy variation for estimating the demand for food stamps.

1979) which eliminated the purchase requirement for all food stamp participants, increasing the subsidy to its current 100% level without changing the maximum benefit amount (USDA Food and Nutrition Service, 2025). As a result, instead of having to pay 35 cents to receive \$1 of food stamps, individuals now could receive that 65 cents of food stamps for free. Notably, the free food stamp reform did not affect the non-monetary participation costs associated with receiving food stamps: both before and after the elimination of the purchase requirement, individuals still had to go to the local food stamp offices to pick up their stamps.[9] The subsidy structure created by the 1979 elimination of the purchase requirement – a 100 percent subsidy on food up to a maximum transfer amount – has remained in place to the current day. Yet there is remarkably little empirical evidence of the impact of the removal of the purchase requirement.[10]

Figure 1 provides a stylized illustration of the consumer choice sets and possible behavior of eligible households under both the current free food stamp regime (blue lines) and the pre-1979 purchase requirement regime (black lines); it also shows the budget set for those who do not enroll (dashed red line).[11] Eligible individuals with income I face a choice of whether or not to participate in the food stamp program and how to allocate their budget between food expenditures (x-axis) and expenditures on all other goods (y-axis).

Under the pre-1979 purchase requirement (black budget set), participating households must pay cash ($\$R \geq 0$) to purchase their food stamps ($\$F \geq 0$); F (known as the "bonus") reflects the amount paid by the government. In other words, the household pays the government $\$R \geq 0$ and in return receives a dollar value of food stamps $\$(R + F)$ from the government.[12] The program parameters F and R together determine the magnitude of the food subsidy $s = \frac{F}{R+F}$ that the individuals receive on the first $(R + F)$ dollars of food purchases, after which they face the full (unsubsidized) marginal price for any additional food (which we normalize to 1).

The limit as the subsidy $s \to 1$ (i.e., as the purchase requirement $R \to 0$) is the contemporary free food stamps program (blue budget set). Under the free food stamps program, enrollees face a 100 percent food subsidy up to F, after which they again face the full (unsubsidized) price of food for any additional food purchases. The elimination of the purchase requirement thus reduced

[9]Although scholars have described the elimination of the purchase requirement as the centerpiece of the 1979 reform, (Berry, 1982; Rosenfeld, 2010), it did include some other changes to the food stamp program. We describe these changes in Appendix A.3 and discuss their implications for the normative analysis in Section 5.

[10]Historical accounts have claimed that eliminating the Food Stamp Program's purchase requirement raised program enrollment, with some estimates suggesting that enrollment rose by 3.6 million households, or roughly 25% (Ohls and Beebout, 1993). Existing empirical analyses of the impact of the elimination of the purchase requirement use only time series (e.g., Morentz, 1980; Ohls and Beebout, 1993) or cross-sectional (e.g., Brown, 1988) variation, and do not consider impacts on outcomes other than enrollment.

[11]The illustration follows that of, e.g., Moffitt (1989) and Hoynes and Schanzenbach (2009) but adds a (monetized) participation (or take up) cost of enrolling in the program (k). For expositional clarity, we graph this participation cost as additive, but in the analysis below we do not impose this (see Section 5).

[12]In principle, the household could pay $0 \leq r \leq R$ for various values of r, and in return receive a dollar value of food stamps $f = r(R + F)/R$ from the government; in practice, however, essentially all chose the maximum allotment $R + F$ (see Appendix B.2).

what enrollees had to pay for food stamps (from $R > 0$ to 0) but did not change the amount of government spending per enrollee; the bonus F was unchanged.

The figure illustrates two other aspects of the program. First, the hypothetical enrolled households it displays are *infra-marginal*: that is, they would have bought more food than their food stamps benefit even in the absence of the program. Evidence from both the contemporary program and the time period of the purchase requirement indicates this was the case for most enrollees (Appendix B). Second, it assumes that participation involves a cost k, which we model as a monetary participation cost.[13] The existence of participation costs is consistent with evidence of incomplete take up both under the original partial subsidy (Richardson, 1979; Congressional Budget Office, 1977) and under the current, 100 percent subsidy (Hoynes and Schanzenbach, 2015). There is also direct empirical evidence that participation costs contribute to incomplete take up of the free food stamps program (Schanzenbach, 2009; Finkelstein and Notowidigdo, 2019).

The figure illustrates three distinct types of households affected by the removal of the purchase requirement: never takers, compliers, and always-takers; because the elimination of the purchase requirement (weakly) expands the budget set, there are no defiers. Person x is a never taker: she chooses sufficiently little food in the absence of the food stamp program that she is better off not participating (under both the purchase requirement and the free stamps program) rather than incurring the participation cost k.[14] Person z is an always taker: she participates under the purchase requirement and continues to participate under the free food stamp program. Person v is a complier: she chose not to participate under the purchase requirement, but participates under the free food stamp program. Our estimate of the enrollment response to the elimination of the purchase requirement will identify the fraction of enrollees in the free stamp program who are compliers.

2 Data

2.1 Data Sources

2.1.1 Administrative Data

Our primary analyses use county-year data from 1961 through 1981 to study the impact of both the initial rollout of the food stamp program and the 1979 elimination of the purchase requirement.

The key outcomes we examine are food stamp enrollment and infant health. Why infant health? Important evidence from the program rollout finds that food stamps improved infant health, and

[13]Participation costs may include the time cost of enrolling and (in the period we are studying) of retrieving food stamps from the local food stamps office each month (Congressional Budget Office, 1977). There may also be psychological participation costs like stigma (Moffitt, 1983; Anders and Rafkin, 2025).

[14]There may also be households who consume a lot more food — even more than what the food stamp program would provide — but still choose not to participate due to sufficiently high participation costs k. These are not depicted simply for visual ease.

this evidence has shaped policy discussions over the design of transfer programs (see introduction). Yet, there are reasons to suspect the impact on infant health may be limited: Although pediatric nutrition and developmental biology link maternal nutritional deficits to adverse infant health outcomes, including low birth weight (e.g., Keats et al., 2021; González-Fernández et al., 2024), other work emphasizes that maternal physiology may buffer fetal growth against short-run fluctuations in maternal nutrition (Thayer et al., 2020); moreover, the U.S. population may not face nutritional deficits severe enough to generate infant-health impacts. Another motivation for studying infant health is that, beyond its intrinsic interest, it is also a useful summary measure of early-life conditions and predicts later-life health, education, and economic outcomes (e.g., Black et al., 2007). Finally, high-quality administrative data make infant health one of the rare outcomes that can be measured consistently during this period and at scale.

Food Stamp Enrollment: 1961–81. Our primary participation measure is county-year counts of the number of participants in the food stamp program, which come from books published by the Food and Nutrition Service. Almond et al. (2011)'s replication package provide these data from 1961-1979; we digitized the data on participant counts for 1980–81 from books that we found in the Cornell libraries (U.S. Deptartment of Agriculture, 1980-1981).[15] We also use data from the Bureau of Economic Analysis' Regional Economic Information System (REIS) (U.S. Department of Commerce, n.d.) to examine the total dollars transferred through the food stamp program in each county-year. We deflate spending measures to $1970 using the CPI-U.

Infant Health: 1968–81. Our outcome variables for infants come from birth-level data that are available nationwide from the National Center for Health Statistics Vital Statistics starting in 1968 (National Center for Health Statistics, 1968-1987).[16] For live births, the data contain birth weight, gestational age, and infant death in their first 12 months. They also contain exact date of birth, state and county of birth, maternal race and marital status. Our primary outcomes are birth weight and the share of births that are low birth weight, defined as birth weight under 2,500 grams. We also report supplemental analyses of gestational age, infant mortality, and birth rates.

1970 Census Data. To form county-level enrollment, spending rates, and poverty rates at the county level, we use the 1970 Census to obtain county-level counts of total population and of the population in poverty.

[15] In robustness analyses, we extend the data through 1984 using data digitized from the same books in later years.

[16] Prior to 1972, states reported outcomes for a random 50% of births; after 1972 some states started reporting 100% of births and by 1985 all states were reporting 100% of births.

2.1.2 Survey Data

We draw on several national surveys to study how demographic characteristics of enrollees are affected by the free food stamp reform and to analyze how enrollees' food purchases compare to food stamp allotments. Relative to the administrative data, the survey data involve smaller sample sizes and do not contain fine geographic identifiers.

Characteristics of Enrollees: 1973–81. For enrollee characteristics, we use annual data from the Panel Study of Income Dynamics (PSID) (Social Research Center, n.d.) from 1973–81 and data from the National Health and Nutrition Examination Survey (NHANES II) (U.S. Department of Health and Human Services et al., n.d.) from 1976–80. The PSID is a nationally representative, annual, longitudinal survey of about 4,800 households with information on (self-reported) demographics (including race, income, age and family structure) and food stamp enrollment. The NHANES II is a one-time survey (conducted over four years) of approximately 20,000 individuals. It over-sampled populations at elevated risk of malnutrition (including low-income individuals). Sampling weights in the NHANES II adjust for the over-sampling to make the survey representative of the national population at its midpoint (March 1978). Appendix C.6 provides more detail.

Food Purchases: 1977–81. For food purchases, we use data from the USDA's Nationwide Food Consumption Survey (NFCS, 1977–1980). This is a survey of approximately 7,500 food-stamp eligible households, about half of whom were enrolled in food stamps at the time of the survey. Appendix B provides more detail.

2.2 Construction of Key Variables

We measure food-stamp *participation* (FS_{ct}) as the food stamp enrollment rate in county c and year t:

$$\text{EnrollmentRate}_{ct} \equiv \text{FS}_{ct} = \frac{\text{Number of Food Stamp Enrollees in County } c \text{ and Year } t}{\text{Number of People in County } c \text{ in 1970}} \tag{1}$$

We proxy for the county-level *eligibility* rate with the share of the people in the county in poverty in the 1970 Census:

$$\text{EligibilityRate}_c = \frac{\text{Number of People in Poverty in County } c \text{ in 1970}}{\text{Number of People in County } c \text{ in 1970}} \tag{2}$$

Putting these together, we define a county's *exposure* to the 1979 free food stamp reform as the share of the county's population that was eligible for food stamps (Equation 2) but not enrolled in them (Equation 1):

$$\text{Exposure}_c \equiv \text{EligibilityRate}_c - \text{EnrollmentRate}_{cz}. \tag{3}$$

All of the components of the exposure measure come from the 1970 census except for the count of individuals enrolled in food stamps, which is defined for a fixed year $z < 1979$. We choose $z = 1972$ to allow us to examine trends across counties with differential exposure for a reasonably long period prior to the 1979 reform. However, we show below that the choice of year is inconsequential. Due to measurement error from combining different data sources in the exposure measure's numerator and denominator, we estimate that 3.2% of population-weighted counties in our baseline sample have negative exposure. We censor these estimates at zero, and for symmetry we censor the top 3.2% of the exposure distribution to the 96.8th percentile.[17]

Our empirical analysis of the free food stamp reform leverages county-level variation in Exposure_c. The intuition behind the exposure design is that counties with a greater share of the population that was eligible for food stamps but had not enrolled had greater scope for an enrollment increase in response to the elimination of the purchase requirement. All else equal, a county's exposure is increasing in the county's eligibility rate (i.e., the poverty rate) and decreasing in its take up rate (i.e., the share of the eligible population enrolled). We show below that in practice the instrument operates primarily through variation in the poverty rate, rather than variation in take up among eligible households.

2.3 Sample Restrictions

In all of the county-year analyses of administrative data, we form balanced panels. This necessitates excluding the approximately 5% of (1970-population-weighted) counties that are missing census data, participation data, or infant health data in any relevant analysis year. For the analysis of the free food stamp reform we use data for county-years 1972–81. We limit our analysis to the approximately three-quarters of population-weighted counties that had introduced food stamp program in 1970 or earlier; this is to ensure that our analysis from 1972 on is limited to counties that are plausibly in "steady state" after the initial rollout. For the program rollout, we analyze county-years from 1968 (the start of the birth weight data) through 1978 (the last year prior to the elimination of the purchase requirement). We follow Almond et al. (2011)'s event study sample and restrict to the approximately 40% of population-weighted counties that rolled out the food stamp program after July 1969. This restriction ensures that we observe pre-period infant health outcomes for each county in the sample.

Our baseline analysis sample for the purchase requirement thus consists of 1,804 counties representing 71% of the population, while our baseline analysis sample for the rollout consists of 1,520 counties representing about 40% of the population. These samples are largely disjoint (Appendix Figure A2). Appendix Table A1 provides more details on the sample restrictions.

[17] Appendix Figure A1 shows the raw exposure measure and how we censored it. In everything else that follows we use the censored exposure measure.

3 Impacts of the Free Stamps Reform

3.1 Empirical Framework

We estimate the following event studies:

$$\mathrm{FS}_{ct} = \sum_{\tau \neq 1978} \delta_\tau \left(\mathbf{1}\{t=\tau\} \times \text{Exposure}_c\right) + \alpha_c^1 + \gamma_t^1 + u_{ct} \tag{4}$$

$$y_{ct} = \sum_{\tau \neq 1978} \beta_\tau \left(\mathbf{1}\{t=\tau\} \times \text{Exposure}_c\right) + \alpha_c^2 + \gamma_t^2 + \varepsilon_{ct} \tag{5}$$

where $\mathbf{1}\{t=\tau\}$ is an indicator equal to one if observation year t corresponds to calendar year τ, with 1978 omitted as the reference year. The parameters α_c are county fixed effects, γ_t are calendar-year fixed effects, and Exposure_c is our time-invariant, county-level exposure measure defined in Equation 3 above. FS_{ct} denotes the county-year food stamp enrollment rate and y_{ct} denotes a county-year infant health outcome. We weight all regression analyses by the 1970 population and report standard errors clustered by county.

The coefficients of interest are the δ_τ's and β_τ's on the interaction of the calendar-year indicators with exposure to the free stamps reform. For ease of interpretation, we often report the average impacts by comparing these coefficients three years before and after the reform:

$$\beta_{\text{avg}} = \frac{1}{3}[(\beta_{79} + \beta_{80} + \beta_{81}) - (\beta_{76} + \beta_{77} + \underbrace{\beta_{78}}_{\equiv 0})] \tag{6}$$

and similarly for δ_{avg}.

The identifying assumption for causal interpretation of the δ_τ's and β_τ's from Equations 4 and 5 as the impact of the free stamps reform is that there are no shocks to outcomes that coincide with the policy change in 1979 and correlate with the county's exposure to this reform. We present evidence supporting the plausibility of this assumption by documenting similar pre-trends in both FS_{ct} and y_{ct} before 1979 in counties with differential exposure. Examination of these pre-trends will also alleviate potential mean-reversion concerns when analyzing the impact of the reform on enrollment rate, since exposure is mechanically decreasing in the enrollment rate (see Equation 3).

Cross-sectional Patterns. Figure 2a shows the distribution of Exposure_c across (population-weighted) counties. Exposure ranges from 0 to 0.2, with a mean of 0.06 and a 0.1 difference between the 10th and 90th percentile counties.[18] The county-level heat map of exposure in Panel (b) reveals greater exposure in the South, consistent with the historical evidence (Citizens' Board of Inquiry, 1968), as well as in the Great Plains region.

[18] Appendix Figure A3 shows that the county-level exposure measure using the 1972 county-level enrollment rate is highly correlated with exposure measures using the 1975 or 1977 enrollment rate.

Figure 3a shows the distribution of the 1978 share of births that are low birth weight across (population-weighted) counties. The share low birth weight ranges from 0 to 0.2, with a mean share of 0.07; moving from the 10th to 90th percentile of the distribution increases the share low birth weight by 0.04 percentage points. The county-level heat map in Panel (b) reveals higher shares of low birth weight births in the southeast and southwest, but again exhibits meaningful within-state variation.

Panel (c) shows that counties with greater exposure to the free stamps reform have a higher share of births with low birth weight. The positive correlation is natural since counties with higher poverty rates have mechanically higher exposure (see Equation 3). However, it raises potential concerns about mean reversion when looking at impacts on birth weight, which we again investigate below based on the pattern of pre-period event study coefficients.

3.2 Results.

Impact on Food Stamp Enrollment. Figure 4 shows impacts on the enrollment rate. Panel (a) shows the raw time series of enrollment rates separately by (population-weighted) quartile of county exposure, where higher quartiles correspond to greater exposure. Before 1979, enrollment-rate trends were moving roughly in parallel across the four quartiles of exposure. In 1979, all counties experienced an uptick in enrollment, with the magnitude of the uptick increasing in quartile of exposure. In particular, counties in the fourth quartile of exposure exhibit a pronounced jump in enrollment relative to the third quartile, which is consistent with the particularly large increase in average exposure between the third quartile (average exposure $= 0.056$) and the fourth quartile (average exposure $= 0.127$). By contrast, average exposure is 0.03 in the second quartile and 0.01 in the first quartile; consistent with our design, these series exhibit little divergence in 1979. Panel (b) shows the δ_τ coefficients from estimating Equation 4 with quartile-of-exposure dummies in place of the continuous Exposure_c measure; the first quartile is the omitted category. Consistent with the raw data in Panel (a), we observe a substantially higher enrollment increase in the fourth quartile.

Panel (c) plots our baseline estimate of the impact on food stamp enrollment, showing the δ_τ coefficients from estimating Equation 4 with the linear exposure measure. As expected from the prior two panels, it shows a sharp jump in enrollment after the elimination of the purchase requirement in more-exposed counties compared to less-exposed counties. Relative to the three years prior to the free food stamp reform, the average increase in the enrollment rate three years after the reform in a county that was completely exposed relative to one that had no exposure was 0.324 (s.e. 0.027). Given average exposure prior to the reform was 0.06 (see Figure 2), the estimates imply that the elimination of the purchase requirement increased average food stamp enrollment by 1.85 percentage points (standard error = 0.15), or about 25 percent relative to an average enrollment rate in 1976-78 of about 7.6 percentage points.[19]

Complier Characteristics. Advocates of the free food stamp reform argued that it would increase food stamp enrollment among the poorest households (AEI, 1977). The time-series survey data are consistent with this.[20] Figure 5a shows that the increase in enrollment was concentrated at the bottom of the income distribution. Panel (b) shows that the average income of enrollees fell precipitously right around the reform, both on an absolute scale and relative to those who were not enrolled. Table 1 summarizes various characteristics of food stamp enrollees before the free food stamp reform (Column 1) and after (Column 2), while column 3 uses these two columns together with standard complier algebra to back out characteristics of compliers; we use a complier share of 0.20 based on the estimated enrollment response.[21] Across a variety of measures, compliers (Column 3) look substantially poorer than always-takers (Column 1). For example, compliers' average equivalized household income (i.e., adjusting for the number of adults and children in the household), is less than half of always-takers' income. Compliers are also more than 5 times more likely to be below 50 percent of the poverty line (39 percent, versus 7 percent for always-takers).[22]

The finding that the free stamps reform drew in lower-income enrollees suggests that we might expect the impact of enrollment on birth weight from the free stamps program to be larger than what was previously found for the program rollout.

Impact on Birth Weight. Figure 6 shows estimates of the impact of the elimination of the purchase requirement on birth weight. The top three panels show results for log birth weight, and the bottom three for share low birth weight. We again move from the raw data by quartile (Panels a and d), to event studies by quartile (Panels b and e), to our baseline event-study estimates using a linear exposure measure (Panels c and f).

The raw data cast doubt on the hypothesis that the free-stamps reform raised infant birth weight. There is no evidence of changes in birth weight in 1979 relative to trend, either in aggregate or in higher-exposure counties relative to lower-exposure counties (Panels a and d). The event studies by quartile also show no discernible impact (Panels b and e).

Our baseline estimates in Panels (c) and (f) show null impacts for the event-study coefficients (i.e., the β_τ's from Equation 5). For log birth weight (Panel c), the difference in the estimated

[19]Such interpretation of magnitudes using the average exposure involves an assumption that the impact on a county with zero exposure would be zero (no "missing intercept"). This assumption is plausible in our setting since zero exposure implies no "slack" in food stamps take-up that could respond to changes in the subsidy.

[20]Appendix C.6 contains more detail on the data and variable definitions, and analyses, as well as additional results.

[21]Specifically, the share of compliers is given by $\sigma_c = 1.85 \div (1.85 + 7.6) \approx 0.20$) where 1.85 is our estimated percentage point increase in enrollment from the elimination of the purchase requirement and 7.6 is the estimated enrollment rate prior to its elimination. Of course, this calculation assumes that, absent the purchase requirement reform, there would have been no change in enrollment. Consistent with this, the enrollment rate in the lowest exposure group in Figure 4a (with average exposure of 0.01) rises slightly after the free food stamp reform but remains similar to pre-reform levels overall.

[22]We defer a discussion of the table's last row depicting social marginal welfare weights until Section 5.

average coefficients post-reform relative to pre-reform is –0.0013 log points (standard error = 0.0040). Thus, in the average county (with an exposure of 0.06), the free stamps reform is associated with a statistically insignificant decline in birth weight of about 0.0078 percent (s.e. 0.024 percent). Likewise, there is little evidence of an impact on the share low birth weight (Panel f). The difference in the estimated average coefficients post reform relative to pre-reform is -0.0013 (s.e. 0.0044), suggesting that the impact on the average county was a statistically insignificant decline in the share low birth weight of 0.0078 percentage points (s.e. 0.026 percentage points).

We also see no evidence of a statistically significant impact of the free food stamp reform anywhere in the birth weight distribution more generally (Figure 7). The only marginally significant impact is a reduction in birth weight in the second-highest birth weight bin — between 4,000 to 4,499 grams — where the point estimate suggests that an increase in exposure of 0.1 is associated with a 0.86% reduction in births between these thresholds.

Identification. The non-experimental nature of the variation in exposure raises questions about the validity of the identifying assumption that, absent the 1979 free food stamp reform, counties with different exposure to that reform would have experienced similar trends in enrollment and birth outcomes. Reassuringly, trends in enrollment rates between higher and lower exposure counties appear fairly flat between 1972 and 1978 for both enrollment (Figure 4 Panel c) and birth outcomes (Figure 6 Panels c and f).

As discussed, mean reversion is a natural concern given our empirical strategy. However the results are not consistent with mean reversion. Mean reversion would produce a *decline* in enrollment in more-exposed counties starting in 1972, the year that our exposure variable is defined. Instead, we see a modest increase in enrollment rates in higher exposure counties in 1972–75, followed by a modest relative fall in 1975–78.

A more general concern is omitted variables bias: county-level exposure could be correlated with changes in confounding factors that are correlated with enrollment or infant health. Because the sharp increase in enrollment coincides with the policy in 1979, any confound would need to produce a large, contemporaneous shock in 1979 to food stamp enrollment that differentially affected high-exposure counties relative to low-exposure counties. In considering potential confounds, we first observed that the identifying variation in the exposure measure primarily comes from counties' poverty rates (i.e., the share eligible) rather than take-up rates (i.e., the share of eligible households who enroll). Specifically, in a mediation test, we find that adding controls for deciles of county take-up rate indicators $\times$ calendar year indicators does not affect the estimated impact of the free food stamp reform on enrollment in Equation 4, but adding controls for deciles of county poverty-rate indicators $\times$ calendar year indicators eliminates the impact of the reform (Appendix Figure A5). This finding is consistent with the time series evidence in Figure 5 and Table 1 that the reform drew in poorer enrollees.

The fact that more exposed counties are poorer raises the concern that any secular force that correlates with both the poverty rate and the enrollment rate could differentially affect outcomes in higher vs. lower-exposure counties. Given that food stamp enrollment is highly counter-cyclical (Appendix Figure A6), variation in local business cycles across counties with different exposure is therefore a potential confound. In Appendix C.1 we therefore proxy for local business cycles using employment in industries that have small shares of workers on food stamps and implement Freyaldenhoven et al. (2019)'s proxy approach to correcting for potential confounds. We find that this procedure purges pre-trends in the food stamp enrollment analysis but has a negligible impact on the magnitude of the estimated impact of the free food stamp reform and does not affect the null results for birth weight measures.

Robustness. We next probe the robustness of the enrollment increase and the null effects on both birth weight measures to a number of alternative specifications. In addition to the local business cycle proxy discussed above, we examine sensitivity to including state-by-year fixed effects which may capture other time-varying confounds at the state level that are correlated with exposure. There is also a well-known, large decrease in births over our study period due to an "echo" from the baby boom (Appendix Figure A7). Since exposure may correlate with county-level differences in cohort effects, we also tried controlling for the size of the 1970 county-level female population of child-bearing age, interacted with calendar-year fixed effects. Finally, because much of the variation in exposure comes from differences in exposure between counties in the third vs. fourth quartile of exposure (recall Figure 4 Panel a), we adopt specifications that restrict the analysis to these two quartiles.

Appendix Figure A8 shows that our estimates survive these tests.[23] Across all alternative specifications the null impacts on log birth weight and share low birth weight persist. The statistically significant increase in enrollment also persists, although in the specification with state-by-year fixed effects, the magnitude declines by about half.[24]

Finally, we note that our analysis stops in 1981 because that year Congress passed three major pieces of food legislation which restricted food stamp eligibility and reduced benefit inflation-indexing (Congressional Quarterly, ed, 1982; Allen, 1983). Nevertheless, Appendix Figure A18 shows results through 1984. We find a *decrease* in birth weight (Panel b) and an increase in share low birth weight (Panel c) in more exposed counties in the later part of the sample. However, the wrong-signed estimate for birth weight is concentrated at the top of the birth weight distribution (Panel d), making it unlikely to reflect the impacts of the 1979 free-stamps reform.

[23] Appendix Figures A10 to A12 present the underlying event studies.

[24] We consider the implications of the smaller estimate with state-by-year fixed effects for the normative analysis in sensitivity checks (Appendix D.4).

Additional Outcomes. We also estimated the impact of the elimination of the purchase requirement on other measures of food stamp participation and infant health. Appendix Figure A13 shows the estimated impact of the free food stamp reform on food stamp spending per capita. Moving from 0 to 100% exposure increased food stamp spending per capita on average by $40 per capita (standard error = $5.4), suggesting an average increase in food stamp spending from the free food stamp reform (given average exposure of 0.06) of $2.4 per capita, or about a 16% increase relative to average per-capita food stamp spending of about $14.50 in 1976-1978. Appendix Figure A9a shows robustness of this estimate across alternative specifications.

We also looked at the impact of the removal of the purchase requirement on four other measures of infant health studied by Almond et al. (2011): birth counts, infant mortality rate, gestational age, and the share of early term births (defined as share of births with a gestation period of less than 37 weeks). These outcomes are of direct interest and can also help in interpreting the birth weight results; changes in births or infant mortality may create compositional changes that complicate interpretation of the birth weight results, while changes in gestational age are a potential pathway by which birth weight may change (the other being higher intrauterine growth conditional on gestational age). Panels (b) through (e) of Appendix Figure A9 summarize the findings.[25] There is no evidence across specifications of an impact of the free stamps reform on infant mortality. There is some evidence that the removal is associated with a decline in births (a somewhat counter-intuitive result) and a decline in the share of births that are pre-term, but these results are not robust across alternative specifications. Overall, these additional analyses do not provide compelling evidence of improvements in infant health associated with the removal of the purchase requirement, although they do not fully reject the possibility of such improvements.

4 IV Impact of Food Stamps on Birth Weight Using Two Reforms

The preceding analysis indicates that there is not a statistically significant impact of the elimination of the purchase requirement on birth weight. Interpretation of the magnitude of these "null results" requires an instrumental variable (IV) framework that relates food stamp enrollment to birth outcomes. The IV framework also facilitates comparisons to the impact of the initial rollout of the food stamp program on birth weight in common units. In this section, we therefore first re-analyze the reduced form impact of the initial introduction of the food stamp program on birth weight, using the staggered, county-by-county rollout design pioneered by Almond et al. (2011). We then analyze both reforms in a common IV framework.

[25] Appendix Figures A13 to A17 present the underlying event studies. Appendix C.5 describes data details and sample restrictions for these additional outcomes.

4.1 Reduced Form Impact of the Food Stamp Rollout on Birth Weight

4.1.1 Estimating equations

Let r denote time relative to when the food stamp program was rolled out in county c, with $r = 0$ denoting the date of the rollout. Because food stamp enrollment data are available by calendar year (t) while infant health outcomes are available for calendar year-months (m), relative time r is either measured in calendar years or in 12-month periods relative to the month of rollout, depending on the outcome.

Because of the staggered implementation and the potential for heterogeneous treatment effects, we implement a stacked event study analysis by constructing and "stacking" (i.e., vertically appending) datasets d for groups of counties sharing the same rollout date $r = 0$. This approach follows Cengiz et al. (2019) and has recent econometric justification in Wing et al. (2024); see also Callaway and Sant'Anna (2021) and Sun and Abraham (2021). Appendix C.2 provides additional details on our implementation, which we summarize briefly here.

In dataset d, for counties that rolled out on that date (the treatment group), we include up to 3 years before and after the rollout (i.e., relative years –3 to +2); for counties that did not rollout on that date (the control group), we include data from any of the corresponding 6 years before their county's rollout. We then stack the datasets and estimate:

$$\mathrm{FS}_{ctd} = \alpha_{cd} + \gamma_{td} + \sum_{\tau \neq -1} \delta_\tau (\mathbf{1}\{r = \tau\}) + \epsilon_{ctd} \tag{7}$$

$$y_{cmd} = \alpha_{cd} + \gamma_{md} + \sum_{\tau \neq -1} \beta_\tau (\mathbf{1}\{r = \tau\}) + \epsilon_{cmd} \tag{8}$$

where FS_{ctd} is the food stamp enrollment rate, y_{cmd} is a birth weight outcome, α_{cd} are county-by-dataset fixed effects and γ_{td} (γ_{md}) are calendar time-by-dataset fixed effects (where calendar time is either year t or year-month m). We cluster our analyses at the county level, and weight the analyses using 1970 population-weighted versions of the weights in Wing et al. (2024).

The coefficients of interest are the δ_τ and β_τ coefficients on the indicators for whether the food stamp program was in effect in county c in relative time r for dataset d. All counties are eventually treated (i.e., eventually adopt the food stamps program). Therefore, the identifying assumption is that the timing of when the county adopts the program is exogenous, and not correlated with other factors that affect enrollment or birth weight. The pattern of pre-period coefficients let us assess the validity of this identifying assumption; it is also investigated in detail in Hoynes and Schanzenbach (2009), Almond et al. (2011), and Hoynes et al. (2016).

4.1.2 Results

Figure 8 shows the estimated impacts of the rollout. Panel (a) indicates that the rollout led to an increase in the enrollment rate of 5.5 percentage points (standard error = 0.2). However we find no impacts on birth weight. The point estimates indicate a statistically insignificant impact of the rollout on log birth weight (Panel b) of -0.04 (standard error = 0.1) and a statistically insignificant impact on share low birth weight (Panel c) of -0.1 (standard error = 0.2). We explore the robustness of these results to a number of alternative specifications and found that they look similar if we use population weights (rather than Wing et al., 2024 weights), use two-way fixed effects (TWFE) rather than stacking, include state $\times$ time fixed effects, or include the 1970 female population of child-bearing age interacted with time fixed effects.[26]

4.2 IV analyses of the impact of food stamps on birth weight

4.2.1 Estimating Equations

We use both reforms as instruments for food stamp enrollment, focusing on birth weight as an outcome.

Rollout. For the rollout analysis, our second-stage estimating equation is:

$$y_{crd} = \lambda \mathrm{FS}_{crd} + \alpha_{cd} + \tau_{md} + e_{cmd} \tag{9}$$

where y_{crd} is a birth weight outcome in county c, relative time r, and dataset d, FS_{crd} denotes the food stamp enrollment rate, α_{cd} are county-by-dataset fixed effects and τ_{md} are calendar month-year-by-dataset fixed effects. We estimate this equation by instrumenting for FS_{crd} with the following first-stage equation:

$$\mathrm{FS}_{crd} = \delta 1(r \geq 0) + \alpha^1_{cd} + \tau^1_{td} + e^1_{ctd} \tag{10}$$

where the excluded instrument $1(r >= 0)$ is an indicator that is 1 if relative time is ≥ 0, and τ_{td} are calendar year-by-dataset fixed effects.

Free Stamps Reform. For the free stamps reform, we limit our analysis to the three years before and after the free food stamp reform (1976-1981). We estimate the second stage equation:

$$y_{ct} = \lambda \mathrm{FS}_{ct} + \alpha_c + \gamma_t + e_{ct}, \tag{11}$$

[26] Appendix Figure A19 summarizes comparison across these alternative specifications, while Appendix Figures A21 to A23 present the underlying event studies. Similarly, Appendix Figure A20 summarizes estimates for additional first stage and infant health outcomes, and Appendix Figures A24 through A28 present the underlying event studies.

by instrumenting for FS_{ct} with the following first stage equation:

$$\text{FS}_{ct} = \delta(\text{Post}_t \times \text{Exposure}_c) + \alpha_c^1 + \gamma_t^1 + u_{ct}. \tag{12}$$

where the excluded instrument $\text{Post}_t \times \text{Exposure}_c$ consists of an indicator that is 1 if calendar year t is 1979–1981 (Post_t) interacted with county-level exposure (Exposure_c).

Both Reforms as Instruments. Finally, we pool the rollout and purchase requirement IV analyses, imposing a homogeneous treatment effects assumption across reforms. Specifically, we append to the datasets described above for the rollout analysis a flat (not stacked) sample of county-years for 1976-1981 (i.e., year -3 to 2 relative to the 1979 reform) and treat this as an additional dataset d. We precision-weight estimates across the two reforms.[27]

Exclusion Restriction. The IV analyses require the exclusion restriction that the reforms only affect birth outcomes via their impacts on food stamps enrollment. For the rollout, this does not require additional assumptions beyond what is needed for the reduced-form analysis of the causal impact of the rollout on birth weight. For the free food stamps reform, however, a potential complication is that if the increase in food subsidy changed the price of food for always-takers, it could potentially affect the food consumption and health of always-takers. In practice, however because always takers were essentially all inframarginal (Appendix B), the elimination of the purchase requirement did not change the marginal price of food faced by enrollees (see Figure 1).

4.2.2 Results

Table 2 Panel (a) presents the IV estimates of the impact of food stamp enrollment on birth weight, for each reform separately and also pooling the analysis across reforms. In Column 1, we show the first stage for each reform, rescaling the free food stamp reform first-stage to reflect the average change in exposure. The remaining columns show the IV estimates for the impact of food stamp on log birth weight (Columns 2 through 4) and share low birth weight (Columns 5 through 7). For both birth weight measures and both reforms, there is no evidence of a statistically significant impact of food stamps, which is unsurprising since the reduced forms are null and the first stages are strong. For the elimination of the purchase requirement, the IV estimate in Column 3 is that enrollment in food stamps reduces birth weight by 0.4% (s.e. 1.2). The 95% CI runs from -2.6% to 1.9%, ruling out that enrolling an individual in food stamps raises birth weight by more than 2%. For the rollout instrument (Column 2), the 95% confidence intervals rule out a 6% increase in birth weight. For the share low birth weight, the IV estimate from the elimination of the purchase

[27] By estimating in the above manner and clustering at the county level, standard errors account for county-level autocorrelation across the reforms.

requirement (Column 6) is a statistically insignificant decline of 0.4 percentage points (standard error = 1.3); for the rollout, the IV point estimate is larger (a 3.5 percentage point reduction in the share low birth weight) but less precise (standard error = 3.7).

Columns 4 and 7 pool both reforms for a precision-weighted estimate of the impact of food stamp enrollment on birth weight. The first stage estimate for the purchase requirement (rescaled based on the average exposure in the sample) is about 35 percent that of the rollout (0.0185 ÷ 0.0523). However, the reduced-form estimates from the free-stamps reform are much more precise; as a result, the precision-weighted estimates are more similar to the free-stamps reform. They indicate that enrollment in food stamps reduces birth weight by a statistically insignificant 0.3%, with 95% confidence intervals ruling out increases of 1.9 percent or more (Column 4). Food stamp enrollment decreases the probability of low birth weight by 0.6 pp, with 95% confidence intervals ruling out declines of more than 3 percentage points (Column 7).

4.3 Birth Weight Impacts by Race

Given the findings from Almond et al. (2011) of larger improvements from the rollout in Black birth weight, we also study impacts on birth weight separately for white and Black infants. For the race-specific analyses of each reform, we further restrict the set of counties used to analyze each reform to the subset of baseline counties that have at least 25 births of that race in every county-year analyzed.[28]

Figure 9 shows reduced-form impacts by race and policy reform. The free stamps reform did not produce statistically significant effects on birth weight for either white or Black infants (Panels a and b). The rollout did not have an impact on white births, but appears to have improved health for Black infants (Panels c and d). For Black infants, the rollout caused a statistically significant decline in the share low birth weight, and a pronounced but statistically insignificant increase in log birth weight. The average effect of the food stamp rollout for Black infant health is 0.56 log points (standard error 0.27) increase in birth weight, and a 1.2 percentage point (standard error = 0.43 pp) decline in the share low birth weight. However, the positive impacts on Black infants are not robust to reasonable changes in the specification, like using simple population weights vs. Wing et al. (2024) weights (Appendix Figure A29). Moreover, across alternative specifications, there is no evidence of improvements in white birth weight (Appendix Figure A30).

Next, we turn to examining IV impacts by race and pooling the results (Table 2, Panels b and

[28]For the free stamps analysis, the sample of counties for white births (1,748 counties) retains over 99% of the population in the baseline sample while the counties represented by the sample of Black births (636 counties) retains almost 80% of the county population in the baseline sample and 96% of the Black population in the baseline sample of counties. For the rollout analysis, the resulting sample of white births includes 1,449 counties that represent more than 99% of the baseline rollout analysis population. The resulting sample of Black births includes 491 counties that represent 74% of the population in the baseline rollout analysis and 91% of the Black population in the baseline rollout analysis.

c). The IV analysis is complicated by the lack of data on race-specific enrollment rates, which are needed for race-specific estimates of the first stage. We therefore impute county-specific, race-specific enrollment rates based on county-year enrollment data and additional state and national data (see details in Appendix C.4). The first-stage impact of each reform on enrollment rates appears to be about five times larger for Black households than whites.

Consistent with the reduced forms, there is no evidence from either reform or from pooling the reforms of a statistically significant impact of food stamps on white birth weight. Combining the reforms, the precision-weighted IV estimates for Black births are also null, but results vary across instruments for one outcome. For Black births, there is evidence from IV analysis of the rollout (but not the free stamps reform) of a statistically significant 9.5 percentage point improvement in birth weight (standard error 3.9) for one of the measures, share low birth weight; results are insignificant for the free-stamps reform (reduction of 1.0 percentage points, standard error 1.8), and for log birth weight across both instruments. The pooled estimate is not statistically significant, for either log birth weight or share low birth weight. For instance, 95% confidence intervals rule out that food-stamps enrollment causes more than a 3.6% increase in birth weight among Blacks.

In Appendix C.3 we discuss in more detail how our reduced form estimates of the impact of the rollout on birth weight compare to Almond et al. (2011)'s estimates. The authors do not show average impacts pooled across races, so our primary estimates are not directly comparable. However, we can compare our estimated impacts by race to theirs. The authors find statistically significant improvements in birth outcomes for white infants, and larger but not statistically significant improvements for Black infants. Qualitatively, we find the same pattern of larger improvements in birth weight for Black infants than white infants. However, our estimates for white infants are not statistically significant while our estimates for Black infants are statistically significant in some specifications but not others.

5 Normative analysis of subsidies for means-tested transfers

The previous analysis found little evidence that food stamp enrollment improves infant health. But food stamps are a transfer to low-income households. Causal impacts on health — or other outcomes — are neither necessary nor sufficient for redistributive transfers to be social welfare improving. Rather, key inputs into transfers' redistributive value are the private value of the transfer to enrollees and enrollees' social marginal welfare weights.[29] In this section, therefore, we sketch a stylized framework that illustrates the key empirical objects that determine how a change in subsidy affects the redistributive value of a transfer. We use our empirical evidence of the enrollment response to the increase in subsidy to estimate these objects. We then empirically evaluate both enrollees' perceived private value of free food stamps and the social-welfare trade-off

[29] Section 3.1 of Banerjee et al. (2024) discusses this point in depth.

between free food stamps and positive cost sharing.

5.1 Conceptual Framework

Setup. A government transfer program provides eligible enrollees some good F (whose market price is normalized to 1) at a subsidy $0 \leq s \leq 1$. That is, the household pays $(1-s)F$ in cash and receives F from the government.

We define the average enrolled household's "normative net private value" (NNPV) from the program per dollar of mechanical government spending (sF) as NNPV(s) $= V(s) + D(s)$; for simplicity we often suppress the dependence on F because we analyze variation in s holding F constant. Enrollees' NNPV in turn includes both their *perceived* net private value V and any distortions $D \lesseqgtr 0$ that lead perceived private valuations to depart from how much a household "should" be willing to pay based on their normative utility. Distortions D can reflect externalities or behavioral biases (internalities).[30] We allow heterogeneous types $\theta \sim G$ and let valuations and distortions vary by θ. Types also have potentially heterogeneous social marginal welfare weights λ_θ. Finally, we denote the social value of enrolling type θ — i.e., the welfare-weighted NNPV — by T_θ, with

$$T_\theta \equiv \lambda_\theta(V_\theta(s) + D_\theta(s)).$$

T_θ captures the private and social value of transferring F per dollar of mechanical cost sF to type θ; conceptually, it corresponds to the numerator of a welfare-weighted MVPF (Hendren and Sprung-Keyser, 2020).

Analyzing an increase in subsidy. We can use this framework to analyze the targeting and welfare properties of an increase in the subsidy Δs for $s_0 \rightarrow s_1$, with $s_1 > s_0$. For simplicity, we assume that there are two types $\theta \in \{a, c\}$, where $\underline{a}$lways-takers take up when $s \geq s_0$ and $\underline{c}$ompliers take up only when $s \geq s_1$. The parameter σ_θ denotes the share of each type under the higher-subsidy regime (with $\sigma_a + \sigma_c = 1$). We also make a "constant proportional value" assumption that $V_a(s_0) = V_a(s_1)$; that is, the net perceived private value of a dollar of mechanical spending under the lower subsidy s_0 is valued the same as a dollar of mechanical spending under the larger subsidy s_1.[31]

[30] An example of an externality could be a failure of efficient intra-household bargaining, in which parents do not internalize the positive health benefits that a transfer such as food stamps might have on their children. An example of an internality could be if households misperceive the value of an in-kind transfer such as food stamps; then, their valuations based on revealed-preference would differ from their normative willingness to pay (Allcott and Taubinsky, 2015; Bernheim and Taubinsky, 2018).

[31] This assumption allows us to adopt the standard perspective that individuals who would enroll at the lower subsidy value the increase in the subsidy by the dollar value of the increased transfer. This is a natural assumption to make with unit demand (e.g., Hendren, 2020; Finkelstein et al., 2019b). See Appendix D.1 for more discussion.

A priori, it is ambiguous whether an increase in the subsidy will improve or worsen targeting, where improving targeting is defined as $T_c(s_1) > T_a(s_1)$. On the one hand, a higher subsidy enrolls individuals with lower net perceived private value. Intuitively, at s_0 compliers reveal themselves as having lower net perceived private value than always-takers since they were not willing to enroll in the program (therefore $V_c(s_0) \leq 0$) but always-takers were (therefore $V_a(s_0) \geq 0$).[32] On the other hand, the higher subsidy may also enroll individuals with higher social value, coming from higher social welfare weights (λ) or higher (positive) distortions (D). For example, the welfare weight λ_c might exceed λ_a if compliers are poorer or are liquidity constrained. On the other hand, if the poorest people always enroll because they are at subsistence and have very high marginal utility of consumption, then $\lambda_a > \lambda_c$.

Whether increasing the subsidy improves targeting (i.e., $T_c(s_1) \lesseqgtr T_a(s_1)$) thus depends on

$$\lambda_a(V_a(s_1) + D_a(s_1)) \lesseqgtr \lambda_c(V_c(s_1) + D_c(s_1)) \tag{13}$$

which is an empirical question.

Of course, the welfare consequences of the subsidy depend not only on its targeting properties but also on its social costs. The social cost of a given subsidy consists of both mechanical costs (sF) which are increasing in the subsidy, and fiscal externalities. Under the simplifying assumption that the types have common fiscal externalities per dollar of mechanical spending ($\text{FE}_a(s) = \text{FE}_c(s) = \overline{\text{FE}}$), we can write the welfare-weighted MVPF of a program with subsidy s_1 and s_0 as:

$$W(s_1) = \frac{\sigma_a\lambda_a(V_a(s_1) + D_a(s_1)) + \sigma_c\lambda_c(V_c(s_1) + D_c(s_1))}{1 + \overline{\text{FE}}}. \tag{14}$$

and

$$W(s_0) = \frac{\lambda_a(V_a(s_0) + D_a(s_0))}{1 + \overline{\text{FE}}} = \frac{\lambda_a(V_a(s_1) + D_a(s_1))}{1 + \overline{\text{FE}}} \tag{15}$$

where we use the constant proportional value assumption (as well as an assumption that the subsidy does not directly affect distortions per mechanical dollar spent) in the last step to equate the per dollar valuation at the lower subsidy s_0 to that at the higher subsidy s_1. The condition for the increase in the subsidy to increase social welfare per dollar of government spending is thus:

$$W(s_1) > W(s_0) \iff \lambda_c(V_c(s_1) + D_c(s_1)) > \lambda_a(V_a(s_1) + D_a(s_1)). \tag{16}$$

That is, the sign of the welfare impact of the subsidy increase depends only on its targeting

[32] Indeed, if s grows an infinitesimal ds, compliers are privately indifferent to the free food stamp reform (i.e., $V_c = 0$) — a standard envelope condition.

properties. Of course, it is possible that an increase in the subsidy (financed by a non-distortionary tax on the average person in society) increases social surplus even if average social welfare per dollar of government spending is reduced; i.e., $W(s_0) > W(s_1) > 1$. In fact, this is what we will find empirically when we analyze the elimination of the purchase requirement.

Graphical depiction. To get a sense for how these forces might play out, we show a graphical framework in Figure 10. The graphs assume $D_\theta = 0$ and $\overline{\text{FE}} = 0$, but allow private valuations V_θ and welfare weights λ_θ to vary within the complier and always-taker groups. The vertical axis shows the net price to obtain \$1 of the benefit ($p = 1 - s$) and the horizontal axis shows the share enrolled q (where $q = 1$ denotes everyone enrolled when $p = 0$) for a given price. The blue Marginal Private Benefit (MPB) curve shows net private value for a free program ($V_\theta(1)$) for each type, where we order types by demand ($V_\theta(1)$) so that the curve is monotonically decreasing. At $p_0 = 1 - s_0$, there are σ_a enrolled households (the always-takers). The average value of the MPB curve from 0 to σ_a shows $V_a(1)$, always-takers' average valuation. Likewise, the average value from σ_a to 1 shows $V_c(1)$, compliers' average valuation. The gray welfare weight curve (WW) plots a possible schedule of welfare weights, ordered again so that $\lambda(q)$ corresponds to the welfare weight of the q-th type to enroll. Multiplying the gray WW curve by the blue MPB curve, we obtain the orange Marginal Social Benefit (MSB). The average value of this MSB curve from 0 to σ_a corresponds to $\lambda_a V_a(1)$, and the average value of this MSB curve from σ_a to 1 corresponds to $\lambda_c V_c(1)$.

The four panels illustrate four possible cases that vary in the shape of the welfare weights (Panels a through c) or the steepness of the marginal private benefit curve (Panel d). For simplicity, we only consider welfare-weight schedules that are weakly increasing, so that compliers always have weakly higher social welfare weights than always-takers; as discussed above, this need not be the case. In Panel (a), the welfare weights are constant across enrollees. Since demand is downward-sloping, the social value of always-takers therefore exceeds the social value of compliers ($\lambda_a V_a(1) > \lambda_c V_c(1)$), and cost-sharing produces more social value per dollar than a free program.[33] Panel (b) allows for welfare weights that are higher for types with lower marginal private values, but still does not reverse the conclusion that $\lambda_a V_a(1) > \lambda_c V_c(1)$. Panel (c) shows that if the welfare weights increase enough as private values decrease, this can reverse the conclusion. Finally, Panel (d) starts with the situation in (b) but flattens the demand curve; this can also deliver the conclusion that $\lambda_c V_c(1) > \lambda_a V_a(1)$. Looking ahead, these graphs show how empirical estimates of the welfare weight schedule and demand (MPB) will determine the social value of imposing a purchase requirement versus making food stamps free.

[33]Naturally, decreasing welfare weights would further widen the advantage of $\lambda_a V_a(1)$ over $\lambda_c V_c(1)$.

Empirical Application: Free Food Stamps. The free-stamps reform closely mimics the conceptual experiment of increasing s, fixing F. Recall that the food stamp program is characterized by a purchase requirement $R \geq 0$ and a bundle of food stamps $R + F$ (Figure 1). Eliminating the purchase requirement (i.e., setting $R \to 0$) increased the subsidy from $s_0 = \frac{F}{R+F} < 1$ to $s_1 = 1$ without changing F.[34]

To move from the framework to our empirical application, we make four main assumptions. First, we assume an exclusion restriction that the free stamps reform was a "pure" subsidy increase, without any additional policy changes. Second, we assume that households had positive disposable income, such that we can learn something about private valuations from their enrollment behavior; as we discuss below, this assumption does not rule out the presence of liquidity constraints. Third, we make a revealed preference assumption, i.e., that $D_\theta = 0$. With no distortions, the condition for an increase in subsidy to improve targeting (see Equation 13) or to increase social welfare per dollar of government spending (see Equation 16) collapses to:

$$W(s_1) > W(s_0) \iff \lambda_c V_c(s_1) > \lambda_a V_a(s_1). \tag{17}$$

Improved targeting on social welfare weights ($\lambda_c > \lambda_a$) is necessary but not sufficient for the increase in subsidy to improve targeting and increase social welfare per dollar of government spending. We revisit each of these three assumptions in Section 5.3 below.

We will use the enrollment response to the free stamps reform to estimate the social marginal welfare weights of each type (λ_a and λ_c) and the average perceived net private valuation of free food stamps for each type ($V_a(1)$ and $V_c(1)$), which is enough to estimate Equation 17. We will also use the enrollment response to estimate the share of individuals enrolled under free food stamps who are always-takers vs. compliers (σ_a and σ_c); this allows us to also estimate the average perceived net private value of food stamps per dollar of mechanical government spending (i.e., $V(1) \equiv \sigma_a V_a(1) + \sigma_c V_c(1)$).

Our fourth assumption, required to provide a demand-curve interpretation of the enrollment response we estimated in Section 3.2, is that a "Law of Demand" holds:

$$\text{household } \theta \text{ enrolls at subsidy } s \iff V_\theta(1) \geq 1 - s. \tag{18}$$

This assumption says that if you are provided a good F for free, your consumer surplus per dollar of the free good exceeds $1 - s$ if and only if you would have bought that good at a per

[34] Notice that the policy therefore changes the total size of the food stamps bundle from $R + F$ to F. To work in units of F and s, rather than levels of R and $R + F$, requires us to modify the constant proportional value assumption above (i.e., $V_a(s_0, F) = V_a(1, F)$) to assume $V_a(s_0, R + F) = V_a(1, F)$. In other words, compared to the constant proportional value assumption in the general framework, here we impose that the per-dollar valuation of free stamps is constant even when the size of the food stamps bundle changes (from $R + F$ to F) as in the actual reform. See discussion in Appendix D.1.

dollar subsidy rate s. This is of course natural to assume, but is not guaranteed in settings where subsidies are nonlinear, which is common for many in-kind transfers.[35] With non-linear subsidies, consumer surplus per dollar of a free good F could be more than $1 - s$, but the household might still not enroll in the program at a subsidy of s. We discuss microfoundations of Equation 18 in Appendix D.1.

5.2 Estimating Net Perceived Private Value.

Subsidy Variation from the Free Food Stamp Reform. Under the Law of Demand, variation in the subsidy identifies the perceived net private value of food stamps by tracing a demand curve. Intuitively, if all we observe is enrollment under free food stamps, the perceived net private valuation for enrollees can be anywhere between 0 (if they are just indifferent between enrolling or not) and 1 (if they have no participation costs from enrolling and value each dollar of food stamps as a dollar of cash): $V(1) \in [0, 1]$. However, if we observe enrollment both under a lower subsidy ($s_0 < 1$) and when free ($s_1 = 1$), we identify distinct, and tighter, bounds on valuations. Valuations tighten *both* for households who would have enrolled even under the lower subsidy (always-takers) and those who only enroll under the free program (compliers). Letting $V_\theta \equiv V_\theta(1)$ to lighten notation, we now have that for always-takers, V_a is bounded from below by $1 - s_0$ rather than 0, since they were willing to enroll at subsidy s_0: $V_a \in [1 - s_0, 1]$. Likewise, for compliers, V_c is now bounded from above by $1 - s_0$ rather than 1, since they were not willing to enroll at subsidy s_0: $V_c \in [0, 1 - s_0]$. Thus observing the enrollment response to an increase in the subsidy from s_0 to 1 allows us to tighten the bounds on the average value of free stamps $V(1) \in [0, 1]$ to $V(1) \in [\sigma_a(1 - s_0),\ \sigma_a + \sigma_c(1 - s_0)]$, where the always takers share σ_a is given by the baseline enrollment under the lower subsidy s_0 and compliers' share σ_c is given by the enrollment response to the increased subsidy.

Figure 11a shows the results. The vertical axis shows the price to obtain \$1 net of take-up costs (equivalently, 1 minus the subsidy). The horizontal axis shows the quantity enrolled at each price. With data from both the free-stamps ($s_1 = 1$) and purchase-requirement ($s_0 = 0.65$) regimes, we observe three points on the demand curve: no enrollment with no subsidy $s = 0$ (since there is no food stamp program this holds by definition), all post-reform enrollees enrolled at a price of 0 ($s_1 = 1$), and $\sigma_a = 1 - \sigma_c = 0.80$ (based on the enrollment response estimated in Section 3.2) enrolled at a price of 0.35 ($s_0 = 0.65$).[36] The nonparametric bounds on the demand curve are

[35] For many in-kind transfers, the individual's decision involves both whether to enroll in the program and how much of the good to buy, with purchases of the good for those enrolled subsidized up to some cap (F), after which enrollees can purchase additional amounts at market prices. This type of subsidy structure can be found not only in the food stamp program but also across a range of in-kind transfer programs, including subsidies for rent, child care, and education.

[36] Relative to other work using revealed-preference to obtain the value of social *insurance* (e.g., Landais and Spinnewijn, 2021), being able to observe this third point at $s = 0$ is a feature of our setting. In the insurance case, willingness to pay has no upper bound. Then, extrapolation into infra-marginal demand involves parametric

given by the shaded blue rectangles. Since our empirical application has a pre-reform subsidy of $s_0 = 0.65$ to a post-reform subsidy of $s_1 = 1$, we know $V_a \in [0.35, 1]$ and $V_c \in [0, 0.35]$. Therefore, using these bounds and our estimates of $\sigma_c = 0.2$ and $\sigma_a = 0.8$, the elimination of the purchase requirement allows us to tighten the bounds from $V(1) \in [0, 1]$ to $V(1) \in [0.28, 0.87]$.

If we further assume that demand is piecewise linear, so that valuations lie at the midpoint of the bounds, we get that $V_a = 0.675$ and $V_c = 0.175$. In other words, the private value for compliers ($V_c = 0.175$) is only about one-quarter that of always-takers. Moreover we estimate an average value of a dollar of free stamps at 58 cents ($V(1) = 0.58$). This candidate demand curve is plotted in the dashed line connecting the diagonals of the rectangles.

Additional Subsidy Variation: 1970 Bonus Reform. In principle, if we could observe demand at all subsidy levels ($s \in [0, 1]$) we could use the above approach to point-identify demand and hence $V(1)$. In practice, we can refine our estimates by exploiting an additional source of price variation. In 1970, the program increased the average subsidy from $s_0 = 0.40$ to $s_1 = 0.65$ (Hoover, 1974). In Appendix D.3, we construct a version of an exposure measure similar to Equation 3, but based on the share eligible for but not enrolled in food stamps in 1968, prior to the 1970 subsidy expansion. Descriptive trends suggest that high-exposure counties experienced a large increase in enrollment relative to low-exposure counties (Appendix Figure A44a). When we estimate an event study, we find that the subsidy increase increased the enrollment rate by about 2 pp in aggregate, with reassuringly flat pre-trends prior to the earlier subsidy expansion (Figure A44b).

Figure 11(b) plots the demand curve for food stamps with the additional data point from the 1970 reform. This additional point lies essentially on the linear demand curve created from 0 to σ_a in panel (a). As a result, using this additional variation has little impact on the value of free food stamps implied when we assume piece-wise linear demand; is it now 59 cents ($V(1) = 0.59$) rather than the 58 cents above. However the additional data point allows us to further tighten the non parametric bounds and the tightening is substantial: from $V(1) \in [0.28, 0.87]$ to $V(1) \in [0.42, 0.77]$. The private value for compliers ($V_c = 0.175$) is still only about one-quarter that of always-takers ($V_a = 0.696$).

Interpretation. The enrollment response to the increase in food subsidy indicates that households act as if they do not value food stamps dollar for dollar. If they did, eliminating the purchase requirement would not affect enrollment. Thus, our evidence of a substantial (25%) enrollment response to the increase in the food subsidy points to the existence of meaningful deadweight loss that dissipates some of the transfer's gross private value. Even if revealed preference fails and households have behavioral bias, it is potentially important that they behave "as if" food stamps are worth only 59 cents per dollar.

assumptions.

Why would households act as if they perceive their value of food stamps to be less than a dollar for dollar? It is not because they literally lack any cash on hand to purchase food stamps; as discussed in Appendix A, the purchase requirement was designed to be no greater than 30% of household monthly income (and indeed was zero for the lowest income individuals). Moreover, Figure 5a shows that the vast majority of compliers have positive incomes. Therefore, either the food stamp program delivers more food than the household would choose to consume out of the equivalent cash, and/or there are participation costs to receiving food stamps.

If all households were infra-marginal at F, food stamps would be fungible with cash and we could interpret the average value of a food stamps dollar being less than dollar-per-dollar ($V_\theta(1) < 1$) as entirely reflecting ordeal costs. As noted earlier, prior research has argued that most food stamp enrollees are inframarginal; we confirm this for our study period using administrative data on food stamp benefit schedules and survey data on food expenditures among food stamp enrollees (see Appendix B.2). This suggests that much of the deadweight loss comes from ordeal costs that dissipate some of the transfer's gross private value.[37]

Implications. An estimate of 41 cents dissipated per dollar of food stamps ($1 - V(1) = 0.41$) suggests that two fifths of food stamps' value is eroded relative to a frictionless cash transfer with $V(1) \equiv 1$. This far exceeds what one gets from monetizing time costs associated with applying — about 2 cents per dollar transferred in Finkelstein et al. (2019b) — but is somewhat more comparable to envelope theorem-based estimates of ordeal costs from benefit take-up elasticities (17 cents per dollar in SNAP in Rafkin et al., 2025). Perhaps most importantly, our findings imply that the U.S. engages in substantially less redistribution than is commonly assumed under standard accounting practices that already treat the program as if it were equivalent to a frictionless cash transfer valued \$1-for-dollar (e.g., Blanchet et al., 2022; Congressional Budget Office, 2025).

Robustness. Our baseline estimate of $V(1) = 0.59$ assumed piecewise linearity of the demand curve, which is naturally a strong assumption. We therefore also consider shape restrictions to relax piecewise-linearity and make less parametric assumptions. Following Kang and Vasserman (2025), we define the *robustness parameter* r to allow demand curves whose slopes within each observed interval may deviate from the piecewise-linear demand slope by up to $r \times 100$ percent in either direction. Intuitively, the parameter r rules out demand curves that are anywhere "too

[37] However, it appears that not all of the deadweight loss is from ordeal costs. While about 93% of always-takers are infra marginal at F, about 55% of compliers are not infra marginal at F (Appendix B.2). Therefore, some of the less than dollar-for-dollar valuation for free food stamps reflects the fact that about 11% of enrollees (55% of compliers $\times$ 20% compliers among those enrolled in free food stamps) have less than dollar-for-dollar preferences for food spending. In addition, households may be infra marginal at F but not infra marginal at the larger bundle $R + F$. Indeed, in Appendix B.2, we show that essentially none of the compliers are infra marginal at $R + F$, which means their enrollment response to the elimination of the purchase requirement could in principle be entirely as a function of preferences.

flat" or "too steep" relative to the observed variation.[38] When $r \to 0$, we recover the piecewise-linear benchmark, under which admissible demand curves are the piecewise-linear functions interpolating the observed points. When $r \to 1$, we recover the nonparametric bounds above, allowing any monotone demand curve consistent with the observed data (the light blue or green regions in Figure A45). Figure 11c shows bounds on the willingness to pay for food stamps $V(1)$ as a function of the robustness parameter r, under three scenarios: without observing data from the purchase requirement (red series), with data from the purchase requirement (blue series), and with data from the purchase requirement and the 1970 reform (green series). The shape restriction reduces the bounds in all three scenarios. The value of food stamps is point identified as the area under the piecewise-linear demand curve when $r = 0$. When $r = 0.5$, the value of food stamps is tightly restricted to lie between [0.53, 0.65] if we include the 1970 bonus reform and [0.48, 0.68] if we exclude it.[39]

5.3 Empirical Welfare Tradeoff: Free Food Stamps vs. Partial Subsidy

5.3.1 Social Marginal Welfare Weights (λ_c, λ_a).

Table 1 already showed that compliers were substantially poorer than always-takers. We convert these estimates into social marginal welfare weights λ_a and λ_c by assuming that households' utility is CRRA in household equivalized income with a risk aversion parameter of 3 (so that marginal utility is income^{-3}). We normalize the social marginal welfare weights by assuming that it is 1 for the average household in society; this is consistent with a utilitarian social planner and the financing of any new spending in the food stamp program via a non-distortionary tax on the average household. We then study the average social marginal welfare weight among enrollees in the survey data, before and after the free-stamps reform.

The last row of Table 1 shows the results. The average social marginal welfare weight for compliers is 6.2, compared to 3.2 for always-takers, where a welfare weight of 1 corresponds to the welfare weight of the average household in society from 1973–1981.[40] One way of interpreting this finding is that compliers have twice as high social marginal utility from a dollar as always-enrollees ($\approx 6.2 \div 3.2$). This suggests that the increase in the subsidy improved targeting on social welfare

[38] More formally, consider an interval of the demand curve defined by the subsidy and take-up tuples $(1 - s_i, q_i)$ and $(1 - s_j, q_j)$. Let $\hat{\beta}$ be the slope of the line connecting these points: $\hat{\beta} = \frac{q_j - q_i}{s_i - s_j}$, with $\underline{\beta}$ and $\overline{\beta}$ being lower and upper bounds on $\hat{\beta}$. Then r is defined by

$$\underline{\beta} = \frac{\beta}{1 - r} \qquad \text{and} \qquad \overline{\beta} = \beta(1 + r). \tag{19}$$

[39] The robustness parameter $r = 0.5$ is reasonable benchmark, as it lies halfway between the polar values of $r = 0$ (piecewise-linear) and $r = 1$ (completely nonparametric).

[40] These estimates of social marginal welfare weights are high, for both always-takers and compliers. This is a standard consequence of CRRA utility, which decreases very quickly in income, and assuming that the budget constraint is closed via a non-distortionary tax on the average person in society.

weights.

Higher welfare weights among compliers relative to always takers are necessary but not sufficient to improve overall targeting (see Equation 17). Qualitatively, the fact that welfare-weights $\lambda_c > \lambda_a$ but private values $V_a > V_c$ shows that there is a trade-off between private and social value in this setting. The ultimate sign and magnitude of the trade-off will inherently depend on society's redistributive preferences.

5.3.2 Welfare Trade-Off

Our preceding estimates of V_c, V_a, λ_a and λ_c imply that social welfare per dollar of government spending is higher under cost sharing than free food stamps. Our framework shows that free stamps produce more welfare per dollar of mechanical expenditure than cost sharing (i.e., $W(1) > W(0.65)$) if and only if $\lambda_c V_c > \lambda_a V_a$ (recall Equation 17). Specifically, compliers' social value $\lambda_c V_c = 6.2 \times 0.18 \approx 1.1$, while always-takers' social value $\lambda_a V_a = 3.2 \times 0.69 \approx 2.2$. In other words, even though compliers' social marginal welfare weights exceed always-takers' ($\lambda_c > \lambda_a$), empirically, we find this is not large enough to offset our estimates of compliers' lower private values ($V_c < V_a$).

Figure 12 depicts these estimates in an empirical version of the framework graph (Figure 10). The blue line on the curve plots our demand estimate; the gray line plots our estimates of welfare weights; and the orange line plots the social value ($\lambda_\theta V_\theta$), where we again sort all estimates in demand space. In the graph, cost-sharing produces more social welfare per dollar than free stamps ($\lambda_a V_a > \lambda_c V_c$ in the orange curve).

Moreover, given our estimate that compliers are 20% of the enrollees under the free-stamps regime ($\sigma_c = 0.2$) we can use the above estimates to calculate $W(1) = \frac{1.98}{1+\text{FE}}$ (see Equation 14). Thus, unless the negative fiscal externality is greater than or equal to the mechanical cost, the free stamps program financed by a non-distortionary tax on the average person in society still improves social welfare ($W(1) > 1$). It just produces less social welfare per dollar of government expenditure than a 65% subsidy.[41]

[41] A likely source of negative fiscal externalities comes from individuals reducing their labor supply (and hence income tax revenue earned) to quality for food stamps. A priori, these negative fiscal externalities are likely low, since the tax-and-transfer rate on low-income households was (just 12.9% according to Hendren and Sprung-Keyser, 2020). Consistent with this, Hendren and Sprung-Keyser (2020) calibrate that the large but imprecise decline in earnings from the rollout of the food stamp program labor-supply estimated by Hoynes and Schanzenbach (2012) implies a negative fiscal externality of just 16 cents (standard error: 70 cents) in tax revenue per dollar of food stamps transferred. In Appendix D.2, we directly estimate the employment effects of the free-stamps reform using a county-by-industry exposure design and find no evidence of impacts on employment. Moreover, longer-run fiscal externalities may be large and positive (Hendren and Sprung-Keyser, 2020; Bailey et al., 2024).

5.3.3 Assumptions and Limitations

Our finding that free food stamps deliver less social welfare per dollar of government spending than the pre-reform subsidy level (i.e., $W(1) < W(0.65)$) requires several additional assumptions beyond what was needed for the positive analysis of the impact of the removal of the purchase requirement on enrollment and infant health. We briefly discuss several potentially important limitations with the above analysis. We focus here on potential policy confounds, liquidity constraints, and violations of revealed preference. Appendix D.4 provides more detail on these analyses, as well as additional sensitivity analyses to including incorporating fiscal externalities, alternative estimates of V, and different estimates of the first stage.

Potential policy confounds. Our welfare analysis assumes that the policy can be modeled as a pure change in subsidy. Importantly, the 1979 reform did *not* affect food stamp transfers (i.e., the bonus F), the transaction costs of procuring food stamps (recipients still had to go to the local food stamp office to pick up food stamps), or benefit levels. The free food stamp reform did, however, make changes to the eligibility rules which may have reduced eligibility and hence enrollment (see Appendix A.3).[42] As we discuss in the Appendix, however, accounting for potential unmodeled eligibility reductions would actually strengthen the normative case for cost-sharing. Finally, another slippage between the framework and our empirical application is that, with the purchase requirement, R was not constant in the population, but instead increased with income (Section 1). As we discuss in the Appendix, this suggests our estimate of $V(1)$ is an upper bound, but also suggests we may under-estimate the social welfare gains from enrolling compliers.

Liquidity Constraints. Liquidity constraints do not pose a *conceptual* threat to our analysis. If households decline to enroll at a positive purchase price $s < 1$ because, due to liquidity constraints, they have a high marginal utility of consumption out of current resources, our estimate of $V(1)$ correctly reflects their private value of food stamps, given their budget and borrowing constraints. That said, the presence of such liquidity constraints might have *empirical* implications for our estimates of λ_θ, since they could create distributional concerns arising from differences in resources across enrollees. More broadly, our estimates of λ_θ are based on a relatively crude time series comparison of enrollee income before and after the elimination of the purchase requirement, together with an assumption about the utility function and may therefore contain several sources of error. In the Appendix we therefore explore alternative social marginal welfare weights and show that under standard parameterizations of social welfare functions we cannot overturn our

[42] Any contamination from changes in eligibility will not affect our positive analysis, of the impact of food stamps on infant health. That analysis requires only the exclusion restriction that the 1979 reform did not affect infant health except through its effects on food stamp enrollment; enrollment effects coming through potential changes in eligibility do not threaten this.

result that the increase in subsidy reduced targeting. However, because compliers are poorer than always-takers, if the social welfare function puts sufficiently high weight on the poorest members of society, the result will be overturned.

As noted, if households have literally zero cash on hand, then we cannot learn anything about their private valuation. However this appears unlikely. The purchase requirement R was designed to be no greater than 30% of household monthly household income (and indeed was zero for the very lowest income individuals).[43] In addition, Appendix B.2 shows that almost half of complier households are inframarginal at F, and hence had cash on hand to buy their groceries. Our interpretation is therefore that few households had truly zero cash on hand, and that liquidity constraints are therefore best addressed via the welfare weights.

Revealed preference. Behavioral biases such as misperceptions or present bias (i.e., distortions D in the framework above) may drive a wedge between perceived net private value (V) and net private value (NNPV). Our normative analysis thus far has assumed that such distortions do not exist. The existence of distortions per se need not alter our conclusions. Indeed, if welfare-weighted distortions are equal ($\lambda_a D_a = \lambda_c D_c$) then they drop out of the welfare condition (16). For instance, if society values all children equally, and food stamp enrollment benefits children in ways that households do not fully internalize, this would not affect our results. We perform calibration exercises in the Appendix to give a sense of what kind of magnitudes would be required to overturn our results. We conclude that either there would need to be perhaps implausibly large distortions for all types, or substantial heterogeneity in distortions, with compliers much more biased than always-takers. For instance, if 0% of always-takers are present biased, at least 62% of compliers would need to be present biased.

6 Conclusion

When in-kind transfers are free, even the neediest households can access the good — but they may not derive much value from doing so. This trade-off arises in many settings, from health technologies in developing countries (Ashraf et al., 2010; Cohen and Dupas, 2010) to setting the subsidy level in U.S. means-tested programs such as health insurance programs for low-income adults, the National School Lunch Program, and financial aid for higher education.

We study these issues empirically in the context of one of the largest means-tested transfer programs in the U.S. — the food stamp program or SNAP. We examine the 1979 reform that transformed this program from its original, partial subsidy design to the current, full subsidy (i.e.,

[43]A related concern is that households' willingness to accept cash for food stamps could be different from their willingness to pay, a kind of income effect when the transfer F is large relative to income. This does not bear on whether W is optimal at a given price, but it can affect the normative interpretation of V, just as Marshallian demand can mislead about welfare when income effects are present.

"free food stamps"). From a positive perspective, we find that the increase from a partial (65%) to a 100% subsidy raised enrollment by about 25%, and that marginal enrollees had substantially lower incomes than those already enrolled. We find no evidence that the enrollment induced by the move to free food stamps improved infant health. We are also unable to reject the null hypothesis that food stamp enrollment induced by the original rollout of the program in the 1960s and 1970s had no impact on infant health on average. Precision-weighted IV estimates that combine both reforms rule out an increase in birth weight from maternal enrollment in food stamps of more than 1.9%.

Our normative analysis suggests the presence of a qualitative trade-off between private and social value. On the one hand, compliers (i.e., marginal enrollees to the free food stamp reform) act as if they have about one-quarter the private value of always-takers. On the other hand, free food stamps substantially improve targeting on social value, since compliers are needier. Whether free food stamps produce more or less welfare per dollar than partially subsidized food stamps thus depends on society's preferences for redistribution, as well as the magnitude (and difference between compliers and always-takers) of any distortions that make enrollees' perceived valuation differ from their normative valuation. In our baseline calibration and several alternative parameterizations, we find that free food stamps produce less welfare per dollar than partially subsidized food stamps.

In addition to the tradeoffs inherent in subsidy design, we draw two other broader conclusions. First, our estimates of null effects on infant health suggest more limited impacts of food stamps on infant health in its first few decades than previously believed (e.g., National Academies of Sciences, Engineering, and Medicine, 2019). An open question is whether or not the increased food stamp enrollment from the 1979 move to free food stamps improved the long-run health and earnings of those who newly received food stamps in childhood — as studies examining the impact of the rollout of the initial, partially subsidized food stamp program have found (Hoynes et al., 2016; Bailey et al., 2024; Bitler and Figinski, 2024; Barr and Smith, 2023). Second, our estimate that individuals behave as if their average net private value of food stamps is 59 cents per dollar (with nonparametric bounds of 42 to 77 cents per dollar) is well below the value of $1-per-dollar assumed by the Congressional Budget Office (Congressional Budget Office, 2025) and by research on distributional national accounts (Piketty et al., 2018; Blanchet et al., 2022; Auten and Splinter, 2024), suggesting there could be meaningful deadweight loss in U.S. transfer programs. Taken together, our findings speak to core debates about both the value and optimal design of means-tested in-kind transfer programs.

References

AEI, "Food Stamp Reform," Legislative Analysis No. 4, 95th Congress, American Enterprise Institute for Public Policy Research, Washington, DC May 1977. ISBN 0-8447-0186-6.

Aizer, Anna, Hilary Hoynes, and Adriana Lleras-Muney, "Children and the US social safety net: Balancing disincentives for adults and benefits for children," *Journal of Economic Perspectives*, 2022, *36* (2), 149–174.

Alatas, Vivi, Ririn Purnamasari, Matthew Wai-Poi, Abhijit Banerjee, Benjamin A Olken, and Rema Hanna, "Self-targeting: Evidence from a field experiment in Indonesia," *Journal of Political Economy*, 2016, *124* (2), 371–427.

Allcott, Hunt and Dmitry Taubinsky, "Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market," *American Economic Review*, 2015, *105* (8), 2501–2538.

Allen, Joyce E., "Food Programs Grow Slowly From 1979 to 1983," *National Food Review*, 1983. Food Situation and Review section. Contact: (202) 475-5119.

Almond, Douglas, Hilary W Hoynes, and Diane Whitmore Schanzenbach, "Inside the war on poverty: The impact of food stamps on birth outcomes," *The review of economics and statistics*, 2011, *93* (2), 387–403.

_ , Janet Currie, and Valentina Duque, "Childhood circumstances and adult outcomes: Act II," *Journal of Economic Literature*, 2018, *56* (4), 1360–1446.

Anders, Jenna and Charlie Rafkin, "The welfare effects of eligibility expansions: theory and evidence from SNAP," *Available at SSRN 4140433*, 2025.

Anderson, Patricia M and Bruce D Meyer, "Unemployment insurance takeup rates and the after-tax value of benefits," *The quarterly journal of economics*, 1997, *112* (3), 913–937.

Ashraf, Nava, James Berry, and Jesse M Shapiro, "Can higher prices stimulate product use? Evidence from a field experiment in Zambia," *American Economic Review*, 2010, *100* (5), 2383–2413.

Associated Press, "Newsom wants to pause health care expansion to low-income immigrants," May 2025. Accessed 2026-01-31.

Auten, Gerald and David Splinter, "Income inequality in the United States: Using tax data to measure long-term trends," *Journal of Political Economy*, 2024, *132* (7), 2179–2227.

Bailey, Martha J, Hilary Hoynes, Maya Rossin-Slater, and Reed Walker, "Is the social safety net a long-term investment? Large-scale evidence from the food stamps program," *Review of Economic Studies*, 2024, *91* (3), 1291–1330.

Banerjee, Abhijit, Rema Hanna, Benjamin A Olken, and Diana Sverdlin Lisker, "Social protection in the developing world," *Journal of Economic Literature*, 2024, *62* (4), 1349–1421.

Barr, Andrew and Alexander A Smith, "Fighting crime in the cradle: The effects of early childhood access to nutritional assistance," *Journal of Human Resources*, 2023, *58* (1), 43–73.

Bernheim, B Douglas and Dmitry Taubinsky, "Behavioral public economics," *Handbook of behavioral economics: Applications and Foundations 1*, 2018, *1*, 381–516.

Berry, Jeffrey M, "Consumers and the hunger lobby," *Proceedings of the Academy of Political Science*, 1982, *34* (3), 68–78.

Bitler, Marianne and Theodore F Figinski, "Long-Run Effects of Food Assistance: Evidence from the Food Stamp Program and Administrative Data," Technical Report, National Bureau of Economic Research 2024.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes, "From the cradle to the labor market? The effect of birth weight on adult outcomes," *The Quarterly Journal of Economics*, 2007, *122* (1), 409–439.

Blanchet, Thomas, Lucas Chancel, and Amory Gethin, "Why is Europe more equal than the United States?," *American Economic Journal: Applied Economics*, 2022, *14* (4), 480–518.

Bronchetti, Erin T, Garret Christensen, and Hilary W Hoynes, "Local food prices, SNAP purchasing power, and child health," *Journal of health economics*, 2019, *68*, 102231.

Brooks, Tricia, Jennifer Tolbert, Anna Mudumala, Amaya Diana, Allexa Gardner, Aubrianna Osorio, and Shoshi Preuss, "Medicaid and CHIP Eligibility, Enrollment, and Renewal Policies as States Resume Routine Operations: Following the Unwinding of the Pandemic-Era Continuous Enrollment Provision," *San Francisco: KFF*, 2025.

_, Lauren Roygardner, Samantha Artiga, Olivia Pham, and Rachel Dolan, "Medicaid and CHIP eligibility, enrollment, and cost sharing policies as of January 2019: Findings from a 50-state survey," *Kaiser family foundation*, 2019.

Brown, Gregory M, "End of purchase requirement fails to change food stamp participation," *Monthly Labor Review*, 1988, *111* (7), 14–18.

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

Capozza, Francesco and Krishna Srinivasan, "Who should get money? Estimating welfare weights in the U.S.," 2026.

Carlson, Steven and Joseph Llobrera, *SNAP is linked with improved health outcomes and lower health care costs*, Center on Budget and Policy Priorities, 2022.

Carrington, William, Molly Dahl, and Justin Falk, "Growth in Means-Tested Programs and Tax Credits for Low-Income Households.," *Congressional Budget Office*, 2013.

Caswell, Julie A. and Ann L. Yaktine, "History, Background, and Goals of the Supplemental Nutrition Assistance Program," in "Supplemental Nutrition Assistance Program: Examining the Evidence to Define Benefit Adequacy," Washington, DC: National Academies Press, 2013.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The effect of minimum wages on low-wage jobs," *The Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

Chen, Alice, Emily Oster, and Heidi Williams, "Why Is Infant Mortality Higher in the United States than in Europe?," *American Economic Journal: Economic Policy*, 2016, *8* (2), 89–124.

Citizens' Board of Inquiry, "Hunger, U.S.A," Technical Report, Citizens' Board of Inquiry into Hunger and Malnutrition in the United States 1968.

Citro, Constance Forbes and Robert T Michael, *Measuring Poverty: A New Approach*, National Academy Press, 1995.

Clarkson, Kenneth W, "Welfare benefits of the food stamp program," *Southern Economic Journal*, 1976, pp. 864–878.

Cohen, Jessica and Pascaline Dupas, "Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment," *The Quarterly Journal of Economics*, 2010, *125* (1), 1–45.

Colas, Mark, Sebastian Findeisen, and Dominik Sachs, "Optimal need-based financial aid," *Journal of Political Economy*, 2021, *129* (2), 492–533.

Congressional Budget Office, "The Food Stamp Program: Income or Food Supplementation?," Technical Report 1977.

__ , "Reconciling the Official Poverty Measure and CBO's Distributional Analysis of Household Income," Technical Report, Congressional Budget Office January 2025. Accessed 2026-01-31.

Congressional Quarterly, ed., "Food Stamp Legislation," in Congressional Quarterly, ed., *Congressional Quarterly Almanac, 1981*, Washington, DC: Congressional Quarterly, 1982. Document ID: cqal81-1173275.

Congressional Research Service, "H.R. 2649—Food Stamp Act: Bill Summary and Status, 95th Congress (1977–1978)," https://www.congress.gov/bill/95th-congress/house-bill/2649 1977. Accessed: 2026-03-24.

Dague, Laura, "The effect of Medicaid premiums on enrollment: A regression discontinuity approach," *Journal of Health Economics*, 2014, *37*, 1–12.

Deshpande, Manasi and Yue Li, "Who is screened out? Application costs and the targeting of disability programs," *American Economic Journal: Economic Policy*, 2019, *11* (4), 213–248.

Dube, Arindrajit, T. William Lester, and Michael Reich, "Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties," *Review of Economics and Statistics*, 2010, *92* (4), 945–964.

East, Chloe N, "The effect of food stamps on children's health: Evidence from immigrants' changing eligibility," *Journal of Human Resources*, 2020, *55* (2), 387–427.

Eckert, Fabian, Ka leung Lam, Atif R. Mian, Karsten Müller, Rafael Schwalb, and Amir Sufi, "The Early County Business Pattern Files: 1946–1974," Working Paper 30578, National Bureau of Economic Research 2022.

__ , Teresa C. Fort, Peter K. Schott, and Natalie J. Yang, "Imputing Missing Values in the US Census Bureau's County Business Patterns," Working Paper 26632, National Bureau of Economic Research 2020.

Econometrica, Inc., "Updated Analysis of Cash Food Expenditures of SNAP Households," Technical Report, Econometrica, Incorporated 2012.

Einav, Liran, Amy Finkelstein, and Heidi Williams, "Paying on the margin for medical care: Evidence from breast cancer treatments," *American Economic Journal: Economic Policy*, 2016, *8* (1), 52–79.

Federal Register, "Federal Register, October 17, 1978," Vol. 43, No. 201, p. 47811 1978. U.S. Government Printing Office.

Finkelstein, Amy and Matthew J Notowidigdo, "Take-up and targeting: Experimental evidence from SNAP," *The Quarterly Journal of Economics*, 2019, *134* (3), 1505–1556.

_ , Nathaniel Hendren, and Erzo FP Luttmer, "The value of medicaid: Interpreting results from the oregon health insurance experiment," *Journal of Political Economy*, 2019, *127* (6), 2836–2874.

_ , _ , and Mark Shepard, "Subsidizing health insurance for low-income adults: Evidence from Massachusetts," *American Economic Review*, 2019, *109* (4), 1530–1567.

Food and Nutrition Service, United States Department of Agriculture, "Participation of State Agencies and Eligible Households," 42 Fed. Reg. 22357 May 1977. (to be codified at 7 C.F.R. pt. 271).

_ , "Participation of State Agencies and Eligible Households," 43 Fed. Reg. 21304 May 1978. (to be codified at 7 C.F.R. pt. 271).

_ , "Certification of Eligible Households," 44 Fed. Reg. 64069 1979. (to be codified at 7 C.F.R. pt. 273).

_ , "Certification of Eligible Households," 45 Fed. Reg. 79743 1980. (to be codified at 7 C.F.R. pt. 273).

Food and Nutrition Service, USDA, "Food Stamp Program Allotments," Federal Register, Vol. 36, No. 14118, July 29, 1971 1971.

_ , "Food Stamp Program Allotments," Federal Register, Vol. 37, No. 97, May 18, 1972 1972.

_ , "Food Stamp Program Allotments," Federal Register, Vol. 38, No. 61, March 30, 1973 1973.

_ , "Food Stamp Program Allotments," Federal Register, Vol. 38, No. 210, November 1, 1974 1974.

_ , "Food Stamp Program Allotments," Federal Register, Vol. 41, No. 27366, July 2, 1976 1976.

_ , "Food Stamp Program Allotments," Federal Register, Vol. 42, No. 58152, Nov 8, 1977 1978.

_ , "Food Stamp Program Allotments," Federal Register, Vol. 43, No. 54199, Nov 21, 1978 1979.

Frederick, Shane, George Loewenstein, and Ted O'donoghue, "Time discounting and time preference: A critical review," *Journal of economic literature*, 2002, *40* (2), 351–401.

Freyaldenhoven, Simon, Christian Hansen, and Jesse M. Shapiro, "Pre-Event Trends in the Panel Event-Study Design," *American Economic Review*, 2019, *109* (9), 3307–3338.

González-Fernández, Doris, Oviya Muralidharan, Paulo A Neves, and Zulfiqar A Bhutta, "Associations of maternal nutritional status and supplementation with fetal, newborn, and infant outcomes in low-income and middle-income settings: An overview of reviews," *Nutrients*, 2024, *16* (21), 3725.

Haller, Andreas and Stefan Staubli, "Measuring the Value of Disability Insurance from Take-Up Decisions," in "116th Annual Conference on Taxation" NTA 2023.

Hastings, Justine and Jesse M Shapiro, "How are SNAP benefits spent? Evidence from a retail panel," *American Economic Review*, 2018, *108* (12), 3493–3540.

Hendren, Nathaniel, "Measuring economic efficiency using inverse-optimum weights," *Journal of public Economics*, 2020, *187*, 104198.

_ and Ben Sprung-Keyser, "A unified welfare analysis of government policies," *The Quarterly journal of economics*, 2020, *135* (3), 1209–1318.

Herd, Pamela and Donald Moynihan, "Administrative Burdens in the Social Safety Net," *Journal of Economic Perspectives*, 2025, *39* (1), 129–150.

Hoover, Dael M., "Current Issues in the Food-Aid Programs," *Congressional Record*, 1974.

Hoynes, Hilary and Diane Whitmore Schanzenbach, "US food and nutrition programs," in "Economics of means-tested transfer programs in the United States, volume 1," University of Chicago Press, 2015, pp. 219–301.

_ , _ , and Douglas Almond, "Long-run impacts of childhood access to the safety net," *American Economic Review*, 2016, *106* (4), 903–934.

Hoynes, Hilary W and Diane Whitmore Schanzenbach, "Consumption responses to in-kind transfers: Evidence from the introduction of the food stamp program," *American Economic Journal: Applied Economics*, 2009, *1* (4), 109–139.

_ and _ , "Safety net investments in children," *Brookings Papers on Economic Activity*, 2018, pp. 89–150.

_ , Leslie McGranahan, and Diane W Schanzenbach, "SNAP and food consumption," 2015.

Hoynes, Hilary Williamson and Diane Whitmore Schanzenbach, "Work incentives and the food stamp program," *Journal of Public Economics*, 2012, *96* (1-2), 151–162.

Illinois Department of Public Aid, "Food Stamps Help You: Buy More Food, Pay Less Money," https://babel.hathitrust.org/cgi/pt?id=uiuo.ark:/13960/t2898gj3c&seq=6 1977. Accessed: 2026-03-27.

Kang, Zi Yang and Shoshana Vasserman, "Robustness measures for welfare analysis," *American Economic Review*, 2025, *115* (8), 2449–2487.

Keats, Emily C, Jai K Das, Rehana A Salam, Zohra S Lassi, Aamer Imdad, Robert E Black, and Zulfiqar A Bhutta, "Effective interventions to address maternal and child malnutrition: an update of the evidence," *The Lancet Child & Adolescent Health*, 2021, *5* (5), 367–384.

Landais, Camille and Johannes Spinnewijn, "The value of unemployment insurance," *The Review of Economic Studies*, 2021, *88* (6), 3041–3085.

Lieber, Ethan MJ and Lee M Lockwood, "Targeting with in-kind transfers: Evidence from Medicaid home care," *American Economic Review*, 2019, *109* (4), 1461–1485.

MacDonald, Maurice, *Food, stamps, and income maintenance*, New York: Academic Press, 1977.

McQuillan, Casey and Brendan Moore, "Incomplete and Endogenous Take-Up of Unemployment Insurance Benefits," *Available at SSRN 5692222*, 2025.

Moffitt, Robert, "An economic model of welfare stigma," *The American economic review*, 1983, *73* (5), 1023–1035.

—, "Estimating the value of an in-kind transfer: The case of food stamps," *Econometrica: Journal of the Econometric Society*, 1989, pp. 385–409.

Montpetit, Sébastien, Luisa Carrer, and Pierre-Loup Beauregard, "A welfare analysis of universal childcare: Lessons from a Canadian reform," 2025.

Morentz, Deborah S, "Effects of Recent Changes in the Food Stamp Program on Participation and Food Consumption of Participating Households in Tompkins County, New York," 1980.

Mullainathan, Sendhil and Eldar Shafir, *Scarcity: Why having too little means so much*, Macmillan, 2013.

Naik, Canishk, "Mental Health and the Targeting of Social Assistance," Technical Report, Working Paper 2025.

National Academies of Sciences, Engineering, and Medicine, *A roadmap to reducing child poverty*, Washington, DC: The National Academies Press, 2019.

National Cancer Institute, "SEER Population Databases," 2026.

National Center for Health Statistics, "Detail Natality," https://www.cdc.gov/nchs/data_access/vitalstatsonline.htm 1968-1987.

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

of Economic Analysis U.S. Department of Commerce, Bureau, "Regional Economic Information System (REIS)," https://apps.bea.gov/regional/downloadzip.html. Accessed: May 22, 2025.

Ohls, James C and Harold Beebout, *The food stamp program: Design tradeoffs, policy, and impacts*, The Urban Insitute, 1993.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman, "Distributional national accounts: methods and estimates for the United States," *The Quarterly Journal of Economics*, 2018, *133* (2), 553–609.

Pukelis, Kelsey, "SNAP Policies and Enrollment Following the COVID-19 Pandemic," in "2024 APPAM Fall Research Conference" APPAM 2024.

Rafkin, Charlie, Adam Solomon, and Evan Soltas, "Self-targeting in US transfer programs," *Available at SSRN 4495537*, 2025.

Reese, Robert B, J Gerald Feaster, and Garey B Perkins, "Bonus Food Stamps and Cash Income Supplements," *Marketing research report*, 1974.

Richardson, Joe, "A Concise History of the Food Stamp Program," Technical Report 79-244 EPW 1979.

Rosenfeld, Sam, "Fed by reform: Congressional politics, partisan change, and the food stamp program, 1961–1981," *Journal of Policy History*, 2010, *22* (4), 474–507.

Samuelson, Paul A, "The pure theory of public expenditure," *The review of economics and statistics*, 1954, pp. 387–389.

Schanzenbach, Diane, *Experimental estimates of the barriers to Food Stamp enrollment*, Institute for Research on Poverty, University of Wisconsin-Madison Wisconsin, 2009.

Shepard, Mark and Myles Wagner, "Do ordeals work for selection markets? Evidence from health insurance auto-enrollment," *American Economic Review*, 2025, *115* (3), 772–822.

Social Research Center, "Panel Study of Income Dynamics, public use dataset."

Spinnewijn, Johannes, "Heterogeneity, demand for insurance, and adverse selection," *American Economic Journal: Economic Policy*, 2017, *9* (1), 308–343.

Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of econometrics*, 2021, *225* (2), 175–199.

Tebaldi, Pietro, Alexander Torgovitsky, and Hanbin Yang, "Nonparametric estimates of demand in the california health insurance exchange," *Econometrica*, 2023, *91* (1), 107–146.

Thayer, Zaneta M, Julienne Rutherford, and Christopher W Kuzawa, "The maternal nutritional buffering model: an evolutionary framework for pregnancy nutritional intervention," *Evolution, medicine, and public health*, 2020, *2020* (1), 14–27.

Trippe, Carole and Daisy Ewell, "An analysis of cash food expenditures of food stamp households," *Washington, DC: Mathematica Policy Research*, 2007.

United States Food and Nutrition Service, *Food Stamp Facts: Allotments and Purchase Requirements*, Washington, DC: United States Government Printing Office, 1975.

U.S. Bureau of Labor Statistics, "Employment-Population Ratio," 2026. FRED Series EMRATIO.

U.S. Congress, "Food Stamp Act of 1970," Pub. L. No. 91-671, 84 Stat. 2048 (1971) 1970.

U.S. Department of Health and Human Services, "A method of imputing length of gestation on birth certificates," Technical Report (PHS) 82-1367, Public Health Service, Office of Health Research, Statistics, and Technology, National Center for Health Statistics 1982.

_ , "Prior HHS Poverty Guidelines and Federal Register References," 2024.

_ **, National Center for Health Statistics, and Centers for Disease Control and Prevention**, "Second National Health and Nutrition Examination Survey (NHANES II), 1976-1980."

U.S. Deptartment of Agriculture, *Food stamp statistical summary of project area operations report*, Washington, D.C.: Food and Nutrition Service, U.S. Dept. of Agriculture, 1980-1981.

U.S. Government Publishing Office, "7 C.F.R. § 271.3," Code of Federal Regulations, Title 7, Part 271, Section 3 1977.

_ , "7 C.F.R. § 271.3," Code of Federal Regulations, Title 7, Part 271, Section 3 1978.

_ , "7 C.F.R. § 273.9," Code of Federal Regulations, Title 7, Part 273, Section 9 1979.

_ , "7 C.F.R. § 273.9," Code of Federal Regulations, Title 7, Part 273, Section 9 1980.

U.S. Senate, "Congressional Record – Senate," Congressional Record September 1969. Debate on S. 2547, Food Stamp Act amendments.

USDA, Economic Research Service, "Supplemental nutrition assistance program (SNAP)–key statistics and research," 2025. Accessed 2026-05-18.

USDA Food and Nutrition Service, "A Short History of SNAP," https://www.fns.usda.gov/snap/history 2025. Accessed: May 22, 2025.

Whitmore, Diane, "What are food stamps worth?," Technical Report 2002.

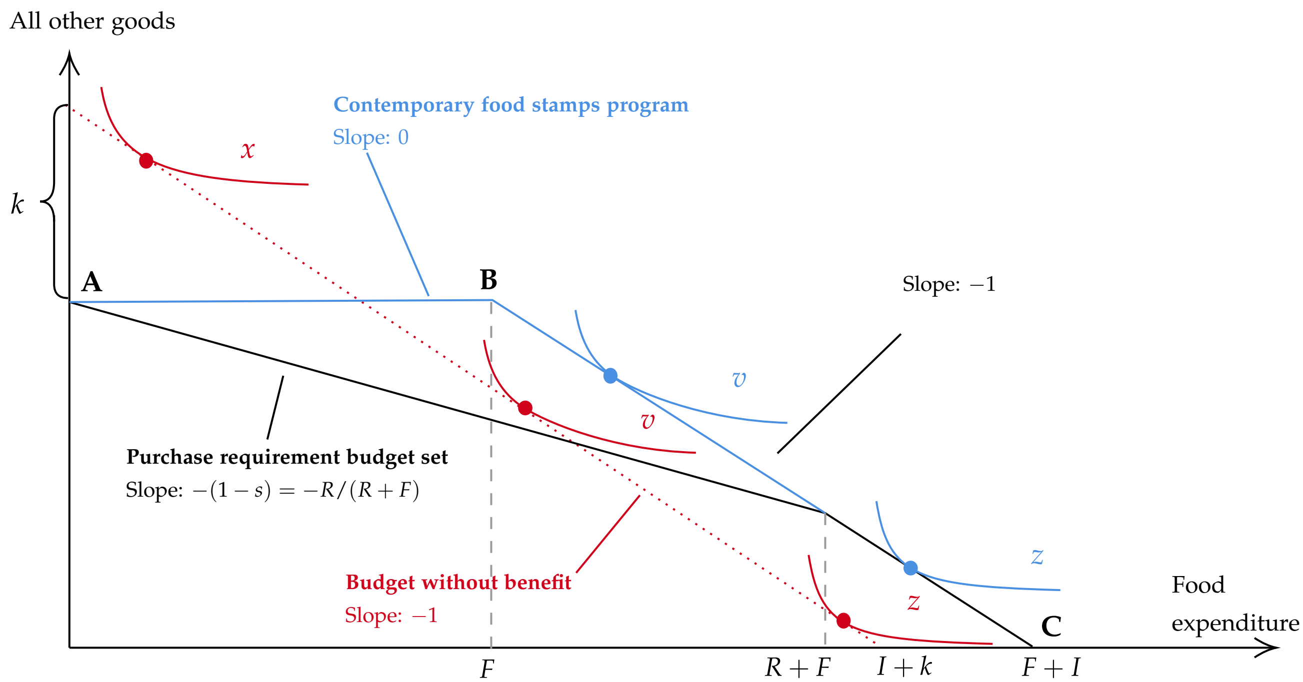
Wing, Coady, Seth M Freedman, and Alex Hollingsworth, "Stacked difference-in-differences," Technical Report, National Bureau of Economic Research 2024.

Wu, Derek and Bruce D Meyer, "Certification and Recertification in Welfare Programs: What Happens When Automation Goes Wrong?," Technical Report, National Bureau of Economic Research 2023.

Yang, Frank, Piotr Dworczak, and Mohammad Akbarpour, "Comparison of screening devices," *Available at SSRN 4456198*, 2025.

7 Figures

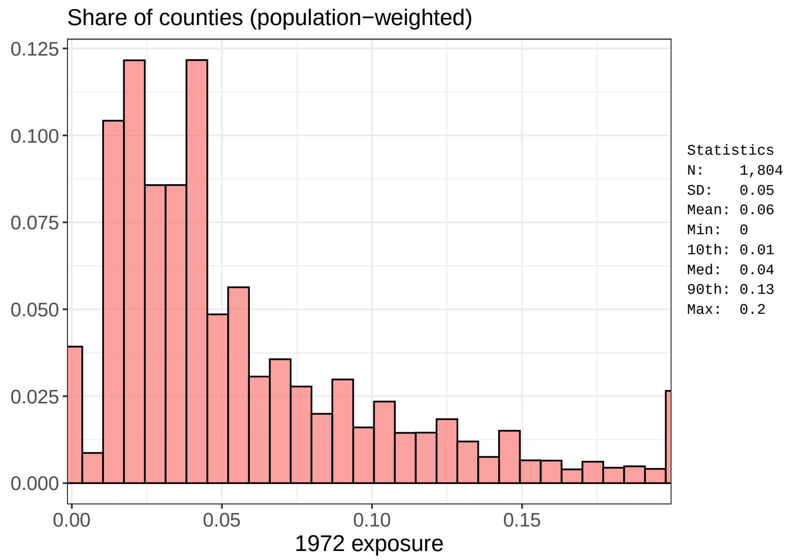
Figure 1: Impact of Eliminating the Purchase Requirement: Graphical Illustration



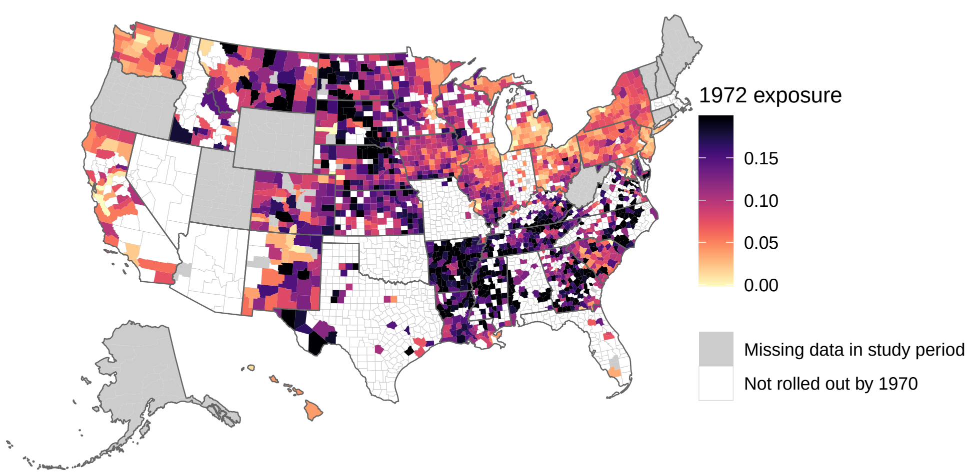
Note: Figure shows budget sets and hypothetical choices for those not enrolled in the food stamp program (red line), those enrolled under the pre-1979 purchase requirement (black line) and those enrolled under the current, free stamps program (blue line). F denotes the food stamp payment by the government (which is the same under the pre- and post-1979 reform) and R denotes the purchase requirement under the pre-1979 regime.

Figure 2: Variation in Exposure to Free Food Stamp Reform

(a)



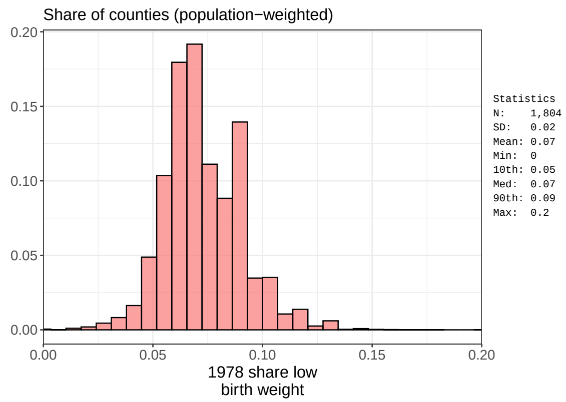
(b)



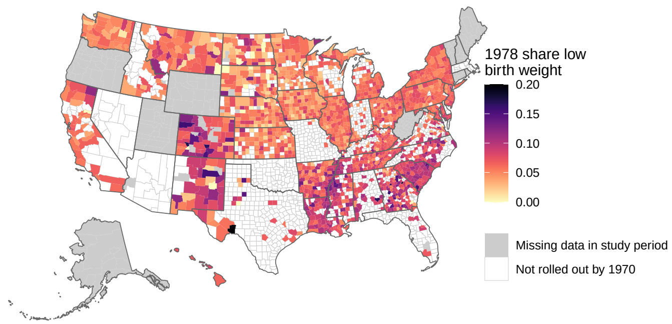
Notes: Figure shows a histogram of exposure across (population-weighted) counties (Panel a) and a heat map of exposure across counties (Panel b). Exposure is measured based on 1972 Food Stamp Program enrollment counts as in Equation 3; negative values are winsorized to zero. Sample is limited to the 1,804 counties in the baseline analysis sample for the free food stamp reform (see Table A1.)

Figure 3: Variation in 1978 County-level Share Low Birth Weight

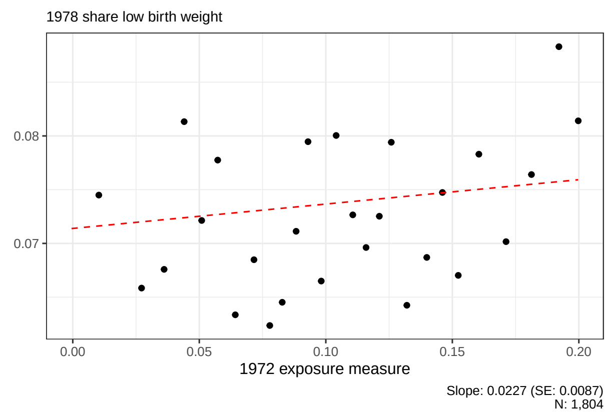
(a)



(b)



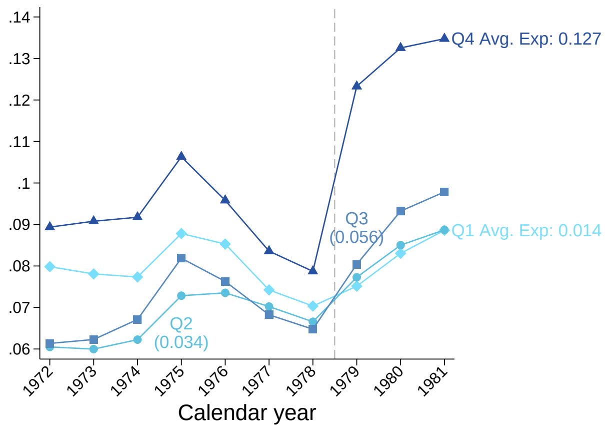
(c)



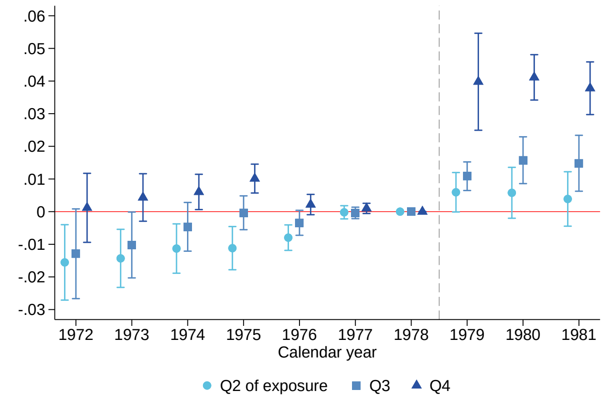
Notes: Figure describes the 1978 county-level share of births < 2,500 grams (low birth weight). Panel (a) shows a population-weighted histogram of share low birth weight, and Panel (b) shows a heat map of share low birth weight across counties. Panel (c) shows a binned scatterplot of 1978 share low birth weight against the 1972 exposure measure, weighted by 1970 county population. A linear fit illustrates the population-weighted correlation between the two county-level measures. Sample is limited to the baseline analysis sample for the free food stamps reform (see Table A1.)

Figure 4: Free Food Stamp Reform: Impact on Enrollment Rate

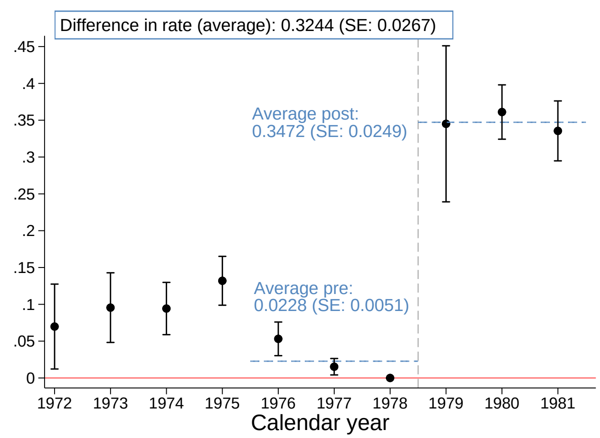
(a) Trends by Quartile of Exposure Measure



(b) Impacts by Quartile of Exposure Measure



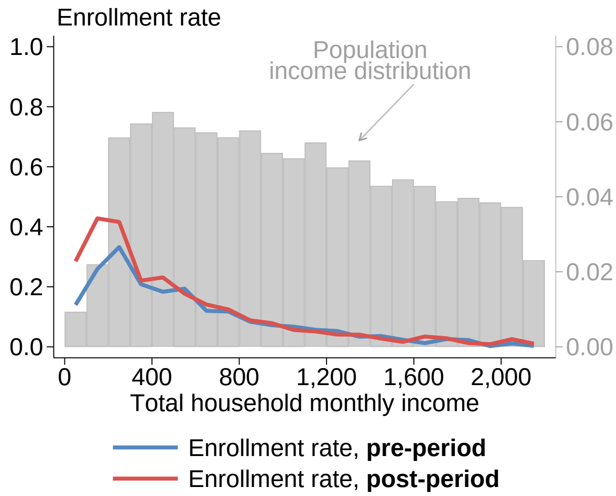
(c) Impact on Food Stamp Program Enrollment



Notes: Panel (a) shows 1970 population-weighted average county-level enrollment rate over time by population-weighted quartile of the exposure measure. Panel (b) shows the results of estimating the first stage Equation 4 using binary cuts of the 1972 exposure measure (where the lowest quartile of exposure comprise the comparison group) for Food Stamp Program enrollment. In Panels (a) and (b), darker shades in the series color scale correspond with increasing quartile of exposure. Panel (c) shows the results of estimating the first stage Equation 4 using the continuous exposure measure. Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level. N = 1,804 counties.

Figure 5: Free Food Stamp Reform Enrollee Incomes

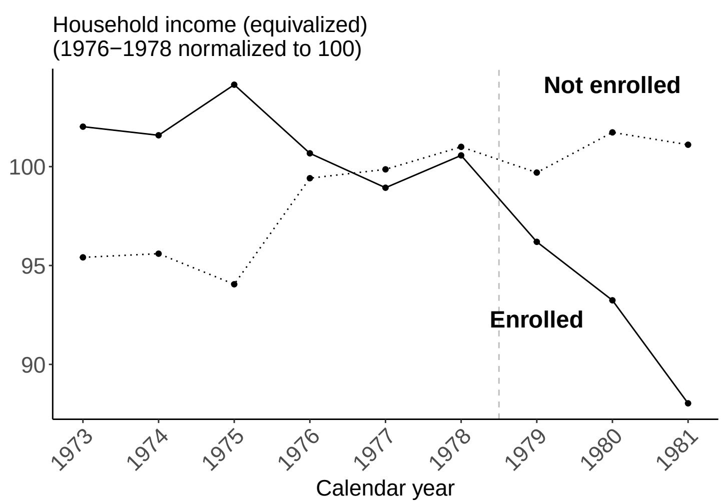
(a)



Percentiles of monthly income among food stamp enrollees:

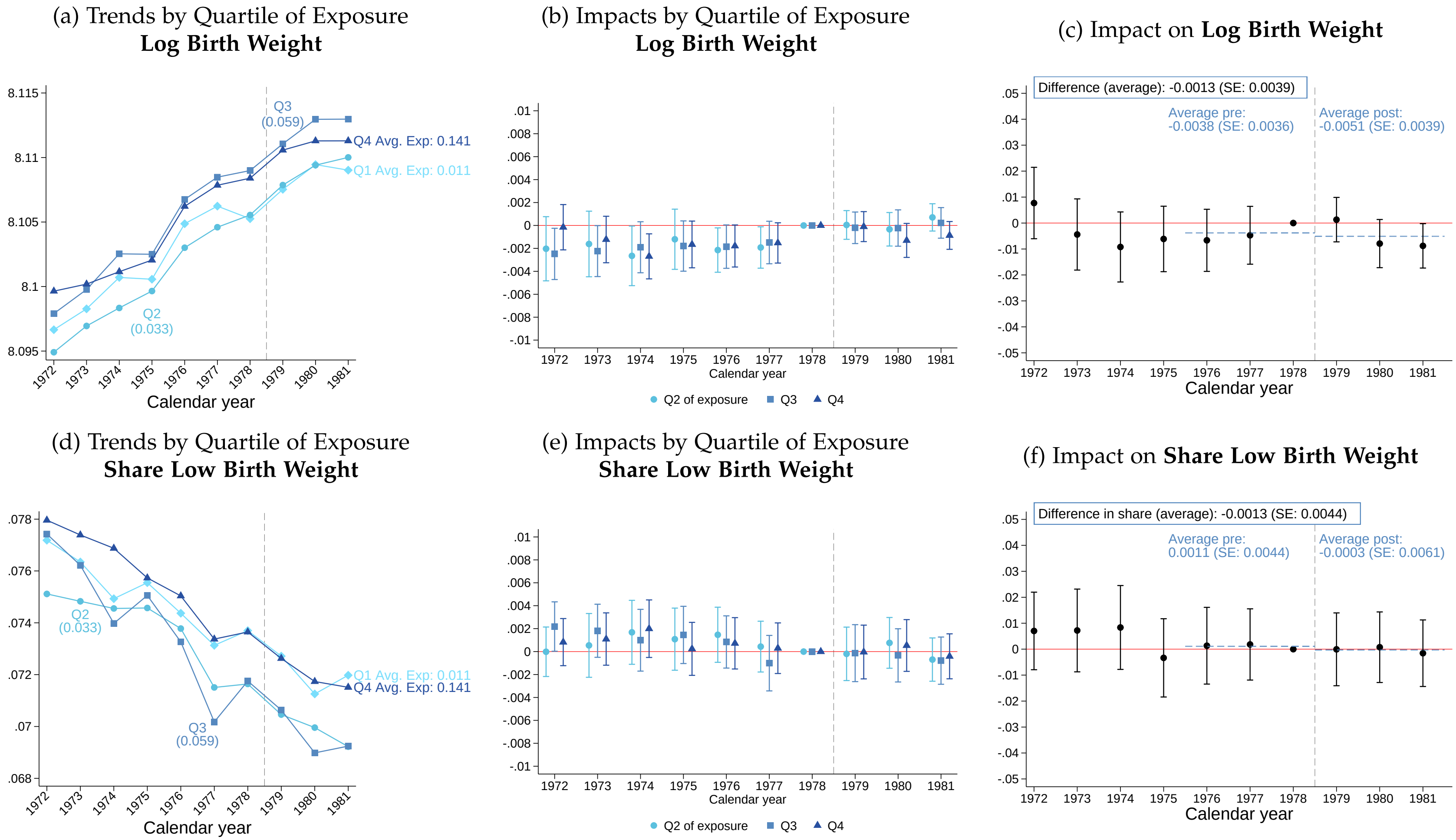
	Pre	*Post*
p25	$298	$261
p50	$501	$444
p75	$809	$751

(b)



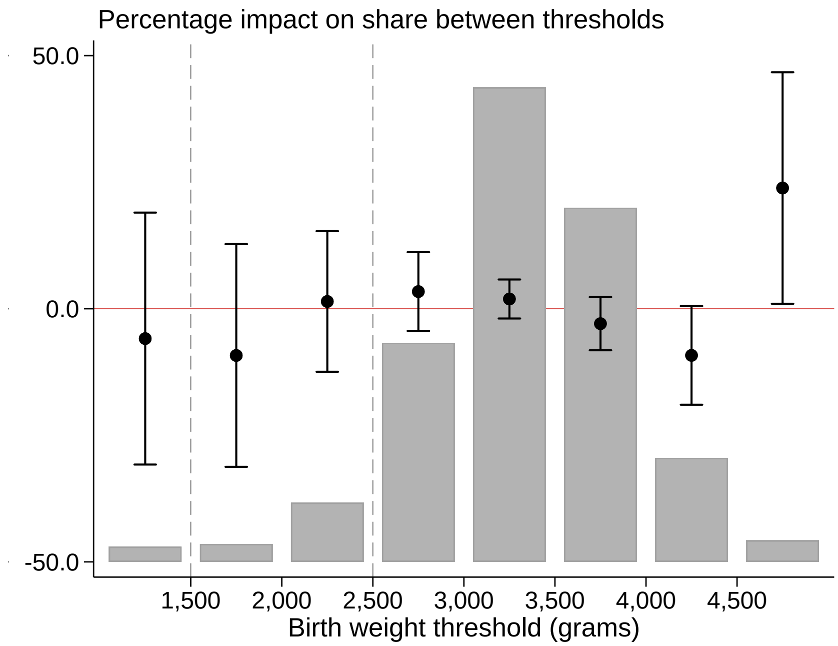
Notes: Panel (a) limits the PSID sample to households with income below the 99th percentile total family income of enrolled households across 1976–1981 (N = 29,990 household-years). The red and blue lines show food stamp enrollment rates in the pre and post period, respectively; specifically, they show the share of PSID household-years who are enrolled in food stamps in each $100-wide bin of total family income. The gray bars depict the share of *total* household-years (i.e., not limited to food stamp enrollees) from 1976–1981 in each total family monthly income bin. All analyses use PSID survey weights. Panel (b) presents average household income equivalized (i.e., adjusting for number of adults and children in the households) based on pooling observations from the PSID and NHANES. All income amounts are in $1979. See Appendix C.6 for details.

Figure 6: Free Food Stamp Reform: Impact on Infant Birth Weight



Notes: Panels (a) and (d) show 1970 population-weighted average county-level over time by population-weighted quartile of the exposure measure. Panels (b) and (e) show the results of estimating the reduced form Equation 5 using binary cuts of the 1972 exposure measure (where the lowest quartile of exposure comprise the comparison group) for Food Stamp Program enrollment. In panels, darker shades in the series color scale correspond with increasing quartile of exposure. Panels (c) and (f) show the results of estimating the reduced form Equation 5 using the continuous exposure measure. Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level. N = 1,804 counties.

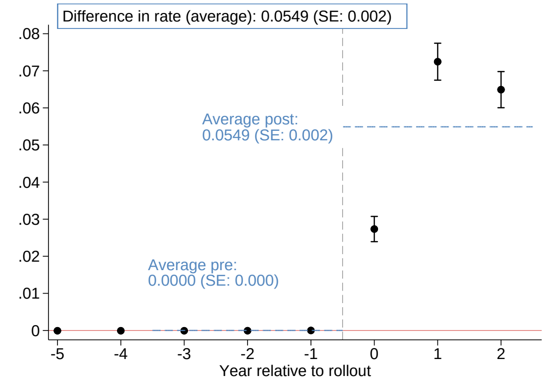
Figure 7: Free Food Stamp Reform: Impacts on Share of Births Within Birth Weight Ranges



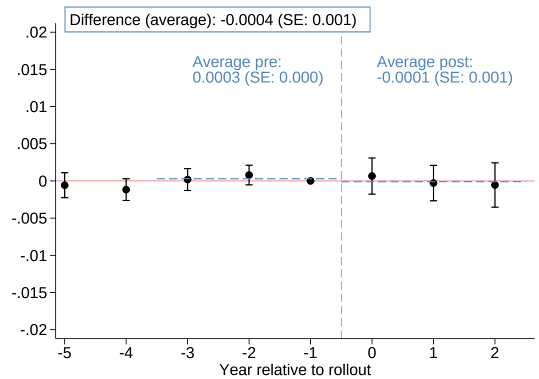
Notes: Figure shows scaled estimates of β_{avg} (as defined in Equation 6) from estimating Equation 5 where the outcome is the share of births that fall within a specified number of grams. Because the share of births in 1976-1978 in various birth weight ranges varies greatly (as shown in the light gray bars), we scale each estimate of β_{avg} by the share of births in 1976-1978 that fall within each range. We define the ranges as inclusive of the lower threshold and exclusive of the upper threshold. For example, we define the first outcome range as "$<$ 1,500 grams" followed by "$\geq$ 1,500 and $<$ 2,000." Similarly, we define the last range as "$\geq$ 4,500." Note that the vertical dashed lines denote the thresholds for very low birth weight ($<$ 1,5000 grams) and low birth weight ($<$ 2,500 grams). Regressions are weighted by 1970 county-level population and standard errors are clustered at the county level. N = 1,804 counties.

Figure 8: Impact of Food Stamps Program Rollout: Event Study

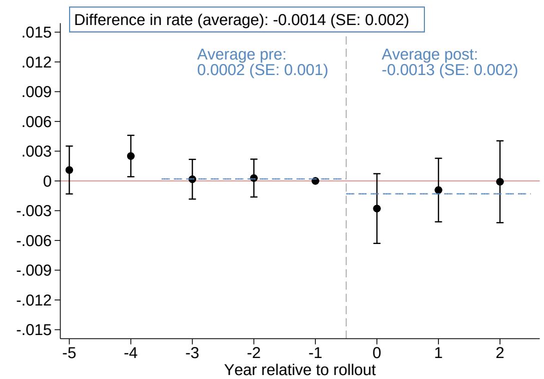
(a) Impact on Food Stamp Program Enrollment



(b) Impact on Log Birth Weight

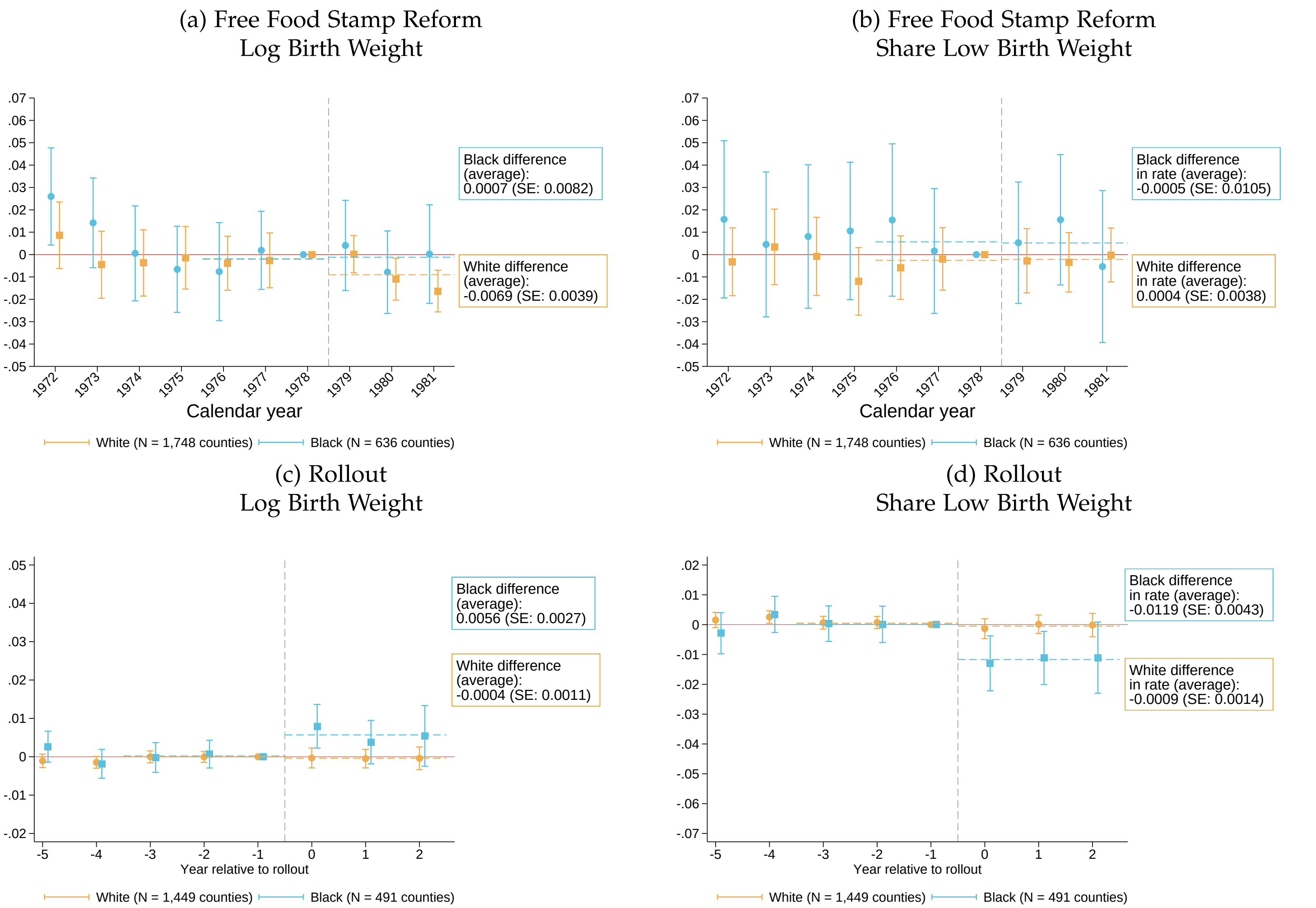


(c) Impact on Share Low Birth Weight



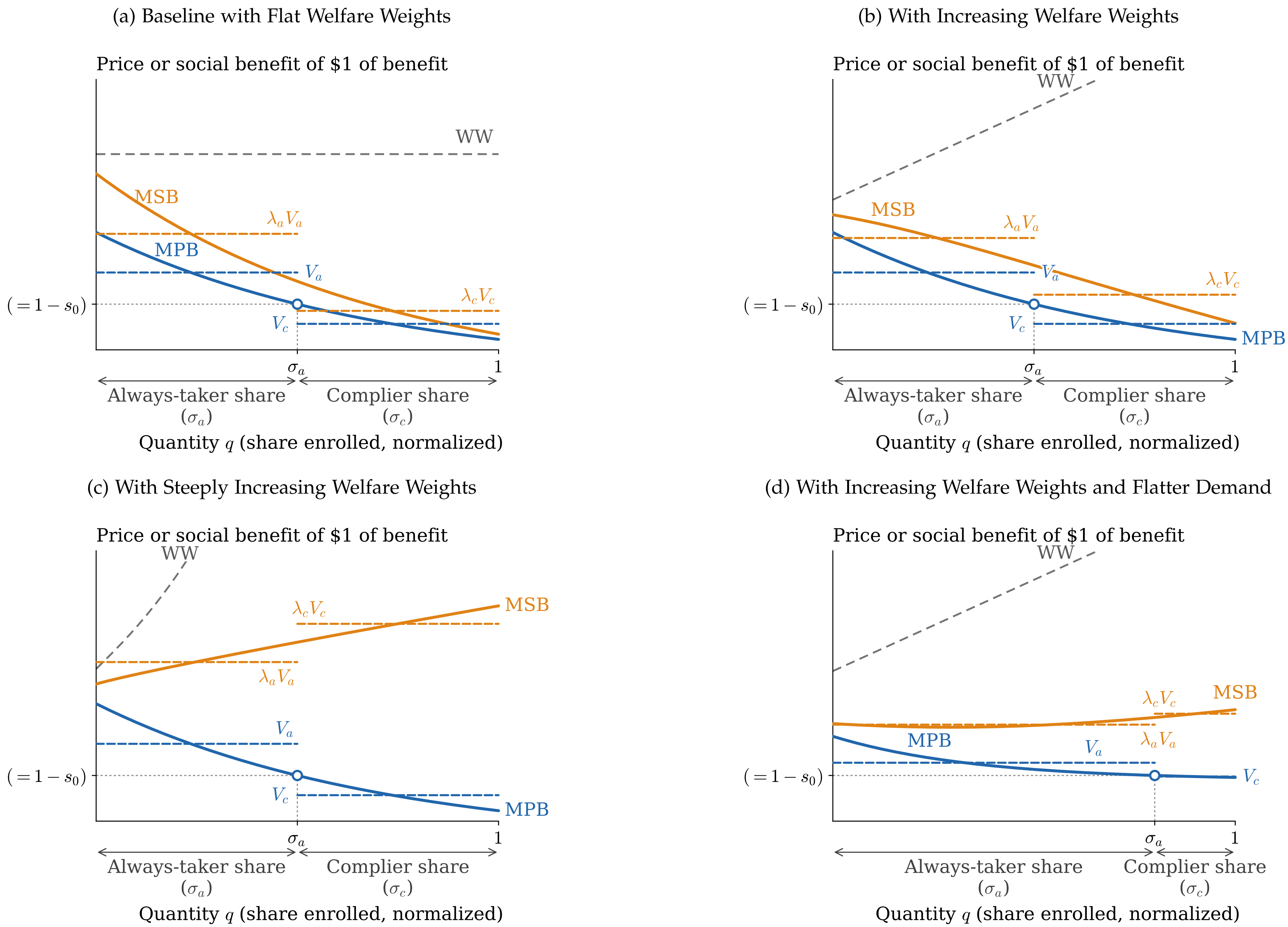
Notes: Figure shows estimates of β_r from Equation 7 in Panel (a) and from Equation 8 in Panels (b) and (c). The regressions are weighted by dataset-normalized weights which are described in Equation 22. Vertical lines denote 95% confidence intervals computed using standard errors clustered at the county level. Since Food Stamp Program enrollment is only reported annually, relative year corresponds to a calendar year for each county's rollout timing in Panel (a). Since birth weight data can be aggregated to the monthly level, relative year is constructed as the twelve months following each county's rollout timing and does not correspond to a calendar year in Panels (b) and (c). N = 1,520 counties.

Figure 9: Impacts on Infant Health Outcomes by Race



Notes: Panels (a) and (b) show the results of estimating the reduced form Equation 5 using a continuous 1972 exposure measure on infant birth weight outcomes, for white and Black births separately. Regressions for Panels (a) and (b) are weighted by 1970 county-race population, and standard errors are clustered at the county level. Panels (c) and (d) show estimates from Equation 8 using birth weight outcomes for white and Black births, separately. The regressions are weighted by dataset-normalized weights which are described above in Equation 22. From the baseline sample for each analysis, the sample in this figure keeps only county-races with 25 or more births in each year relevant to the analyses (1972–1981 for the free food stamps reform, 1970–1975 for rollout). Vertical lines denote 95% confidence intervals computed using standard errors clustered at the county-race level. Sample size, in terms of counties, is denoted in each panel legend.

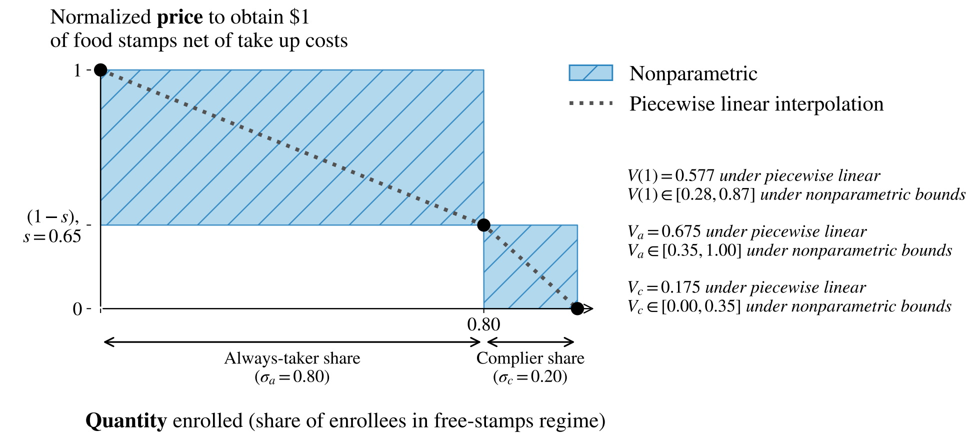
Figure 10: Illustration of Conceptual Framework



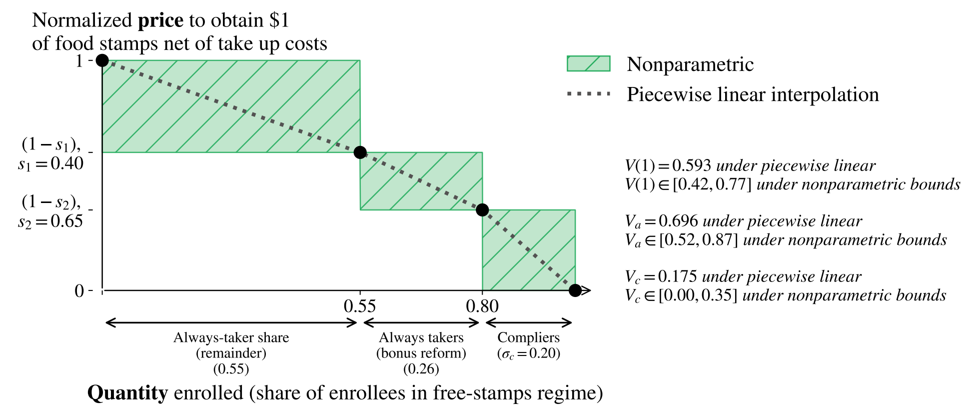
Note: Figures present the Marginal Private Benefit curve in solid blue (demand, corresponding to perceived private values of a fully subsidized transfer $V_\theta(1)$), Welfare Weights in gray (λ_θ), and Marginal Social Benefit curves in orange ($\lambda_\theta V_\theta(1)$). Types θ are sorted by demand for the in-kind good. The dashed horizontal blue and orange lines represent the average value of the curve between the line's endpoints. Comparing the average value of the orange line up to σ_a (the always-taker share) against the average value of the orange line from σ_a to 1 (the complier share) indicates whether moving to free transfer is desirable ($\lambda_a V_a(1) \lesseqgtr \lambda_c V_c$). Panel (a) presents a situation with flat welfare weights, such that λ_θ is constant. Panel (b) presents a situation with increasing welfare weights. Panel (c) presents a situation with even more steeply increasing welfare weights. Panel (d) returns to the situation in Panel (b), but presents a flatter demand curve.

Figure 11: The Value of Food Stamps: Demand Curves

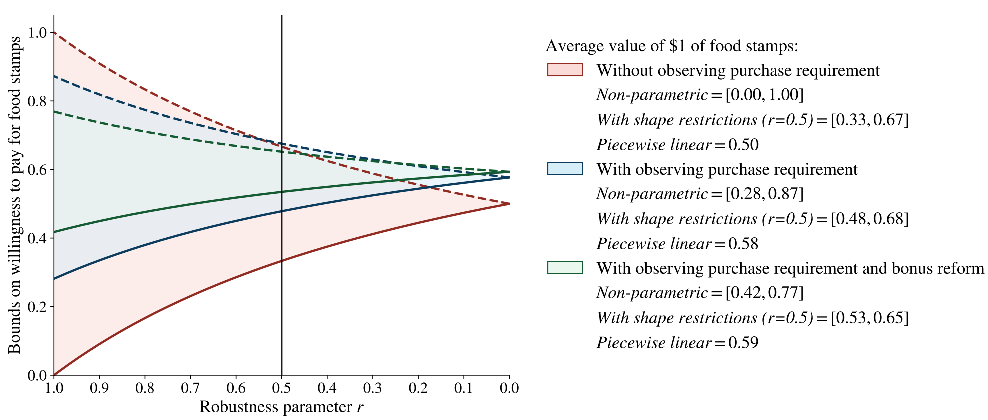
(a) With Purchase Requirement Only



(b) With Purchase Requirement and Bonus Reform

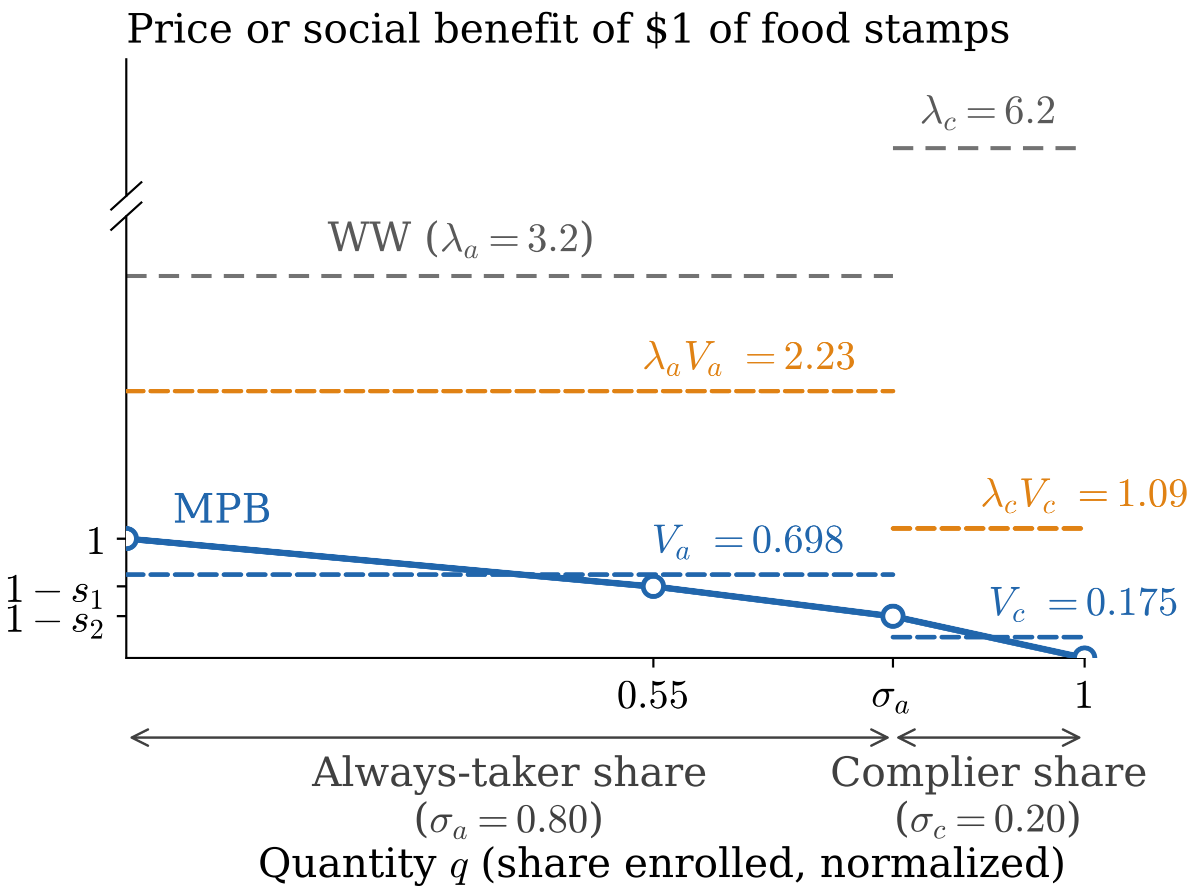


(c) Robustness to Parametric Restrictions



Note: Panels (a) and (b) show points on the demand curve for food stamps. The points $(0, 1)$ and $(1, 0)$ reflect the assumptions that no one would enroll in food stamps absent any subsidy, and that individuals observed enrolling must weakly prefer participation. The points with a subsidy less than 100% come from the take-up response to the elimination of the purchase requirement and the 1970 reform in the magnitude of the subsidy. We linearly interpolate between these points to form an estimate of the average value of food stamps among all enrollees. In Panel (c), we use the framework in Kang and Vasserman (2025) to consider robustness to parametric choices on the demand curve. When the robustness parameter $r = 1$, we show the fully nonparametric bounds on the demand curves under monotonicity. When $r = 0$, we show the linear interpolations. When $r \in (0, 1)$, demand curves whose slopes within each observed interval may deviate from the piecewise-linear demand slope by up to $r \times 100$ percent in either direction (Equation 19).

Figure 12: The Welfare Value of Free Stamps vs. Cost Sharing: Empirical Version of Figure 10



Note: Figure presents an empirical version of Figure 10. The solid blue line represents the Marginal Private Benefit (demand curve, V_θ). The dashed blue line represents the average value of the Marginal Private Benefit curve between the horizontal line's endpoints. The gray dashed line represents the Welfare Weights curve (λ_θ). The orange dashed line represents the average value of the Marginal Social Benefit curve obtained by multiplying the blue and gray lines.

Table 1: Household Characteristics of Food Stamp Enrollees

Variable	(1) Pre-period (always takers) (1976-1978)	(2) Post-period average (1979-1981)	(3) Complier average (Estimated from Agg FS)
Average annual household income ($1979)	8,187	7,285	3,594
	(189)	(163)	(1,135)
Equivalized income	4,108	3,797	2,525
	(76)	(78)	(487)
Share below 50% poverty level	0.066	0.129	0.389
	(0.009)	(0.008)	(0.066)
Share below 25% poverty level, given below 50%	0.160	0.233	0.535
	(0.038)	(0.030)	(0.217)
Social marginal welfare weight	3.220	3.801	6.179
	(0.077)	(0.072)	(0.528)
Households (N)	3,957	3,599	

Notes: Table reports characteristics of households enrolled in food stamps in the pre-period (Column 1) and the post period (Column 2). Column 3 reports mean outcomes among compliers based on the estimated complier share of 0.20. Standard errors are shown in parentheses; in Column 3 they are calculated using the delta method, assuming zero covariance between the pre-reform mean, the post-reform mean, and the estimate of the complier share (whose estimation error affects the estimated standard error). Data are from the NHANES and PSID. Annual household income indicates the total annual income of the household (including transfer income) in 1979 dollars. Equivalized income adjusts for the number of adults and children in the household. Social marginal welfare weights are defined as the inverse cube of winsorized (to the 5th percentile) equivalized income, divided by the weighted mean of this inverse cube all household-years 1976-1981. See Appendix C.6 for additional details. Note that "Share below 25% poverty level, given below 50%" reports a statistic for a subset of the reported count of households. Specifically, N = 548 households in the pre-period and N = 621 households in the post-period.

Table 2: IV Analysis of Food Stamp Participation on Birth Weight

		Ln(birthweight)			Share low birthweight		
	(1) First stage	(2) Rollout	(3) Free FS Reform	(4) Pooled	(5) Rollout	(6) Free FS Reform	(7) Pooled
Panel A: All racegroups							
Post rollout	0.0523						
	(0.0015)						
Post free FS reform	0.0185						
	(0.0015)						
Participation rate		0.0062	-0.0041	-0.0034	-0.0351	-0.0041	-0.0062
		(0.0278)	(0.0120)	(0.0114)	(0.0370)	(0.0134)	(0.0127)
First-stage F		1008.5	147.9	575.8	1008.5	147.9	575.8
p-value (joint test of instr.)	<0.0001						
Hansen J p-value				0.732			0.429
Baseline Mean	0.0253	8.102	8.106	8.104	0.0732	0.073	0.073
Observations	3351086	3340081	10824	3350905	3340081	10824	3350905
Counties	2853	1520	1804	2853	1520	1804	2853
Panel A: Black sample							
Post rollout	0.1353						
	(0.0046)						
Post free FS reform	0.0421						
	(0.0050)						
Participation rate (imputed)		0.0512	0.0123	0.0154	-0.0949	-0.0103	-0.0171
		(0.0263)	(0.0120)	(0.0111)	(0.0386)	(0.0175)	(0.0163)
First-stage F		733.9	69.9	401.8	733.9	69.9	401.8
p-value (joint test of instr.)	<0.0001						
Hansen J p-value				0.176			0.044
Baseline Mean	0.0746	8.033	8.040	8.036	0.1324	0.128	0.131
Observations	1000367	996411	3816	1000227	996411	3816	1000227
Counties	998	491	636	998	491	636	998
Panel B: White sample							
Post rollout	0.0250						
	(0.0006)						
Post free FS reform	0.0075						
	(0.0010)						
Participation rate (imputed)		-0.0260	-0.0518	-0.0495	-0.0286	0.0055	0.0024
		(0.0642)	(0.0309)	(0.0286)	(0.0824)	(0.0285)	(0.0270)
First-stage F		1790.3	58.7	921.4	1790.3	58.7	921.4
p-value (joint test of instr.)	<0.0001						
Hansen J p-value				0.716			0.696
Baseline Mean	0.0173	8.110	8.122	8.114	0.0664	0.060	0.064
Observations	3224153	3213482	10488	3223970	3213482	10488	3223970
Counties	2752	1449	1748	2752	1449	1748	2752

Notes: Sample includes relative years -3 to +2 for Food Stamp Program rollout and years 1976-1981 for the 1979 reform. Table shows precision-weighted estimates of δ from Equations 12 and 10 in Column 1 for both races pooled, then each race-specific sample, using a dataset which pools the free food stamp reform and rollout datasets. Columns 4 and 7 show precision-weighted estimates of λ from Equations 11 and 9 for both races pooled, then each race-specific sample, using a dataset which pools the free food stamp reform and rollout datasets. Columns 3 and 6 presents λ from Equation 11 for the flat free food stamp reform datasets alone, and then Columns 2 and 5, λ from Equation 9 using the stacked rollout datasets alone. The regressions are weighted by county population (county-race population for race-specific estimates) in 1970 for free food stamp reform county-years and stacked dataset weights according to Equation 22 for rollout county-years. The weights within each sample (rollout and free food stamp reform) have been re-scaled to sum to 1. All columns relay baseline estimates. Standard errors are clustered at the county level. Note that the "Post reform" impact is the exposure $*$ post coefficient re-scaled by the average 1972 exposure within the sample. The first-stage p-value reports the joint significance test of the excluded instruments. The Hansen J p-value reports the overidentification test of instrument exogeneity.

Appendices for Online Publication

Contents

A Food Stamp Program: History of Major Reforms

A.1 Early Program History

In 1939, the first version of the food stamp program took the form of a food surplus redistribution program to combat the effects of unemployment. Those receiving benefits purchased orange stamps to spend on their normal food purchases, and in addition to each $1 of orange stamps, they would receive 50 cents of blue stamps to spend on items deemed to be surplus. By 1943, World War II efforts eliminated the surpluses that had partially motivated the program, but over the four years of program activity, about 20 million people in the U.S. received benefits (Caswell and Yaktine, 2013). In February 1961, the Kennedy administration began piloting a series of food stamp programs across 43 counties and cities; this pilot program still required participants to purchase coupons to spend on food but no longer limited food purchases to surplus goods. In January 1964, President Johnson institutionalized the program through the Food Stamp Act of 1964 which delineated its purpose as improving nutrition for the poor. Although the funding and benefits for the program were set federally, the legislation left states to decide eligibility and operational schemes (USDA Food and Nutrition Service, 2025). Coupon allotments – i.e., the amount of coupons a family could purchase (or $R + F$ in the notation of Section 1) – varied by household size, income, and region of the country, with separate benefits schedules (for both the allotment ($R + F$) and the purchase requirement (R) for the North and the South (Richardson, 1979). Benefits were left largely unchanged for the remainder of the decade. In early 1969, the mean subsidy rate, in terms of bonus (F) as a share of the allotment ($F + R$), was 40% (Hoover, 1974).

Fueled by concerns of low program participation, benefits were increased, effective January 1970. One nationalized benefits schedule took the place of the region-specific ones, the purchase requirements were decreased, and the bonus was increased; as a result, the mean subsidy rate jumped to 63% (Hoover, 1974; Richardson, 1979). These changes to the benefits schedules were part of a larger set of amendments to the Food Stamp Act signed into law in 1971. The amendments introduced regular indexing of allotments to account for inflation and nationalized eligibility standards (except for Alaska, Hawaii, and territories) (Richardson, 1979), including establishing the categorical eligibility of households enrolled in public assistance programs (U.S. Congress, 1970). The amendments also prohibited the purchase requirement from exceeding 30% of income, and for the first time, the purchase requirement was eliminated for low-income households with monthly income below $30 (Richardson, 1979). Legislation in 1973 mandated that the Food Stamp Program be rolled out across all counties by July 1974. These amendments also increased the frequency of indexing food stamp allotments to twice per year and ensured access for certain groups, such as those in treatment or rehabilitation centers. The number of program participants continued to grow throughout the early 1970s, reaching 18.5 million people in 1976 (USDA Food and Nutrition Service, 2025).

A.2 The status quo prior to the 1979 reform.

By the mid-1970s, the program's structure was based on the goal that every household should have the opportunity to obtain a nutritionally adequate diet, and that households should not have to contribute more than about 30% of their own income towards monthly food consumption; the government would therefore make up the shortfall between the household's own income and the minimum cost of that nutritionally adequate diet (Congressional Budget Office, 1977; Richardson, 1979). The dollar value of that diet was called the household's monthly allotment; it was based on the U.S. Department of Agriculture's (USDA) Thrifty Food Plan, and was constant across household income but varied based on household size to account for economies of scale that allowed for bulk purchases in larger families. The allocation of the allotment ($F+R$) between the participants (via the purchase requirement R) and the government (via a transfer or "bonus" (F)) depended on household income, with the purchase requirement increasing in income and the government transfer decreasing in income (Richardson, 1979).

Participating households would pay the government $0 \leq r \leq R$ (where R denotes the maximum participant purchase allowed) for various values of r, and in return receive a dollar value of food stamps $f = r(R+F)/R$ from the government. The total dollar value of food stamps was split between the individual contribution (r) and the government contribution rF/R up to a maximum government transfer of F that would occur if the household paid the maximum purchase requirement ($r = R$). More specifically, households could choose $r \in (0, 0.25R, 0.5R, 0.75R, R)$ (Congressional Budget Office, 1977; Richardson, 1979). Households could purchase any of these shares of their monthly coupon allotment, up to the entire allotment, at one or multiple instances of issuance in that month (U.S. Senate, 1969). In practice, however, as discussed in Section 1, participating households essentially all chose the maximum allotment $R+F$. The food subsidy the household received varied with household income. Specifically, while as mentioned above, the total monthly allotment ($F+R$) - i.e., the dollar value of a nutritionally adequate diet - was the same across households of a given size, the split in the allotment between purchase requirement (R) and government transfer or bonus (F) varied with household income. The subsidy was 100 percent for the poorest households who had no purchase requirement ($R = 0$), and increased by about \$1 with every \$3 increase in income (Congressional Budget Office, 1977). [44] At the highest income eligibility level for a family of four the monthly purchase requirement (\$150) was more than five times higher than the monthly bonus (\$24), defined as the allotment (\$174) minus the purchase requirement (Appendix Figure A33).

[44] Transfers from other government programs such as Aid for Families with Dependent Children (AFDC) counted as income so that AFDC recipients - who were over two-fifths of household recipients in 1975 - still had to pay the purchase requirement (Congressional Budget Office, 1977).

To apply for food stamps, households contacted their local Department of Public Aid and provided documentation of income and household expenses including rent, utility bills, savings, and medical bills. Generally, the person who managed household money and did the shopping was encouraged to be the applicant, though provisions were in place to allow an authorized representative to carry out the application and purchasing. If the household was deemed eligible and approved, they received a document in the mail called "Authorization to Purchase" with their purchase requirement (R) and bonus (F) amounts, as well as the necessary identification to present in the grocery store. Households would take their documentation to designated sales locations like post offices, banks, or currency exchanges to purchase food stamps. There was an opportunity to order food stamps by mail for certain households (Illinois Department of Public Aid, 1977).

A.3 The 1979 reform: Elimination of the purchase requirement

The 1979 reform eliminated the purchase requirement for everyone (i.e., set $R = 0$ for all eligible households) but left the government transfer – i.e., the bonus F essentially unchanged; Appendix Figure A34, Panel (a) shows that the bonus increased annually (reflecting the inflation indexation put into effect in 1971), but that 1979 was no different in this respect than prior years.

As a result of the 1979 reform, the food stamp subsidy ($s = \frac{F}{R+F}$) became 100%. This entailed a larger increase in the subsidy for higher-income households, as shown in Appendix Figure A34, Panel (b). Prior to the free food stamp reform, the purchase requirement was increasing with income (Appendix Figure A33), so that the subsidy declined with income; the reform eliminated the purchase requirement for everyone (Appendix Figure A35), so that all eligible households received a full subsidy (100%).

Changes in Eligibility. Although the elimination of the purchase requirement was the centerpiece of the 1979 reform, it also altered eligibility rules (Berry, 1982; Rosenfeld, 2010). Specifically, it lowered the maximum net income level allowed for eligibility, and also reduced the deductions that a family could take in calculating net income; both changes reduced eligibility among higher income individuals. It also eliminated categorical eligibility for recipients of cash welfare (general welfare, AFDC, and SSI); after the 1979 reform, these individuals needed to apply to the food stamp program separately, which may have decreased enrollment among lower income individuals (Richardson, 1979).

We looked directly at the elimination of welfare program-based categorical eligibility. Categorical eligibility mainly came through participation in AFDC, SSI, and General Assistance. Evidence from the PSID suggests that this channel is quantitatively small. First, Appendix Figure A36 shows that the share of households participating in both the Food Stamp Program and AFDC is small — about 5%, compared with about 8% participating in the Food Stamp Program only — and remains flat over time. Second, in Appendix Figure A37 Panel (a), we classify households in every year using the pre-reform categorical eligibility concept: households that would have only been eligible for food stamps due to categorical eligibility under the pre-1979 rules. Among these households, food stamp participation remains stable and stays above 40% even after 1979. Thus, most households that would previously have qualified through categorical eligibility continued to obtain benefits after the free food stamp reform. Third, Appendix Figure A37 (Panel b) shows that the share of food stamp recipients who appear to qualify only through pre-1979 categorical eligibility falls after 1979, but only modestly. Overall, these facts suggest that removing categorical eligibility was not a major driver of the free food stamp reform's effects. If it were, one would expect much larger shifts in joint Food Stamp–AFDC participation or in the share of recipients appearing to rely solely on categorical eligibility.

Ordeals. The 1979 reform did not change the basic "ordeals" associated with receiving food stamps — individuals still had to go to the local food stamp offices to pick up stamps. The amendments did include several provisions aimed at increasing access among eligible individuals. For example, it allowed the use of mail or telephone applications, but in most cases, home visits were then often required for certifying a household's eligibility. The free food stamp reform required local stamps offices to provide outreach and nutrition education materials, as well as access to bilingual personnel and materials for applicants (Federal Register, 1978). While the timing and requirements of the elimination of the purchase requirement were uniformly applied at the federal level, the implementation of these non-benefit-related provisions were left to the states, and in most cases, specific service areas and certification offices (Federal Register, 1978). It is unclear how quickly these accessibility changes were implemented or what impact they had, if any.

B Food Purchases Relative to Food Stamp Benefits

Most food stamp enrollees are infra-marginal; in the absence of the food stamp program they would have purchased more food than their food stamp benefit F. In this section we first summarize the existing evidence behind this claim—both in the contemporary food stamp program and in the 1970s while the purchase requirement was in place. We then present additional evidence from analyses of food spending from the 1977 - 1980 Surveys of Food Consumption in Low Income Households.

B.1 Existing evidence of inframarginality

Contemporary free food stamps program. Estimates from the contemporary program suggest that 5 to 25 percent of enrollees would have spent less on food in the absence of the program than they received in food stamps (Trippe and Ewell, 2007; Econometrica, 2012; Hoynes et al., 2015; Hastings and Shapiro, 2018). Specifically, survey data of food expenditures for households on SNAP suggest that between 15 and 25 percent of participants are not inframarginal (Trippe and Ewell, 2007; Econometrica, 2012; Hoynes et al., 2015) while transaction-level data from a large US grocery retailer on SNAP-eligible spending by households who ever use SNAP in months that they are not on SNAP suggest that 5 percent are not infra-marginal (Hastings and Shapiro, 2018).

Estimates from the 1970s purchase requirement. Estimates from the time period of the purchase requirement suggest that only about 6 to 8 percent of enrollees would have spent less on food in the absence of the program than they received in food stamps (Reese et al., 1974; Clarkson, 1976; MacDonald, 1977; Moffitt, 1989).

B.2 Analyses of Surveys of Food Consumption (1977–80)

We combine survey data on the household food spending and demographic characteristics of food-stamp eligible households with benefit schedules from the Federal Register. This allows us to directly analyze food expenditures relative to food stamps benefits immediately prior to the elimination of the purchase requirement (1977–78) and immediately after it (1979–80).

Surveys of Food Consumption in Low Income Households The survey data consist of the 1977–78 and 1979–80 Surveys of Food Consumption in Low Income Households, which were part of the larger Nationwide Food Consumption Survey (NFCS). Each was a multistage, stratified probability sample of households in the 48 contiguous United States that were either receiving food stamps or eligible to receive food stamps between November 1977 and March 1978. The two surveys (henceforth collectively referred to as NFCS) contain information for 7,625 low-income households. Household-level survey weights are included.

We form monthly food expenditure measures from NFCS reports of the typical amount that they spent on food and non-alcoholic beverages at the grocery store or similar establishments (including purchases made with food stamps) over the three months prior to the survey.

Before the free stamps reform, households who were enrolled in food stamps were also asked whether they purchased the full amount of food stamps available to them in the prior month. If they did not, they were asked the fraction of food stamps that they did purchase (options were "One quarter," "One half," and "Three quarters").

Finally, for each household, the NFCS survey contains a binary indicator for whether the household was on food stamps at the time of their response, the number of individuals in the household, the household's income from a variety of sources (e.g., income from employment vs. social security), and information on the households expenses (e.g., rent and utilities).

Data on food stamp benefit schedules. Benefit schedules defined bonus amounts F_{ih} according to household size h and net income i and (during the time of the purchase requirement) the purchase requirement R_{ih}. Appendix Figures A33 and A35 provide examples. We use the benefits schedules published in the Federal Register from 1977 through 1980. Benefits schedules were updated one to two times each calendar year. For simplicity, we use a single schedule from each year: May 1977, May 1978, November 1979, and December 1980 (Food and Nutrition Service, United States Department of Agriculture, 1977, 1978, 1979, 1980).

Measuring F and R before and after the elimination of the purchase requirement Both before and after the free stamps reform, benefits schedules were based on the Thrifty Food Plan (TFP), a minimum monthly food budget needed to afford a nutritionally adequate diet. Transfer amounts (i.e., the bonus F) were designed to let households afford the TFP without spending more than 30% of their net income, and varied by household size. We denote TFP_{ht} to indicate the value of the TFP for a household of size h in year t.

During the time of the purchase requirement, benefits were defined such that a household of size h in year $t < 1979$ with net income i would buy a coupon allotment of food stamps worth $\overline{F}_{ht}$ from the government at a price of R_{iht}. (The value of the coupon allotment $\overline{F}_{ht}$ was determined by TFP_{ht}.) This resulted in a bonus $F_{iht} = \overline{F}_{ht} - R_{iht}$.

After the elimination of the purchase requirement, benefits were redefined such that a household of size h in year $t \geq 1979$ with net income i would directly receive a bonus $F_{iht} = \text{TFP}_{ht} - \lfloor 0.3 * i \rfloor$ (there was a minimum bonus of \$10 for eligible one- and two-person households). For some analyses, we will additionally define a *counterfactual* purchase requirement R^*_{iht} as the purchase requirement that the household would have had in 1978 (i.e., R_{ih1978}), inflation-adjusted back to the year that the household was observed in.

Net income, defined in the Code of Federal Regulations (U.S. Government Publishing Office, 1977, 1978, 1979, 1980), was the household's total gross income minus a set of deductions. Prior to 1979, the deductions were a percentage of employment income, federal and state income tax, medical expenses, dependent care expenses, unanticipated expenses, tuition and mandatory fees from educational institutions, court-ordered child support and alimony payments, and shelter costs (e.g., rent, utilities). After 1979, the deductions were simplified to a standard deduction, a percentage of earned income, dependent care expenses, and shelter costs. The NFCS contains both pre- and post-tax income, as well as information on dependent care expenses and shelter costs. Thus, we were able to partially implement the pre-1979 deductions and fully implement the post-1979 deductions. Notably, fully implementing the deductions in one period but not the other may lead to imbalanced measurement error in net income between the two periods. Therefore, we use the pre-period deductions to estimate net income in all years. Then, we use this estimated net income to assign benefit amounts to each NFCS household.

Defining and measuring inframarginality. We define two kinds of inframarginality. A household is considered *strongly inframarginal* if it spends more than $R_{iht} + F_{iht}$ (or, after the elimination of the purchase requirement, $R^*_{iht} + F_{iht}$) on food. It is considered *weakly inframarginal* if it spends more than F_{iht} on food. Strong inframarginality thus implies weak inframarginality but the converse is not true.

Our estimate of net income for determining R_{iht}, R^*_{iht}, and F_{iht} likely comes with unavoidable measurement error inherent to the NFCS data. This error propagates into our estimation of "weak inframarginality" (i.e., share with food spending more than F_{iht}) and into our estimation of "strong inframarginality" after the elimination of the purchase requirement (i.e., share with food spending more than $R^*_{iht} + F_{iht}$). However, because $\overline{F}_{ht} = R_{iht} + F_{iht}$ depended only on household size, this error does not affect our assessment of strong inframarginality under the purchase requirement.

Results. Several pieces of evidence suggest that households are inframarginal. First, under the purchase requirement, food stamps participants had the option to purchase either the full amount or a fraction of the food stamps available to them in a given month. Specifically, participants who chose to purchase a fraction could purchase either one quarter, one half, or three quarters of their allotment for the corresponding fraction of their purchase requirement. Appendix Figure A47 shows the distribution of the percentage of food stamps purchased for households enrolled in food stamps. The vast majority of food stamps enrollees purchased their entire coupon allotment. This is consistent with pre-1979 food stamp enrollees all being strongly infra-marginal.

A second piece of evidence on inframarginality comes from food expenditure patterns of enrollees. Appendix Figure A48 shows the distribution of monthly food expenditures relative to the household-specific $R + F$ for enrollees in the pre-reform period (top panel) and the post-reform period (bottom panel). Specifically, the bottom panel shows the distribution of monthly food expenditures relative to the household specific $R + F$ in the post-reform period based on the *counterfactual* R^* equal to the TFP minus the household's bonus F.

Appendix Table A10 combines these pre-period and post-period means with standard complier algebra to infer characteristics of always-takers (Column 1) and compliers (Column 3). The first two rows report the share who are weakly inframarginal (food spending at least F) and the share who are strongly inframarginal (food spending at least $R + F$). We also report the estimates using a cutoff of 90% of the estimand (F or $R + F$) as well as 110%. This is motivated by two factors.

First, our estimate of food expenditures and of food stamp bonuses F (which depend on income) naturally come with (potentially substantial) measurement error. Second, if there is bunching at the kink of the non-linear budget set at $R + F$ in the pre-period or F in the post period, this could overstate the set of inframarginal enrollees.

The findings in Column 1 confirm the prior findings (discussed above) that prior to the elimination of the purchase requirement, most households enrolled in food stamps were (weakly) infra-marginal: 93% of households on food stamps had monthly food purchases in excess of their food stamp bonus F, with this estimate ranging from 88 to 95% as we vary the cutoff. In addition, it suggests that most of these households were also *strongly* infra marginal: 63% of households on food stamps had monthly food purchases in excess of their total allotment $R + F$, with this estimate ranging from 49 to 83 percent as we vary the cutoff. Combined with the evidence in Appendix Figure A47 that most households enrolled in food stamps purchased their entire food stamp allotment, we view this evidence as consistent with most of the pre-reform enrollees being not only weakly but also strongly infra-marginal.

Column 3 shows similar calculations for compliers, using standard complier algebra, together with our estimated complier share of 0.20 (see Section 3.2). The results suggest that 45% of compliers were weakly inframarginal, with that estimate ranging from 44 to 57% as we vary the cutoff. We also estimate that the share of strongly inframarginal compliers is quite low: -0.21 (0.20). The point estimate is below zero, and the 95% confidence interval excludes values above 0.18. The implication is that compliers could value food stamps below $1-per-dollar transferred, since food stamps distort their purchases toward more food consumption than otherwise. A caveat is that the calculation of the counterfactual purchase requirement R^* introduces another source of potential measurement error.

C Empirical Appendix

C.1 Freyaldenhoven et al. (2019) Correction

To address possible confounds that could be correlated with both exposure to the free food stamp reform and with the enrollment rate or birth outcomes, we use the proxy approach in Freyaldenhoven et al. (2019). This requires taking an explicit stance on the potential unobserved confound and then identifying a proxy variable that follows the same time path as the confound but which is itself unaffected by the policy change.

As discussed in the main text, food stamp enrollment rates are highly counter-cyclical and since counties with different exposure have very different poverty rates, it is plausible that they may be experiencing different local business cycles. We therefore use as a proxy variable p_{ct} employment in industries without many workers on food stamps. Employment in such *unexposed* industries can proxy for the role of local business cycles but is plausibly unaffected by the purchase requirement reform (which could potentially affect employment and earnings among food-stamp eligible individuals (Hoynes and Schanzenbach, 2012). The Freyaldenhoven et al. (2019) procedure adjusts the focal variable — such as the food stamp enrollment rate or infant health — using leads of the proxy as instruments. Graphically, this procedure corresponds to subtracting the proxy from the focal variable. The identifying assumption is that absent the free food stamp reform, food stamp enrollment or infant health would move proportionally with employment in unexposed industries, and that the free food stamp reform did not affect unexposed industries' employment. This assumption is supported if pre-trends in enrollment or infant health follow the same time pattern of unexposed employment.

Construction of proxy variable. We proxy for local business cycles of employment with log employment in unexposed industries, defined as industries that likely have small shares of workers on food stamps. To construct this county-year measure (Proxy_{ct}), we use data from County Business Patterns, specifically the datasets containing employment counts by Eckert et al. (2020) and Eckert et al. (2022).[45] We first identify the industries likely to be highly exposed to the food stamp reform. We follow Dube et al. (2010) in identifying low-wage sectors as retail, food services, and accommodations, and we sum county-level employment across these sectors for each year. We also add employment in agriculture, forestry, and fisheries in our definition of low-wage sectors, because these are important low-wage sectors in the U.S. south. We then compute employment in unexposed industries as employment across all industries, less employment across these exposed, low-wage industries.[46]

[45] For confidentiality reasons, employment counts are masked for many county-industry-years. In Eckert et al. (2020), the authors describe the process they used to impute masked employment counts.

[46] For each county-year, we sum the employment counts across any specified low-wage sectors observed. This means that for some county-years, employment counts in some low-wage industries with missing data may treated as zero in

Implementation. Our second-stage equation relates the proxy variable to outcomes of interest (either food stamp enrollment or infant health):

$$y_{ct} = \lambda \text{Proxy}_{ct} + \sum_{\tau \neq 1978,1975} \beta_\tau (\mathbf{1}\{t=\tau\} \times \text{Exposure}_c^z) + \alpha_c + \gamma_t + \varepsilon_{ct} \tag{20}$$

which we estimate by instrumenting for Proxy_{ct} with the following first stage equation:

$$\text{Proxy}_{ct} = \pi \left(\text{Post}_{t+1}^s \times \text{Exposure}_c^z\right) + \sum_{\tau \neq 1978,1975} \phi_\tau \left(\mathbf{1}\{t=\tau\} \times \text{Exposure}_c^z\right) + \alpha_c^1 + \gamma_t^1 + u_{ct} \tag{21}$$

The excluded instrument is the lead of the policy variable defined as $\text{Post}_{t+1}^s \times \text{Exposure}_c^z$, where Post_t^s is an indicator that is 1 if calendar year t is 1979 or later. We normalize both 1978 and 1975 (relative years -1, -4) to be equal to each other and zero. For both of the above equations, we cluster at the county level and weight by 1970 county population.

Appendix Figure A38 presents the results, following the template in Freyaldenhoven et al. (2019). Panel (a) plots the event study in the proxy variable, running the same regression as in Equation 5 with log count employed in unexposed industries (i.e. Proxy_{ct}) as the outcome variable. Consistent with a recovering macroeconomy in more exposed (i.e., higher poverty) counties, employment rises in the years leading up to 1976. Because food stamp enrollment is countercyclical, this force could plausibly increase enrollment in these years and thus generate the 1972–76 pre-trends in Figure 4.

Panel (b) runs the same event study as in Panel (a) but normalizes the δ_τ event study coefficients shown in Panel (a) (based on estimating Equation 5), and compares them (in green) to the δ_τ coefficients shown in Figure 4c (shown in blue). This process involves a normalization, so that the $\delta_{\tilde{\tau}}$ coefficient corresponding to 1975 when Equation 5 is estimated for the outcome Proxy_{ct} equals the $\delta_{\tilde{\tau}}$ coefficient corresponding to 1975 when Equation 5 is estimated for the outcome FS_{ct} (i.e. the food stamp enrollment rate). Graphically, this corresponds to "flipping" the series in Panel (a) and "stretching" it, pinning the series so the green series and blue series in Panel (b) are equal at 1975.[47] The figure shows that the proxy closely tracks the pre-1979 pattern in enrollment, but not the enrollment jump in 1979.

Finally, Panel (c) shows our adjusted first-stage estimates. Graphically, the adjusted series in black is created by subtracting the rescaled proxy from the event study coefficients (green series, Panel (b)) estimated for food stamp enrollment (blue series, Panel (b)). This adjustment eliminates the pre-trend while leaving the post-period effect largely unchanged. The adjusted first-stage

the calculation of unexposed employment used as a proxy measure. This differs from the construction of the exposed industry employment outcome used as an outcome (not a proxy) in Appendix Figures A41 and A42 where we require employment be observed across all low-wage sectors.

[47] In the Freyaldenhoven et al. (2019) framework, the researcher must choose a period to line up the proxy and outcome; this period is then omitted in adjusted estimates. We normalize the series so that the proxy and outcome coincide in 1975, the year in which the pre-trend is largest.

estimate in Panel (c) is 0.3141 (s.e. 0.050), which is economically and statistically indistinguishable from the unadjusted first-stage estimate of 0.3257 (s.e. 0.027) in Figure 4. In sum, macroeconomic forces account for the modest pre-trend in the first stage, but not for the jump in 1979.

We perform the same analysis using the unexposed employment proxy for birth weight outcomes and additional infant health outcomes in Panel (b) of Appendix Figures A11 through A17 and find minimal change to the reduced form estimates.

C.2 Empirical Analysis of the Rollout: Additional Details

Stacked Dataset Weights. The stacked difference-in-differences weights described by Wing et al. (2024) aim to make control/treatment observations comparable across datasets so that each dataset contributes to overall estimates proportionally to its share of treated/control.

Control and treatment weights, at the dataset d and relative year r level, are constructed as:

$$w_{dr} = \begin{cases} \dfrac{Pop_d^{\text{treat}} / Pop_{\text{tot}}^{\text{treat}}}{Pop_{dr}^{\text{treat}} / Pop_r^{\text{treat}}}, & \text{if county–time is treated,} \\ \dfrac{Pop_d^{\text{treat}} / Pop_{\text{tot}}^{\text{treat}}}{Pop_{dr}^{\text{ctrl}} / Pop_r^{\text{ctrl}}}, & \text{if county–time is control.} \end{cases} \tag{22}$$

The subscripts tot denote the total across relative years.

Wing et al. (2024) show that these weights produce a population weighted ATT. Our specific weights differ slightly from Wing et al. (2024) (Table 1, Row 2) because we incorporate both population weights and the stacking weights, so we use populations of treated or control in the denominator.

Two-way Fixed Effects. The canonical way to estimate the impact of the rollout is via two-way fixed effects (TWFE). For counties c, years t, months m, and relative year r, we estimate the first stage and reduced form:

$$x_{ct} = \alpha_c + \tau_t + \sum_{\tau \neq -1} \beta_\tau (\mathbf{1}\{r = \tau\}) + \epsilon_{ct} \tag{23}$$

$$y_{cm} = \alpha_c + \tau_m + \sum_{\tau \neq -1} \beta_\tau (\mathbf{1}\{r = \tau\}) + \epsilon_{cm} \tag{24}$$

where x_{ctr} is a measure of the food stamp program (e.g., enrollment rate); y_{cmr} is an outcome (e.g., birth weight); α_c is a county fixed effect (or county-by-race) and τ_t, τ_m is a year- or year-month fixed effect; and P_{cr} is an indicator if the food stamp program was in effect in county c at relative period r.

Since all counties are eventually treated, there is a standard time-period-cohort multi-collinearity problem. To identify relative time indicators β_r, for the TWFE estimator we adopt the practice of binning these indicators at +5 and –5. Economically, this restriction imposes that coefficients are constant far from the 1979 free food stamp reform.

To address concerns about treatment effect heterogeneity in staggered rollout designs (Callaway and Sant'Anna, 2021; Sun and Abraham, 2021), we construct and stack datasets d for groups of counties sharing the same rollout date $r = 0$. This approach follows Cengiz et al. (2019) and has recent econometric justification in Wing et al. (2024).

C.3 Comparison to Almond et al. (2011).

Overview. Almond et al. (2011) (hereafter AHS) estimate the effect of the county-level rollout of the food stamp program on birth weight. They find that the rollout increased birth weight by 2.64 grams (standard error = 0.896) for whites and 4.12 grams (standard error = 2.32) for Blacks, corresponding to an increase of roughly 0.08% and 0.13% of mean birth weight, respectively. They also estimate that rollout reduced the share of low birth weight births by 0.06 percentage points (standard error = 0.03) for whites and 0.16 percentage points (standard error = 0.1) for Blacks.[48] They do not report aveage impacts across both races, which is our focus.

Our re-analysis of the impact of the rollout on birth weight by race qualitatively replicates AHS when we adopt their specifications. Moreover, like AHS, our specifications tend to find smaller magnitudes of impacts on whites than Blacks. That said, statistical significance and magnitudes of our by-race estimates differ from AHS. We therefore investigated why we reach these different findings, starting from our replication of their results and toggling changes in researcher choices until we arrive at our estimates.

Our re-analysis of the impact of the rollout on birth weight largely follows AHS, but differs in some implementation choices and also incorporates some advances in the econometrics of event studies that have occurred since their paper was published. Specifically, our approach differs from theirs in the following main ways:

- AHS use a flat panel dataset of county-quarters and a two-way fixed effects estimation strategy. We use a stacked dataset estimation strategy.
- AHS often include time-varying county-level controls and/or linear time trends in county-level, time-invariant characteristics in their specifications, while we do not.
- Our samples differ slightly. AHS use all county-quarters in the period 1968 through 1977, except Alaska. We limit the sample to a panel of counties that rolled out in July 1969 or later to ensure all counties in the sample have an observable pre-period in the birth weight data, which begins in 1968; AHS make a similar restriction for their event study analysis (their Figure 5).

[48] See Almond et al. (2011), Table 1, Columns 2 and 6.

- AHS keep county-race-quarters with at least 25 births; we keep county-race cells that have at least 25 births in every year included in the rollout analysis.
- AHS define births as treated if the food stamp program was in place three months prior to birth (to proxy for the food stamp program having been in place by the third trimester of gestation); we define all births after the food stamp program is in place as treated.
- AHS weight their regression analyses by birth counts in each county-race-quarter; we use the time-invarient stacked weights recommended by Wing et al. (2024). The weights that we use are not relevant for AHS's TWFE design, but end up being important when comparing differences in results.

From the analyses below, we conclude that differences between our race-specific findings and AHS's arise from a combination of choices involving controls, weights, and the use of a stacked dataset rather than flat panel TWFE design.

Successful Replication of AHS. To explore the consequences of these differences, we start in Tables A3 and A4 by using publicly available data and the replication package provided by AHS to replicate their original results for birth weight and share low birth weight respectively. Specifically, we implement AHS's estimating Equation 1 — which we reproduce below — separately (as they do) for white and Black births:

$$Y_{ct} = \alpha + \delta \text{FSP}_{ct} + \beta \left(\text{CB60}_c \times t \right) + \gamma X_{ct} + \eta_c + \delta_t + \varepsilon_{ct} \tag{25}$$

where Y_{ct} is an outcome (e.g., birth weight), η_c is a county fixed effect, and δ_t is a time fixed effect. CB60_c are 1960 county characteristics (specifically share Black population, share urban, share farmland, and logged population), which they interact with a linear time trend. The X_{ct} vector includes annual county-level controls from the Bureau of Economic Analysis' Regional Economic Information System (REIS), including real per capita transfers and the log of real annual county per capita income. The key variable of interest, following AHS, FSP_{ct} is an indicator that takes the value of 1 if the Food Stamp Program is in place at the beginning of the quarter prior to birth; they do this to proxy for program exposure in the third trimester of pregnancy. Our replication exercise makes the same sample restrictions and weighting choices as in AHS, including limiting to data on births in years 1968–1977 and weighting regressions by birth count.

Columns 1 through 4 of Tables A3 and A4 reproduce the estimates reported in AHS; Columns 5 through 8 report the results of our replication exercise. We successfully replicate AHS. Specifically, Column 1 estimates Equation 25 exactly, and Columns 2, 3, and 4 add state-specific linear time trends ($\lambda_s \times t$), state by year fixed effects (μ_{st}), and county-specific linear time trends ($\lambda_c \times t$), respectively and separately as AHS do in their Table 1. Our replication estimates closely (if not exactly) reproduce the original AHS estimates, deviating from their numbers by less than 20% (and usually by considerably less). One reason our replication attempt is not numerically identical may be small differences in cleaning and aggregating the birth certificate data. Alternatively, there may also be other differences in the use and availability of certain AHS control variables.

Comparison of our results and AHS. Having replicated AHS, we then proceed in Tables A5 and A6 to interrogate the role of the differences in samples and specifications. In both tables, Columns 1 through 4 report the results of our replication exercise shown in Columns 5 through 8 of Tables A3 and A4. Subsequent columns then toggle the changes described above. There are many different orders in which we could apply the various differences between our baseline and theirs; we choose one that seemed a natural, cumulative progression.

Column 5 uses the baseline AHS specification (Equation 25) but removes time-varying county-level covariates, leaving only the time and county fixed effects. We maintain the specification choice to exclude these covariates in all subsequent columns. Column 6 limits the sample to counties that rolled out in July 1969 or later and again maintains this restriction in all subsequent columns.

Starting in Column 7, we switch to the stacked dataset estimation approach that we use for the rollout (Section 4.1.1). We show the role of the different estimation approach in two steps. First, in Columns 7 and 8, we adapt the AHS estimating Equation 25 to a stacked dataset specification (with only county and time fixed effects):

$$y_{cmd} = \alpha_{cd} + \tau_{md} + \delta FSP_{cmd} + \epsilon_{cmd} \tag{26}$$

where y_{cmd} is an outcome (e.g., birth weight), α_{cd} is a county-by-dataset fixed effect, and τ_{md} is a unique month-by-dataset fixed effect. When we transition to the stacked data approach, we also switch from the quarter-year fixed effects used by AHS to the month-year-dataset fixed effects in our baseline specification. Until we switch to our treatment definition in Column 13, we keep AHS's treatment definition where $\text{FSP}_{cmd} = 1$ if the Food Stamp Program is in place at the beginning of the quarter prior to birth. Between Columns 7 and 8, we introduce baseline stacked dataset weights (per Wing et al. (2024), according to Equation 22).[49]

Then, in Columns 9–14 we estimate effects for each relative year in an equation reflecting the baseline reduced-form event-study equation (as in Equation 8):

$$y_{cmd} = \alpha_{cd} + \gamma_{md} + \sum_{\tau \neq -1} \beta_\tau (\mathbf{1}\{r = \tau\}) + \epsilon_{cmd} \tag{27}$$

where the terms are identical to that of Equation 26 except for the introduction of relative year

[49] The weights' relative year is constructed from the onset of treatment as defined by AHS, when $\text{FSP}_{ct} = 1$.

indicators in place of the post period program exposure indicator. We now include data from years 1968–78 and report a linear combination of relative year coefficients, constructed to estimate an average difference across relative years -3 through +2:

$$\beta_{\text{avg}} = \frac{1}{3}\left[(\beta_0 + \beta_1 + \beta_2) - (\beta_{-3} + \beta_{-2} + \underbrace{\beta_{-1}}_{\equiv 0})\right]. \tag{28}$$

For Columns 9–12, relative year is constructed from the onset of treatment as defined by AHS, i.e., from three months after initial program rollout. For the final two columns, relative year is defined relative to the actual month of program rollout as in our main analysis. Columns 9 and 10 estimate Equation 27 on a stacked dataset limited only by county-rollout timing. Columns 11 through 14 limit the counties included to reflect our baseline sample for the race-specific analyses.[50] Finally, Columns 13–14 define relative year as relative to month of Food Stamp Program rollout, instead of from third-trimester program exposure as defined by AHS. Between each pair of columns, we toggle between birth count weights (as in AHS) and the stacked dataset weights (as in our analysis).

In general, we find larger impacts on Black birth weight and smaller impacts on white birth weight than AHS. Looking across the specifications, a few broad patterns appear. First, removing time-varying county-level controls (Column 5) decreases estimated impacts on whites relative to AHS. Second, using the stacked dataset analyses (Column 7 on) tends to increase estimated impacts for both race groups, as does using the stacked dataset weights. By contrast, factors like the treatment-date definition and the sample restriction to July 1969 or later appear less important.

Ultimately, our baseline estimate of the impact of rollout of the food stamp program shows a statistically significant decrease in the share of low birth weight among Black births and an insignificant and economically small impact on white births (Appendix Table A6, Column 14). The estimate for Blacks is not robust to changes in the sample or specification (e.g., Column 13). More generally, as we discussed in Section 4.3, in some specifications but not others, there is evidence of a statistically significant improvement in Black birth weight. By contrast, there is no evidence in any specification of improvements in white birth weight or overall birth weight (Appendix Figures A19 Panels (b) and (c), A29, and A30).

Finally, since AHS do not estimate effects for a logged birth weight outcome, Appendix Table A7 estimates the AHS specification for this outcome of interest and proceeds with the same stepwise changes to the specification described above. Specifically, Column 1 in Appendix Table A7 corresponds to Column 3 in Appendix Tables A3 and A6, and Columns 2–11 correspond to

[50]These limitations include ensuring participation and birth weight are observed in each year 1970–1975 and dropping county-race-year cells with fewer than 25 births. Finally, we also ensure that after these limitations, the county-race has at least one year observed in the range -3 to +2 as these are the relative years included in the linear combination estimate.

the earlier Columns 5–14.

C.4 Imputing Race Specific Enrollment Rates

We do not directly observe race-specific county-year enrollment rates. Instead, we construct them using county-year overall enrollment rates, race-specific county-level population, race-specific poverty rates by state, and national race-specific take-up rates. The key assumption is that the ratio of national take-up rates for non-whites to whites applies uniformly across counties.

Recall from Section 2 that the county-year food stamp enrollment rate is

$$\mathrm{FS}_{ct} = \frac{e_{ct}}{\mathrm{pop}_c}, \tag{29}$$

where e_{ct} is the number of people enrolled in food stamps in county c in year t and pop_c is the 1970 county population. This can be decomposed as

$$\mathrm{FS}_{ct} = \mathrm{FST}_{ct} \times \mathrm{pov}_c, \tag{30}$$

where $\mathrm{FST}_{ct} = e_{ct}/p_c$ is the take-up rate (enrollees as a share of the poor) and $\mathrm{pov}_c = p_c/\mathrm{pop}_c$ is the 1970 poverty rate.

Our goal is to construct race-specific enrollment rates FS_{rct} for $r \in \{w, b\}$ (white and non-white). The strategy proceeds in three steps.

Step 1: Estimate the Relative Take-up Ratio δ. Using the PSID, we calculate national race-specific take-up rates annually from 1970 through 1981 (Appendix Figure A39, Panel a). These are defined as the number of PSID households of race r enrolled in food stamps divided by the number of PSID households of race r with income below the census needs standard.[51] We define the relative take-up ratio as

$$\delta_t = \frac{\mathrm{FST}_{bt}}{\mathrm{FST}_{wt}}. \tag{31}$$

Since δ_t is approximately constant over time (Appendix Figure A39, Panel b), we use its average value of $\delta = 1.6$.

Step 2: Approximate Race-specific County Poverty Counts. We approximate the race-specific county population in poverty as

$$p_{cr} = \mathrm{pov}_{s(c)r} \times \mathrm{pop}_{cr}, \tag{32}$$

[51]The census needs standard is a year-specific poverty threshold established by the Census Bureau, based on sex of head of household, family size, number of children, and age of householder.

where $\text{pov}_{s(c)r}$ is the state-level race-specific poverty rate from the census and pop_{cr} is the race-specific county population, calculated using county-level shares of population by race group and county population from the census.

Step 3: Back out Race-specific Take-up and Enrollment Rates. The aggregate county-year take-up rate can be written as a weighted average of race-specific take-up rates:

$$\text{FST}_{ct} = \frac{\text{FST}_{cbt}\, p_{cb} + \text{FST}_{cwt}\, p_{cw}}{p_{cb} + p_{cw}}. \tag{33}$$

Substituting $\text{FST}_{cbt} = \delta\, \text{FST}_{cwt}$ and solving yields

$$\text{FST}_{cwt} = \text{FST}_{ct} \cdot \frac{p_{cb} + p_{cw}}{\delta\, p_{cb} + p_{cw}}, \tag{34}$$

$$\text{FST}_{cbt} = \text{FST}_{ct} \cdot \frac{p_{cb} + p_{cw}}{p_{cb} + p_{cw}/\delta}. \tag{35}$$

Finally, race-specific county-year enrollment rates are

$$\text{FS}_{rct} = \text{FST}_{rct} \times \text{pov}_{cr}. \tag{36}$$

C.5 Additional infant health outcomes

We discuss additional data details and sample restrictions for our measures of gestational age, infant mortality, and birth counts.

Gestational age. As in Chen et al. (2016), we limit gestation observations to births with at least 22 weeks gestation. We set an upper limit at less than 49 weeks gestation to deal with outliers that are more likely errors. This censors only the highest half percentile of gestation values. We set all gestational ages reported outside that range to missing. In addition, in 1981, NCHS began imputing gestation for births missing the day of the mother's last menstrual period (LMP) using the accurately recorded day of LMP of the most recent birth record with the same months of gestation and same 500-gram interval of birth weight; we set these imputed birth-level gestation values as missing as well. Finally, as shown in Appendix Figure A40, the Department of Health and Human Services reports the table showing the year that each state began reporting the full day, month, and year of LMP. This consistent reporting would ensure a more accurate gestation calculation. For state-years that do not report LMP, we treat all gestation values as missing.

Infant mortality. Infant mortality rate is the count of deaths before 12 months per 1,000 births reported in the birth files.

Log birth counts. Prior to 1972, only 50% of birth records are reported in the birth files. Starting in 1972, some states began processing the records of 100% of births occurring within that state. Each birth record contains a weight factor which indicates whether the birth occurred in a state-year reporting 50% or 100% of births. We use these record-level weights to account for these reporting differences, allowing births in cohorts reporting half of births to be counted as two births.

Sample restrictions. Across all of these additional infant health outcomes, we apply additional sample restrictions to the baseline analysis samples defined in Appendix Table A1. Specifically, for each of these outcomes, we further drop counties if they are missing data for the relevant analysis years (i.e., 1972-1981 for the free food stamps reform analysis, and 1970-1975 for the rollout); for log infant mortality and log birth counts, we drop the county if the non-logged outcome is zero in any of the relevant years. For the free food stamps reform analysis, these samples are 78.44% (1,359 counties), 95.72% (1,145 counties) and 100% of the (population-weighted) 1,804-county baseline sample for the gestation outcomes, infant mortality, and birth counts, respectively. For the rollout analysis, the limited samples constitute 66.68% (932 counties), 95.45% (1,067 counties) and 100% of the (population-weighted) 1,520-county baseline sample for the gestation outcomes, infant mortality, and birth counts, respectively.

C.6 Complier Characteristics

To estimate complier characteristics, we harmonize observations from two separate surveys: the Panel Study of Income Dynamics (PSID) and the National Health and Nutrition Examination Survey II (NHANES II). The PSID began in 1968 as an annual survey of sample households and is nationally representative. We use PSID years 1973 through 1981. The CDC initiated the NHANES in 1976 as a cross sectional survey of the civilian, noninstitutionalized national population. Survey responses were collected over the span of 5 years, until its conclusion in 1980. Thus, the number of observations in each year of the NHANES is inconsistent, and the data in any given year are not a nationally representative sample. The NHANES oversamples populations vulnerable to malnutrition such as low-income, pre-school, and elderly individuals. Sampling weights address this oversampling. Importantly, both surveys report information on income, household composition, household weights, and Food Stamp Program enrollment that we employ to examine complier characteristics. Appendix Table A9 describes the surveys' observations across years by survey and enrollment status.

The PSID and NHANES each have an annual measure of total household income which includes taxable earnings of all household members as well as cash transfers. We convert to montly income by dividing annual income by 12. To calculate equivalized household income, we adjust total household income for the number of adults and children using the equivalence scale in Citro and Michael (1995) and impose a household income floor of the 5th percentile of total income in 1970. We calculate social marginal welfare weights as as the inverse cube of winsorized (to the 5th percentile) equivalized income, divided by the weighted mean of this inverse cube for all household-years 1976-1981. We compare the federal poverty level (FPL) from U.S. Department of Health and Human Services (2024) to total family income, for each household size and year, to construct the share with income below 50% of the FPL, and 25% of the FPL.

To obtain precision weighted estimates of the average household characteristics of food stamp enrollees each year, we estimate the following equation on the combined PSID and NHANES data on food stamp enrollees in years 1973–1981:

$$y_{htd} = \sum_{p \neq 1978} \alpha_p \cdot \mathbf{1}\{p = t\} + \gamma \text{NHANES}_d + \varepsilon_{htd} \tag{37}$$

where the α_p coefficients capture the differential change in outcomes for food stamp enrollees in each year p, y_{htd} is an outcome variable (e.g., equivalized income) in household h and year t, from dataset d, NHANES_d is an indicator for the observation coming from the NHANES dataset, making PSID is the omitted group. Survey weights are normalized to sum to the count of enrolled households within each dataset-year. Standard errors are clustered at the household level.

To obtain precision weighted estimates of the average household characteristics for enrollees in the pre-period and post-period, we estimate the following equation on the combined data for Food Stamp Program enrollees in years 1976–1981:

$$y_{htd} = \alpha + \beta \cdot \text{NHANES}_d + \delta \cdot \text{Post}_t + \varepsilon_{htd} \tag{38}$$

where α is the pre-period mean (normalized to PSID levels via the NHANES indicator), δ is the average change from between periods, and Post_t is equal to 1 if year is 1979 or later. Survey weights are normalized to sum to the count of enrollees within each dataset-period (pre- or post-period), and standard errors are clustered at the household level.

To obtain complier means in Table 1, we combine estimates from Equation 38 together with our estimate of the complier share of $\sigma_c = 0.2$ (see Section 3.2) to estimate complier means as:

$$\bar{y}_{\text{complier}} = \alpha + \frac{\delta}{\sigma_c} \tag{39}$$

where α is the pre-period mean from Equation 38 and δ is the post-period coefficient (also in Equation 38. Standard errors for the complier means are estimated using the delta method, assuming zero covariance between δ and σ_c, but accounting for covariance between α and δ which come from the same regression on pooled survey data in Equation 38. This accounts for estimation uncertainty in both the first stage estimates, and the period-specific survey means.

Results Appendix Figure A46 summarizes the results for equivalized income. The data indicate that after the elimination of the purchase requirement, mean equivalent household income falls among enrollees; trends among those not enrolled in food stamps are less pronounced.

D Welfare Analysis

D.1 Discussion of Welfare Assumptions.

D.1.1 Constant Proportional Value

Section 5.1 imposed the "constant proportional value" assumption

$$\text{(A1, CPV for always-takers)} \qquad V_a(s_0, F) = V_a(s_1, F), \tag{40}$$

which we modify for our empirical setting in Section 5.1 (taking into account that the subsidy bundle changes size) to:

$$\text{(A1}'\text{)} \qquad V_a(s_0, R+F) = V_a(s_1, F). \tag{41}$$

A1 is equivalent to assuming that the welfare impact of the free food stamp reform on always-takers operates exclusively through a change in the subsidy. Thus, the subsidy may mechanically change how much of the bundle the government finances, but it does not independently affect households' valuations of the benefit F or their non-price costs of participation.

For intuition, consider a standard setting with unit demand and no enrollment costs. Always-takers purchase under either subsidy rate and are therefore inframarginal. By envelope logic, a change in the subsidy affects their welfare only through the direct change in their budget constraint, so welfare moves one-for-one with the subsidy. Constant proportional value A1$'$ makes an analogous restriction in the case of changing bundle sizes. Because government spending per enrollee is F under both regimes, eliminating the purchase requirement must leave always-takers' net private value per dollar of spending unchanged. The assumption therefore rules out additional impacts of the free food stamp reform on always-takers' valuations or non-price participation costs, such as stigma.

To see the restriction on primitives, adopt notation from Section 5.2. Specifically, let tildes $\widetilde{\cdot}$ denote levels rather than per-sF units. Then, $\widetilde{V}_\theta \equiv V_\theta \cdot sF$. We define $\widetilde{X}_\theta(R+F)$ as willingness to pay for $R+F$ in food, and $\widetilde{k}_\theta(s, R+F)$ as the transaction or ordeal costs associated with purchasing F in food stamps, with

$$\widetilde{V}_\theta(s, R+F) \equiv \widetilde{X}_\theta(R+F) - \widetilde{k}_\theta(s, R+F) - R. \tag{42}$$

Now, define the net value of the additional R of stamps provided under the purchase requirement as

$$\Delta_a(s, R+F) := \widetilde{V}_a(s, R+F) - \widetilde{V}_a(1, F) \tag{43}$$

so that

$$\Delta_a(s_0, R+F) = \left[\widetilde{X}_a(R+F) - \widetilde{X}_a(F) - R\right] - \left[\widetilde{k}_a(s_0, R+F) - \widetilde{k}_a(1, F)\right]. \tag{44}$$

Then

$$V_a(s_0, R+F) = V_a(1, F) + \frac{\Delta_a(s_0, R+F)}{F}, \tag{45}$$

so constant proportional value is equivalent to $\Delta_a(s_0, R+F) = 0$. In particular, if always-takers are inframarginal in the additional R of stamps, so that they value those stamps dollar-for-dollar (meaning that $\widetilde{X}_a(R+F) - \widetilde{X}_a(F) - R = 0$), the assumption reduces to

$$\widetilde{k}_a(s_0, R+F) = \widetilde{k}_a(1, F). \tag{46}$$

Thus, in this case, the assumption says that ordeal costs are bundle-invariant (additive) for always-takers, so eliminating the purchase requirement does not alter ordeal costs.

D.1.2 Law of Demand

To estimate perceived net private value in Section 5.2 we impose what we call a "Law of Demand" (18). We now discuss why we need to make this assumption and provide microfoundations.

Why Do We Need the Assumption? The key issue is that observed enrollment under the *purchase requirement* does not rank households' values for the *free program*. For example, suppose always-takers face additive ordeal costs and enroll under the purchase requirement but barely prefer to do so. Then their value for free food stamps may be only slightly above zero (since their value per dollar of free stamps is equal to their value per dollar under the purchase requirement, by the constant proportional value assumption). If compliers' ordeal costs respond to the subsidy, a complier who does not enroll under the purchase requirement could then value the free program more than an always-taker. The Law of Demand rules out this possibility and ensures that always-takers lie weakly above compliers in their value for free food stamps.

More fundamentally, this issue arises whenever there is a nonlinear subsidy on consumption of some good up to a cap; that is, it is not specific to our setting where the bundle size $R+F$ changes. Fix the benefit cap F and vary only the subsidy rate s. Purchases up to F are subsidized, while purchases above F face the market price. For these households who consume less than F absent the subsidy, enrollment can change the quantity of the good consumed, so the take-up decision reveals willingness to pay for the additional amount of the good induced by the subsidy. Because of the nonlinear nature of the subsidy, price variation does not directly reveal the average value per dollar of F. This identification issue arises with top-ups under nonlinear subsidies (Einav et al., 2016), and applies broadly to settings like child-care, education, or rent subsidies.

Microfoundations. Maintaining the notation in Appendix D.1.1, define net value per dollar of the benefit bundle (but gross of the purchase price) as

$$Y_\theta(s, B) := \frac{\widetilde{X}_\theta(B) - \widetilde{k}_\theta(s, B)}{B}. \tag{47}$$

By definition, a household enrolls if and only if $Y_\theta(s, B) \geq 1 - s$. Moreover, $Y_\theta(1, F) = V_\theta(1, F)$.

For compliers, assume a flavor of the constant proportional gross value assumption but on Y (rather than V):

$$\text{(A2, } Y\text{-CPV for compliers)} \qquad Y_c(s_0, R + F) = Y_c(1, F). \tag{48}$$

Because compliers enroll when stamps are free but not under the purchase requirement, A2 implies

$$0 \leq V_c(1, F) < 1 - s_0. \tag{49}$$

For always-takers, constant proportional value and enrollment under the purchase requirement imply only $V_a(1, F) \geq 0$. We now impose a monotonicity condition which says that always-takers value food stamps weakly more than compliers:

$$\text{(A3, monotonicity)} \qquad \inf_{\omega_a \in \Omega_a} V_a(1, F; \omega_a) \geq \sup_{\omega_c \in \Omega_c} V_c(1, F; \omega_c), \tag{50}$$

where Ω_θ is the set of preference primitives consistent with type θ's observed behavior.[52] A3 is an intuitive sorting condition that rules out always-takers from having lower values of free stamps than compliers.

Since (without further information) complier values can approach $1 - s_0$ from below, this condition implies

$$V_a(1, F) \geq 1 - s_0 > V_c(1, F). \tag{51}$$

That is, A2 and A3 together imply the Law of Demand (18) for the free-stamps reform. We impose analogous microfoundations to allow other price variation (i.e., the bonus reform) to identify other points on the demand curve demand.

[52]Here ω_θ indexes realizations of preferences that the econometrician cannot rule out given type θ's observed behavior and the preceding assumptions. We can likewise interpret ω_θ as indexing (known) within-type heterogeneity under an additional assumption that there are types whose demand approaches the $1 - s_0$ threshold.

Plausibility of A2. For intuition about the plausibility of A2, we consider whether it would hold under different cases. We can rewrite A2 as

$$\frac{\widetilde{X}_c(R+F) - \widetilde{k}_c(s_0, R+F)}{R+F} = \frac{\widetilde{X}_c(F) - \widetilde{k}_c(1,F)}{F}. \tag{52}$$

Equivalently,

$$\frac{\widetilde{X}_c(F)}{F} - \frac{\widetilde{X}_c(R+F)}{R+F} = \frac{\widetilde{k}_c(1,F)}{F} - \frac{\widetilde{k}_c(s_0, R+F)}{R+F}. \tag{53}$$

Thus, any change in preferences over gross food value per dollar across the two bundles must be matched by the same change in ordeal costs per dollar.

First, suppose compliers are strongly inframarginal:

$$\frac{\widetilde{X}_c(R+F)}{R+F} = \frac{\widetilde{X}_c(F)}{F} = 1. \tag{54}$$

Equation 53 then requires ordeal costs to be proportional to bundle size:

$$\frac{\widetilde{k}_c(s_0, R+F)}{R+F} = \frac{\widetilde{k}_c(1,F)}{F}. \tag{55}$$

This contrasts with the bundle-invariant (additive) ordeal costs implied by constant proportional value for strongly inframarginal always-takers.

Second, suppose compliers are weakly but not strongly inframarginal (which Appendix B.2 finds evidence of):

$$\widetilde{X}_c(F) = F, \qquad \widetilde{X}_c(R+F) < R+F. \tag{56}$$

A2 then requires ordeal costs per dollar to be higher for the smaller F-bundle:

$$\frac{\widetilde{k}_c(1,F)}{F} - \frac{\widetilde{k}_c(s_0, R+F)}{R+F} = 1 - \frac{\widetilde{X}_c(R+F)}{R+F} > 0. \tag{57}$$

In particular, A2 can hold with additive ordeal costs, $\widetilde{k}_c(s_0, R+F) = \widetilde{k}_c(1,F)$, if the increase in average food value for the larger bundle exactly equals the resulting increase in ordeal costs per dollar.

Finally, suppose compliers are not inframarginal even at F, so that $\widetilde{X}_c(F) < F$. A sufficient condition for Equation 53 in this case is that *both* gross valuation and ordeal costs scale proportionally with bundle size:

$$\frac{\widetilde{X}_c(R+F)}{R+F} = \frac{\widetilde{X}_c(F)}{F}, \qquad \frac{\widetilde{k}_c(s_0, R+F)}{R+F} = \frac{\widetilde{k}_c(1,F)}{F}, \tag{58}$$

which then implies that

$$\frac{\widetilde{X}_c(R+F)}{\widetilde{X}_c(F)} = \frac{\widetilde{k}_c(s_0, R+F)}{\widetilde{k}_c(1,F)} = \frac{R+F}{F}. \tag{59}$$

Intuitively, Equation 59 says that in this case both preferences and ordeal costs satisfy the same proportional valuation in bundle size.

D.2 Fiscal Externalities: Earnings and Unemployment

D.2.1 Empirical Estimates

Our main estimates assumed FE$= 0$, in part motivated by small FEs in Hendren and Sprung-Keyser (2020). We complement these estimates by examining the impact of the free-stamps policy and the rollout on employment outcomes. If food stamps reduce employment, that could suggest important fiscal externalities.

Data and Sample. We use county-by-industry employment and earnings data from the County Business Patterns. For confidentiality, some local data are suppressed, and some outcomes use Eckert et al. (2020) and Eckert et al. (2022)'s imputed CBP data that infers CBP county-industry employments or payrolls in suppressed periods from bounds on the number of establishments. In addition to looking at employment data across industries, we focus on industries with enough workers on food stamps that they could be plausibly affected by changes in take up. Exposed industries, chosen because they employ large shares of low-wage workers, are: agriculture, forestry, fisheries, retail, food services, and accommodations (see Appendix C.1 for rationale on the specific industry choices).

We restrict each sample to counties with data for that outcome in all years of the analysis, forming balanced panels outcome-by-outcome. When studying the elimination of the purchase requirement, we focus on counties that are in the primary analysis sample and have employment outcome data in all years 1972–1981. For the rollout, we use counties in the primary rollout analysis sample and have employment outcome data in relative years –3 to $+2$. We form balanced panels outcome-by-outcome. Because industry-specific data are not available for all counties in all years, the balanced panel for exposed-industry outcomes drops some smaller counties. Nevertheless, 87% of the population resides in the approximately 600 counties that enter the balanced panel for the purchase requirement analysis. For the rollout analysis of exposed-industry employment, the resulting sample consists of 437 counties, 78% of the baseline rollout sample and 29% of 1970 U.S. population.

Results. We find no impact of food stamps on employment. The small impact of food stamps enrollment on employment is consistent with our calibration of the fiscal externality of 0.

For the purchase requirement, we use the event-study specification in the main text (Equation 4), but with CBP employment data as outcomes. We find relatively tight nulls, consistent with the above calibration (Appendix Figure A41). Panels (a) and (c) show the first stages for the sample of counties where we have log employment data, and Panels (b) and (d) show the reduced forms for log employment in all industries and log employment in exposed industries. Recalling that the difference in exposure between the 10th- and 90th-percentile counties is about 0.1, food stamp exposure reduces employment by -0.22% (s.e. 1.1%) and increases employment in exposed industries by 1.3% (s.e. 1.2%). We can therefore rule out large employment declines of 2% overall and 1.1% in exposed industries. Next, we examine impacts of food stamps rollout using the stacked specification (Equation 8). In Appendix Figure A42, we find tight null results; for employment, we find that rollout reduced employment by 0.8% (s.e. 0.7%) and raised employment in exposed industries by 0.4% (s.e. 1%).

D.3 Bonus Reform

We provide more detail on our empirical analysis of the 1970 "bonus reform" used in Section 5.2. In 1968, Congress modified the program to increase the subsidy starting in January 1970; although this subsidy increase was (confusingly) part of the set of amendments signed into law in 1971, the administration enacted the significant changes to benefits in January 1970 (Richardson, 1979; Hoover, 1974).

Specifically, the 1970 reform raised the "bonus" amount F, thus increasing the subsidy rate $\frac{F}{R+F}$. Because (see Appendix B) enrollees at this time were essentially all inframarginal (i.e., would have purchased more food than F even in the absence of food stamps), the bonus reform functioned equivalently to a subsidy increase. Published tables on the average subsidy rate indicate that the subsidy changed from 40% before the bonus reform to about 65% (specifically, 63.2%) after the bonus reform (Hoover, 1974).

The bonus reform occurred at a national level. Therefore, to study its on enrollment, we use an exposure design similar to Section 3. Because the bonus reform occurred in a period during the food stamps rollout, we restrict our analysis to counties in the baseline free food stamp reform analysis that had rolled out the food stamp program by 1967 or earlier. This restricts our analysis to 730 counties, approximately 52 percent of population-weighted counties in the baseline free food stamp reform analysis sample.

We use the same measure of county-level exposure (i.e., Eligibility rate $-$ Enrollment rate) as before (see Equation 3) but now measure the enrollment county in 1968 (i.e., $z = 1968$). Appendix Figure A43 shows the distribution of our exposure measure. It is winsorized at the 1970 population-weighted 1st (0.035) and 99th percentiles (0.332). The average 1968 exposure measure among these counties is 0.102 (standard deviation 0.055).

To estimate the impact of the 1970 subsidy increase on the enrollment rate, we estimate an augmented version of of Equation 4. Specifically we estimate:

$$FS_{ct} = \sum_{s \neq 1969} \delta_s \left(\mathbf{1}\{t = s\} \times \text{Exposure}_c^z \right) + \alpha_c^1 + \gamma_t^1 + \lambda_{o(c)t}^1 + u_{ct} \tag{60}$$

where $\lambda_{o(c)t}$ are rollout-cohort-by-calendar-year fixed effects, with $o(c)$ denoting the year of county c's rollout year.

Results. We first present visual evidence in Appendix Figure A44 Panel (a), akin to Figure 4 (Panel a) showing trends in enrollment rates by quartile of exposure, from 1967–72. We see that all four quartiles of exposure have a jump in enrollment in 1970. Consistent with an enrollment response, a moderate gap emerges between the top three quartiles and the first quartile that was not present before the bonus reform.[53]

Appendix Figure A44 Panel (b) shows the results from estimating the event study in Equation 60. Between 1967 and 1969, we see a relatively flat pre-trend. In 1970, enrollment jumps up in more exposed relative to less exposed counties and remains elevated through the end of the analysis in 1972. The point estimate implies that going from 0 to 1 exposure units is associated with a 24 (standard error = 2) percentage point increase in the enrollment rate. Multiplying by the average exposure of 0.102 produces a 2.4 percentage point ($\approx 0.2383 \times 0.102 \times 100$) increase in enrollment from the increase in the subsidy from 40 to 65%.

D.4 Sensitivity analysis

We provide more detail behind the sensitivity analyses of our welfare results discussed in section5.3 that free-stamps produce less welfare per dollar of mechanical expenditure than cost sharing (i.e., $W(1) < W(0.65)$). Because this finding in turn stems from the fact that welfare-weighted valuations are more than twice as high for always-takers than for compliers (i.e., $\lambda_a V_a = 2.2 > 1.1$ compared to $\lambda_c V_c = 1.1$), we approach our analysis of sensitivity to various economic and econometric assumptions by asking whether there are potentially unmodeled or unaccounted for factors that could double the welfare-weighted valuations of compliers relative to always-takers. Appendix Table A8 reports and compares estimates of social welfare for compliers ($W_c \equiv \frac{\lambda_c(V_c+D_c)}{1+\text{FE}_c}$), for always-takers ($W_a \equiv \frac{\lambda_a(V_a+D_a)}{1+\text{FE}_a} = W(0.65)$) and $W(1) = \sigma_a W_a + \sigma_c W_c$ under alternative assumptions. The first row summarizes the baseline estimates.

Distortions (D_θ). Rows 2 through 4 consider distortions (D_c and D_a) which our baseline analysis assumed to be zero. Row 2 allows for type-specific distortions to equal type-specific valuations (i.e., $V_c = D_c$, $V_a = D_a$); given our estimates of $\lambda_c = 6.2 > \lambda_a = 3.2$, we continue to find

[53] We expect even the bottom quartile to have some increase in enrollment because its exposure is nonzero (about 0.05).

$W(1) < W(0.65)$. Rows 3 and 4 therefore investigate what would be required in terms of unmodeled distortions to reverse our conclusion that $W(0.65) > W(1)$. Row 3 shows that if we assume equal distortions in levels (i.e., $D = D_a = D_c$) then our results hold as long as $D < 0.39$, or households undervalue food stamps by less than 43 cents per dollar. This would imply severe distortions for compliers in percent terms (almost 250% of their revealed net private value), since they have much lower revealed net private value than always-takers.

We also considered distortions only for compliers, since Spinnewijn (2017) argues that due to selection into enrollment, distortions should be greater among compliers than always-takers; intuitively, always-takers (who choose to enroll even with the purchase requirement) must be selected on having relatively low distortions. Row 4 shows that if always-takers have zero distortions ($D_a = 0$), then our results hold as long as $D_c < 0.187$, or about equal to their revealed net private value (recall $V_c = 0.18$).

To gauge the plausibility of the magnitudes of these distortions, we consider what they would require in a stylized model of a particular behavioral bias. We consider present focus, motivated by the idea that households may fail to take up food stamps because the hassle of enrolling today looms larger psychologically than the benefit of using them in the future. Imagine a two-period setting in which future food stamp benefits have a normative net present value of \$1 minus ordeal costs, which accrue contemporaneously. Types are now two-dimensional (θ, η) where $\theta \in \{a, c\}$ (as before) and $\eta \in \{p, n\}$ denotes whether the individual is *p*resent focused or *n*eoclassical. Neoclassical households have present-bias parameter $\beta = 1$ and we assume based on estimates from the laboratory literature that present-focused households have $\beta = 0.7$ (e.g., Frederick et al., 2002). Distortions are given by $D = (1 - \beta)$. Therefore letting π_θ denote the share of type θ who are present-focused, $D_\theta = \pi_\theta(1 - \beta)$. With these parameters, it is not possible — even if everyone were present biased with $\beta = 0.7$ — to obtain that $D_a = D_c = 0.39$ (the level of distortions needed to overturn our results if distortions are equal for compliers and always-takers); rather we would need everyone to be present-biased with $\beta = 0.61$. Alternatively, to get $D_a = 0$ and $D_c = 0.187$ (what we need to overturn results if only compliers experience distortions), requires always-takers are all neoclassical ($\pi_a = 0$) while 62% of compliers are present focused.

Welfare Weights (λ_θ). Rows 5 through 8 consider whether alternative welfare weights could overturn our findings. Several considerations motivate this exploration. First, estimation of our baseline welfare weights λ_c and λ_a required an assumption about the utility function in order to transform the observed differences in household equivalized income into welfare weights. Second, our estimate of the lower income of compliers relative to always-takers should be interpreted cautiously, as secular trends could confound the time-series analysis of changes in the characteristics of food stamp enrollees. Third, the existence of liquidity constrained households could suggest higher social welfare weights.

Our results are robust to a range of alternative welfare weights. For example, even if we doubled the complier welfare weight (Row 5), or assume much greater redistributive preferences by allowing a coefficient of relative risk aversion of 50 (Row 6), we cannot reverse our results. One way we can reverse our results is to adopt a "step-function" type welfare function in which the social planner places positive and equal weight only on people whose annual income is below some threshold (and zero weight on people whose annual incomes exceed this threshold). In Row 7, we show that if the social planner places positive and equal weight only on households whose annual income is less than 50% FPL (and zero weight on people whose annual incomes exceed this threshold), we can obtain $W_c > W_a$. More generally, the maximal FPL threshold κ such that $W_c \geq W_a$ in this class of social welfare functions is $\kappa = 63.1\%$ of the FPL (Row 8). Such social welfare functions are of course possible, but would be counterfactual to both the weights implied by U.S. policy (Hendren, 2020) and household preferences in redistribution experiments (Capozza and Srinivasan, 2026).

Net Private Value (V_θ) . Our estimates of V_θ come from interpolating bounds on demand. In Rows 9 through 11, we therefore consider how alternative assumptions on these demand curves would affect results. We consider changing the valuations to be at the upper bound (Row 9) or lower bound (Row 10) of the nonparametric bounds in Figure 11; these shift $V(1)$ to either 0.77 or 0.42 but cannot reverse the comparison of free stamps to cost sharing.

We can reverse the conclusion if we assume compliers have valuations V_c at their upper bound, and always-takers' valuation at their lower bound. This would say that all compliers and all always-takers are just indifferent, but in the opposite direction (Row 11).[54]

First Stage (σ) . Shifting shares of types (σ_θ) identified from the magnitude of the first stage cannot affect the comparison of $W(0.65)$ and $W(1)$, but it can affect $V(1)$. In Rows 12 and 13, we consider the smallest and largest first stages that we estimate in Figure A8 (either including state-by-year fixed effects or interacting the log of the female population age 15–44 with calendar year), which moves $V(1)$ to 0.63 or 0.59.

Fiscal Externalities (FE) . We assumed zero fiscal externalities. Fiscal externalities from earnings responses must be small because taxes on low-income households are only about 12.5% (at least, under the assumptions in Hendren and Sprung-Keyser, 2020); moreover, we found little evidence on earnings impacts in Appendix D.2. If we assign both parties a conservative fiscal externality of 16 cents (which comes from the earnings impact in Hendren and Sprung-Keyser, 2020), that

[54] V_a need not mechanically equal V_c because the additional data point from the bonus reform implies their valuations are strictly larger than 0.35.

cannot flip the comparison of W_a to W_c (Row 14).[55] We use a large fiscal externality to see whether a larger FE can change results; the final FE in Hendren and Sprung-Keyser (2020) incorporating other longer-run forces is actually smaller at 5 cents.

Row 15 shows that if we assign *only* always-takers the fiscal externality of 16 cents, that cannot flip the results either. Breaking our result would require fiscal externalities beyond labor supply (as in e.g., Bailey et al. (2024)), and they need to primarily accrue to always-takers. Row 16 shows that, if compliers have zero FE, the always-taker FE would need to be at least \$1.07 to overturn our results.

Potential policy confounds. Although the elimination of the purchase requirement was the centerpiece of the 1979 reform, as discussed in Appendix A, the reform also made changes that reduced eligibility, particularly for higher incomes. Accounting for potential unmodeled eligibility reductions would actually strengthen the normative case for cost-sharing. If eligibility restrictions reduced the share of eligible enrollees but did not affect the characteristics of eligible individuals, this would imply that our estimated first stage enrollment response under-estimates the enrollment response to an increase in subsidy; in other words, the increase in subsidy alone would create more compliers with $\lambda_c V_c < \lambda_a V_a$ than we estimated, thus further reducing $W(1)$ relative to $W(0.65)$ (see Equations 14 and 15). In fact, in practice, the new restrictions had more bite at higher incomes and thus reduced the average income of eligible households (see Appendix A.3). If reductions in eligibility among higher income individuals contribute to our estimates of greater poverty among compliers, this suggests that the increase in subsidy per se improved targeting on social value less than we estimate. In sum, these forces would push us to overestimate both λ_c and $V(1)$.

In addition, we have abstracted from the fact that under the purchse requirement, R was not constant but increased with income. This has two potential implications. First, it suggests our estimate of $V(1)$ is likely an upper bound: If average R is negatively correlated with exposure — because high-poverty counties faced lower average purchase requirements — then our estimated first stage is smaller than the first stage from a constant-R reform. A larger counterfactual first stage would imply an even lower value of $V(1)$. Second, this same force could lead us to understate λ_c. If the poorest households faced the smallest change in effective subsidy from the elimination of the purchase requirement, then they were less likely to be induced into take up by the actual reform than they would have been under a constant-R reform. The compliers we observe would therefore be less poor, and receive lower welfare weights, than the compliers induced by a constant-R reform. While these forces are potentially important, overturning our

[55] The estimate of 16 cents comes from a calibration of Hendren and Sprung-Keyser (2020), building off the earnings estimates in Hoynes and Schanzenbach (2012). The estimate is imprecise; the standard error on the 16 cent estimate is at least 70 cents. This comes from rescaling the estimated earnings response in Hoynes and Schanzenbach (2012) of \$219 with a standard error that is 4.4 times larger at \$916).

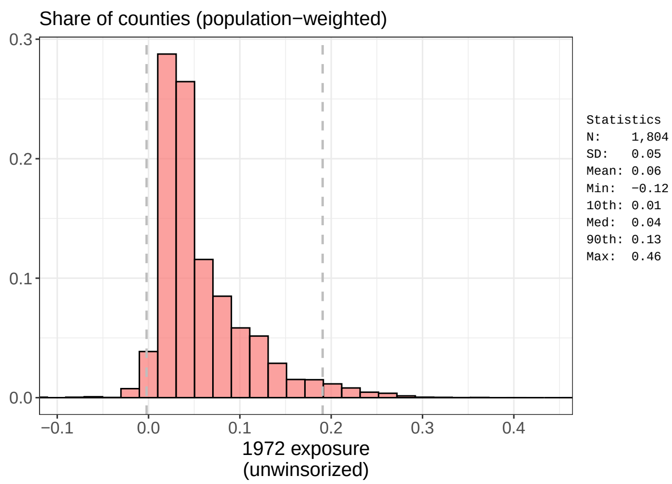
conclusion would require a large correction: our estimates would need to understate λ_c by roughly a factor of two.[56]

[56] A related issue is that some households had a $0 purchase requirement. However, this was only a small share: in the 1976–81 PSID we estimate that less than 1% of 4-person enrolled households would have had a $0 purchase requirement.

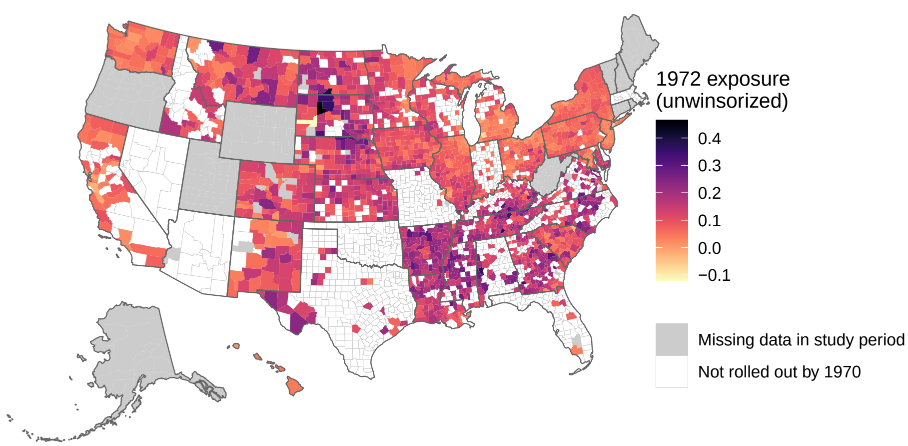
E Appendix Figures and Tables

Figure A1: Variation in Exposure (Unwinsorized)

(a)

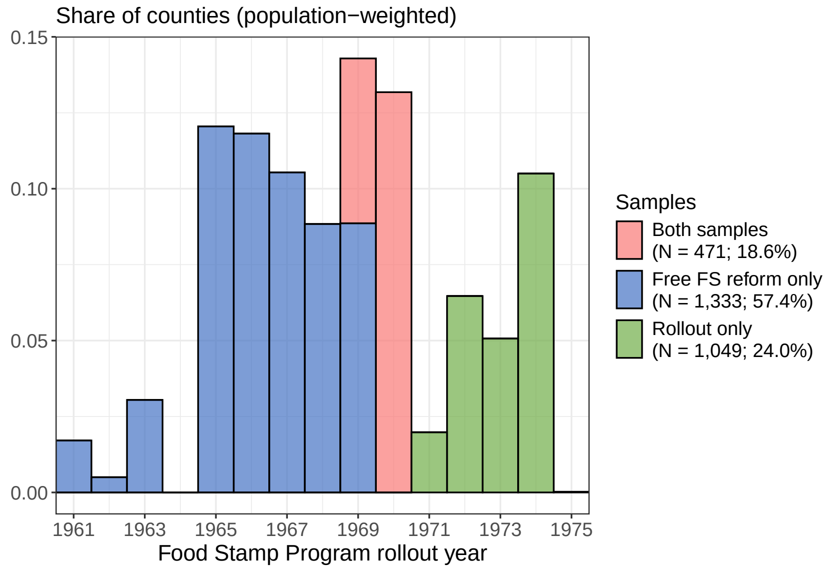


(b)



Notes: Figure shows a histogram of exposure across (population-weighted) counties (Panel a) and a heat map of exposure across counties (Panel b). Exposure is measured based on 1972 Food Stamp Program enrollment counts as in Equation 3, and analysis is limited to the counties that had rolled out their food stamp program by 1970. The figure shows the raw (i.e., uncensored) exposure measure. Our baseline exposure measure (eg. Figure 2 is censored to the population-weighted percentile of 0 at the bottom of the distribution and the corresponding share at the top of the distribution, ±3.2% as indicated by the vertical dashed lines in Panel (a).

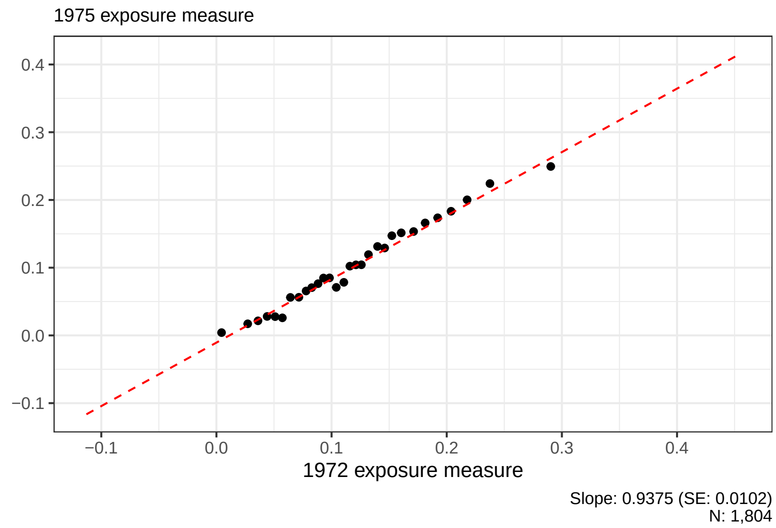
Figure A2: Included Analysis Counties, by Year of Food Stamp Program Rollout



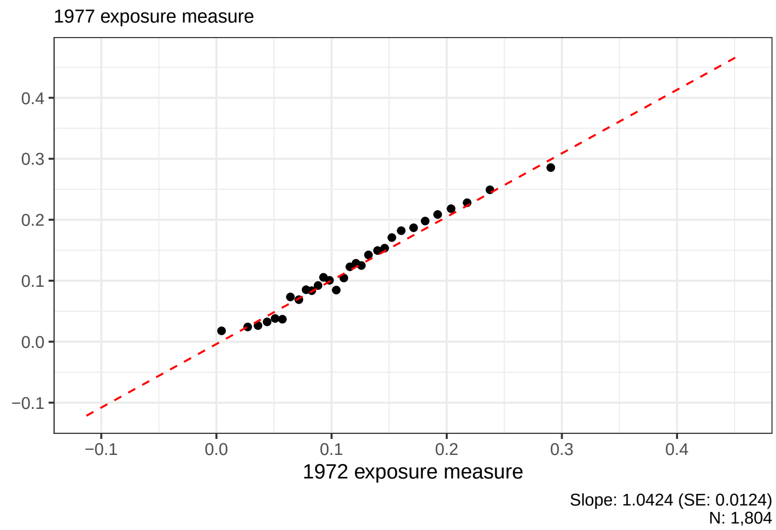
Notes: Figure shows population-weighted histogram of analyzed counties across rollout years. Counties are shaded according to which analysis sample they are included in. The analysis sample for the free food stamp reform includes counties that rolled out food stamps in 1970 or earlier. The 1970 county population-weighted share of counties in each group are found in the figure legend. The analysis sample for the rollout analysis includes counties that rolled out in July 1969 or later.

Figure A3: Correlation between Year-specific Exposure Measures

(a)



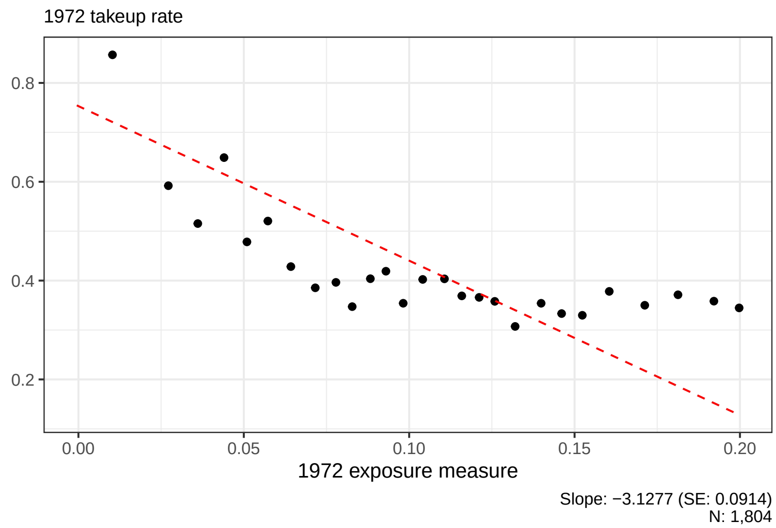
(b)



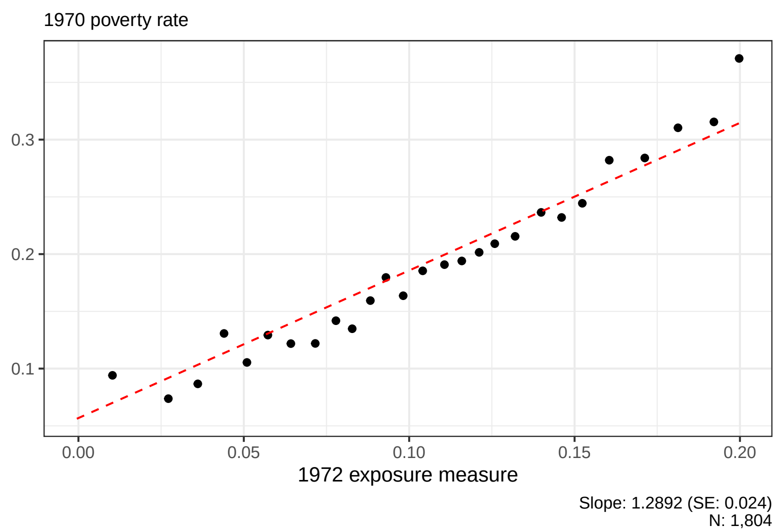
Notes: Figure shows a binned scatterplot of alternative year-specific exposure measure against the 1972 exposure measure, as described in Equation 3 and depicted in Appendix Figure A1, weighted by 1970 county population. Panel (a) describes the correlation of the 1972 exposure measure with the 1975 exposure measure, and Panel (b), the 1977 exposure measure. Each point represents the average 1975, or 1977, exposure within a bin of 1972 exposure values. A linear fit illustrates the population-weighted correlation between the two exposure measures. The sample is limited to the baseline analysis sample for the free food stamp reform (see Appendix Table A1).

Figure A4: Exposure Measure Correlated with Poverty Rate and Take-up Rate

(a) Correlation Between Exposure and Take Up
(Share of Eligible Individuals Enrolled)



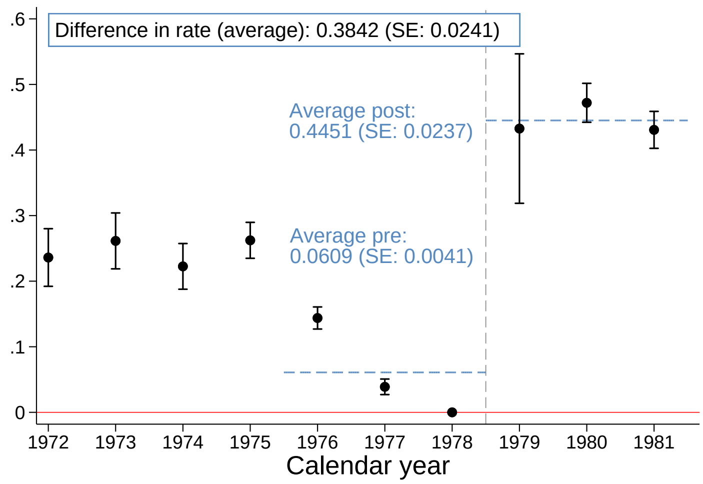
(b) Correlation Between Exposure and Poverty Rate



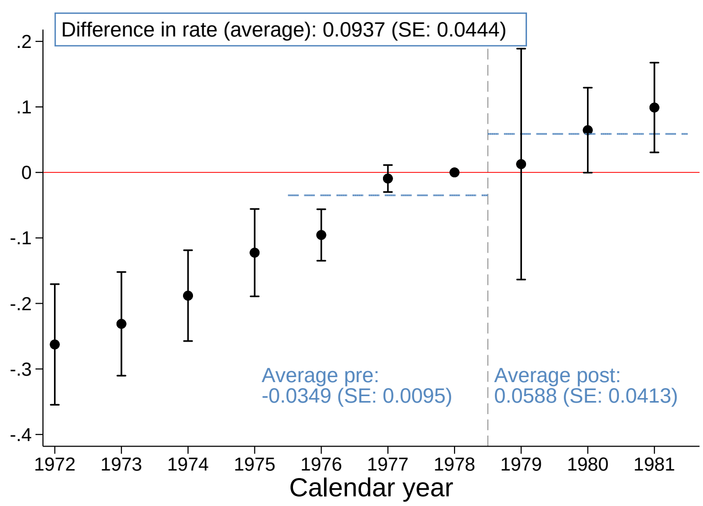
Notes: Figure shows a binned scatterplot of the 1972 exposure measure, as defined by Equation 3, against take-up rate rate (Panel a) (i.e., enrollment rate among those eligible, defined as those below the poverty line) and poverty rate (Panel b), weighted by 1970 county population. Each point represents the average poverty or take-up rate within a bin of 1972 exposure values. A linear fit illustrates the population-weighted correlation between the two measures. The sample is limited to the baseline analysis sample for the free food stamp reform (see Appendix Table A1).

Figure A5: Impact of Free Food Stamp Reform on Enrollment, with Additional Covariates

(a) County's Take-up Rate Decile × Year Fixed Effects

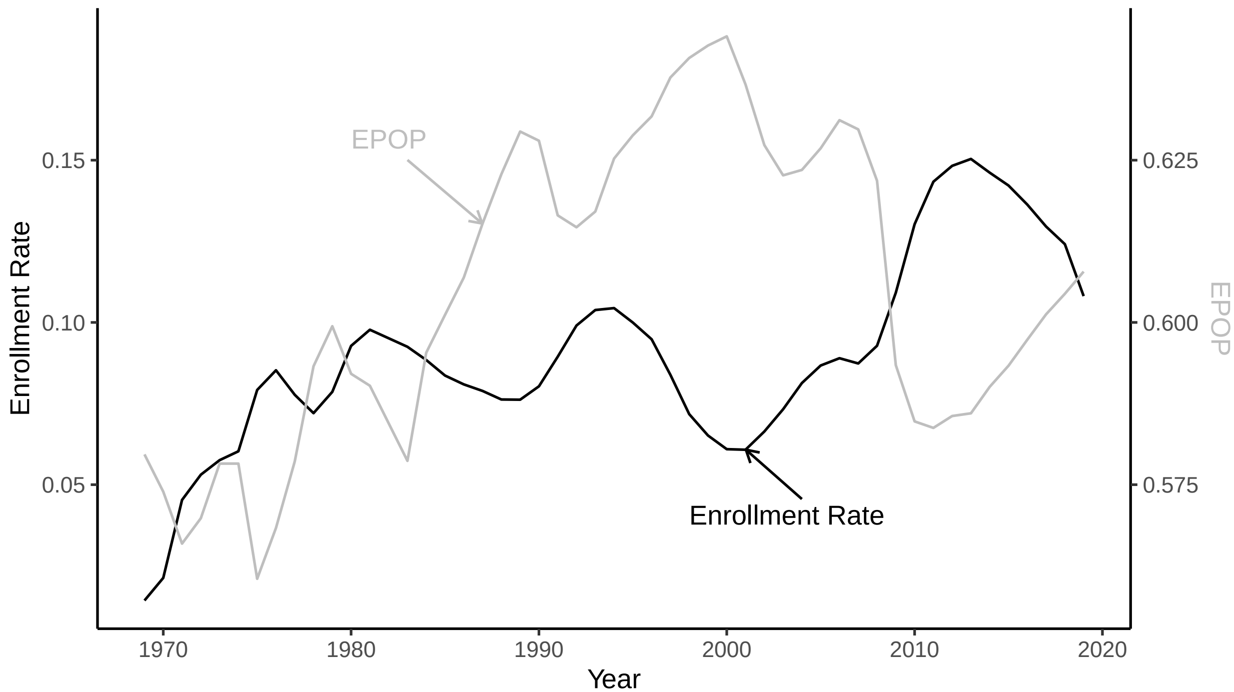


(b) County's Poverty Rate Decile × Year Fixed Effects



Notes: Figure shows the results of estimating Equation 4 with additional controls. Panel (a) adds controls for the county's take-up rate interacted with year fixed effects. The take-up rate is defined as the number of individuals enrolled as a share of the number of individuals eligible to enroll; we use the number of individuals enrolled in the county in 1972 and proxy for the number eligible with the number of people in poverty in the county in 1970. Panel (b) adds controls for the county's decile of the 1970 poverty rate, interacted with year fixed effects. Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level. N = 1,804 counties.

Figure A6: National Trends in Food Stamp Enrollment Rate and Employment-to-Population Ratio



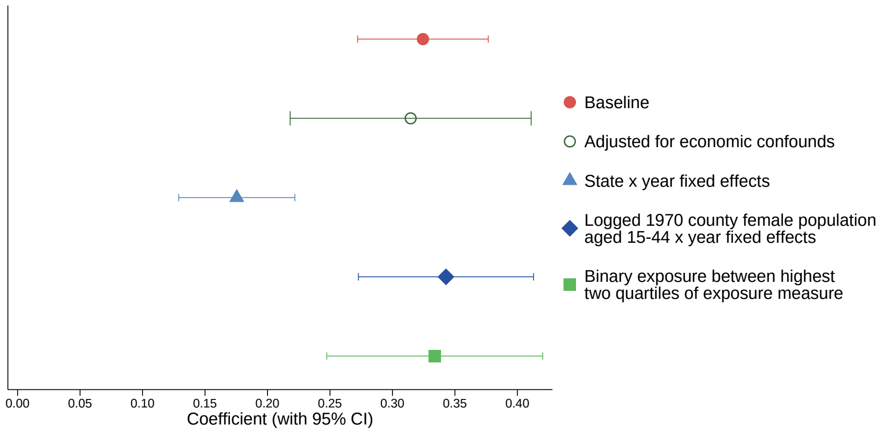
Notes: Figure shows national trends in the food stamp enrollment rate and the employment-to-population ratio (EPOP) over the half century from 1969 to 2019. Data on the national headcount of food stamp enrollees comes from the USDA (USDA Food and Nutrition Service, 2025), data on U.S. population comes from the SEER database (National Cancer Institute, 2026), and data on EPOP comes from the U.S. Bureau of Labor Statistics via the St. Louis Federal Reserve's FRED database (U.S. Bureau of Labor Statistics, 2026). Food Stamp Program enrollment rate is plotted in black and EPOP in gray.

Figure A7: National Trend in Total Births

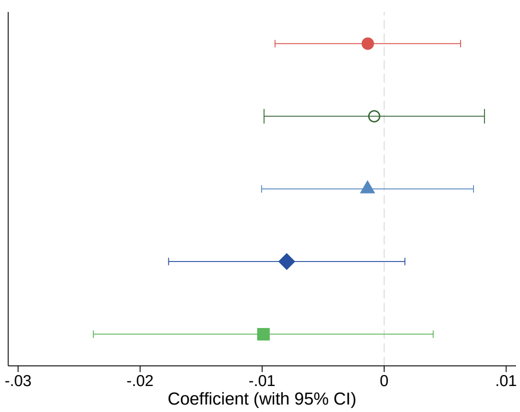
Notes: Figure shows the number of births nationally (across all states except Alaska).

Figure A8: Free Food Stamp Reform: Impacts Across Specifications

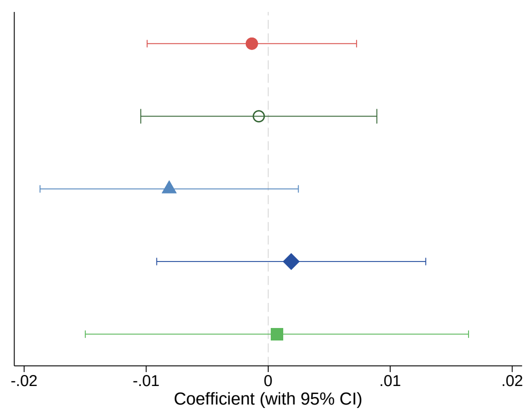
(a) Food Stamp Program Enrollment



(b) Log Birth Weight

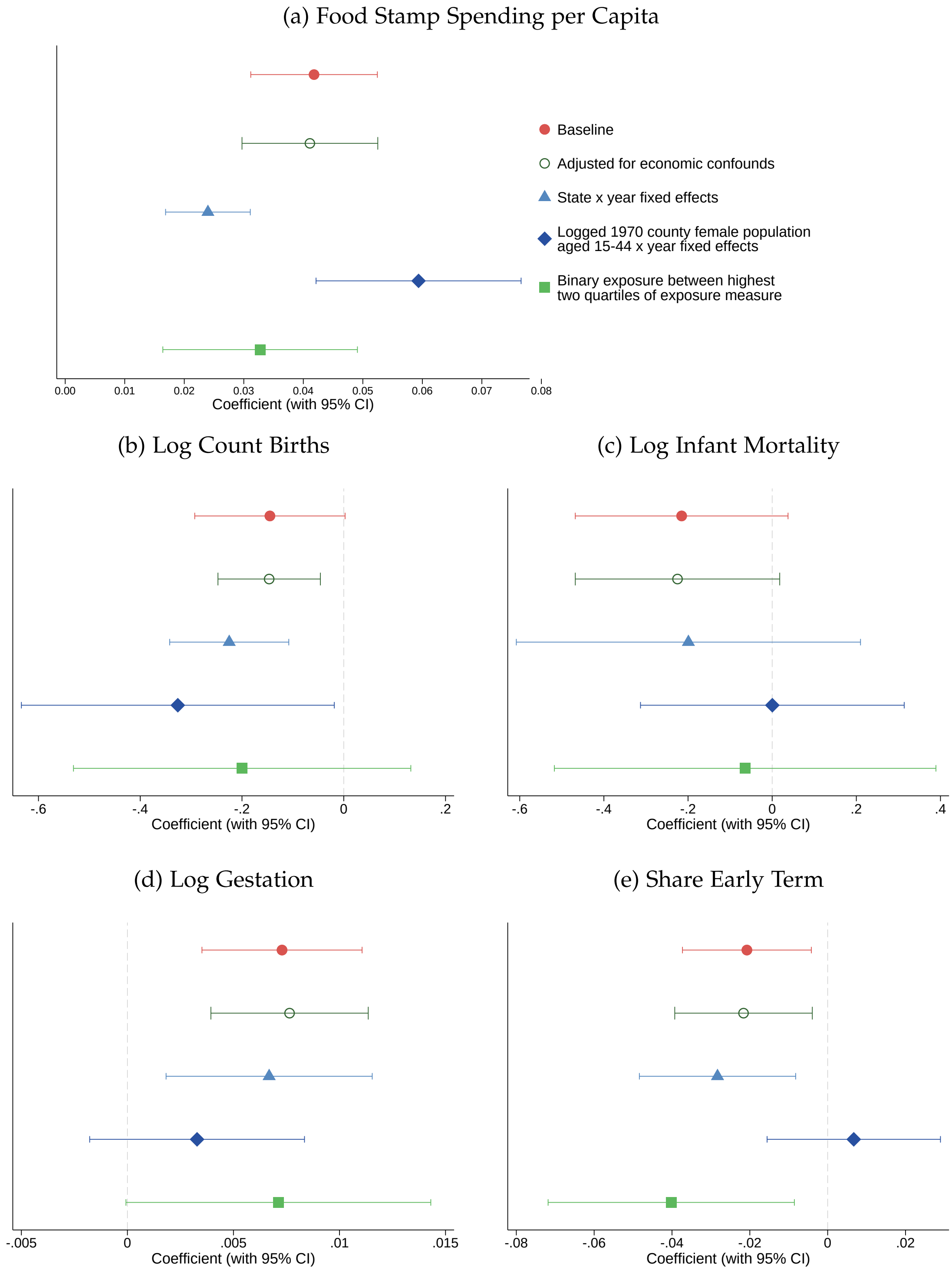


(c) Share Low Birth Weight



Notes: Figure presents estimates of β_{avg} (as defined in Equation 6) from estimating Equation 5 (or in the case of Panel (a) of δ_{avg} from estimating Equation 4) from the baseline specification (i.e., Equations 4 and 5) and four alternative specifications. The "adjusted for economic confounds" specification implements the proxy approach for economic confounds from Freyaldenhoven et al. (2019), using county-year data on log employment counts in industries unlikely to have many workers in food stamps as a proxy for the local business cycle; Appendix C.1. The "state $\times$ year" and "Logged 1970 county female population aged 15–44 $\times$ year" specifications add those respective fixed effects. The "binary exposure between highest two quartiles" estimates Equation 5 using a binary exposure measure for counties in quartile four of the exposure measure (comparison group quartile three); note that the estimates from this specification have been rescaled (divided by the difference in average exposure in Q4 and Q3) to make them comparable to the estimates using a continuous exposure measure. Across specifications, estimates are weighted by 1970 county population and cluster standard errors at the county level. Appendix Figures A10 through A12 present the underlying event studies for all of these specifications.

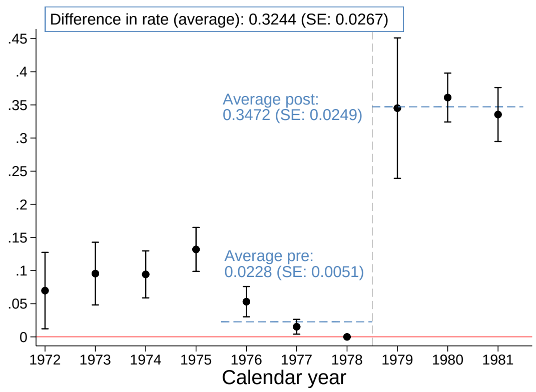
Figure A9: Free Food Stamp Reform: Impacts on Supplemental Outcomes



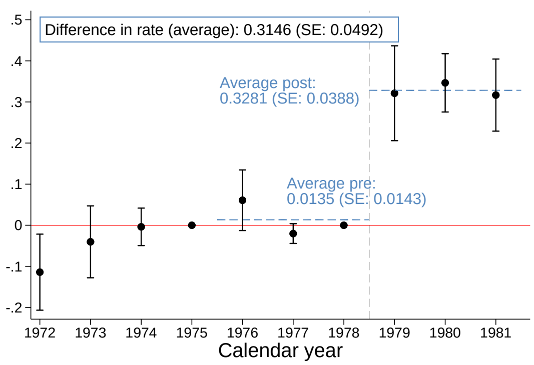
Notes: Figure presents estimates of β_{avg} (as defined in Equation 6) from estimating Equation 4 in Panel (a) for spending per capita (thousands of 1970 dollars) or Equation 5 in all other panels, the baseline specification (in Equation 5) and four alternative specifications. The "adjusted for economic confounds" specification implements the proxy approach for economic confounds from Freyaldenhoven et al. (2019), using county-year data on log employment counts in industries unlikely to have many workers in food stamps as a proxy for the local business cycle; Appendix C.1. The "state $\times$ year" and "Logged 1970 county female population aged 15–44 $\times$ year" specifications add those respective fixed effects. The "binary exposure between highest two quartiles" estimates Equations 4 and 5 using a binary exposure measure for counties in quartile four of the exposure measure (comparison group quartile three); note that the estimates from this specification have been rescaled (divided by the difference in average exposure in Q4 and Q3) to make them comparable to the estimates using a continuous exposure measure. Across specifications, estimates are weighted by 1970 county population and cluster standard errors at the county level. Appendix Figures A13 through A17 present the underlying event studies.

Figure A10: Impact of Free Food Stamps Reform on Enrollment Rate: Alternative Specifications

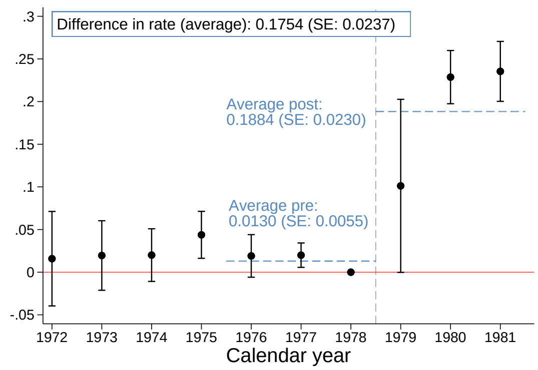
(a) Baseline



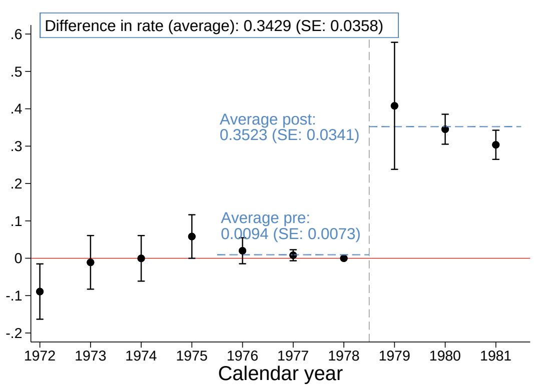
(b) Adjusted for Confounds Using Unexposed Employment



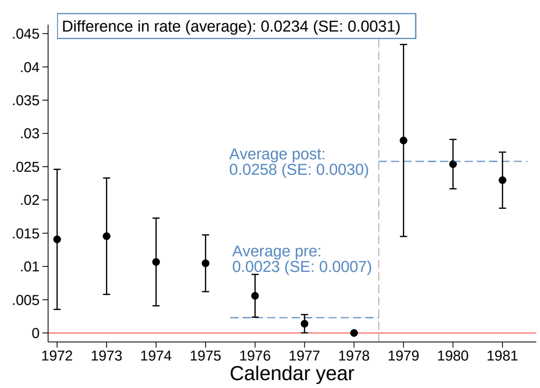
(c) State × Year Fixed Effects



(d) Log Population Women Aged 15–44 × Year Fixed Effects



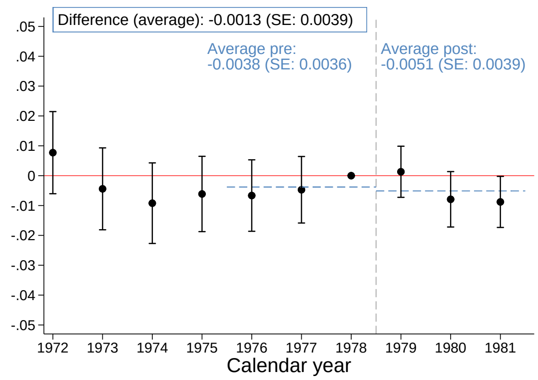
(e) Binary Exposure Measure (Q4 vs. Q3)



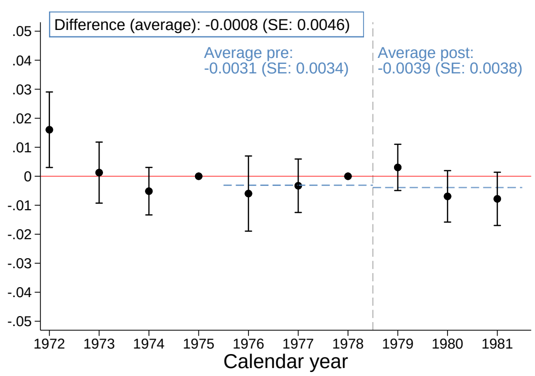
Notes: Figure shows the results of estimating Equation 5 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A8 for more details on each alternative specification. N = 1,804 counties for the baseline sample in Panel (a). N = 1,801 counties in Panel (b) and N = 1,785 counties in Panel (d) due to missing data on employment and population counts respectively. N = 1,803 counties in Panel (c) due to absorption of within-state variation. Panel (e) is limited to counties in the top half of the exposure distribution (N= 1,613). Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level.

Figure A11: Impact of Free Food Stamps Reform on Log Birth Weight: Alternative Specifications

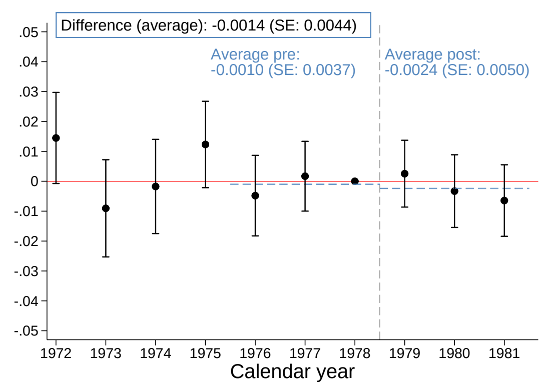
(a) Baseline



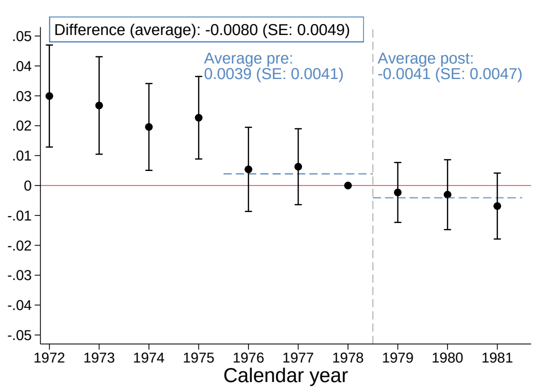
(b) Adjusted for Confounds Using Unexposed Employment



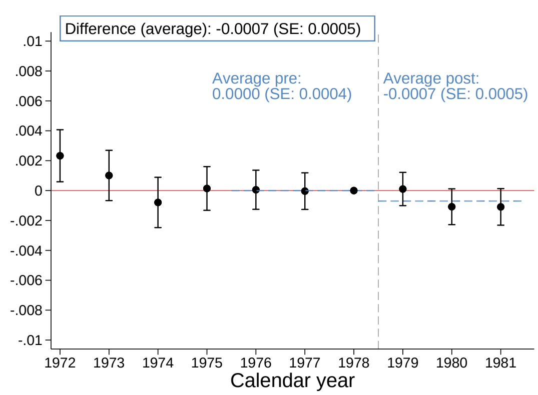
(c) State × Year Fixed Effects



(d) Log Population Women Aged 15–44 × Year Fixed Effects



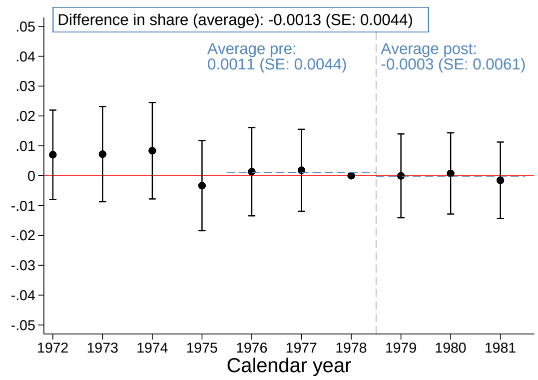
(e) Binary Exposure Measure (Q4 vs. Q3)



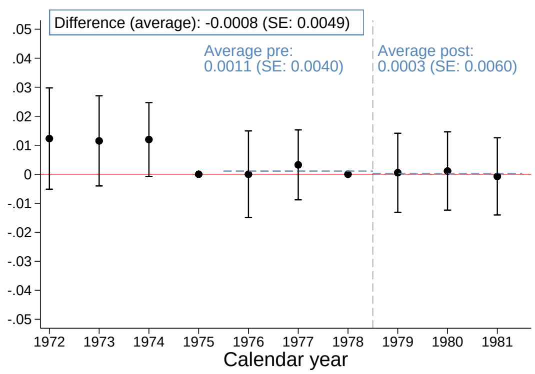
Notes: Figure shows the results of estimating Equation 5 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A8 for more details on each alternative specification. N = 1,804 counties for the baseline sample in Panel (a). N = 1,801 counties in Panel (b) and N = 1,785 counties in Panel (d) due to missing data on employment and population counts respectively. N = 1,803 counties in Panel (c) due to absorption of within-state variation. Panel (e) is limited to counties in the top half of the exposure distribution (N = 1,613 counties). Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level.

Figure A12: Impact of Free Food Stamps Reform on Share Low Birth Weight: Alternative Specifications

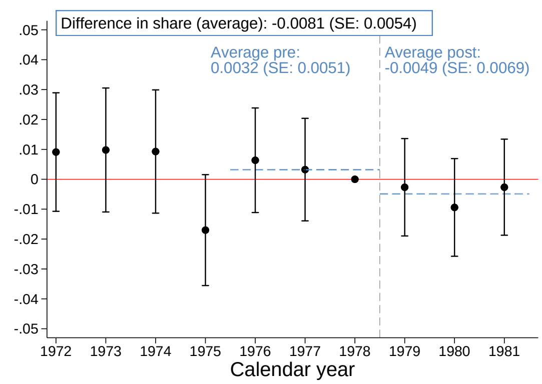
(a) Baseline



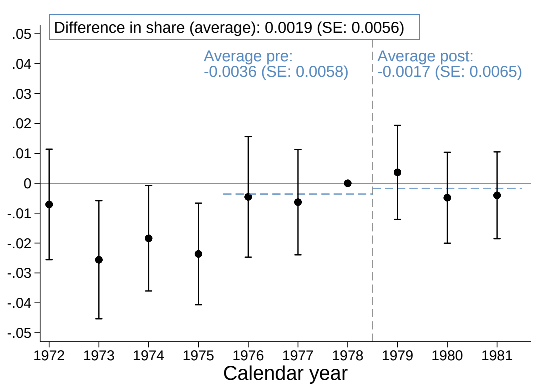
(b) Adjusted for Confounds
Using Unexposed Employment



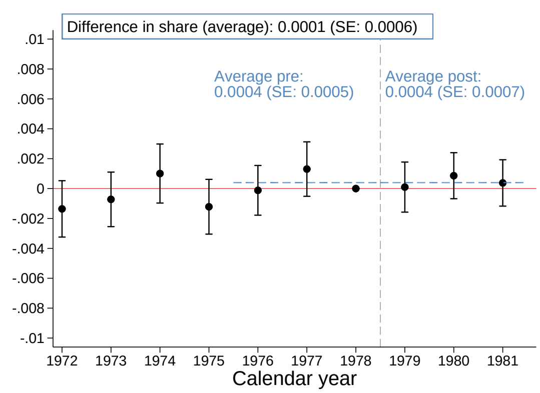
(c) State × Year Fixed Effects



(d) Log Population Women Aged 15–44
× Year Fixed Effects



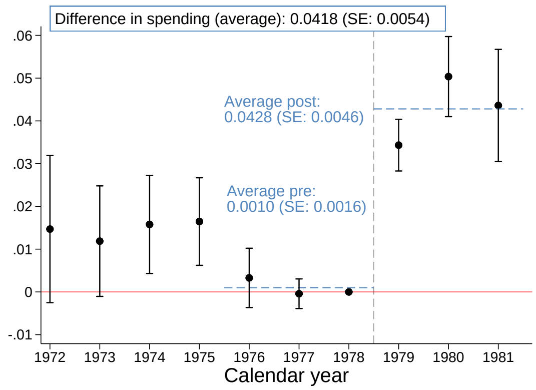
(e) Binary Exposure Measure (Q4 vs. Q3)



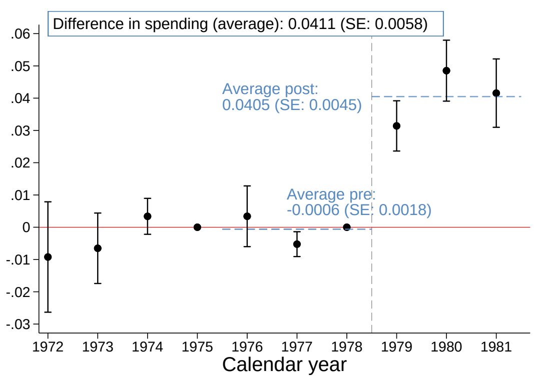
Notes: Figure shows the results of estimating Equation 5 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A8 for more details on each alternative specification. N = 1,804 counties for the baseline sample in Panel (a). N = 1,801 counties in Panel (b) and N = 1,785 counties in Panel (d) due to missing data on employment and population counts respectively. N = 1,803 counties in Panel (c) due to absorption of within-state variation. Panel (e) is limited to counties in the top half of the exposure distribution (N = 1,613 counties). Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level.

Figure A13: Impact of Free Food Stamps Reform on Food Stamp Spending Per Capita

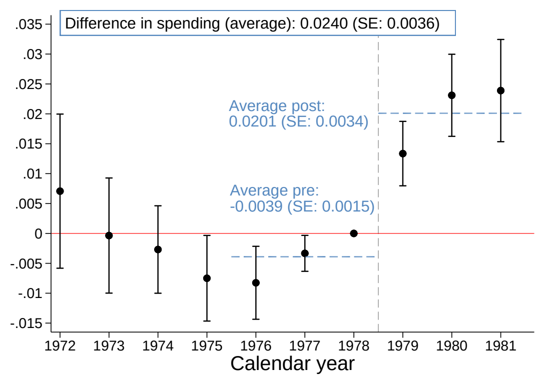
(a) Baseline



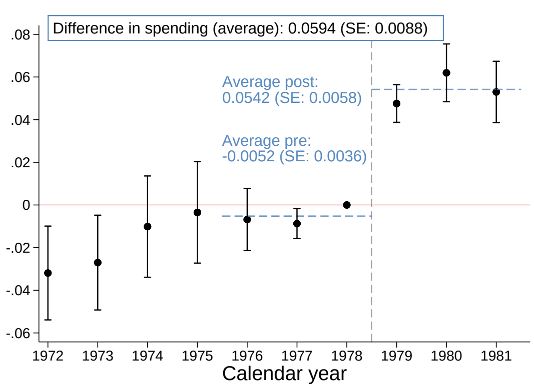
(b) Adjusted for Confounds
Using Unexposed Employment



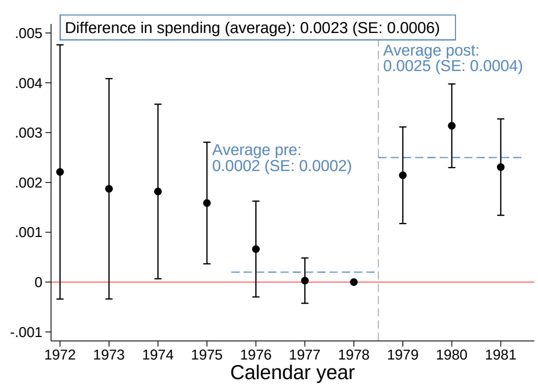
(c) State × Year Fixed Effects



(d) Log Population Women Aged 15–44
× Year Fixed Effects



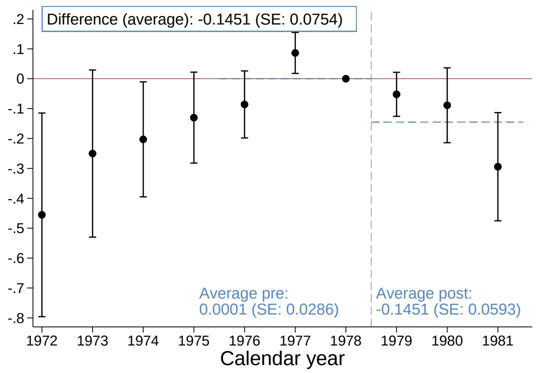
(e) Binary Exposure Measure (Q4 vs. Q3)



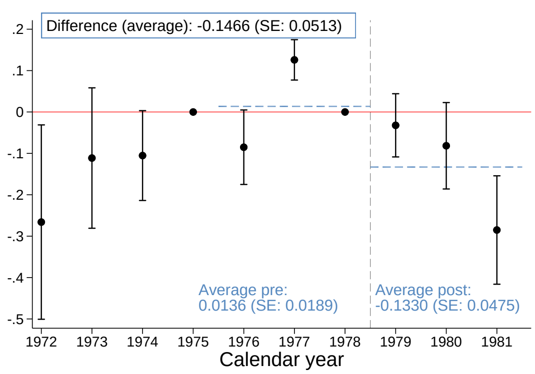
Notes: Figure shows the results of estimating Equation 4. Dependent variable is county-year food stamp spending per capita (in thousands of 1970 dollars). We show results for the baseline specification (Panel a) and four alternative specifications (Panels b through e); see notes to Appendix Figure A8 for more details on each alternative specification. N = 1,774 counties for the baseline sample in Panel (a). N = 1,771 counties in Panel (b) and N = 1,765 counties in Panel (d) due to missing data on employment and population counts respectively. N = 1,773 counties in Panel (c) due to absorption of within-state variation. Panel (e) is limited to counties in the top half of the exposure distribution (N = 1,584 counties). Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level.

Figure A14: Impact of Free Food Stamps Reform on Log Count of Births: Alternative Specifications

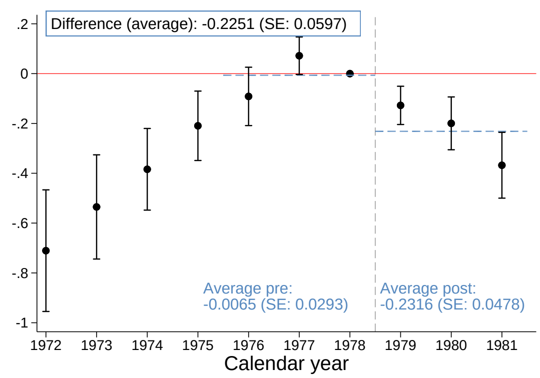
(a) Baseline



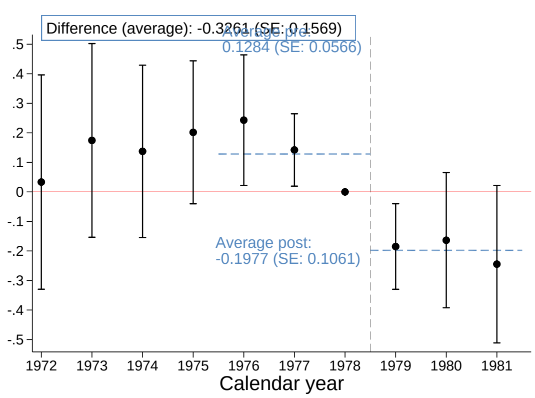
(b) Adjusted for Confounds Using Unexposed Employment



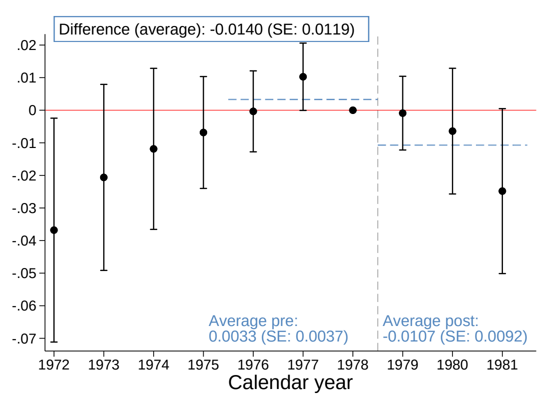
(c) State × Year Fixed Effects



(d) Log Population Women Aged 15–44 × Year Fixed Effects



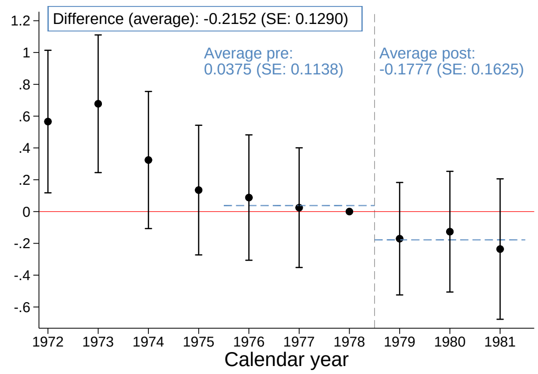
(e) Binary Exposure Measure (Q4 vs. Q3)



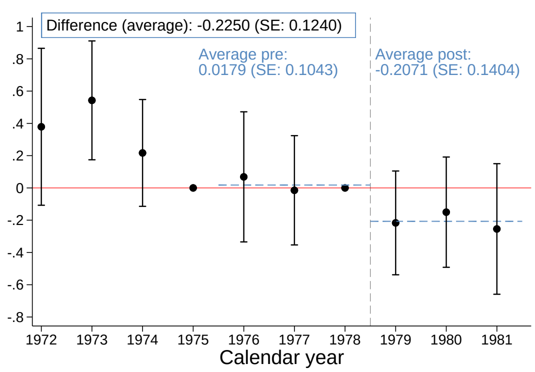
Notes: Figure shows the results of estimating Equation 5 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A8 for more details on each alternative specification. N = 1,804 counties for the baseline sample in Panel (a). N = 1,801 counties in Panel (b) and N = 1,785 counties in Panel (d) due to missing data on employment and population counts respectively. N = 1,803 counties in Panel (c) due to absorption of within-state variation. Panel (e) is limited to counties in the top half of the exposure distribution (N = 1,613 counties). Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level.

Figure A15: Impact of Free Food Stamps Reform on
Log Infant Mortality Rate (Deaths per 1,000 Births):
Alternative Specifications

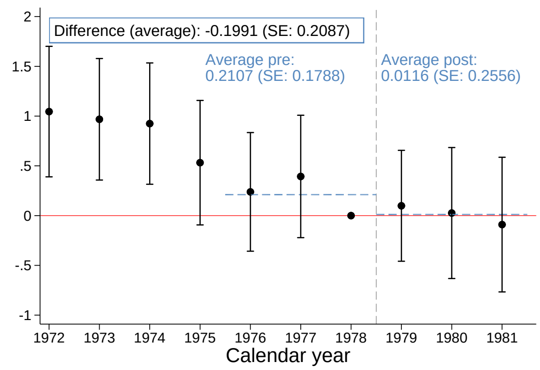
(a) Baseline



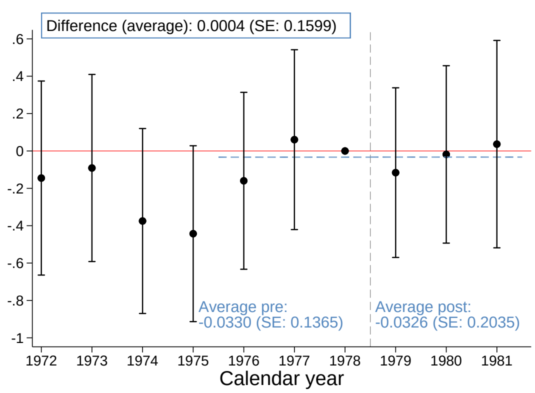
(b) Adjusted for Confounds
Using Unexposed Employment



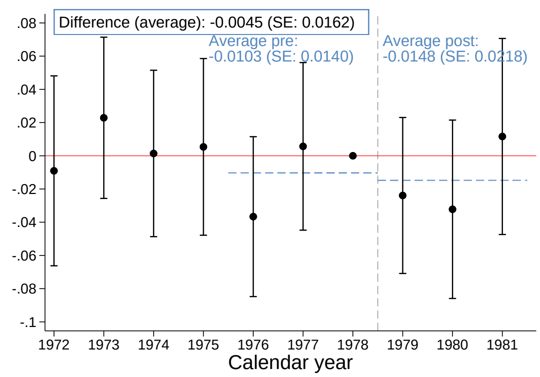
(c) State × Year Fixed Effects



(d) Log Population Women Aged 15–44
× Year Fixed Effects



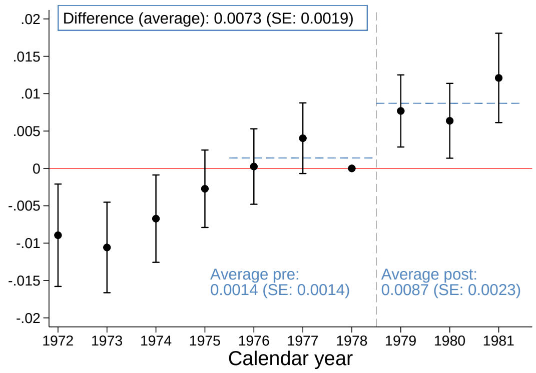
(e) Binary Exposure Mmeasure (Q4 vs. Q3)



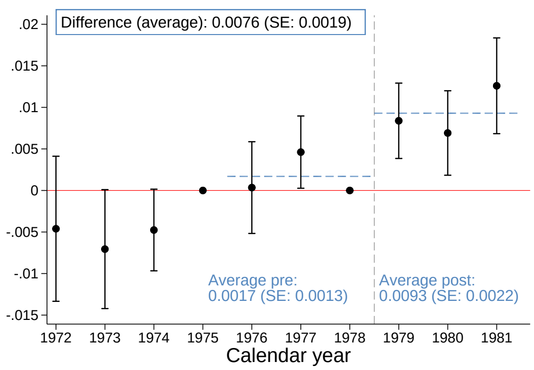
Notes: Figure shows the results of estimating Equation 5 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A8 for more details on each alternative specification. See Appendix C.5 for outcome-specific limitations applied to the free food stamp reform baseline sample. N = 1,145 counties for the baseline estimates in Panel (a). N = 1,144 counties in Panel (b) and N = 1,129 counties in Panel (d) due to missing data on employment and population counts respectively. N = 1,144 counties in Panel (c) due to absorption of within-state variation. Panel (e) is limited to counties in the top half of the exposure distribution (N = 972 counties). Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level.

Figure A16: Impact of Free Food Stamps Reform on Log Gestational Age:
Alternative Specifications

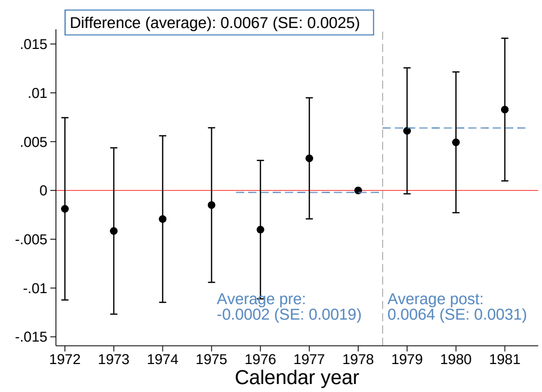
(a) Baseline



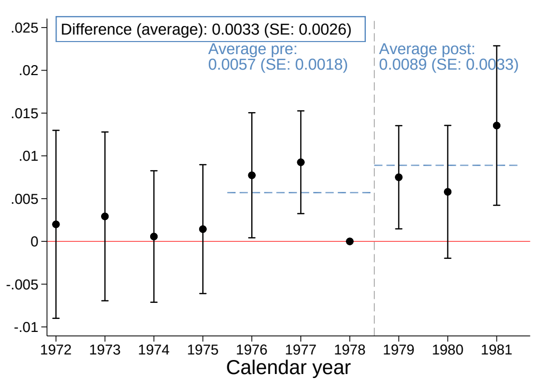
(b) Adjusted for Confounds
Using Unexposed Employment



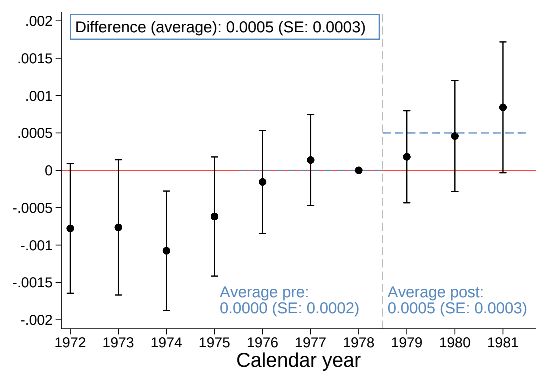
(c) State × Year Fixed Effects



(d) Log Population Women Aged 15–44
× Year Fixed Effects



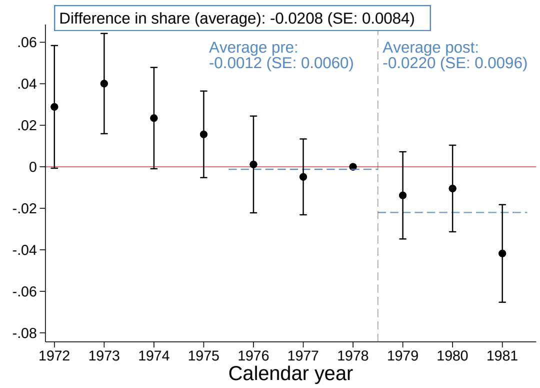
(e) Binary Exposure Measure (Q4 vs. Q3)



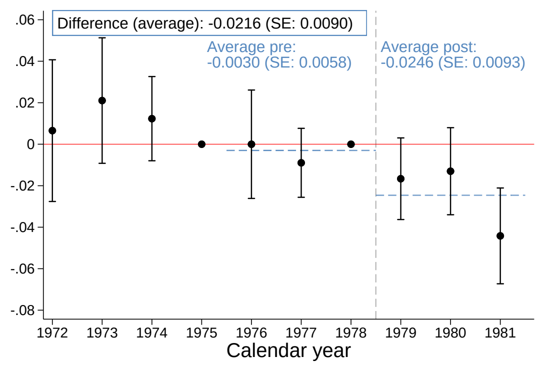
Notes: Figure shows the results of estimating Equation 5 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A8 for more details on each alternative specification. See Appendix C.5 for outcome-specific limitations applied to the free food stamp reform baseline sample. N = 1,359 counties for the baseline estimates in Panel (a). N = 1,357 counties in Panel (b) and N = 1,352 counties in Panel (d) due to missing data on employment and population counts respectively. N = 1,358 counties in Panel (c) due to absorption of within-state variation. Panel (e) is limited to counties in the top half of the exposure distribution (N = 1,193 counties). Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level.

Figure A17: Impact of Free Food Stamps Reform on Share Pre-Term (< 37 weeks): Alternative Specifications

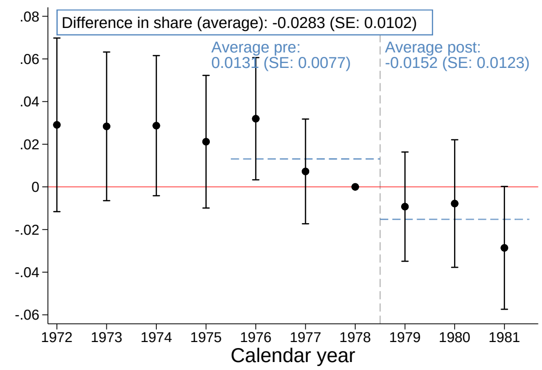
(a) Baseline



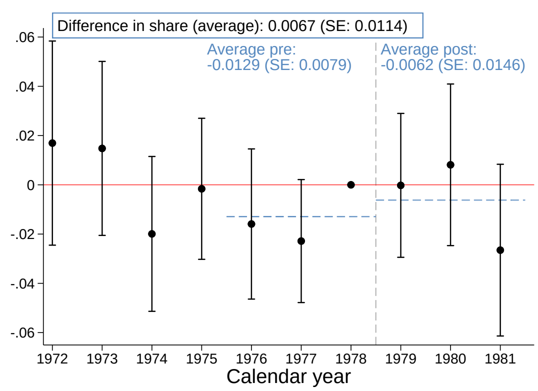
(b) Adjusted for Confounds Using Unexposed Employment



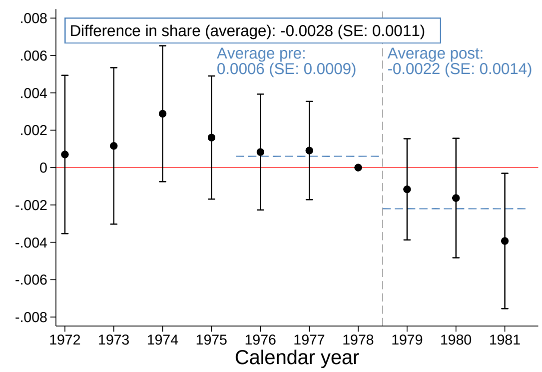
(c) State × Year Fixed Effects



(d) Log Population Women Aged 15–44 × Year Fixed Effects



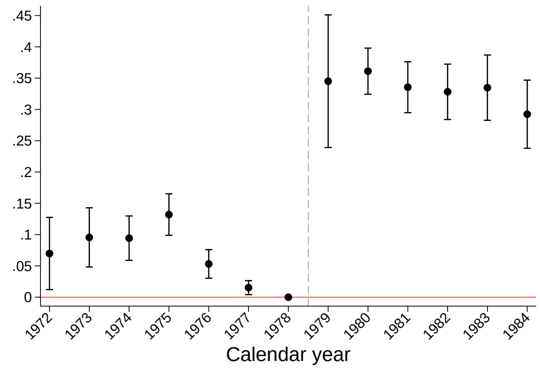
(e) Binary Exposure Measure (Q4 vs. Q3)



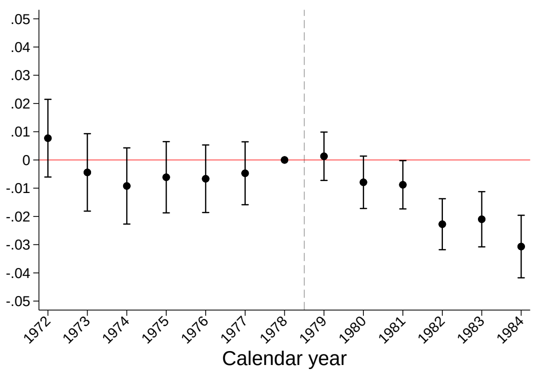
Notes: Figure shows the results of estimating Equation 5 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A8 for more details on each alternative specification. See Appendix C.5 for outcome-specific limitations applied to the free food stamp reform baseline sample. N = 1,359 counties for the baseline estimates in Panel (a). N = 1,357 counties in Panel (b) and N = 1,352 counties in Panel (d) due to missing data on employment and population counts respectively. N = 1,358 counties in Panel (c) due to absorption of within-state variation. Panel (e) is limited to counties in the top half of the exposure distribution (N = 1,193 counties). Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level.

Figure A18: Percentage Impacts of Later Post-reform Period on Share within Birth Weight Ranges

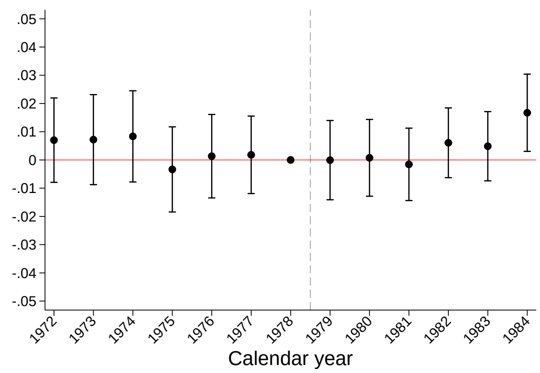
(a) Food Stamp Program Enrollment



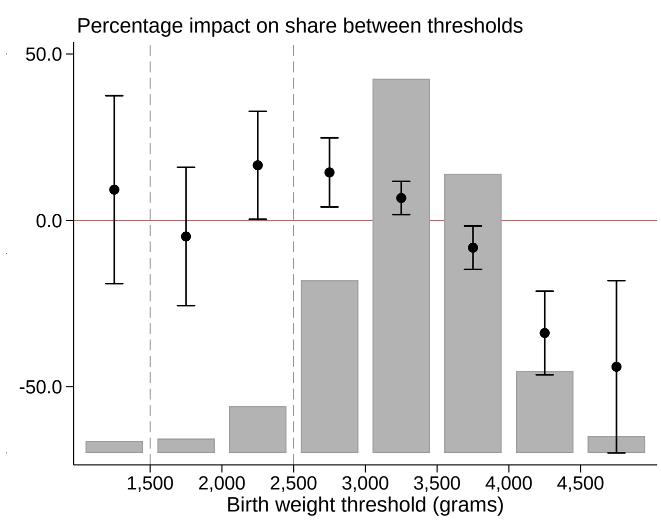
(b) Log Birth Weight



(c) Share Low Birth Weight



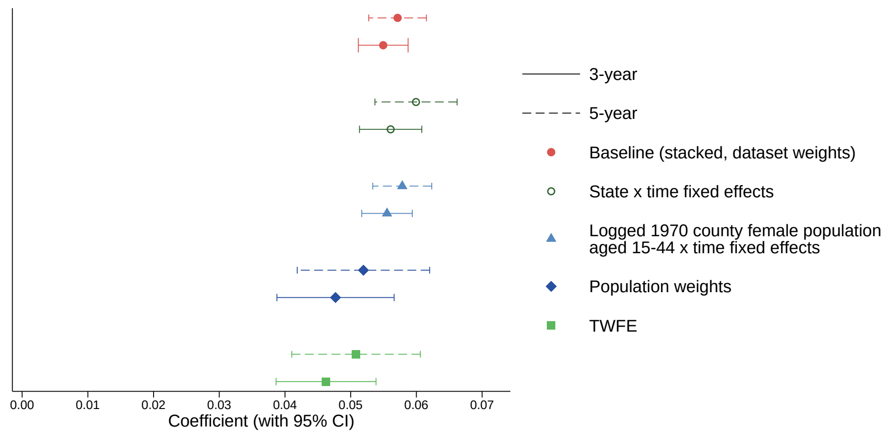
(d) Percentage Impacts of Later Post-reform Period on Share within Birth Weight Ranges



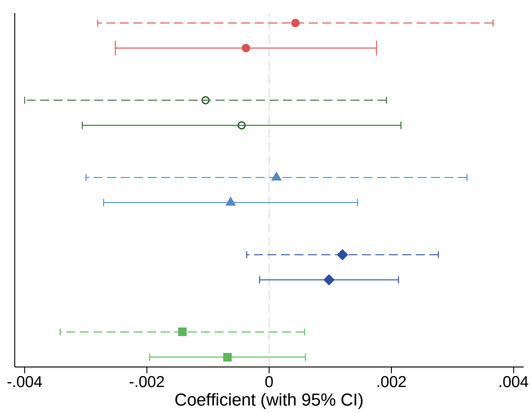
Notes: Panels (a), (b), and (c) show the results of estimating Equations 4 and 5 on enrollment and birth weight outcomes using a continuous 1972 exposure measure for an extended post-reform period. Panel (d) shows the scaled results of estimating the reduced form Equation 5 using a continuous 1972 exposure measure on infant birth weight ranges (left y-axis) and the share of births across the period 1976-1978 that fall within each range (relative shares depicted by gray bars). Estimates are scaled by outcome means across the period 1976-1978. Therefore, the estimates show the percentage impact of the later post-reform period (1982-1984) on the fraction of births in the county-year cell that fall within each range of specified number of grams. We define the ranges as inclusive of the lower threshold and exclusive of the upper threshold. For example, we define the first outcome range as "< 1,500 grams" followed by "≥ 1,500 and < 2,000." Similarly, we define the last range as "≥ 4,500." Note that the vertical dashed lines denote the thresholds for very low birth weight (< 1,5000 grams) and low birth weight (< 2,500 grams). The sample is limited to the baseline analysis sample for the free food stamp reform (see Appendix Table A1). Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level. N = 1,804 counties.

Figure A19: Rollout Impacts Across Specifications
Pooled Births

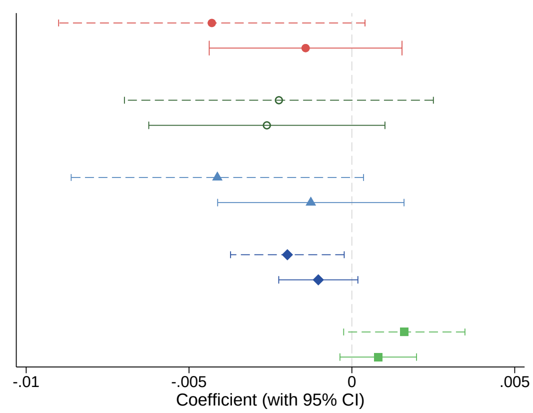
(a) Food Stamp Program Enrollment



(b) Log Birth Weight



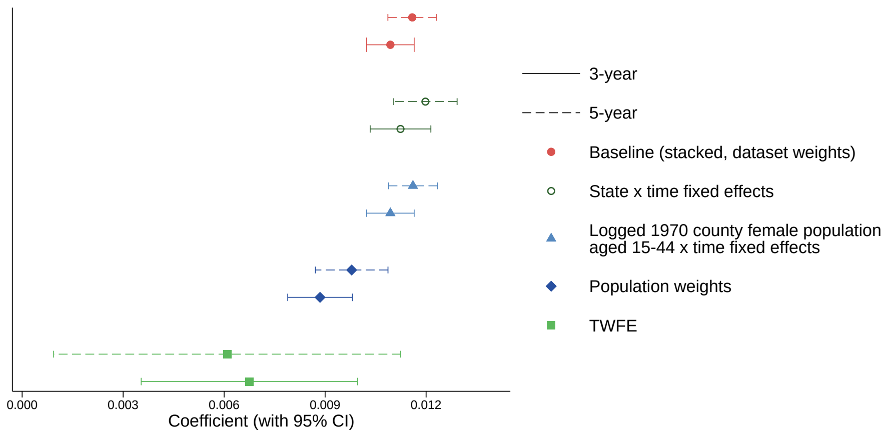
(c) Share Low Birth Weight



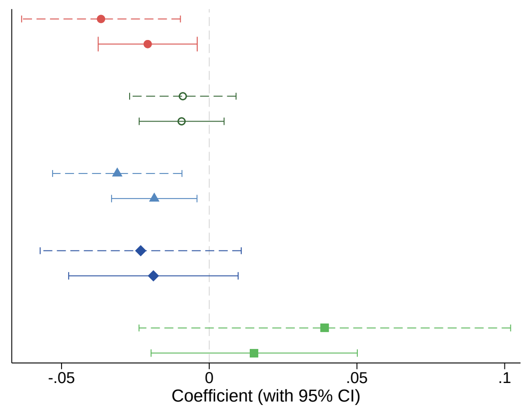
Notes: Figure compiles the 3-year and 5-year difference coefficients (as in Equation 28) and 95% CIs from the baseline specification and four alternative specifications for pooled races. Across specifications, estimates we cluster standard errors at the county level. The baseline specification estimates the event study specification from Equations 7 and 8 using a stacked dataset and dataset weights. The "population weights" specification estimates Equation 7 and 8 for an outcome of interest, using the baseline stacked specification with 1970 county population weights. The "state $\times$ time" and "Logged 1970 county female population aged 15–44 $\times$ time specifications add those respective fixed effects, where time is year or unique year-month depending upon data frequency. The "TWFE" estimates Equations 23 and 24 using a flat dataset, population weights, and binning at relative years beyond -5 and +5. Appendix Figures A21 through A23 present the underlying event study estimates.

Figure A20: Rollout Impacts Across Specifications
Pooled Births, Supplemental Outcomes

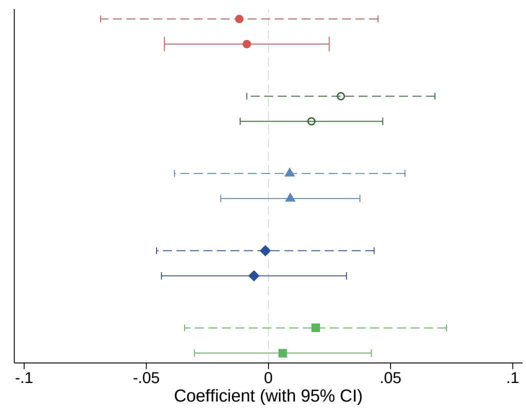
(a) Food Stamp Program Spending per Capita



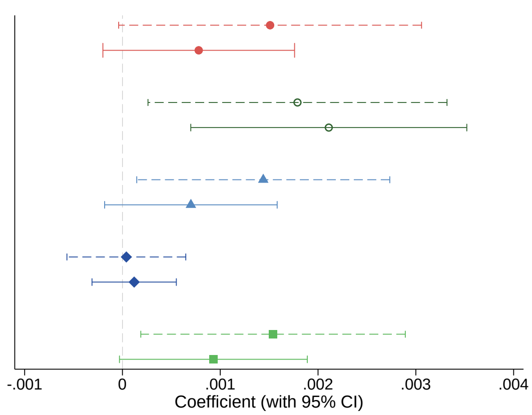
(b) Log Count Births



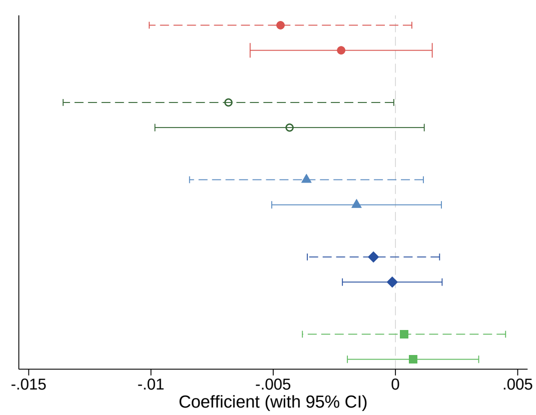
(c) Log Infant Mortality



(d) Log Gestation



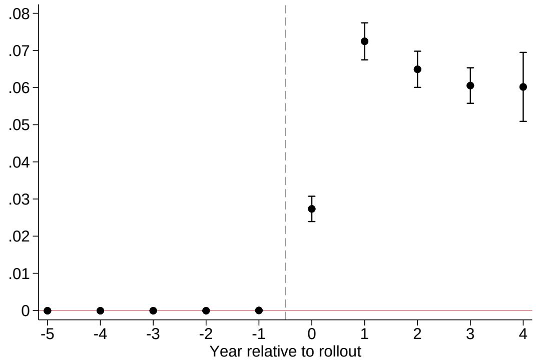
(e) Share Early Term



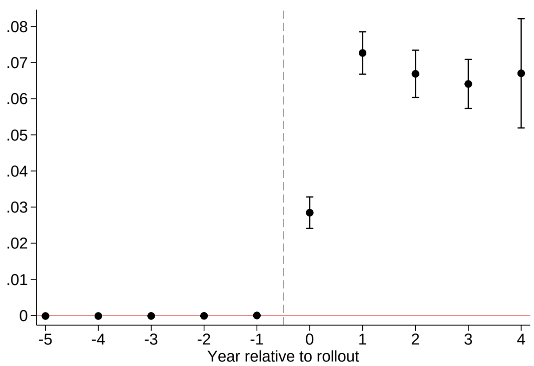
Notes: Figure compiles the 3-year and 5-year difference coefficients (as in Equation 28) and 95% CIs from the baseline specification and four alternative specifications for pooled races. Across specifications, we cluster standard errors at the county level. The baseline specification estimates the event study specification from Equations 7 and 8 using a stacked dataset and dataset weights. The "population weights" specification estimates Equations 7 and 8 for an outcome of interest, using the baseline stacked specification with 1970 county population weights. The "state × time" and "Logged 1970 county female population aged 15–44 × time specifications add those respective fixed effects, where time is year or unique year-month depending upon data frequency. The "TWFE" estimates Equations 23 and 24 using a flat dataset, population weights, and binning at relative years beyond -5 and +5. Appendix Figures A24 through A28 present the underlying event study estimates.

Figure A21: Rollout Impacts on Food Stamp Program Enrollment
Across Specifications

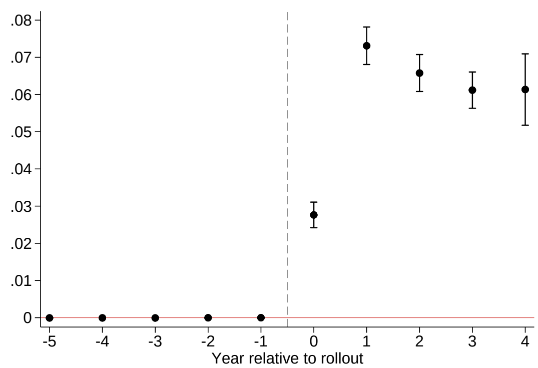
(a) Baseline



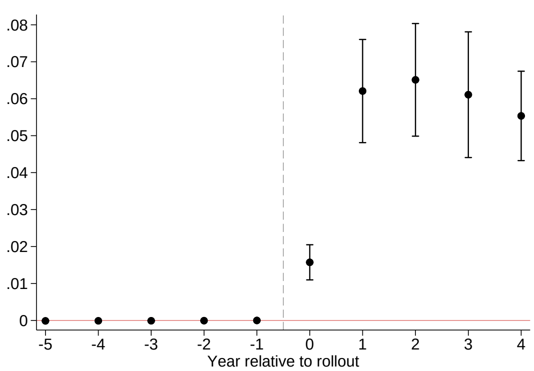
(b) State × Year Fixed Effects



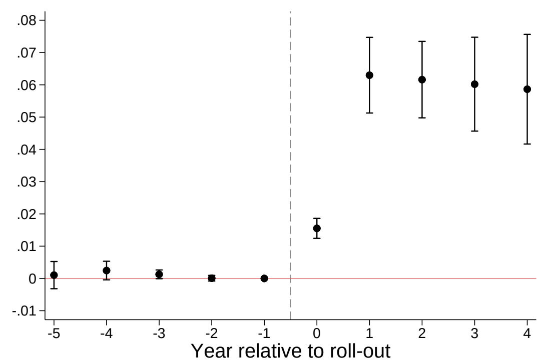
(c) Log Population Women Aged 15–44
× Year Fixed Effects



(d) 1970 County Population Weights



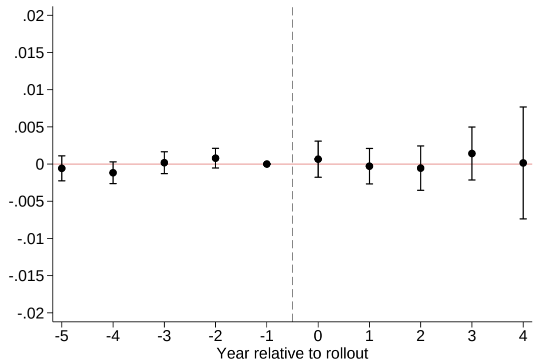
(e) TWFE



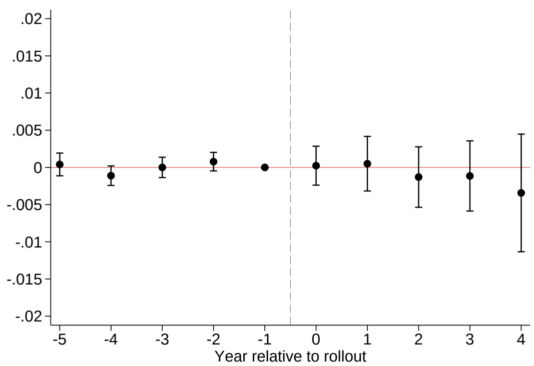
Notes: Figure shows the results of estimating β_r from Equation 7 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A19 for more details on each alternative specification. N = 1,520 counties for the baseline estimates in Panel (a), as well as for population-weighted estimates in Panel (d) and TWFE estimates in Panel (e). N = 1,506 counties in Panel (c) due to missing data on population counts. N = 1,519 counties in Panel (b) due to absorption of within-state variation. Standard errors are clustered at the county level.

Figure A22: Rollout Impacts on Log Birth Weight
Across Specifications

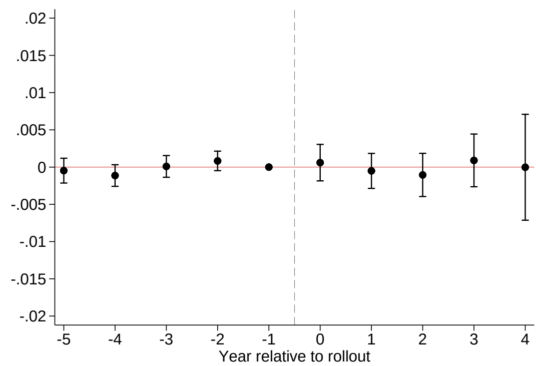
(a) Baseline



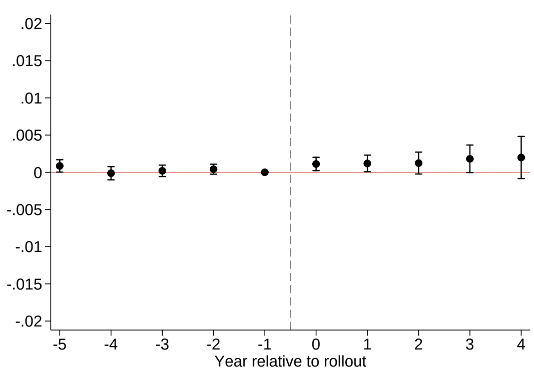
(b) State × Year Fixed Effects



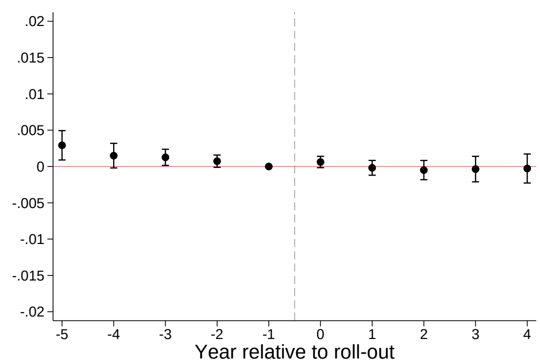
(c) Log Population Women Aged 15–44
× Year Fixed Effects



(d) 1970 County Population Weights



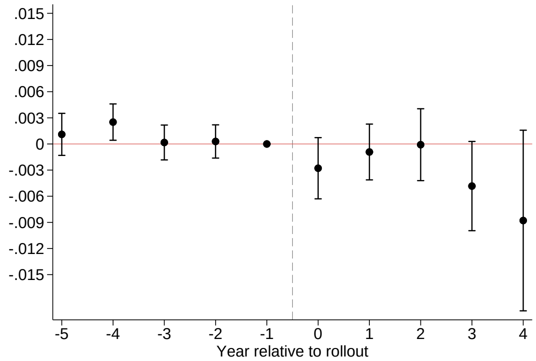
(e) TWFE



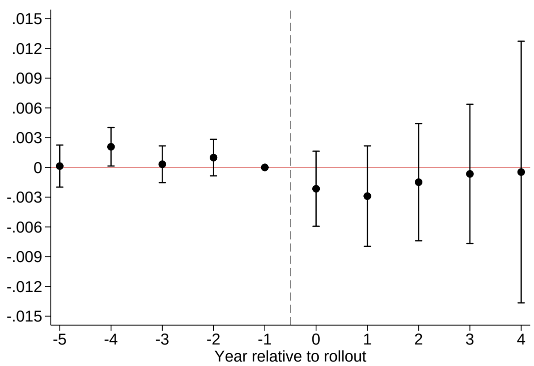
Notes: Figure shows the results of estimating β_r from Equation 8 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A19 for more details on each alternative specification. N = 1,520 counties for the baseline estimates in Panel (a), as well as for population-weighted estimates in Panel (d) and TWFE estimates in Panel (e). N = 1,506 counties in Panel (c) due to missing data on population counts. N = 1,519 counties in Panel (b) due to absorption of within-state variation. Standard errors are clustered at the county level.

Figure A23: Rollout Impacts on Share Low Birth Weight
Across Specifications

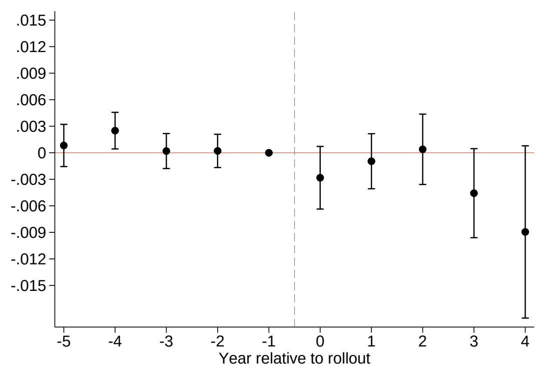
(a) Baseline



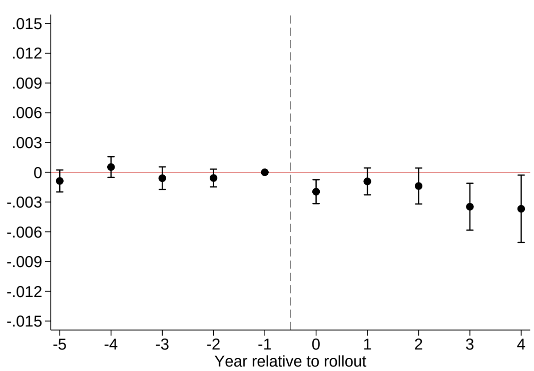
(b) State × Year Fixed Effects



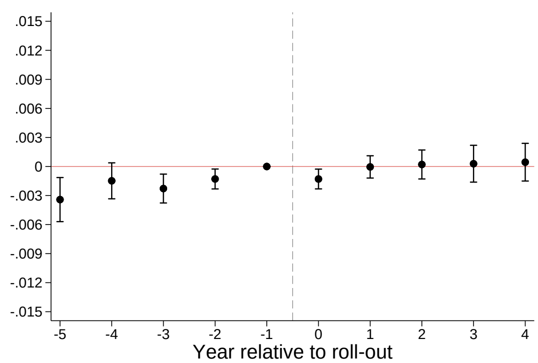
(c) Log Population Women Aged 15–44
× Year Fixed Effects



(d) 1970 County Population Weights



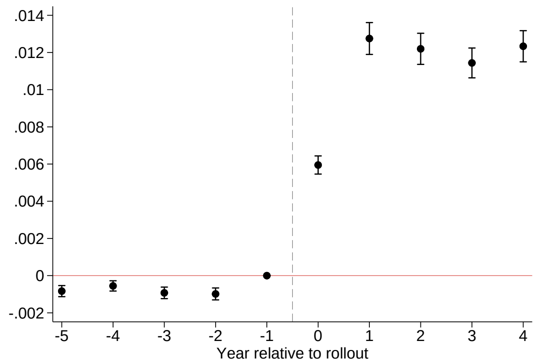
(e) TWFE



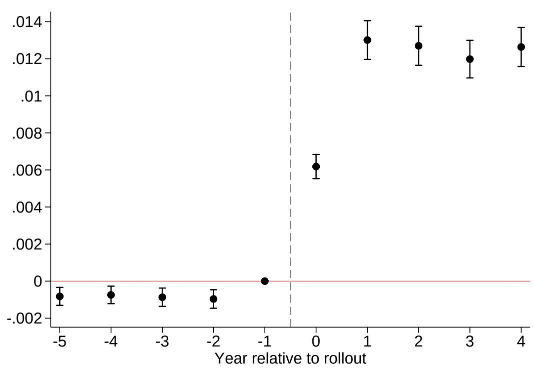
Notes: Figure shows the results of estimating β_r from Equation 8 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A19 for more details on each alternative specification. N = 1,520 counties for the baseline estimates in Panel (a), as well as for population-weighted estimates in Panel (d) and TWFE estimates in Panel (e). N = 1,506 counties in Panel (c) due to missing data on population counts. N = 1,519 counties in Panel (b) due to absorption of within-state variation. Standard errors are clustered at the county level.

Figure A24: Rollout Impacts on Food Stamps Spending per Capita Across Specifications

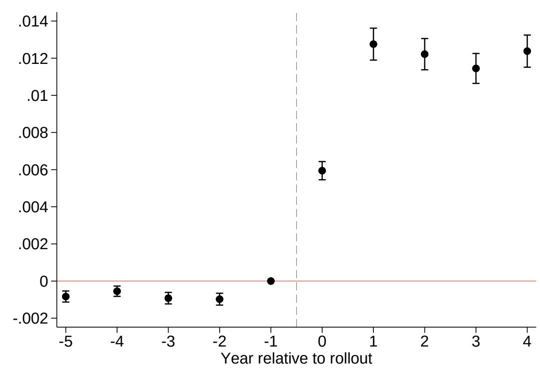
(a) Baseline



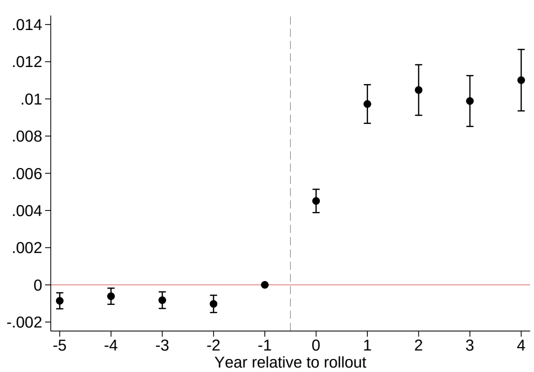
(b) State × Year Fixed Effects



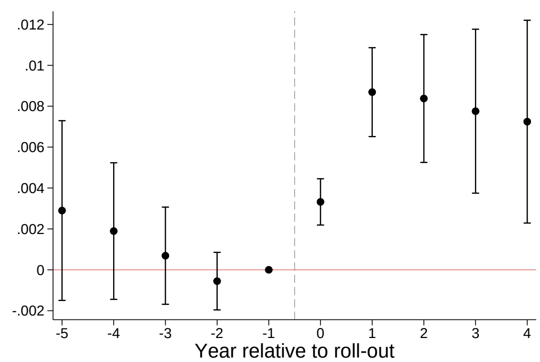
(c) Log Population Women Aged 15–44 × Year Fixed Effects



(d) 1970 County Population Weights



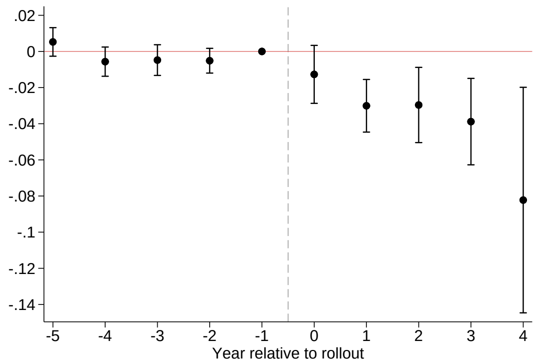
(e) TWFE



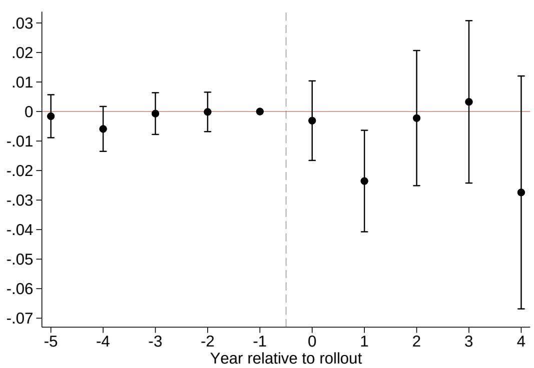
Notes: Figure shows the results of estimating β_r from Equation 8 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A19 for more details on each alternative specification. See Appendix C.5 for outcome-specific limitations applied to the rollout analysis baseline sample. N = 1,490 counties for the baseline estimates in Panel (a), as well as for population-weighted estimates in Panel (d) and TWFE estimates in Panel (e). N = 1,485 counties in Panel (c) due to missing data on population counts. N = 1,488 counties in Panel (b) due to absorption of within-state variation. Standard errors are clustered at the county level.

Figure A25: Rollout Impacts on Log Count of Births
Across Specifications

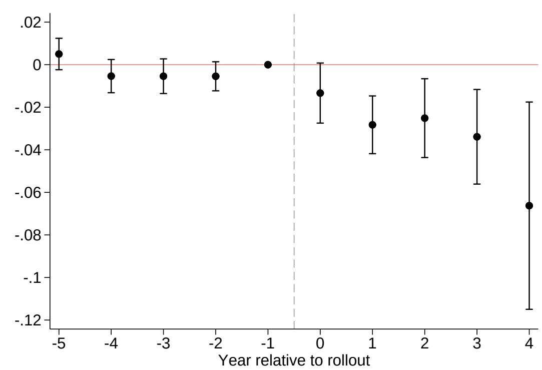
(a) Baseline



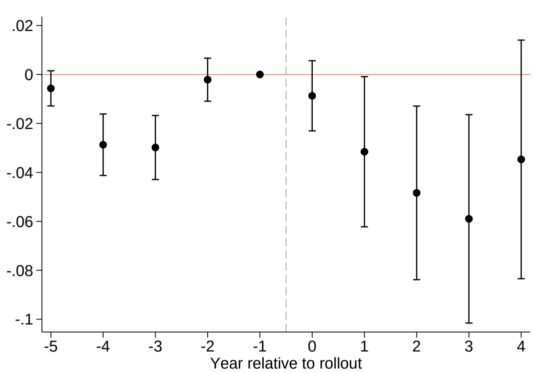
(b) State × Year Fixed Effects



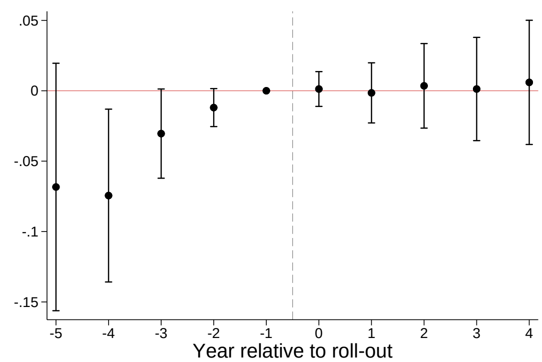
(c) Log Population Women Aged 15–44
× Year Fixed Effects



(d) 1970 County Population weigWts



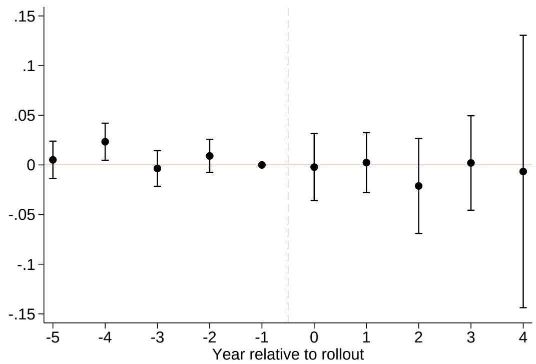
(e) TWFE



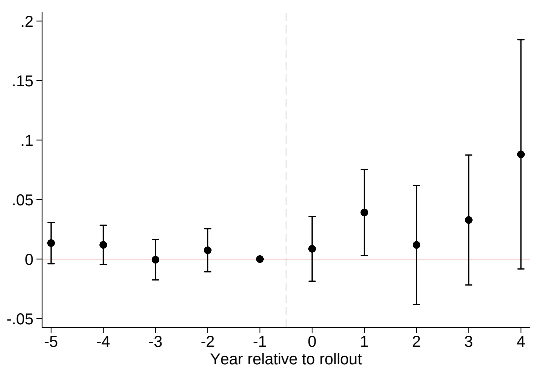
Notes: Figure shows the results of estimating β_r from Equation 8 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A19 for more details on each alternative specification. See Appendix C.5 for outcome-specific limitations applied to the rollout analysis baseline sample. N = 1,520 counties for the baseline estimates in Panel (a), as well as for population-weighted estimates in Panel (d) and TWFE estimates in Panel (e). N = 1,506 counties in Panel (c) due to missing data on population counts. N = 1,519 counties in Panel (b) due to absorption of within-state variation. Standard errors are clustered at the county level.

Figure A26: Rollout Impacts on Log Infant Mortality Rate
(Deaths < 12 Months per 1,000 Births)
Across Specifications

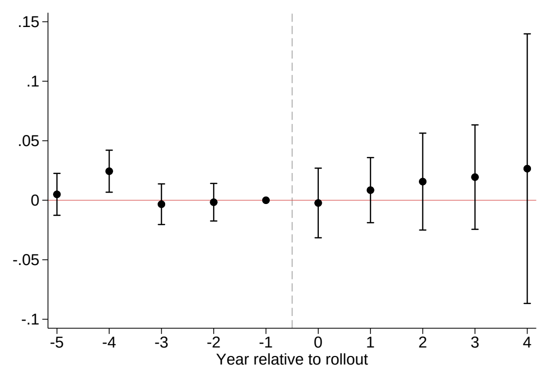
(a) Baseline



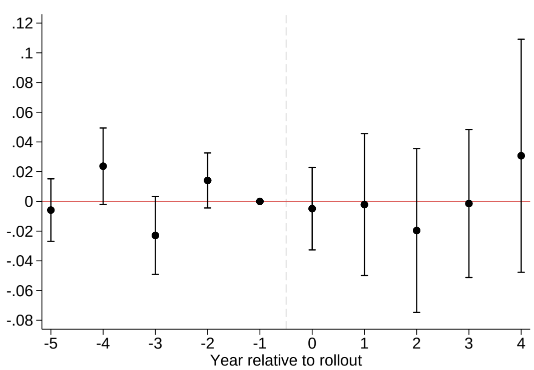
(b) State × Year Fixed Effects



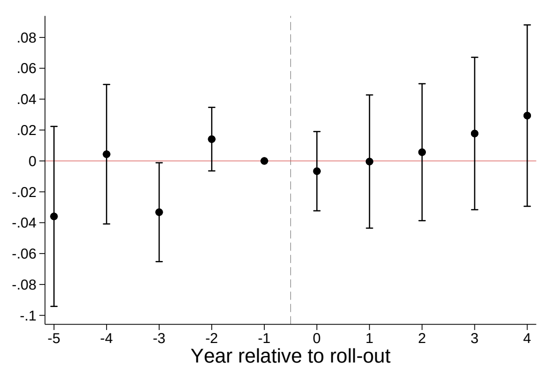
(c) Log Population Women Aged 15–44
× Year Fixed Effects



(d) 1970 County Population Weights



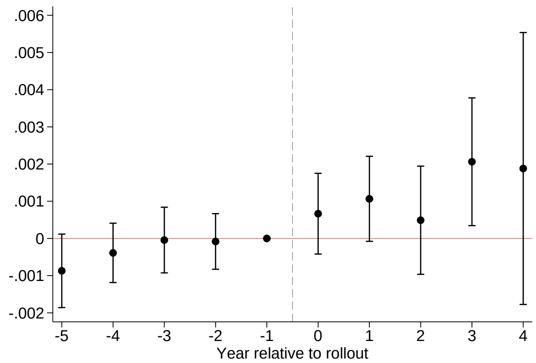
(e) TWFE



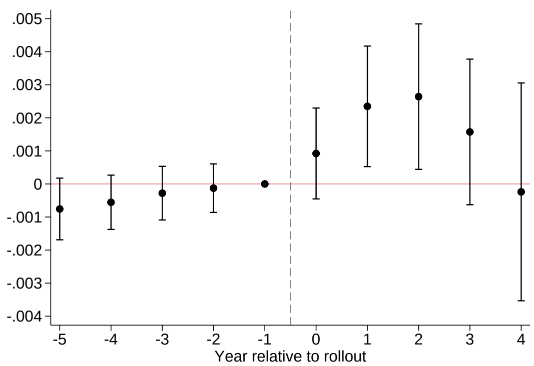
Notes: Figure shows the results of estimating β_r from Equation 8 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A19 for more details on each alternative specification. See Appendix C.5 for outcome-specific limitations applied to the rollout analysis baseline sample. N = 1,067 counties for the baseline estimates in Panel (a), as well as for population-weighted estimates in Panel (d) and TWFE estimates in Panel (e). N = 1,055 counties in Panel (c) due to missing data on population counts. N = 1,065 counties in Panel (b) due to absorption of within-state variation. Standard errors are clustered at the county level.

Figure A27: Rollout Impacts on Log Gestation
Across Specifications

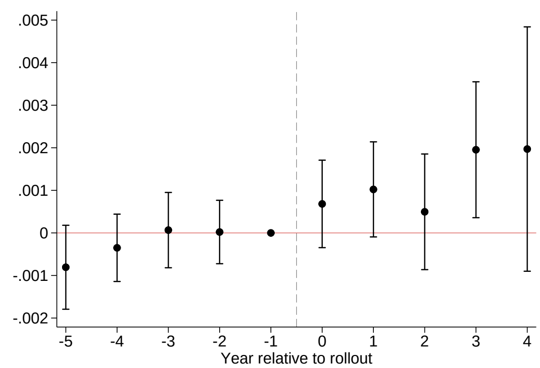
(a) Baseline



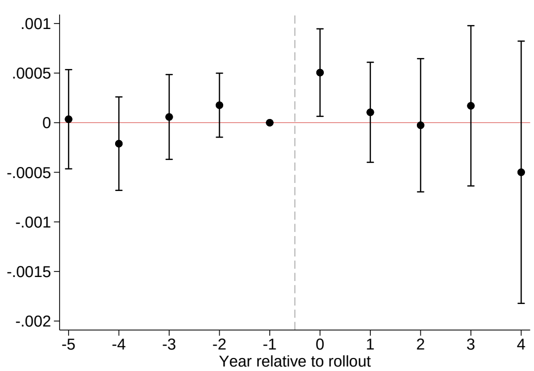
(b) State × Year Fixed Effects



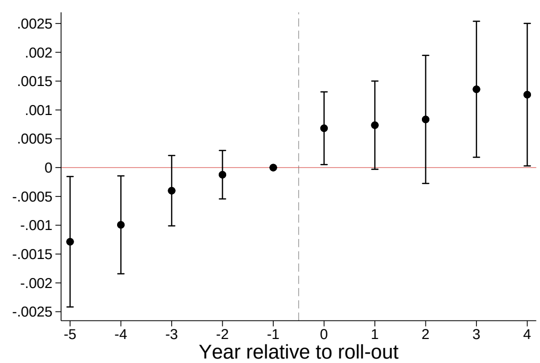
(c) Log Population Women Aged 15–44 × Year Fixed Effects



(d) 1970 County Population Weights



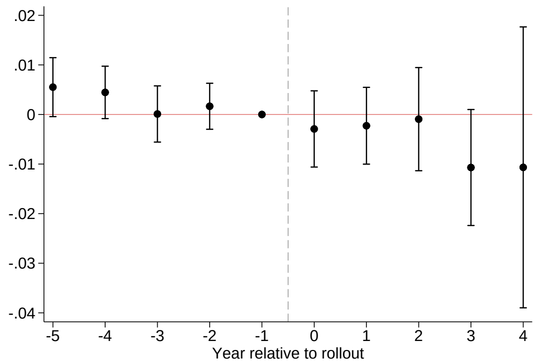
(e) TWFE



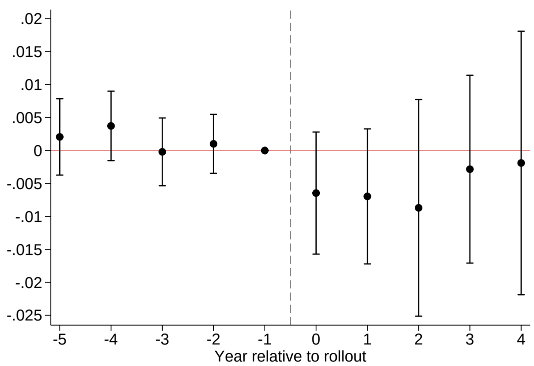
Notes: Figure shows the results of estimating β_r from Equation 8 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A19 for more details on each alternative specification. See Appendix C.5 for outcome-specific limitations applied to the rollout analysis baseline sample. N = 932 counties for the baseline estimates in Panel (a), as well as for population-weighted estimates in Panel (d) and TWFE estimates in Panel (e). N = 931 counties in Panel (c) due to missing data on population counts. N = 931 counties in Panel (b) due to absorption of within-state variation. Standard errors are clustered at the county level.

Figure A28: Rollout Impacts on Share Early Term ($<$ 39 weeks gestation) Across Specifications

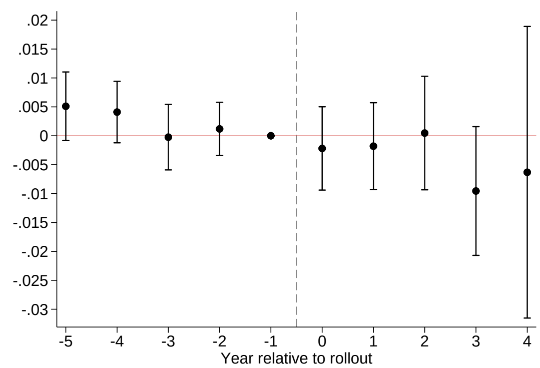
(a) Baseline



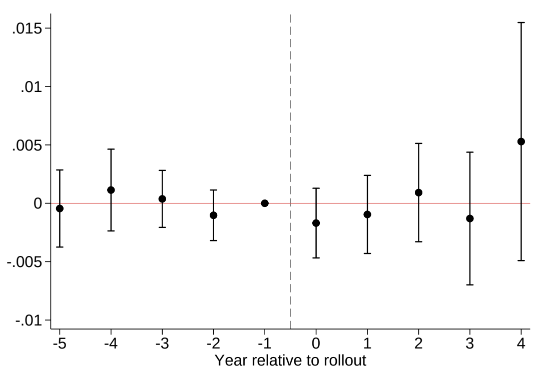
(b) State $\times$ Year Fixed Effects



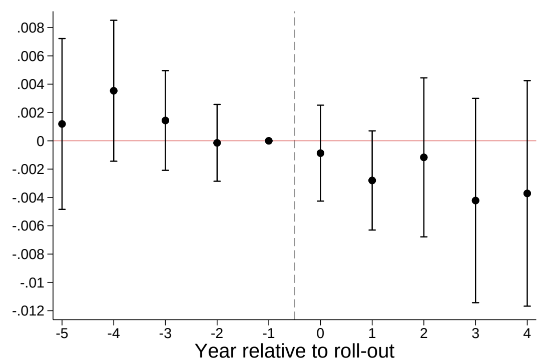
(c) Log Population Women Aged 15–44 $\times$ Year Fixed Effects



(d) 1970 County Population Weights



(e) TWFE

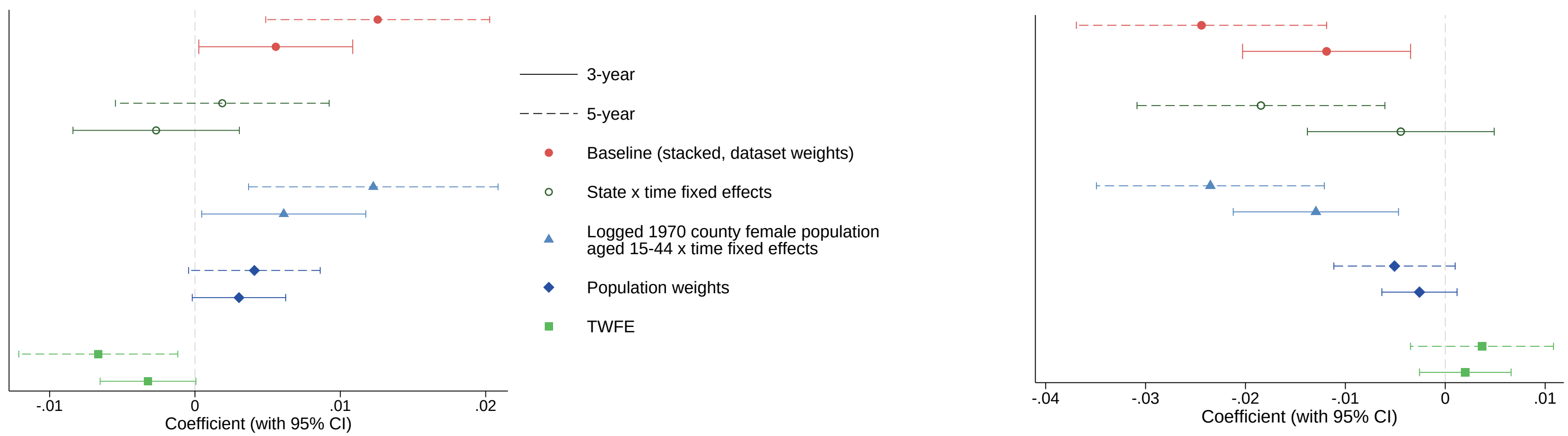


Notes: Figure shows the results of estimating β_r from Equation 8 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A19 for more details on each alternative specification. See Appendix C.5 for outcome-specific limitations applied to the rollout analysis baseline sample. N = 932 counties for the baseline estimates in Panel (a), as well as for population-weighted estimates in Panel (d) and TWFE estimates in Panel (e). N = 931 counties in Panel (c) due to missing data on population counts. N = 931 counties in Panel (b) due to absorption of within-state variation. Standard errors are clustered at the county level.

Figure A29: Rollout Impacts Across Specifications
Black Births

(a) Log Birth Weight

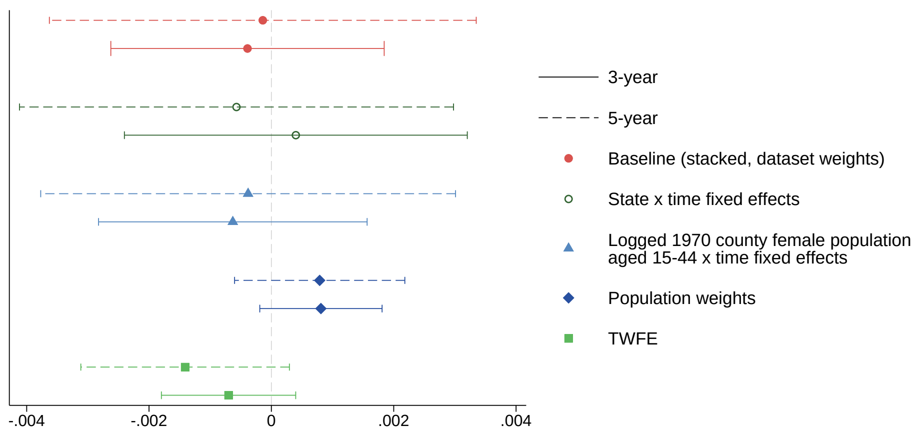
(b) Share Low Birth Weight



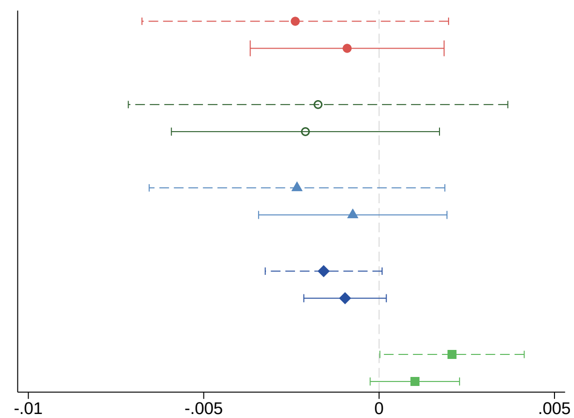
Notes: Figure compiles the 3-year and 5-year difference coefficients (as in Equation 28) and 95% CIs from the baseline specification and four alternative specifications for a sample of Black births. Across specifications, estimates we cluster standard errors at the county level. The baseline specification estimates the event study specification from Equations 7 and 8 using a stacked dataset and dataset weights that incorporate race-specific county population. The “population weights” specification estimates Equations 7 and 8 for an outcome of interest, using the baseline stacked specification with 1970 county race population weights. The “state × time” and “Logged 1970 county female population aged 15–44 × time specifications add those respective fixed effects, where time is year or unique year-month depending upon data frequency. The “TWFE” estimates Equations 23 and 24 using a flat dataset, race-county population weights, and binning at relative years beyond -5 and +5. Appendix Figures A31 through A32 present the underlying event study estimates.

Figure A30: Rollout Impacts Across Specifications
White Births

(a) Log Birth Weight



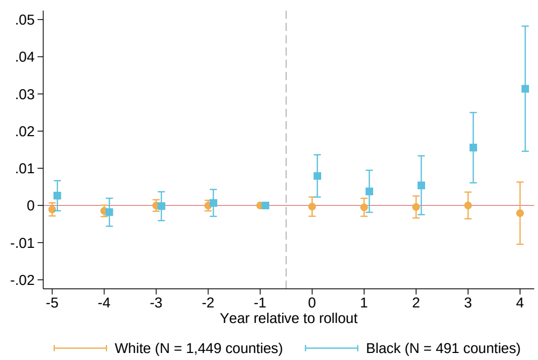
(b) Share Low Birth Weight



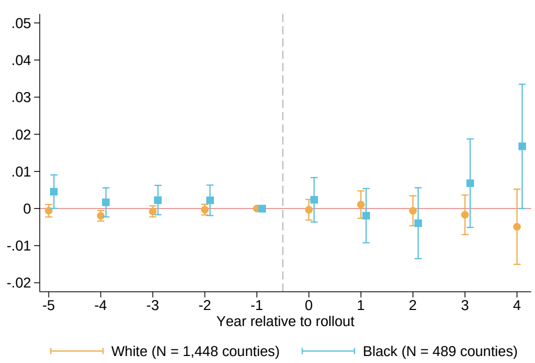
Notes: Figure compiles the 3-year and 5-year difference coefficients (as in Equation 28) and 95% CIs from the baseline specification and four alternative specifications for a sample of white births. Across specifications, estimates we cluster standard errors at the county level. The baseline specification estimates the event study specification from Equations 7 and 8 using a stacked dataset and dataset weights that incorporate race-specific county population. The "population weights" specification estimates Equations 7 and 8 for an outcome of interest, using the baseline stacked specification with 1970 county race population weights. The "state × time" and "Logged 1970 county female population aged 15–44 × time specifications add those respective fixed effects, where time is year or unique year-month depending upon data frequency. The "TWFE" estimates Equations 23 and 24 using a flat dataset, race-county population weights, and binning at relative years beyond -5 and +5. Appendix Figures A31 through A32 present the underlying event study estimates.

Figure A31: Rollout Impacts on Log Birth Weight
Across Specifications by Race

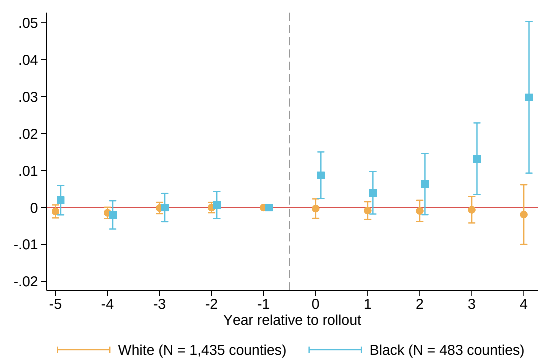
(a) Baseline



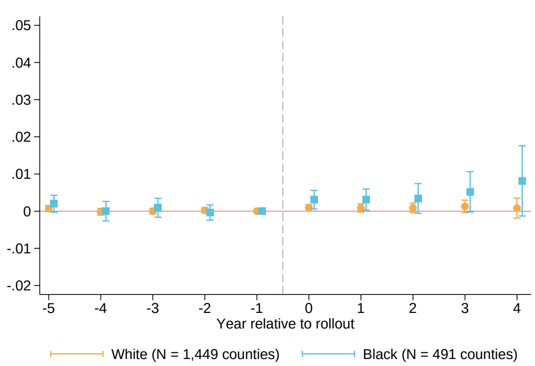
(b) State × Year Fixed Effects



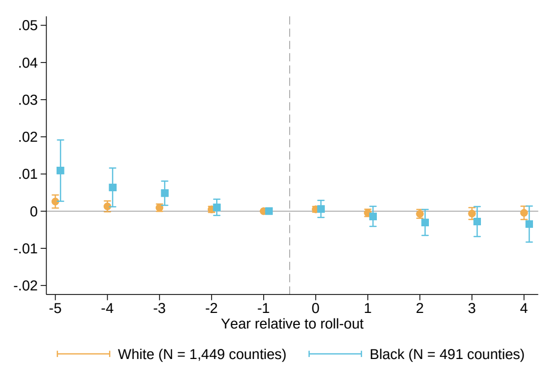
(c) Log Population Women Aged 15–44
× Year Fixed Effects



(d) 1970 County Population Weights



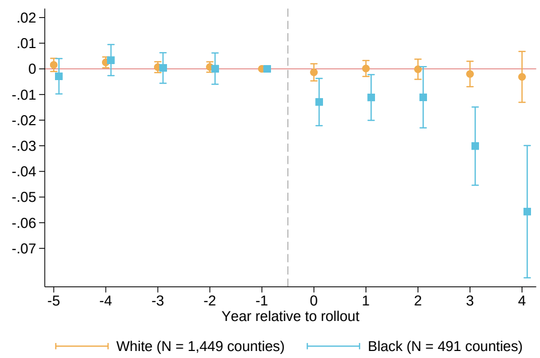
(e) TWFE



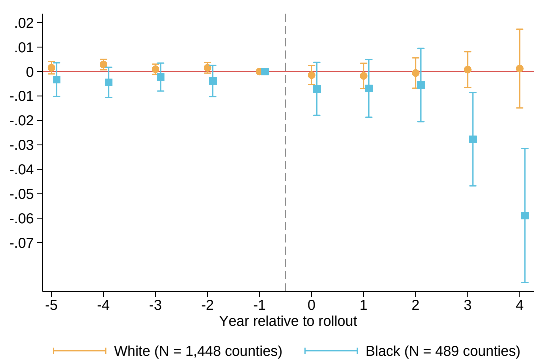
Notes: Figure shows the results of estimating β_r from Equation 8 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A19 for more details on each alternative specification. The baseline rollout sample is additionally limited to counties that have at least 25 race-specific births observed across years 1970–1975. Standard errors are clustered at the county level.

Figure A32: Rollout Impacts on Share Low Birth Weight
Across Specifications by Race

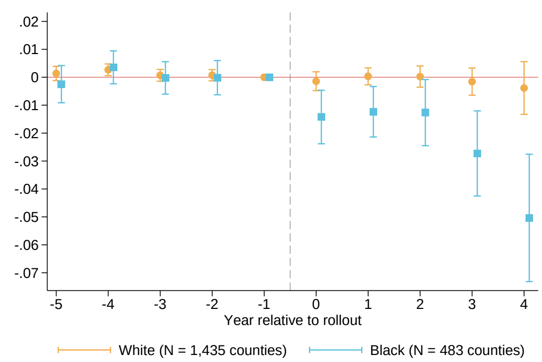
(a) Baseline



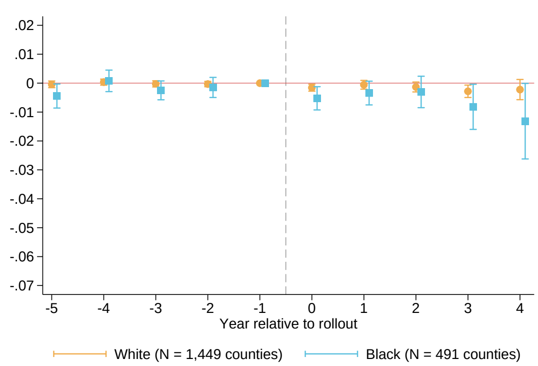
(b) State × Year Fixed Effects



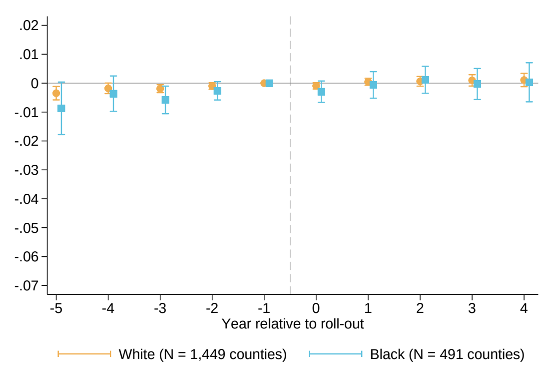
(c) Log Population Women Aged 15–44
× Year Fixed Effects



(d) 1970 County Population Weights



(e) TWFE



Notes: Figure shows the results of estimating β_r from Equation 8 for the baseline specification (Panel a) and four alternative specifications (Panels b through e). See notes to Appendix Figure A19 for more details on each alternative specification. The baseline rollout sample is additionally limited to counties that have at least 25 race-specific births observed across years 1970–1975. Standard errors are clustered at the county level.

Figure A33: Benefits Table, January 1978 (Before the Elimination of the Purchase Requirement)

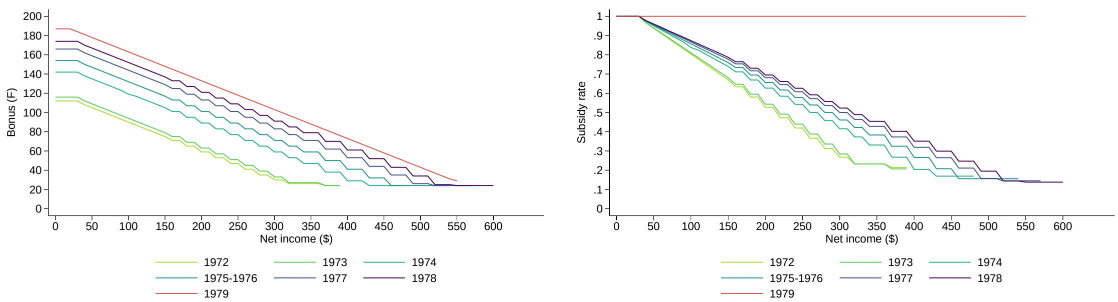
Monthly coupon allotments and purchase requirements—48 States and the District of Columbia

Monthly net income	For a household of—							
	1 person	2 persons	3 persons	4 persons	5 persons	6 persons	7 persons	8 persons
	The monthly coupon allotment is—							
	$52	$96	$138	$174	$206	$248	$274	$314
	And the monthly purchase requirement is—							
$0 to $19.99	0	0	0	0	0	0	0	0
$20 to $29.99	$1	$1	0	0	0	0	0	0
$30 to $39.99	4	4	$4	$4	$5	$5	$5	$5
$40 to $49.99	6	7	7	7	8	8	8	8
$50 to $59.99	8	10	10	10	11	11	12	12
$60 to $69.99	10	12	13	13	14	14	15	16
$70 to $79.99	12	15	16	16	17	17	18	19
$80 to $89.99	14	18	19	19	20	21	21	22
$90 to $99.99	16	21	21	22	23	24	25	26
$100 to $109.99	18	23	24	25	26	27	28	29
$110 to $119.99	21	26	27	28	29	31	32	33
$120 to $129.99	24	29	30	31	33	34	35	36
$130 to $139.99	27	32	33	34	36	37	38	39
$140 to $149.99	30	35	36	37	39	40	41	42
$150 to $169.99	33	38	40	41	42	43	44	45
$170 to $189.99	39	44	46	47	48	49	50	51
$190 to $209.99	40	50	52	53	54	55	56	57
$210 to $229.99	42	56	58	59	60	61	62	63
$230 to $249.99	42	62	64	65	66	67	68	69
$250 to $269.99	42	68	70	71	72	73	74	75
$270 to $289.99		74	76	77	78	79	80	81
$290 to $309.99		76	82	83	84	85	86	87
$310 to $329.99		76	88	89	90	91	92	93
$330 to $359.99		76	94	95	96	97	98	99
$360 to $389.99			103	104	105	106	107	108
$390 to $419.99			112	113	114	115	116	117
$420 to $449.99			120	122	123	124	125	126
$450 to $479.99			120	131	132	133	134	135
$480 to $509.99				140	141	142	143	144
$510 to $539.99				149	150	151	152	153
$540 to $569.99				150	159	160	161	162
$570 to $599.99				150	168	169	170	171
$600 to $629.99					177	178	179	180
$630 to $659.99					178	187	188	189
$660 to $689.99					178	196	197	198
$690 to $719.99						205	206	207
$720 to $749.99						214	215	216
$750 to $779.99						216	224	225
$780 to $809.99						216	233	234
$810 to $839.99						216	238	243
$840 to $869.99							238	252
$870 to $899.99							238	261
$900 to $929.99							238	270
$930 to $959.99								274
$960 to $989.99								274
$990 to $1,019.99								274
$1,020 to $1,049.99								274

Notes: Congress passed a new national benefits schedule every 6 months to index benefits to food price inflation. This schedule took effect in January 1978. Every household of the same size received the same value stamps allotment, but higher income households paid higher purchase requirements for that allotment. Thus, higher income households received smaller bonuses (i.e., the actual benefit value/transfer from the government), because the bonus is the allotment, at the top of the table, minus the household's purchase requirement, which is based on household income bracket. This table is from Federal Register Vol 42 Iss 215.

Figure A34: Benefit Schedules Transfers and Subsidy Rate

(a) Food Stamp Bonus (F) (b) Food Stamp Subsidy Rate (F/F+R)



Notes: Panel (a) shows the bonus/transfer, in dollars, across benefits schedules in years 1972-1979 for a family of four. Panel (b) shows the subsidy rate across these years. Income bins have not been adjusted and thus the figures reflect the benefits received by a family of four within each income bin in terms of nominal dollars. For example, in Panel (a), the point at net income $300 shows the bonus, in dollars, received by a family of four with monthly net income between $290.00-$299.99. Food stamp bonus schedules by net income are constructed from annual allotment tables published by the U.S. Department of Agriculture's Food and Nutrition Service in the Federal Register for 1971–1974 and 1976–1979 (see Food and Nutrition Service, USDA (1971, 1972, 1973, 1974, 1976, 1978, 1979)), supplemented by United States Food and Nutrition Service (1975). These documents report the official Food Stamp coupon allotment tables used by caseworkers to determine household benefits. Note that the published 1979 benefits schedules specify benefits in $3-wide monthly net income bins (instead of $10-wide).

Figure A35: Benefits Table, January 1979 (After the Elimination of the Purchase Requirement)

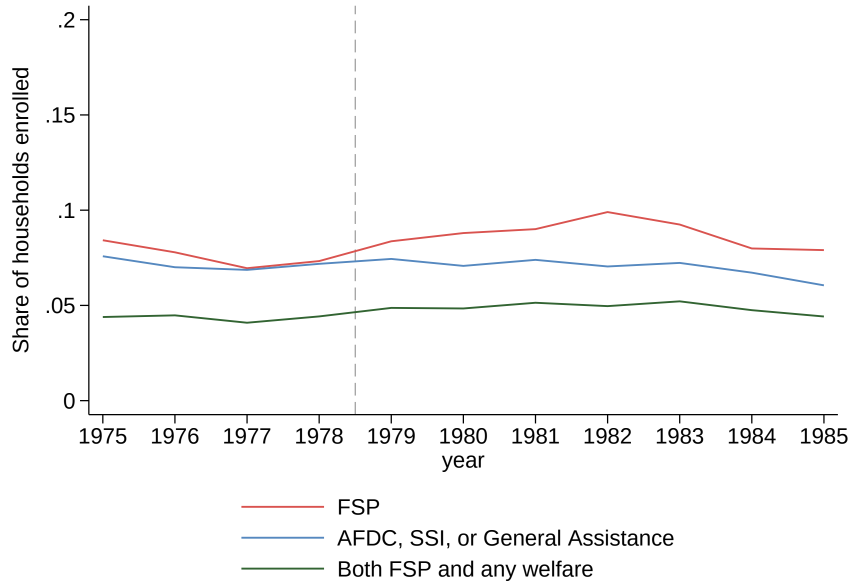
January 1, 1979 - Basis of Coupon Issuance - 1964 Act
48 States and District of Columbia

TABLE I

Monthly Net Income	Bonus by Household Size							
	One Person	Two Persons	Three Persons	Four Persons	Five Persons	Six Persons	Seven Persons	Eight Persons
0 - 19.99	58.00	106.00	152.00	192.00	228.00	274.00	302.00	346.00
20 - 29.99	57.00	105.00	152.00	192.00	228.00	274.00	302.00	346.00
30 - 39.99	54.00	102.00	148.00	188.00	223.00	269.00	297.00	341.00
40 - 49.99	52.00	99.00	145.00	185.00	220.00	266.00	294.00	338.00
50 - 59.99	50.00	96.00	142.00	182.00	217.00	263.00	290.00	334.00
60 - 69.99	48.00	94.00	139.00	179.00	214.00	260.00	287.00	330.00
70 - 79.99	46.00	91.00	136.00	176.00	211.00	257.00	284.00	327.00
80 - 89.99	44.00	88.00	133.00	173.00	208.00	253.00	281.00	324.00
90 - 99.99	42.00	85.00	131.00	170.00	205.00	250.00	277.00	320.00
100 - 109.99	40.00	83.00	128.00	167.00	202.00	247.00	274.00	317.00
110 - 119.99	37.00	80.00	125.00	164.00	199.00	243.00	270.00	313.00
120 - 129.99	34.00	77.00	122.00	161.00	195.00	240.00	267.00	310.00
130 - 139.99	31.00	74.00	119.00	158.00	192.00	237.00	264.00	307.00
140 - 149.99	28.00	71.00	116.00	155.00	189.00	234.00	261.00	304.00
150 - 169.99	25.00	68.00	112.00	151.00	186.00	231.00	258.00	301.00
170 - 189.99	19.00	62.00	106.00	145.00	180.00	225.00	252.00	295.00
190 - 209.99	13.00	56.00	100.00	139.00	174.00	219.00	246.00	289.00
210 - 229.99	10.00	50.00	94.00	133.00	168.00	213.00	240.00	283.00
230 - 249.99	10.00	44.00	88.00	127.00	162.00	207.00	234.00	277.00
250 - 269.99	10.00	38.00	82.00	121.00	156.00	201.00	228.00	271.00
270 - 289.99	10.00	32.00	76.00	115.00	150.00	195.00	222.00	265.00
290 - 309.99		26.00	70.00	109.00	144.00	189.00	216.00	259.00
310 - 329.99		20.00	64.00	103.00	138.00	183.00	210.00	253.00
330 - 359.99		20.00	58.00	97.00	132.00	177.00	204.00	247.00
360 - 389.99		20.00	49.00	88.00	123.00	168.00	195.00	238.00
390 - 419.99			40.00	79.00	114.00	159.00	186.00	229.00
420 - 449.99			31.00	70.00	105.00	150.00	177.00	220.00
450 - 479.99			22.00	61.00	96.00	141.00	168.00	211.00
480 - 509.99			18.00	52.00	87.00	132.00	159.00	202.00
510 - 539.99				43.00	78.00	123.00	150.00	193.00
540 - 569.99				34.00	69.00	114.00	141.00	184.00
570 - 599.99				25.00	60.00	105.00	132.00	175.00
600 - 629.99				24.00	51.00	96.00	123.00	166.00
630 - 659.99				24.00	42.00	87.00	114.00	157.00
660 - 689.99					33.00	78.00	105.00	148.00
690 - 719.99					28.00	69.00	96.00	139.00
720 - 749.99					28.00	60.00	87.00	130.00
750 - 779.99					28.00	51.00	78.00	121.00
780 - 809.99						42.00	69.00	112.00
810 - 839.99						33.00	60.00	103.00
840 - 869.99						32.00	51.00	94.00
870 - 899.99						32.00	42.00	85.00
900 - 929.99						32.00	36.00	76.00
930 - 959.99							36.00	67.00
960 - 989.99							36.00	58.00
990 - 1,019.99							36.00	49.00
1,020 - 1,049.99								40.00
1,050 - 1,079.99								40.00
1,080 - 1,109.99								40.00
1,110 - 1,139.99								40.00
1,140 - and up								40.00

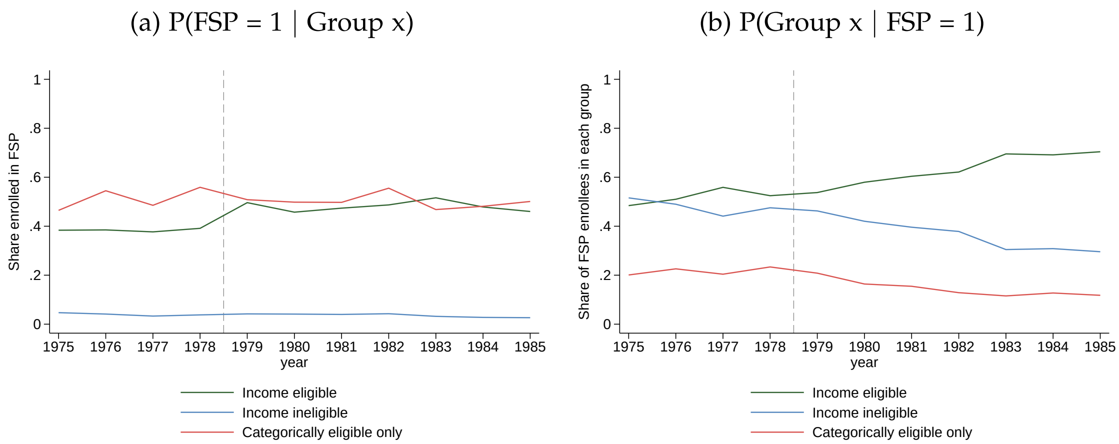
Notes: Congress passed a new national benefits schedule every 6 months to index benefits to food price inflation. This schedule took effect in January, 1979. After the elimination of the purchase requirement, the dollar value of food stamps a household received could be read off after knowing the household's size and income bracket. The benefit amount a household received was roughly equivalent to that household's bonus (allotment minus purchase requirement) under the purchase requirement system (see Figure A33). This table is from Federal Register Vol 43 Iss 225.

Figure A36: Share of Households Enrolled in the Food Stamp Program, AFDC



Notes: Figure shows the share of all households in the PSID with Food Stamp transfers greater than zero, AFDC or SSI or General Assistance transfers greater than zero, and Food Stamp and any of these welfare transfers greater than zero. Measures are weighted by PSID household weights.

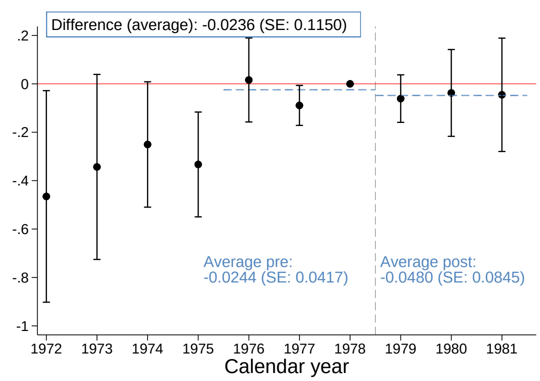
Figure A37: Take up and Food Stamp Program Enrollee Composition by Simulated Eligibility Group



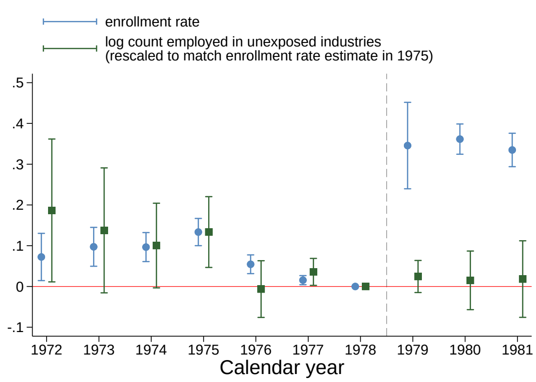
Notes: Figure uses simulated Food Stamp Program eligibility based upon household-level income data in the 1976-1981 PSID for households smaller than 9 people and the 1978 eligibility scheme (in terms of deductions and categorical eligibility), as well as the maximum allowable monthly net income in 1978. Specifically, to simulate eligibility, we calculate monthly net income (in 1978 dollars) by subtracting an earned income deduction according to the 1978 income deductions (10% of wages from labor up to a maximum of $30) from total family income. Then, we designate households with monthly net income below the household size-specific maximum allowable monthly net income for 1978 eligibility as "income eligible," and households with net income above this threshold as "income ineligible." Additionally, we designate households with positive AFDC, SSI, or General Assistance transfers and with net income above the threshold as "categorically eligible only." Panel (a) shows the share of household-years in each group that enroll in Food Stamp Program. Panel (b) shows the share of Food Stamp-enrolled household-years in the PSID that fall into each group.

Figure A38: Impact of Free Food Stamp Reform on Enrollment, Adjusting for Economic Confounds

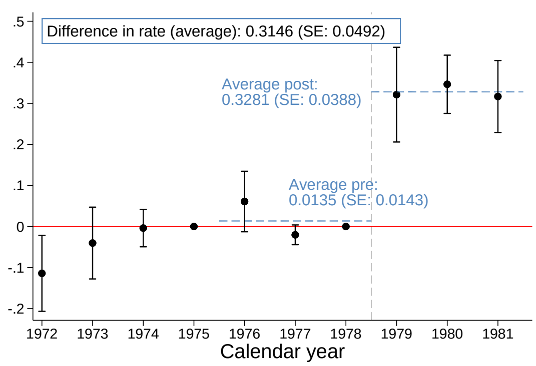
(a) Log Employment Count in Unexposed Industries



(b) Overlaying Enrollment Rate and Rescaled Employment



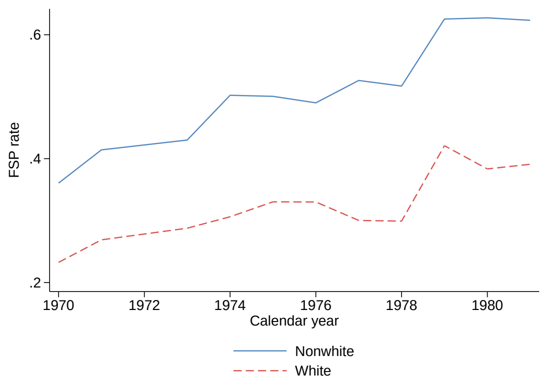
(c) Enrollment Rate



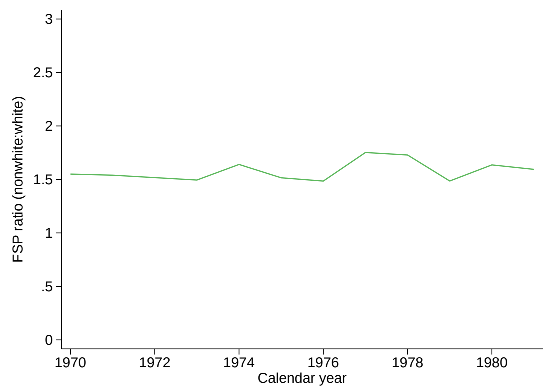
Notes: Figure shows results of implementing the proxy approach for economic confounds from Freyaldenhoven et al. (2019), with employment in industries unlikely to have many workers on food stamps used as a proxy for the local business cycle. Panel (a) shows the results of estimating Equation 5 for the outcome log of employment in unexposed industries, which will be our definition of Proxy_{ct} for Panels (b) and (c). The blue series in Panel (b) shows the standard first stage impact on the enrollment rate (Equation 4). The green series rescales the proxy estimate in Panel (a), normalizing so that the value in 1975 for the proxy equals the value in in 1975 for the first stage. Panel (c) shows the results of estimating Equation 20 with the dependent variable of the food stamp enrollment rate, using the lead of the free food stamp reform-exposure "treatment" as an instrument for the employment component associated with the unobserved confounder as described by Equation 21. All regressions are weighted by 1970 county-level population. Standard errors are clustered at the county level. The sample consists of counties in the baseline free food stamp reform sample with available data on employment counts. N =1,801 counties.

Figure A39: PSID: Food Stamp Program Take-up

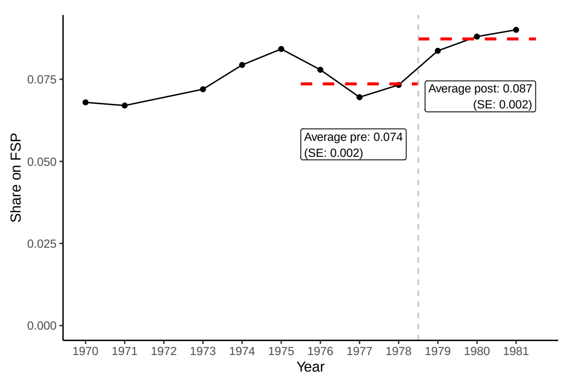
(a) Take-up Rate by Race (PSID).



(b) PSID Food Stamp Program Take-up Ratio by Race



(c) PSID Food Stamp Program Enrollment over Time



Notes: Panel (a) plots the national, annual take-rate rate of food stamp for whites and non whites separately, with take up defined as the share of the eligible population that is enrolled. Panel (b) plots the ratio of those race-specific take-up rates (nonwhite:white) in each year between 1970 and 1981. Panel (c) plots the average enrollment across the PSID sample.

Figure A40: NCHS Table of Last Mentral Period (LMP) Reporting

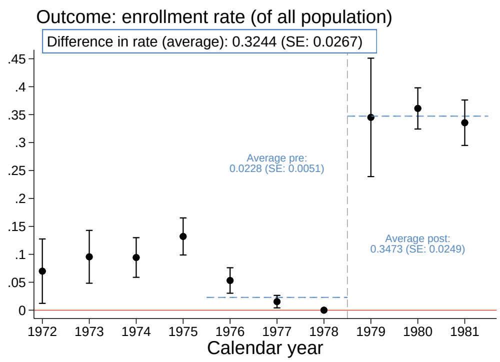
Appendix table. States reporting date last normal menstrual period began: United States, 1968-78

State	1978	1977	1976	1975	1974	1973	1972	1971	1970	1969	1968
Alabama	X	X	X								
Alaska	X	X	X	X	X	X	X	X	X	X	X
Arizona	X	X	X	X	X	X	X	X	X	X	X
Arkansas	X										
California	X	X	X	X	X	X	X	X	X	X	X
Colorado	X	X	X	X	X	X	X	X	X	X	X
Connecticut											
Delaware	X	X	X	X	X	X	X				
District of Columbia	X	X	X	X	X	X	X	X	X	X	X
Florida	X	X	X	X	X	X	X	X	X		
Georgia	X	X	X	X	X	X					
Hawaii	X	X	X	X	X	X	X	X	X	X	X
Idaho	X										
Illinois	X	X	X	X	X	X	X	X	X	X	X
Indiana	X	X	X	X	X	X	X	X	X	X	X
Iowa	X	X	X	X	X	X	X	X	X	X	X
Kansas	X	X	X	X	X	X	X	X	X	X	X
Kentucky	X	X	X	X	X	X	X	X	X	X	X
Louisiana	X	X	X	X	X	X	X	X	X	X	X
Maine	X	X	X	X	X	X	X	X	X	X	X
Maryland	X	X	X	X	X	X	X	X	X	X	X
Massachusetts	X	X	X							X	
Michigan	X	X	X	X	X	X	X	X	X	X	X
Minnesota	X	X	X	X	X	X	X	X	X	X	X
Mississippi	X	X	X	X	X	X	X	X	X	X	X
Missouri	X	X	X	X	X	X	X	X	X	X	X
Montana	X	X	X	X	X	X	X	X	X	X	X
Nebraska	X	X	X	X	X	X	X	X	X	X	X
Nevada	X	X	X	X	X	X	X	X	X	X	X
New Hampshire	X	X	X	X	X	X	X	X	X	X	X
New Jersey	X	X	X	X	X	X	X	X	X	X	X
New Mexico											
New York	X	X	X	X	X	X	X	X	X	X	X
North Carolina	X	X	X	X	X	X	X	X	X	X	X
North Dakota	X	X	X	X	X	X	X	X	X	X	X
Ohio	X	X	X	X	X	X	X	X	X	X	X
Oklahoma	X	X	X	X	X	X	X	X	X	X	X
Oregon	X	X	X	X	X	X	X	X	X		
Pennsylvania	X										
Rhode Island	X	X	X	X	X	X	X	X	X	X	X
South Carolina	X	X	X	X	X	X	X	X	X	X	X
South Dakota	X	X	X	X	X	X	X	X	X	X	X
Tennessee	X	X	X	X	X	X	X	X	X	X	X
Texas											
Utah	X	X	X	X	X	X	X	X	X	X	X
Vermont	X	X	X	X	X	X	X	X	X	X	X
Virginia	X										
Washington	X	X	X	X	X	X	X	X	X	X	X
West Virginia	X	X	X	X	X	X	X	X	X	X	X
Wisconsin	X										
Wyoming	X	X	X	X	X	X	X	X	X	X	X

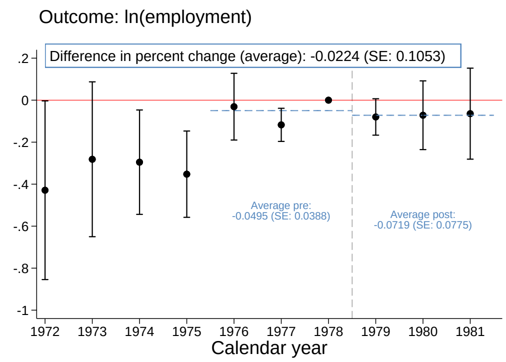
Notes: Table denotes, for each state, which years report LMP for gestation period calculation. Note that (not reported in this table) Connecticut began reporting LMP in 1981, Texas by 1980, and New Mexico did not report LMP in the years included in the analysis (U.S. Department of Health and Human Services, 1982).

Figure A41: 1979 Reform Impacts on Imputed Employment

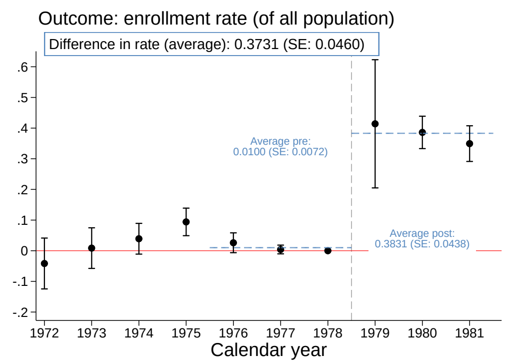
(a) First Stage: Impact on Food Stamp Program Enrollment (All Industries)



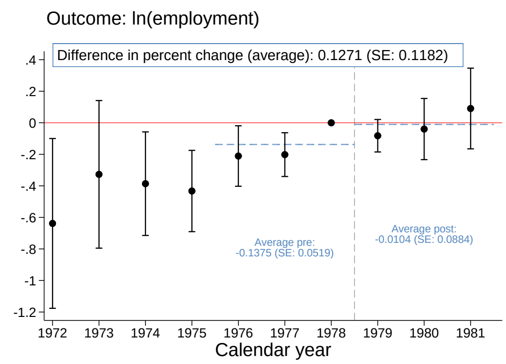
(b) Reduced Form: Impact on Ln(Employment) (All Industries)



(c) First Stage: Impact on Food Stamp Program Enrollment (Low-wage Industry Subset)



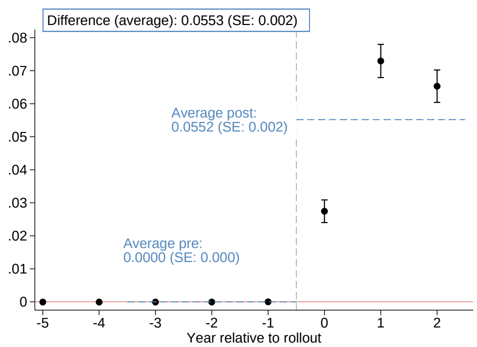
(d) Reduced Form: Impact on Ln(Employment) (Low-wage Industry Subset)



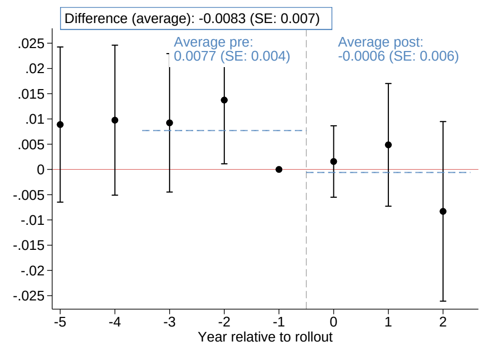
Notes: Panels (a) and (c) show the results of estimating the first stage Equation 4. Panels (b) and (d) show the results of estimating the reduced form Equation 5. Both estimations use a continuous 1972 exposure measure. The logged employment outcome comes from the imputed County Business Patterns datasets made by Eckert et al. (2020) and Eckert et al. (2022). The low-wage industry subset comprises 1972 Standard Industrial Classification codes 07−− (agricultural services, forestry, and fisheries), 7000 (hotels and other lodging places), and 52−− (retail trade). For both sets of industries (all and low-wage subset), the sample is limited to counties that rolled out Food Stamp Program by 1970 and that have both Food Stamp Program participation data and non-zero employment data in every year 1972-1981. For the low-wage industry sample, only county-years with employment data from each of the relevant SIC codes are included. Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level. N = 1,804 counties in Panels (a) and (b), and N = 687 counties in Panels (c) and (d).

Figure A42: Impact of Food Stamps Program Rollout on Employment

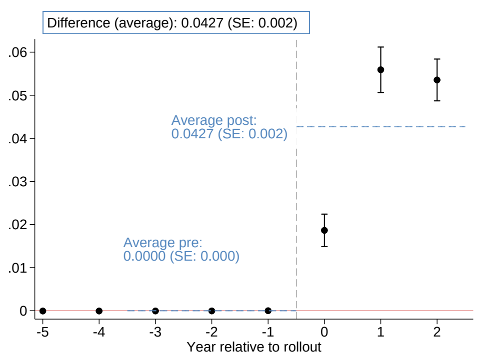
(a) First Stage: Food Stamp Program Enrollment



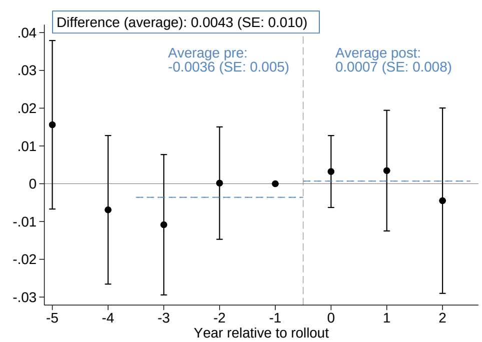
(b) Ln(Employment) (All Industries)



(c) First Stage: Food Stamp Program Enrollment

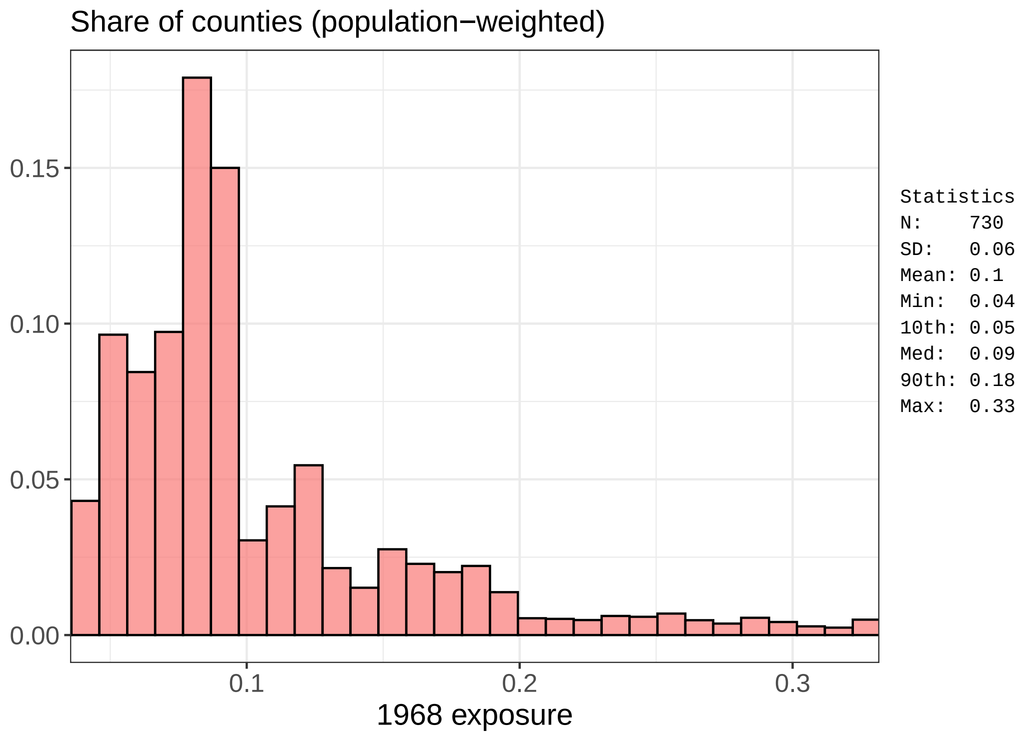


(d) Ln(Employment) (Low-wage Industries)



Notes: Figure show estimates of β_r from Equation 7 in Panels (a) and (c) and from Equation 8 in Panel (b) and (d). The regressions are weighted by dataset-normalized weights which are described above in Equation 22. Since Food Stamp Program enrollment and CBP data are both constructed as annual measures, relative year corresponds to a calendar year for each county's rollout timing across panels. Sample takes the baseline rollout analysis sample and further limits to counties with the CBP outcome of interest observed in each relative year -3 to +2. The logged employment outcome comes from the imputed County Business Patterns datasets made by Eckert et al. (2020) and Eckert et al. (2022). The low-wage industry subset comprises 1972 Standard Industrial Classification codes 07−− (agricultural services, forestry, and fisheries), 7000 (hotels and other lodging places), and 52−− (retail trade). For the low-wage industry sample, only county-years with employment data from each of the relevant SIC codes are included, resulting in a smaller sample of counties. Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level. N = 1,509 counties in Panels (a) and (b), and N = 437 counties in Panels (c) and (d).

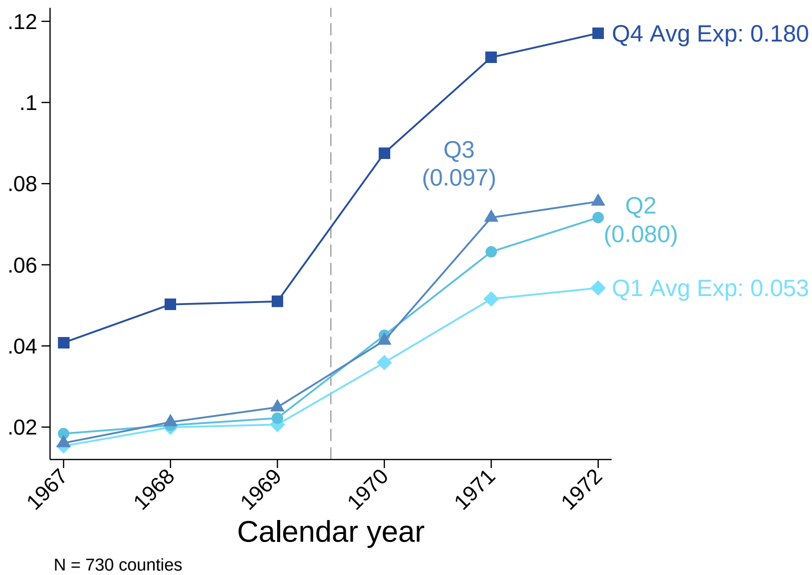
Figure A43: 1968 Exposure Measure



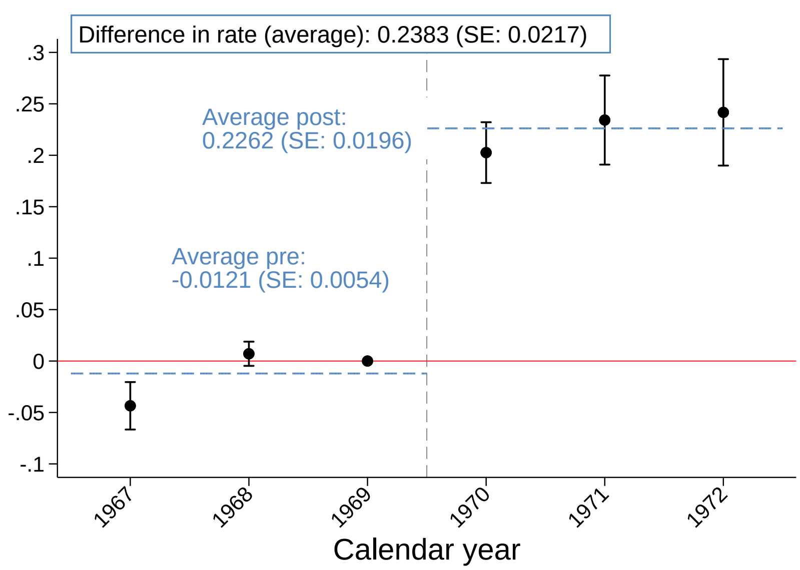
Notes: Figure shows 1970 population-weighted distribution of the county-level 1968 exposure measure, according to Equation 3. The exposure measure is winsorized at the 1st and 99th percentiles, 0.035 and 0.332 respectively. Sample is limited to the subset of the free food stamp reform analysis baseline sample counties where Food Stamps Program was rolled out 1967 or earlier. (N= 730 counties)

Figure A44: 1970 Bonus Reform: Enrollment Rate

(a) Trends by Quartile



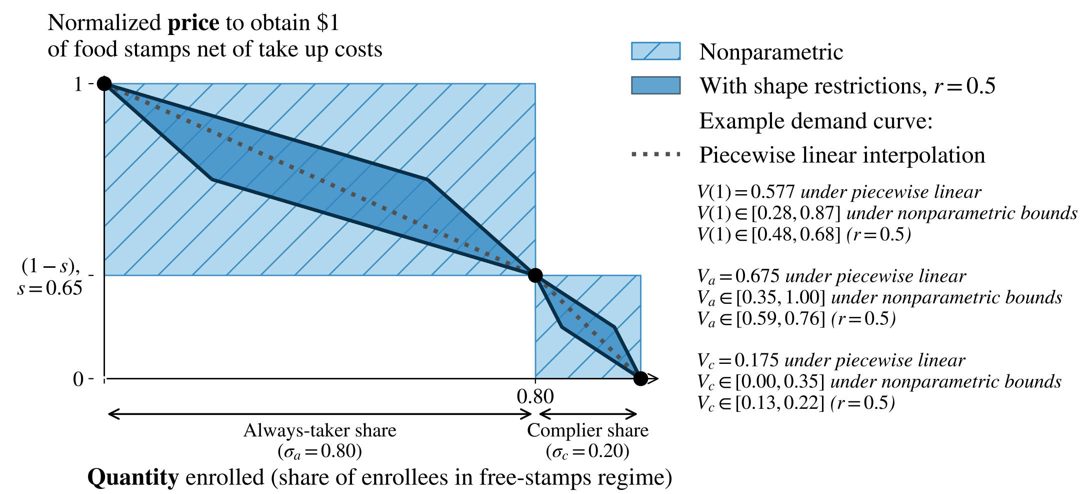
(b) Impact on Enrollment



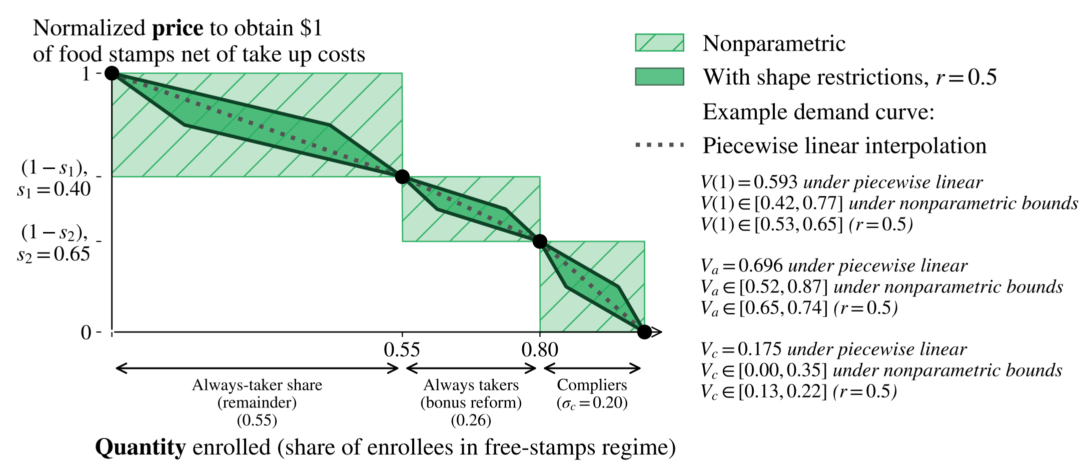
Notes: Panel (a) shows 1970 population-weighted average county-level Food Stamp enrollment over time by population-weighted quartile of the 1968 exposure measure. Darker shades in the series color scale correspond with increasing quartile of exposure. The average exposure measure within each quartile is displayed within series labels. Panel (b) shows the results of estimating Equation 60 which looks at the impact of the 1970 Bonus Reform on the food stamp enrollment rate, using 1968 exposure measure, and controls for years-since-rollout-by-calendar year fixed effects. Sample is limited to the subset of the free food stamp reform analysis baseline sample counties where Food Stamps Program was rolled out 1967 or earlier. Standard errors are clustered at the county level. (N= 730 counties)

Figure A45: Bounds on Value of Food Stamps

(a) With Purchase Requirement Only

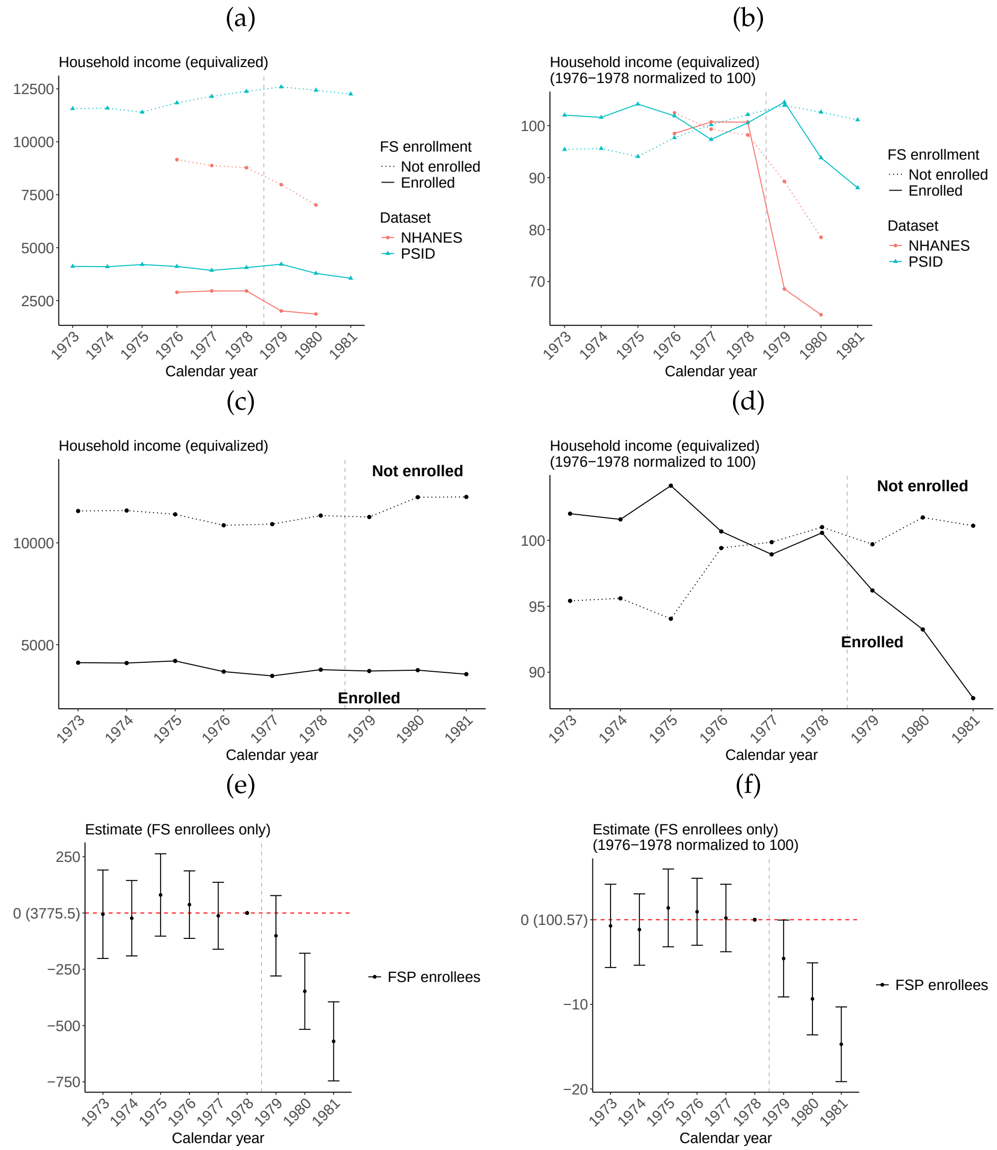


(b) With Purchase Requirement and Bonus Reform



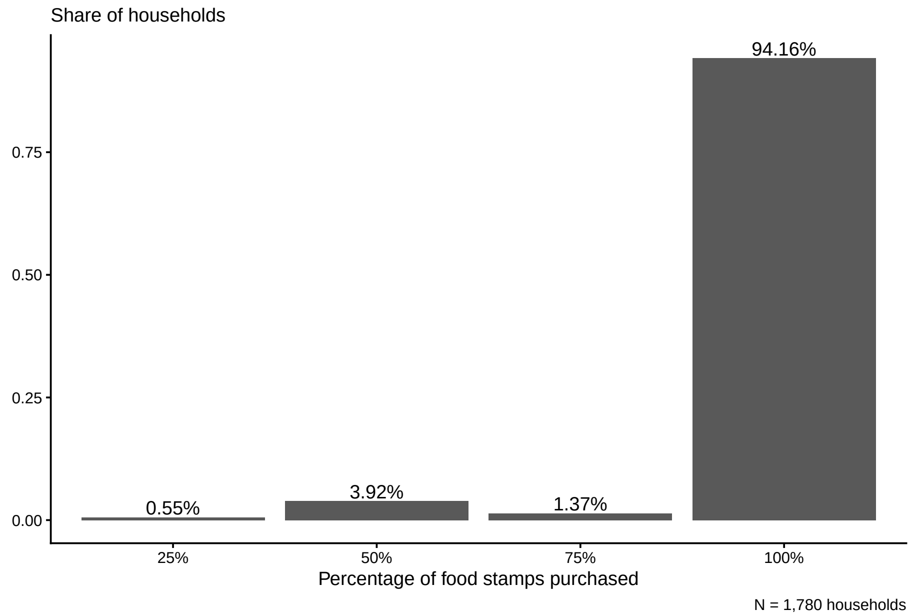
Note: See notes to Figure 11. The dark shaded rhombuses correspond to semi-parametric demand-curve regions using $r=0.5$ from Kang and Vasserman (2025) (see footnote 38 for the definition of r).

Figure A46: Food Stamp Program-enrolled Household Characteristics in the PSID and NHANES



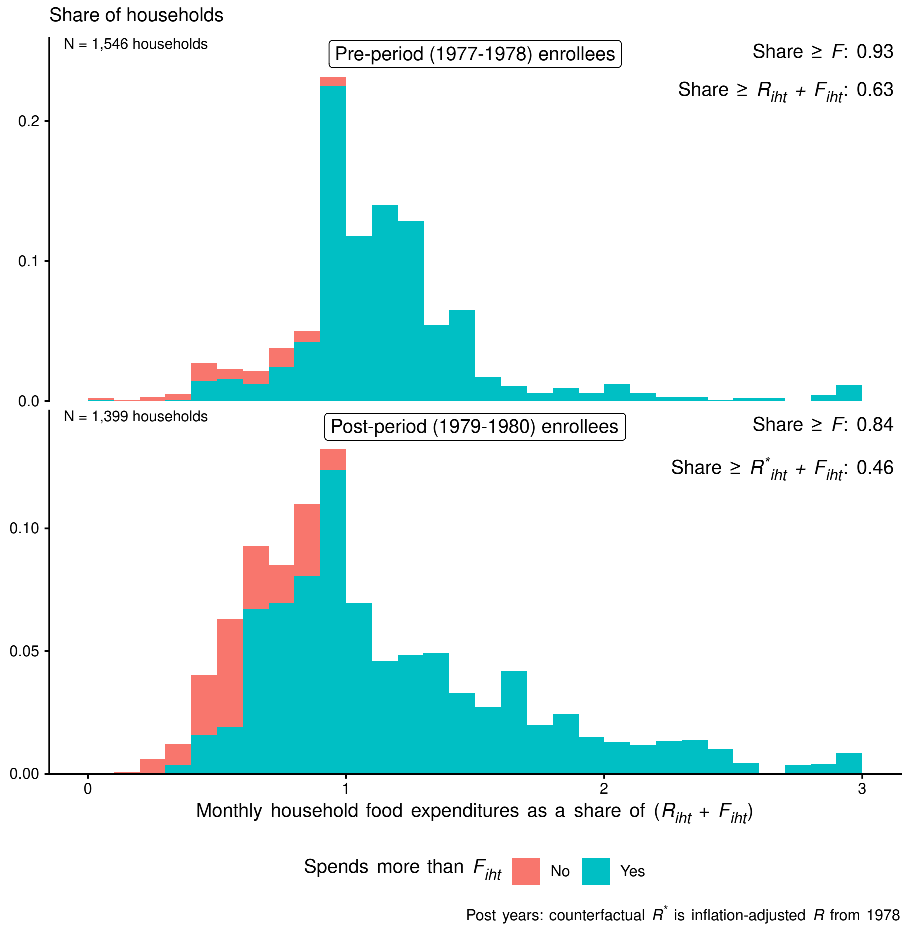
Notes: Figure presents household characteristics from a dataset pooling observations from the PSID and NHANES. Panels (a) through (d) show weighted means of household characteristics over time between groups defined by enrollment status. Panels (e) and (f) show α_p from Equation 37 with the error bars indicating the 95% confidence interval (from standard errors clustered at household level). All estimates use given survey weights normalized to sum to the count of enrolled households within dataset-year, where dataset is PSID or NHANES.

Figure A47: NFCS: Distribution of the Percentage of Food Stamps Purchased



Notes: Figure shows the percentage of the food stamps benefit that households purchased in the month prior to responding to the 1977-1978 Nationwide Food Consumption Survey. Shares are weighted by the NFCS weights. The sample is restricted to households that were surveyed in 1977-1978 and were enrolled in food stamps at the time of their response. See Appendix B.2 for more details on the data.

Figure A48: NFCS: Food Spending Normalized to Inflation-adjusted $R + F$



Notes: Figure shows a histogram of monthly food spending among food stamp enrollees expressed as a share of $R_{iht} + F_{iht}$, winsorized to 3. Colors indicate whether monthly food spending exceeded the bonus F that households were entitled to. Sub-panels are split into the time period the period prior to the elimination of the purchase requirement ('pre period', 1977-1978) and after its elimination ('post period', 1979-1980). All data are from the Nationwide Food Consumption Survey's survey of low income households of individuals who were either receiving food stamps or eligible to receive them; all analyses use survey weights. In the post period, benefits are assigned using net income as it was defined in the pre-period and $R + F$ is determined using the counterfactual R, as described in Appendix B.2. The sample is restricted to households of size 1-8 that are enrolled in food stamps and for which the income and food spending variables are defined.

Table A1: Baseline Sample Restrictions

Restriction	Counties (N)	Share of 1970 population (%)
All counties	3,143	100%
Exclude counties missing census, participation, or health data in all years	2,873	95.32%
(a) Free food stamp reform baseline sample		
Limit to counties that rolled out by 1970	1,812	71.43%
Exclude counties missing participation or birthweight data in any year 1972–1981	1,804	71.13%
(b) Rollout baseline sample		
Limit to counties that rolled out in/after July 1969	1,534	41.37%
Exclude counties missing participation or birthweight data in any year 1970–1975	1,520	39.99%

Notes: Table outlines the sample restrictions imposed upon the full set of U.S. county equivalents. Appendix Table A2 lists the counties that are dropped in the first row (i.e., because they are missing some data element in all years).

Table A2: County Equivalents Dropped from or Aggregated within Analysis Samples

State		**County-equivalent (count)**	**Reason**
Entire states that lack of county-level participation (217 county equivalents)			
Alaska		29	Lacks rollout timing data
Maine		16	Lacks county-level participation
Oregon		36	Lacks county-level participation
New Hampshire		10	Lacks county-level participation
Rhode Island		5	Lacks county-level participation
Utah		29	Lacks county-level participation
Vermont		14	Lacks county-level participation
West Virginia		55	Lacks county-level participation
Wyoming		23	Lacks county-level participation
State	**FIPS**	**County equivalent**	**Reason**
Participation data reporting errors			
Georgia	13183	Long County	Reporting error
Idaho	16065	Madison County	Reporting error
Indiana	18139	Rush County	Reporting error
New Mexico	35021	Harding County	Reporting error
Virginia	51800	Suffolk City	Reporting error
Additional dropped county equivalents			
Arizona	04012	La Paz County	Lacks census data (founded 1983)
Hawaii	15005	Honolulu County	Lacks census data
Montana	30113	Yellowstone National Park	Lacks census data (dissolved 1997)
Colorado	08014	Broomfield County	Lacks census data (founded 2001)
Massachusetts	25001	Barnstable County	Lacks FSP participation data
Massachusetts	25003	Berkshire County	Lacks FSP participation data
Massachusetts	25007	Dukes County	Lacks FSP participation data
Massachusetts	25011	Franklin County	Lacks FSP participation data
Massachusetts	25015	Hampshire County	Lacks FSP participation data
Massachusetts	25017	Middlesex County	Lacks FSP participation data
Massachusetts	25019	Nantucket County	Lacks FSP participation data
Massachusetts	25021	Norfolk County	Lacks FSP participation data
Massachusetts	25023	Plymouth County	Lacks FSP participation data
Minnesota	27077	Lake of the Woods County	Lacks FSP participation data
Missouri	29186	Ste. Genevieve County	Lacks FSP participation data
Connecticut	09013	Tolland County	Lacks FSP participation data
Connecticut	09015	Windham County	Lacks FSP participation data
Florida	12043	Glades County	Participation data aggregated despite different rollout timing
Florida	12051	Hendy County	Participation data aggregated despite different rollout timing
Colorado	8005	Arapahoe County	Participation data aggregated despite different rollout timing
Colorado	8001	Adams County	Participation data aggregated despite different rollout timing
Colorado	8035	Douglas County	Participation data aggregated despite different rollout timing
Colorado	8035	Douglas County	Participation data aggregated despite different rollout timing
Colorado	8119	Teller County	Participation data aggregated despite different rollout timing
Colorado	8041	El Paso County	Participation data aggregated despite different rollout timing
Colorado	8019	Clear Creek County	Participation data aggregated despite different rollout timing
Colorado	8047	Gilpin County	Participation data aggregated despite different rollout timing
Colorado	8049	Grand County	Participation data aggregated despite different rollout timing
Colorado	8097	Pitkin County	Participation data aggregated despite different rollout timing
Colorado	8057	Jackson County	Participation data aggregated despite different rollout timing
Participation data aggregated with other counties			
Virginia	51580	Covington City	Aggregate county
	51005	Alleghany County	
Virginia	51041	Chesterfield County	Aggregate county
	51570	Colonial Heights City	
Virginia	51059	Fairfax County	Aggregate county
	51600	Fairfax City	
	51610	Falls Church City	
Virginia	51163	Rockbridge County	Aggregate county
	51530	Buena Vista City	
	51678	Lexington City	
Virginia	51153	Prince William County	Aggregate county

	51683	Manassas City	
	51685	Manassas Park City	
Virginia	51790	Staunton City	Aggregate county*
	51015	Augusta County	
	51820	Waynesboro City	
Virginia	51019	Bedford County	Aggregate county
	51515	Bedford City	
Virginia	51081	Greensville County	Aggregate county
	51595	Emporia City	
Virginia	51199	York County	Aggregate county
	51735	Poquoson City	
Virginia	51161	Roanoke County	Aggregate county
	51775	Salem City	
Virginia	51083	Halifax County	Aggregate county
	51780	South Boston City	
Virginia	51177	Spotsylvania County	Aggregate county*
	51630	Fredericksburg City	
Virginia	51195	Wise County	Aggregate county*
	51720	Norton City	
New Mexico	35039	Rio Arriba County	Aggregate county
	35028	Los Alamos County	
New Mexico	35061	Valencia County	Aggregate county
	35006	Cibola County	
Colorado	8051	Gunnison County	Aggregate county
	8053	Hinsdale County	
Colorado	8105	Rio Grande County	Aggregate county
	8079	Mineral County	
Montana	30065	Musselshell County	Aggregate county
	30069	Petroleum County	
Nebraska	31101	Keith County	Aggregate county
	31005	Arthur County	
Nebraska	31105	Kimball County	Aggregate county
	31007	Banner County	
Nebraska	31115	Loup County	Aggregate county
	31071	Garfield County	
North Dakota	38033	Golden Valley County	Aggregate county
	38007	Billings County	
North Dakota	38011	Bowman County	Aggregate county
	38087	Slope County	
South Dakota	46071	Jackson County	Aggregate county
	46131	Washabaugh County	
South Dakota	46015	Brule County	Aggregate county
	46017	Buffalo County	
New York	36061	New York City	Aggregate county
	36005	Bronx County	
	36047	Kings County	
	36081	Queens County	
	36085	Richmond County	

Notes: Table lists counties dropped from all analysis samples in the first row of Appendix Table A1 because they are missing census, participation, or health data in all years. Counties without complete census data lack population and/or poverty population information for 1960 and/or 1970. Food Stamp Program participation data aggregates New York City boroughs, so we aggregate census population data similarly. Therefore, Bronx, Kings, Queens, and Richmond are not tracked as individual county equivalents but their population information is summed into one county equivalent.

Table A3: Birth Weight – AHS and Replication

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	AHS				Replication			
White:								
Average FSP (0/1)	2.039	2.635	2.089	2.175	2.055	2.465	2.076	2.143
(SE)	(0.947)	(0.896)	(1.039)	(0.975)	(0.948)	(0.910)	(1.030)	(1.026)
Observations	97785	97785	97785	97785	96846	96846	96787	96787
Mean of dependent variable	3350	3350	3350	3350	3351	3351	3350	3350
Black:								
Average FSP (0/1)	3.454	4.120	5.466	1.665	3.479	3.648	4.637	1.207
(SE)	(2.660)	(2.317)	(2.579)	(2.330)	(2.752)	(2.340)	(2.557)	(2.370)
Observations	27374	27374	27374	27374	26967	26967	26909	26911
Mean of dependent variable	3097	3097	3097	3097	3097	3097	3097	3097
Controls included:								
1960 CCDB × linear time	X	X	X		X	X	X	
REIS controls	X	X	X	X	X	X	X	X
County per capita real income	X	X	X	X	X	X	X	X
Year-quarter fixed effects	X	X	X	X	X	X	X	X
County fixed effects	X	X	X	X	X	X	X	X
State × linear time		X				X		
State × year fixed effects			X				X	
County × linear time				X				X

Notes: Columns 1 through 4 report the results reported in AHS (their Table 1 Columns 1-4). Columns 5 - 8 report our replication results. There, each parameter is from a separate regression of the outcome variable on the food stamp implementation dummy as in Equation 25, with the additional controls as indicated. The treatment is assigned as of three months prior to birth (proxy for beginning of the third trimester). The sample includes race-county-quarter for years including 1968 through 1977 where cells with fewer than 25 births are dropped. In addition to the fixed effects, controls include 1960 county variables (log of population, percentage of land in farming, percentage of population Black, urban) each interacted with a linear time trend, per capita county transfer income (public assistance, medical care, and retirement), and county real per capita income. Estimates are weighted using the number of births in the cell and are clustered on county. Standard errors are in parentheses. Note that race group is determined by the race of the mother in each birth observation.

Table A4: Share Low Birth Weight – AHS and Replication

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	AHS				Replication			
White:								
Average FSP (0/1)	-0.0006	-0.0006	-0.0006	-0.0006	-0.0006	-0.0006	-0.0005	-0.0006
(SE)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0004)
Observations	97785	97785	97785	97785	96846	96846	96787	96787
Mean of dependent variable	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
Black:								
Average FSP (0/1)	-0.0015	-0.0016	-0.0019	-0.0009	-0.0016	-0.0015	-0.0016	-0.0009
(SE)	(0.0010)	(0.0010)	(0.0012)	(0.0012)	(0.0010)	(0.0011)	(0.0012)	(0.0012)
Observations	27374	27374	27374	27374	26967	26967	26909	26911
Mean of dependent variable	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13
Controls included:								
1960 CCDB $\times$ linear time	X	X	X		X	X	X	
REIS controls	X	X	X	X	X	X	X	X
County per capita real income	X	X	X	X	X	X	X	X
Year-quarter fixed effects	X	X	X	X	X	X	X	X
County fixed effects	X	X	X	X	X	X	X	X
State $\times$ linear time		X				X		
State $\times$ year fixed effects			X				X	
County $\times$ linear time				X				X

Notes: Columns 1 through 4 report the results reported in AHS (their Table 1 Columns 5-8). Columns 5 - 8 report our replication results. There, each parameter is from a separate regression of the outcome variable on the outcome variable on the food stamp implementation dummy as in Equation 25, with the additional controls as indicated. See Appendix Table A3 for more details on sample and included fixed effects.

Table A5: Birth Weight, AHS Specification to Stacked Dataset Estimates

	AHS Replication				Stacked Dataset									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
A: White														**Baseline**
FSP	2.055	2.465	2.076	2.143	0.694	1.271	2.156	-0.221	2.557	-0.572	2.395	-0.661	2.589	-0.797
(SE)	0.934	0.897	1.030	1.026	0.985	1.380	1.618	3.107	1.609	2.795	1.716	3.532	1.699	3.788
Observations	96787	96787	96787	96787	97027	48504	5735861	4523507	5888053	4675699	14379194	5341844	14379194	5341844
Mean of dependent variable	3350	3350	3350	3350	3350	3349	3336	3336	3337	3338	3332	3333	3332	3333
B: Black														
FSP	3.479	3.648	4.637	1.207	4.523	3.876	6.609	17.922	6.000	11.120	9.587	16.573	9.745	17.633
(SE)	2.704	2.300	2.557	2.370	4.143	3.190	5.784	7.494	5.459	7.215	4.685	7.419	4.870	7.800
Observations	26911	26911	26909	26911	27083	11645	1372747	969672	1406268	1003193	4738470	1629833	4738470	1629833
Mean of dependent variable	3097	3097	3097	3097	3097	3097	3098	3098	3099	3099	3103	3103	3103	3103
Specification details:														
Estimating equation	25	25	25	25	25	25	26	26	27	27	27	27	27	27
Rollout July 1969 or later only						X	X	X	X	X	X	X	X	X
Linear combination									X	X	X	X	X	X
Birth count weights	X	X	X	X	X	X	X		X		X		X	
Baseline weights								X		X		X		X
Baseline limitations											X	X	X	X
Baseline treatment definition													X	X
Controls included:														
Time, County FEs	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1960 CCDB × linear time	X	X	X											
REIS controls	X	X	X	X										
County per capita real income	X	X	X	X										
State × linear time		X												
State × year fixed effects			X											
County × linear time				X										

Table reports estimates from Equations 25, 26, and 27. Treatment is assigned as of three months prior to birth (proxy for beginning of the third trimester) in Columns 1–12 and as of Food Stamp Program rollout in Columns 13–14 (the "baseline treatment definition"). Columns 1–4 show specifications analogous to Columns 5–8 in the replication of AHS in Table A4. In these columns, additional controls are included as AHS specified. Specifically, these controls include 1960 county variables in Columns 1–3 (log of population, percentage of land in farming, percentage of population Black, urban) each interacted with a linear time trend, per capita county transfer income in Columns 1–4 (public assistance, medical care, and retirement), and county real per capita income in Columns 1–4. Subsequent columns toggle on changes in specification or sample. "Linear combination" means we take an average of per-period event study coefficients. "Baseline weights" mean we use the weights recommended by Wing et al. (2024); see Appendix C.2. In Columns 1–6, the estimation sample is race-county-quarter for years including 1968 through 1977 where cells with fewer than 25 births are dropped. Starting in Column 7, we restructure the sample to account for different estimation approaches. Here, we form the sample at the race-county-month due to the stacking based upon unique rollout month. Similarly, quarter-year and county fixed effects are included across Columns 1–6, and including and after Column 7, for stacked dataset regressions, fixed effects are month-year-by-dataset and county-by-dataset. In Column 9 and after, the estimate is a linear combination (average difference) of 3-year pre-/post-period coefficients in place of the program indicator. Specifically, we include all county-years 1968 through 1978 and only combine the coefficients for years -3 to +2 for the estimate. In Column 11 and after, we apply additional sample restrictions that construct our "baseline sample." These limitations include ensuring each race-county has participation and birth weight observed in each year 1970–1975 and at least one observed year in relative years -3 to +2 for inclusion in linear combination estimates. We also move to dropping race-county-years, instead of race-county-quarters, with fewer than 25 births at Column 11. Across all columns, standard errors are all clustered on county and are in the row directly below the estimate, and the race group is determined by the race of the mother in each birth observation.

Table A6: Share Low Birth Weight, AHS Specification to Stacked Dataset Estimates

	AHS Replication				Stacked Dataset									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
A: White														**Baseline**
FSP	-0.00059	-0.00056	-0.00055	-0.00058	-0.00038	-0.00032	-0.00075	-0.00058	-0.00061	-0.00012	-0.00087	-0.00041	-0.00094	-0.00091
(SE)	0.00027	0.00029	0.00030	0.00038	0.00031	0.00042	0.00058	0.00102	0.00058	0.00098	0.00060	0.00132	0.00060	0.00141
Observations	96787	96787	96787	96787	97027	48504	5735861	4523507	5888053	4675699	14379378	5341914	14379378	5341914
Mean of dependent variable	0.0646	0.0646	0.0646	0.0646	0.0646	0.0645	0.0670	0.0669	0.0667	0.0667	0.0676	0.0675	0.0676	0.0675
B: Black														
FSP	-0.00158	-0.00154	-0.00162	-0.00087	-0.00226	-0.00080	-0.00042	-0.00394	0.00059	-0.00062	-0.00130	-0.01130	-0.00192	-0.01188
(SE)	0.00102	0.00104	0.00117	0.00121	0.00114	0.00159	0.00188	0.00406	0.00193	0.00392	0.00183	0.00417	0.00180	0.00429
Observations	26911	26911	26909	26911	27083	11645	1372747	969672	1406268	1003193	4739194	1630081	4739194	1630081
Mean of dependent variable	0.1305	0.1305	0.1305	0.1305	0.1305	0.1308	0.1321	0.1321	0.1318	0.1317	0.1321	0.1320	0.1321	0.1320
Specification details:														
Estimating equation	25	25	25	25	25	25	26	26	27	27	27	27	27	27
Rollout July 1969 or later only						X	X	X	X	X	X	X	X	X
Linear combination									X	X	X	X	X	X
Birth count weights	X	X	X	X	X	X	X		X		X		X	
Baseline weights								X		X		X		X
Baseline limitations											X	X	X	X
Baseline treatment definition													X	X
Controls included:														
Time, County FEs	X	X	X	X	X	X	X	X	X	X	X	X	X	X
1960 CCDB × linear time	X	X	X											
REIS controls	X	X	X	X										
County per capita real income	X	X	X	X										
State × linear time		X												
State × year fixed effects			X											
County × linear time				X										

Table reports estimates from Equations 25, 26, and 27. Treatment is assigned as of three months prior to birth (proxy for beginning of the third trimester) in Columns 1–12 and as of Food Stamp Program rollout in Columns 13–14 (the "baseline treatment definition"). Columns 1–4 show specifications analogous to Columns 5–8 in the replication of AHS in Table A4. In these columns, additional controls are included as AHS specified. Specifically, these controls include 1960 county variables in Columns 1–3 (log of population, percentage of land in farming, percentage of population Black, urban) each interacted with a linear time trend, per capita county transfer income in Columns 1–4 (public assistance, medical care, and retirement), and county real per capita income in Columns 1–4. Subsequent columns toggle on changes in specification or sample. "Linear combination" means we take an average of per-period event study coefficients. "Baseline weights" mean we use the weights recommended by Wing et al. (2024); see Appendix C.2. In Columns 1–6, the estimation sample is race-county-quarter for years including 1968 through 1977 where cells with fewer than 25 births are dropped. Starting in Column 7, we restructure the sample to account for different estimation approaches. Here, we form the sample at the race-county-month due to the stacking based upon unique rollout month. Similarly, quarter-year and county fixed effects are included across Columns 1–6, and including and after Column 7, for stacked dataset regressions, fixed effects are month-year-by-dataset and county-by-dataset. In Column 9 and after, the estimate is a linear combination (average difference) of 3-year pre-/post-period coefficients in place of the program indicator. Specifically, instead of limiting to years 1968 through 1977, we include all county-years and only combine the coefficients for years -3 to +2 for the estimate. In Column 11 and after, we apply additional sample restrictions that construct our "baseline sample." These limitations include ensuring each race-county has participation and birth weight observed in each year 1970–1975 and at least one observed year in relative years -3 to +2 for inclusion in linear combination estimates. We also move to dropping race-county-years, instead of race-county-quarters, with fewer than 25 births at Column 11. Across all columns, standard errors are all clustered on county and are in the row directly below the estimate, and the race group is determined by the race of the mother in each birth observation.

Table A7: Ln(Birth Weight), AHS Specification to Stacked Dataset Estimates

				Stacked Dataset							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
A: White											**Baseline**
FSP	0.00062	0.00022	0.00040	0.00067	-0.00006	0.00079	-0.00019	0.00073	-0.00025	0.00078	-0.00039
(SE)	0.00031	0.00030	0.00042	0.00049	0.00092	0.00049	0.00084	0.00052	0.00107	0.00052	0.00114
Observations	96787	97027	48504	5735861	4523507	5888053	4675699	14379194	5341844	14379194	5341844
Mean of dependent variable	8.1163	8.1163	8.1160	8.1109	8.1109	8.1113	8.1115	8.1087	8.1089	8.1087	8.1089
A: Black											
FSP	0.00154	0.00150	0.00127	0.00206	0.00593	0.00184	0.00355	0.00308	0.00495	0.00317	0.00556
(SE)	0.00083	0.00134	0.00104	0.00191	0.00253	0.00178	0.00245	0.00151	0.00255	0.00157	0.00270
Observations	26909	27083	11645	1372747	969672	1406268	1003193	4738470	1629833	4738470	1629833
Mean of dependent variable	8.0374	8.0373	8.0372	8.0356	8.0356	8.0358	8.0359	8.0343	8.0343	8.0343	8.0343
Specification details:											
Estimating equation	25	25	25	26	26	27	27	27	27	27	27
Rollout July 1969 or later only			X	X	X	X	X	X	X	X	X
Linear combination						X	X	X	X	X	X
Birth count weights	X	X	X	X		X		X		X	
Baseline weights					X		X		X		X
Baseline limitations								X	X	X	X
Baseline treatment definition										X	X
Controls included:											
Time, County FEs	X	X	X	X	X	X	X	X	X	X	X
1960 CCDB × linear time	X										
REIS controls	X										
County per capita real income	X										
State × year fixed effects	X										

Table reports estimates from Equations 25, 26, and 27. Treatment is assigned as of three months prior to birth (proxy for beginning of the third trimester) in Columns 1–9 and as of Food Stamp Program rollout in Columns 10–11 (the "baseline treatment definition"). Column 1 shows specifications analogous to Columns 5 in the replication of AHS in Table A4. In this column, additional controls are included as AHS specified. Specifically, these controls include 1960 county variables (log of population, percentage of land in farming, percentage of population Black, urban) each interacted with a linear time trend, per capita county transfer income (public assistance, medical care, and retirement), and county real per capita income. Subsequent columns toggle on changes in specification or sample. "Linear combination" means we take an average of per-period event study coefficients. "Baseline weights" mean we use the weights recommended by Wing et al. (2024); see Appendix C.2. In Columns 1–3, the estimation sample is race-county-quarter for years including 1968 through 1977 where cells with fewer than 25 births are dropped. Starting in Column 4, we restructure the sample to account for different estimation approaches. Here, we form the sample at the race-county-month due to the stacking based upon unique rollout month. Similarly, quarter-year and county fixed effects are included across Columns 1–3, and including and after Column 4, for stacked dataset regressions, fixed effects are month-year-by-dataset and county-by-dataset. In Column 6 and after, the estimate is a linear combination (average difference) of 3-year pre-/post-period coefficients in place of the program indicator. Specifically, we include all county-years 1968 through 1978 and only combine the coefficients for years -3 to +2 for the estimate. In Column 8 and after, we apply additional sample restrictions that construct our "baseline sample." These limitations include ensuring each race-county has participation and birth weight observed in each year 1970–1975 and at least one observed year in relative years -3 to +2 for inclusion in linear combination estimates. We also move to dropping race-county-years, instead of race-county-quarters, with fewer than 25 births at Column 8. Across all columns, standard errors are all clustered on county and are in the row directly below the estimate, and the race group is determined by the race of the mother in each birth observation.

Table A8: Welfare Valuation

	Shares of types		Welfare weights		Private valuations		Distortions		Fiscal externalities		Avg demand	Welfare valuations	
	σ_a (1)	σ_c (2)	λ_a (3)	λ_c (4)	V_a (5)	V_c (6)	D_a (7)	D_c (8)	FE_a (9)	FE_c (10)	V(1) (11)	W_a (12)	W_c (13)
1. **Baseline estimate**	**0.804**	**0.196**	**3.22**	**6.18**	**0.696**	**0.175**	**0**	**0**	**0**	**0**	**0.593**	**2.24**	**1.08**
Distortions (D_θ)													
2. Distortion equal to valuation	0.804	0.196	3.22	6.18	0.696	0.175	0.696	0.175	0	0	0.593	4.48	2.16
3. Minimal distortion to overturn results	0.804	0.196	3.22	6.18	0.696	0.175	0.391	0.391	0	0	0.593	3.50	3.50
4. Minimal complier distortion to overturn results	0.804	0.196	3.22	6.18	0.696	0.175	0	0.187	0	0	0.593	2.24	2.24
Welfare Weights (λ_θ)													
5. Double complier welfare weights	0.804	0.196	3.22	12.36	0.696	0.175	0	0	0	0	0.593	2.24	2.16
6. CRRA parameter = 50	0.804	0.196	3.30	10.88	0.696	0.175	0	0	0	0	0.593	2.30	1.90
7. Welfare weight = 1 iff income $<$ 50% FPL	0.804	0.196	2.31	13.66	0.696	0.175	0	0	0	0	0.593	1.61	2.39
8. Welfare weight = 1 iff income $<$ 63.1% FPL	0.804	0.196	2.84	11.40	0.696	0.175	0	0	0	0	0.593	1.97	2.00
Net Private Value (V_θ)													
9. Upper bound for V_a, V_c (with bonus reform)	0.804	0.196	3.22	6.18	0.871	0.350	0	0	0	0	0.769	2.81	2.16
10. Lower bound for V_a, V_c (with bonus reform)	0.804	0.196	3.22	6.18	0.520	0	0	0	0	0	0.418	1.67	0
11. Lower bound for V_a (with bonus), upper bound for V_c	0.804	0.196	3.22	6.18	0.520	0.350	0	0	0	0	0.486	1.67	2.16
First Stage (σ)													
12. Smallest first stage (state×year fixed effects)	0.883	0.117	3.22	6.18	0.696	0.175	0	0	0	0	0.635	2.24	1.08
13. Largest first stage (log female population 15–44×year)	0.795	0.205	3.22	6.18	0.696	0.175	0	0	0	0	0.589	2.24	1.08
Fiscal Externalities (FE)													
14. Conservative FE from Hoynes and Schanzenbach (2012) and Hendren and Sprung-Keyser (2020)	0.804	0.196	3.22	6.18	0.696	0.175	0	0	0.160	0.160	0.593	1.93	0.93
15. Conservative FE, always-takers only	0.804	0.196	3.22	6.18	0.696	0.175	0	0	0.160	0	0.593	1.93	1.08
16. Necessary FE to overturn result, always-takers only	0.804	0.196	3.22	6.18	0.696	0.175	0	0	1.071	0	0.593	1.08	1.08

Note: The first row presents our baseline estimate; there, social marginal welfare weights are estimated from the PSID and constructed by assuming CRRA utility with risk aversion parameter of 3, private valuations assume piecewise-linear demand with valuations at the midpoint of identified bounds with the observation of the bonus reform, and distortions and fiscal externalities or assumed to be zero for both compliers and always-takers. We present sensitivity analyses varying the first-stage estimate, welfare weights, and distortion assumptions.

- $W_\theta = \frac{\lambda_\theta(V_\theta + D_\theta)}{1+\mathrm{FE}_\theta}$ corresponds to the welfare ratio for type θ. W_a is equivalent to $W(0.65)$ from Equation 15.
- Rows 2–4 make assumptions about the distortions. Row 2 assume distortion equals private valuations. Rows 3 and 4 assume the minimal distortion necessary to overturn the result.
- Rows 5–8 make different assumptions about welfare weights. Row 5 doubles compliers' welfare weights λ_c. Row 6 re-estimates λ_c if we use CRRA utility with a risk aversion parameter of 50. Rows 7–8 estimate λ_c using a social welfare function that places a weight of 1 only on households whose annual family income is below 50% FPL (Row 7) or below 63.1% FPL (Row 8) and 0 otherwise; Row 8 shows the highest cutoff that would suffice to overturn our result under this class of welfare weights.
- Rows 9–11 assume different estimates of V_a and V_c that are consistent with revealed preference in Figure 11.
- Rows 12 and 13 make different assumptions about σ_c, based on robustness specifications for the first-stage (Figure A8).
- Our primary estimate for fiscal externalities is 0. In Rows 14 and 15, we use a fiscal externality of 0.16 (Hoynes and Schanzenbach, 2012; Hendren and Sprung-Keyser, 2020). Row 16 presents the always-taker fiscal externality sufficient to overturn the result if compliers have zero FE.

Table A9: Number of Observations, PSID and NHANES

Dataset	FS enrollment	1973	1974	1975	1976	1977	1978	1979	1980	1981
NHANES	Not enrolled	0	0	0	4,712	3,906	4,481	4,692	635	0
NHANES	Enrolled	0	0	0	526	525	353	413	40	0
PSID	Not enrolled	4,786	4,859	4,928	5,113	5,341	5,524	5,518	5,551	5,677
PSID	Enrolled	730	865	933	893	812	848	1,014	1,068	1,064

Table reports the count of observations by dataset-enrollment status across years for the NHANES and PSID. Note that observations are individuals for NHANES and households for PSID.

Table A10: Estimates of Inframarginality

Variable	(1) Pre-period mean (always-takers) (1977-1978)	(2) Post-period mean (1979-1980)	(3) Complier mean (Estimated from Aggregate FS)
Panel A: Share weakly inframarginal			
Food spending $\geq F$	0.93	0.84	0.45
	(0.01)	(0.02)	(0.10)
Food spending $\geq 0.9 \cdot F$	0.95	0.88	0.57
	(0.01)	(0.02)	(0.10)
Food spending $\geq 1.1 \cdot F$	0.88	0.79	0.44
	(0.02)	(0.02)	(0.13)
Panel B: Share strongly inframarginal			
Food spending $\geq R + F$	0.63	0.46	-0.21
	(0.04)	(0.03)	(0.20)
Food spending $\geq 0.9 \cdot (R + F)$	0.83	0.59	-0.37
	(0.02)	(0.03)	(0.17)
Food spending $\geq 1.1 \cdot (R + F)$	0.49	0.39	-0.00
	(0.04)	(0.03)	(0.21)

Notes: Table shows the estimated share of compliers meeting certain food spending criteria based on data from the NFCS and the federal register, and using standard complier algebra based on the estimated complier share of 0.20 (see Section 3.2). The pre- and post-period shares are weighted means of an indicator for the given variable in the 1977-1978 Nationwide Food Consumption Survey. In the post period, benefits are assigned using net income as it was defined in the pre-period and $R + F$ is determined using the counterfactual R, as described in Appendix B.2. The sample is restricted to households in the NFCS for which the income, household size, and food spending variables are defined. Standard errors are shown in parentheses and are calculated using the delta method, assuming zero covariance between the pre-reform mean, the post-reform mean, and the estimate of the complier share (whose estimation error affects the estimated standard error). See Appendix B.2 for more details.