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

Amy Finkelstein Matthew Notowidigdo Charlie Rafkin

October 29, 2026

Abstract

We analyze the food stamp program both as a potential investment in children's health capital and as a redistributive tool by studying the 1979 reform that made food stamps free. We leverage geographic variation in exposure to the reform, which raised the program's food subsidy from 65% to 100%. This raised enrollment by 1.85 percentage points (standard error = 0.15), or about 25%, with new enrollees substantially poorer than existing participants. From an investment perspective, we do not find evidence of improvement in infant health, and 95% confidence intervals rule out gains in birth weight from enrolling in food stamps exceeding 2%. From a redistribution perspective, the enrollment response implies that the average recipient of free food stamps acts as if they value them at 59 cents per dollar, with nonparametric bounds of 42–77 cents. Despite this substantial dissipation in transfer value relative to a frictionless cash transfer, under standard redistributive preferences the free food stamp program increases social welfare because of its targeting to very low-income households. The move to free stamps from a partially subsidized program improved targeting on social value but worsened 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, Marianne Bitler, Eric Budish, Manasi Deshpande, Keith Ericson, Robert Greenstein, Hilary Hoynes, Lee Lockwood, Sarah Miller, Jesse Shapiro, and Mark Shepard 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, 2016; Carrington et al., 2013). This $100-billion-a-year program provides food assistance to 40 million Americans, about 35% of whom are children (U.S. Department of Agriculture, Economic Research Service, 2025). Two central rationales for the food stamp program are that it makes a social investment in children's health capital and that it redistributes to low-income recipients (Hoynes and Schanzenbach, 2018).

An important and perhaps underappreciated aspect of the food stamp program is that, in its first two decades, it provided only a partial food subsidy. Under the original "purchase requirement," households had to pay the program in cash to get food stamps in return; by 1979, the price was about 35 cents per food stamp dollar for the average household, meaning the program served as a 65% subsidy on food. A policy reform in 1979 made food stamps free, like they are today: participants pay nothing, and food stamps serve as a 100% subsidy on food up to the maximum government benefit. The reform did not change the government benefit provided by food stamps. For example, a household no longer had to pay $35 to receive $100 in food stamps; instead, they received $65 in free food stamps (Congressional Research Service, 1977; Richardson, 1979).

This major redesign of the food stamp program provides an opportunity to enhance our understanding of both its investment and redistributive properties. Our analyses of both properties leverage the increase in enrollment from making food stamps free, exploiting geographic variation in exposure to the reform for identification. Prior to the reform, counties varied substantially in the share of the population that was eligible for — but not enrolled in — food stamps, and therefore in the scope for new enrollment in response to the increase in subsidy. Using this exposure design, we estimate that the move to free food stamps raised enrollment by 1.85 percentage points (standard error = 0.15), or about 25 percent. New enrollees were substantially poorer — with about 60% of the average household income of those who enrolled under the purchase requirement.

From the perspective of investing in child health, prior research finds that food stamps improve health in infancy, childhood, and adulthood. This work has been very influential in both academic and policy debates over the safety net (e.g., National Academies of Sciences, Engineering, and Medicine, 2019). Yet, the evidence on food stamps' impacts on health — as well as on other outcomes such as adult educational attainment, employment, earnings, and crime — comes almost entirely from a single natural experiment: the program's county-by-county rollout in the 1960s and 1970s, when the purchase requirement was still in place (Almond et al., 2011; Hoynes et al., 2016; Barr and Smith, 2023; Bitler and Figinski, 2024; Bailey et al., 2024). It has been difficult to find other sources of quasi-experimental variation in this uniform national program to study its impacts on child health.[1]

The free-stamps reform thus provides a valuable opportunity to learn about food stamps' infant-

[1]Two important exceptions are papers that use, respectively, 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). Both analyses examine and find improvements in child health; East (2020) also studies and finds improvements in infant birth weight. Researchers have used other sources of variation as instruments for food stamp enrollment — including federally legislated changes in benefits, state-level changes that affect access, the transition to electronic benefits, and household exit due to failure to recertify (e.g., Hastings and Shapiro, 2018; Ganong and Liebman, 2018; Melvin and Smith, 2022; Song, 2022) — but have not used these to study impacts on infant or child health, likely reflecting limitations of available data linkages.

health investment properties in a period when the program more closely resembled its current form. Following Almond et al. (2011), our primary analysis is of birth weight, an important measure of infant health (Black et al., 2007). We also examine impacts on infant mortality and gestational age. For comparison, and to combine multiple natural experiments for greater precision, we also revisit the program rollout as another instrument for enrollment.

We do not find evidence that food stamp enrollment improves birth weight on average, whether we instrument for enrollment using the 1979 free food stamp reform, the initial rollout, or both. IV estimates from the free food stamp reform find no statistically significant impact of food stamp enrollment on a continuous measure of birth weight or on the share of births that are low birth weight. Point estimates indicate that food stamp enrollment *reduces* birth weight by a statistically insignificant 0.4%, with the 95% confidence interval ruling out an increase in birth weight of more than 2%. IV estimates from the initial rollout also show no statistically significant impact on either birth weight or share low birth weight.[2] Our confidence intervals from the free-stamps reform are about half as large as those using the rollout variation. Precision-weighted IV estimates that combine both reforms under a homogeneous treatment-effects assumption remain wrong-signed for birth weight (–0.3%), and the 95% confidence interval can rule out that food stamps increase birth weight by more than 1.9%. Precision-weighted estimates for the share with low birth weight are also null (point estimate of –0.7 percentage points) but less powered (standard error = 1.3). To contextualize magnitudes, during our study period, Whites had about 8% higher birth weight and a 6 percentage point lower risk of being low birth weight than Blacks.

Looking more broadly at other measures of infant health and at subpopulations, we find two suggestive but fragile pieces of evidence that food stamps may have positive effects on some margins of infant health. In the free-stamps reform, food stamps reduce rates of preterm birth, and in the rollout, food stamps produce statistically and economically significant improvements in birth weight for Blacks (consistent with Almond et al., 2011). However, neither result is present in the other reform, and both results come with caveats, discussed below, that raise questions about their robustness.

At the same time, these analyses do not speak to the redistributive rationale for the food stamp program. Food stamps' causal effects on infant health — or indeed on other outcomes — are neither necessary nor sufficient for the program to redistribute effectively. The program's redistributive value depends on recipients' private value and their social welfare weight (Lieber and Lockwood, 2019). In the second part of the paper, therefore, we use our empirical estimates of the free-stamps reform's impacts on enrollment to infer recipients' perceived private value of the program, and we use a standard social welfare function to translate recipients' incomes into social welfare weights. This allows us to characterize the food stamp program's redistributive value under both free stamps and the purchase requirement.

Estimating the private value of goods when prices are not observed is a well-known challenge for welfare analysis of in-kind government programs (e.g., Samuelson, 1954; Finkelstein et al., 2019a; Autor et al., 2019; Landais and Spinnewijn, 2021). A valuable feature of the free-stamps reform is that it created quasi-experimental variation in the price of enrolling in the program, allowing us to

[2] Almond et al. (2011) do not report effects on the average infant (only by race), so these results are not directly comparable to their findings. We successfully replicate their results by race (see Appendix B.6).

trace out a "demand curve" for food stamps. We incorporate additional price variation by using an earlier reform that raised the average subsidy from 40 to 65% and, we estimate, raised enrollment by 2.4 percentage points (standard error = 0.22). We interpret the estimated demand curve as reflecting recipients' perceived private value of the transfer, net of enrollment costs ("ordeals"). If food stamps were fungible with cash and involved no ordeals, then making food stamps free would not have raised enrollment, since the government transfer did not change. Seen this way, the estimated first-stage enrollment response to the increase in subsidy indicates that households act as if they value food stamps at less than dollar-for-dollar.

The demand curves imply that enrollees act as if their average private value of free food stamps is 59 cents per dollar, with nonparametric bounds of 42 to 77 cents. Despite the dissipation of about two-fifths of food stamps' private value, we find that, due to enrollees' very low incomes (and hence high social welfare weights), the free stamps program still raises social surplus relative to no program. Intuitively, deadweight loss in free food stamps reduces — but does not fully eliminate — the program's redistributive value. The 41 cents of deadweight loss per dollar of government spending on food stamps reflects a combination of "ordeals" (that is, hassle costs or stigma) and distorted consumption from the in-kind nature of the transfer. Since estimates from both the 1970s and the current time period suggest that most food stamp enrollees are "inframarginal," i.e., they buy more food than their food stamps benefit (e.g., Hoynes et al., 2015b; Hastings and Shapiro, 2018; MacDonald, 1977; Moffitt, 1989), we interpret the deadweight loss estimate as largely reflecting ordeals.

In our final normative analysis, we consider the social welfare tradeoffs between zero cost-sharing (free stamps) and positive cost-sharing (purchase requirement). Basic economic logic points toward imposing a positive price on food stamps to improve targeting on perceived *private* value; partial subsidies screen out households with low willingness to pay for the transfer. However, in the context of transfer programs, eligible households may have mechanically low willingness to pay because they have little disposable income. Positive prices can therefore reduce targeting on *social* value by screening out households with the highest social value from redistribution. Whether higher or lower subsidies in food stamps would be socially desirable is thus an empirical question.[3]

Our demand curve estimates imply that compliers to the free food stamp reform act as if they have about one-quarter the private value of always-takers. However, they are also substantially poorer. Under standard redistributive preferences, we find that positive cost-sharing delivers greater social welfare per dollar of government spending than free stamps; compliers' lower private valuations outweigh their higher welfare weights.

Naturally, these normative conclusions require additional assumptions beyond what is needed for the positive analysis of the impact of food stamps on infant health. They do not require abstracting from potential liquidity constraints; as we discuss, our normative analyses explicitly incorporate

[3]This question is also practically relevant: there is 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).

liquidity constraints, which can affect households' willingness to pay and are reflected in their welfare weights as is standard in the Marginal Value of Public Funds (MVPF) literature (Hendren and Sprung-Keyser, 2020). But they do require calibrating the extent of redistributive preferences; sufficiently strong tastes for redistribution can reverse the conclusion that free stamps deliver lower social welfare per dollar of government spending than positive cost sharing, while sufficiently weak tastes for redistribution can reverse the conclusion that free stamps increase social welfare relative to no program. They also require taking a stand on behavioral biases that could create a wedge between enrollees' perceived valuation revealed by their enrollment response and their actual, normative valuation. Our baseline analyses abstract from potential behavioral biases, and these may well be operative, although the sign of their impact on our normative conclusions is a priori ambiguous. We therefore conduct a range of sensitivity analyses, including characterizing the redistributive preferences and distributions of behavioral biases required to overturn our normative findings. Our main welfare conclusions — that a free stamps program increases social welfare relative to no program, but that a program with positive cost-sharing yields higher social value per dollar of government spending than free stamps — both prove surprisingly difficult to reverse.

Contribution. Our paper relates to several distinct literatures. First, our findings of limited impacts on infant health challenge the case for food stamps as an infant health investment for the average enrolled birth. This is particularly noteworthy in light of earlier evidence that the rollout of food stamps improved infant health, and the substantial attention that finding has received in both the academic literature (e.g., Almond et al., 2018; Hoynes and Schanzenbach, 2018; Aizer et al., 2022) and in policy discussions of safety-net investments in children (e.g., National Academies of Sciences, Engineering, and Medicine, 2019; Carlson and Llobrera, 2022). Interestingly, our null results for food stamps' impacts on infant birth weight differ from a handful of studies of the much smaller Women, Infants, and Children (WIC) program,[4] which find — using the staggered program rollout or family-fixed-effects designs — that enrollment in WIC causes a statistically significant 0.3% to 1.6% increase in average birth weight (Hoynes et al., 2011; Rossin-Slater, 2013; Sonchak, 2016; Chorniy et al., 2020; Bitler et al., 2025). This raises the possibility that nutrition programs that more tightly restrict eligible purchases may be better poised to improve infant health.[5] There may also be longer-run benefits from food stamp receipt on other child and adult outcomes, as well as potentially larger impacts for well-targeted demographic subgroups. Using the free food stamp reform to study these margins will be an important direction for future work.

Second, our normative analyses of the value of food stamps and the costs of ordeals belong to a tradition of using revealed preferences to recover private valuations of social insurance or

[4]WIC is much smaller than food stamps, both in overall size (today providing $8 billion in assistance to about 7 million people) and its average benefit amount for enrollees (U.S. Department of Agriculture, Economic Research Service, 2026).

[5]Studies have also found positive impacts of means-tested cash transfers on infant and young child health in many but not all settings. Evidence of positive impacts comes from transfers like the EITC (Strully et al., 2010; Hoynes et al., 2015a), COVID-19 payments (Ruffini, 2023), and child payments targeted to pregnant women (González and Trommlerová, 2022; Reader, 2023; Agarwal et al., 2026), as well as transfer programs in developing countries (Barber and Gertler, 2008; Amarante et al., 2016). However, other papers have not found positive impacts, including studies of AFDC cash transfers (Currie and Cole, 1993) and randomized cash assistance from the negative income tax experiments (Price and Song, 2026); in addition, Hawkins et al. (2025) find no impact on infant mortality rate of substantial post-birth cash transfers through SSI.

redistributive programs. Just a few of these studies use, as we do, identifying price variation (e.g., Finkelstein et al., 2019b; Landais and Spinnewijn, 2021). More often, this research uses variation in other program features such as benefit sizes or ordeals (e.g., Anderson and Meyer, 1997; Haller and Staubli, 2023; Haller et al., 2024; McQuillan and Moore, 2025; Naik, 2025; Cook and Meckel, 2026), and must therefore rely on stronger assumptions to recover private valuations, since the identifying variation is not denominated in dollars. Our findings complement earlier revealed-preference estimates of food stamps' value coming from labor-supply responses to the program rollout (Hendren and Sprung-Keyser, 2020) and surveys on willingness to pay in the 1990s (Whitmore, 2002; Bailey et al., 2024), which yield private values of 62 and 80 cents per dollar, respectively. The revealed-preference approach has well-appreciated strengths and limitations; it provides a principled starting point for normative analysis, but relies on strong assumptions about optimizing behavior which, as noted above, we investigate carefully in sensitivity tests.

Third, we contribute to a lively literature on the targeting of government programs, and particularly whether non-monetary ordeal costs in transfer programs screen out households with low private value (Nichols and Zeckhauser, 1982) or, instead, exclude those with high social value (Mullainathan and Shafir, 2013; Herd and Moynihan, 2025).[6] From a conceptual perspective, we explore the normative implications of a different policy tool (prices) than the ordeals literature, which often takes as a starting point that governments cannot allocate in-kind transfers through prices (e.g., Yang et al., 2025). Our framework nonetheless features a similar normative tradeoff between targeting on private and social value, as in Lieber and Lockwood (2019)'s analysis of the targeting of in-kind home care benefits compared to cash and Shepard and Wagner (2025)'s analysis of ordeals in health insurance. In this way, our paper relates to recent research studying optimal subsidies (i.e., optimal cost-sharing) for other in-kind transfer programs, including health insurance (e.g., Finkelstein et al., 2019b; Tebaldi et al., 2023), higher education (e.g., Colas et al., 2021; Londoño-Vélez et al., 2025), and child care (e.g., Montpetit et al., 2025). From an empirical perspective, the welfare costs of ordeals for enrollees have proven challenging to estimate. Prior research has conducted back-of-the-envelope calculations of ordeal costs based on applicants' value of time (Deshpande and Li, 2019; Finkelstein and Notowidigdo, 2019; Anders and Rafkin, 2026; Cholli and Wu, 2026). Our demand-curve estimates from the free-stamps reform yield much larger welfare costs of ordeals than such calibrations, which implies that the food stamp program performs less redistribution than is commonly assumed under standard accounting practices that currently treat the program as if it were equivalent to a frictionless cash transfer valued dollar-for-dollar (e.g., Blanchet et al., 2022; Congressional Budget Office, 2025).

The rest of the paper proceeds as follows. Section 1 summarizes the food stamp program and the 1979 free-stamps reform. Section 2 describes the data. Section 3 presents reduced-form evidence of the impact of the free-stamps reform on enrollment, complier characteristics, and infant health. Section 4 presents IV analyses of the impact of food stamps on birth weight, using both the initial program rollout and the free-stamps reform as instruments for enrollment. Section 5 performs normative analyses of the program's redistributive properties. The last section concludes.

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

1 Background: The 1979 Free Food Stamp 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] 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 policy was known as the "purchase requirement" and was in effect for the first two decades of the program's existence. The purchase requirement per food stamp dollar increased with household income and initially averaged about 60 cents (i.e., a 40% subsidy).

Starting in the late 1960s, growing recognition of widespread malnutrition in the United States led advocates to call for eliminating the purchase requirement, which they argued posed a barrier to enrollment for the poorest eligible households (Citizens' Board of Inquiry, 1968). In response, in 1970, the average subsidy was raised to about 65% (Hoover, 1972; 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 1979), which repealed the original Food Stamp Act of 1964 and eliminated the purchase requirement for all food stamp participants, increasing the subsidy to its current 100% level without changing benefit amounts (USDA Food and Nutrition Service, 2025). As a result, instead of having to pay 35 cents to receive $1 of food stamps (and a 65-cent food stamp transfer), individuals now could receive that 65-cent food stamp transfer for free. Notably, the free food stamp reform did not systematically affect the non-monetary enrollment 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 $I + k$ face a choice of whether or not to

[7] The information in this section draws heavily on Richardson (1979), Congressional Budget Office (1977), and USDA Food and Nutrition Service (2025). Appendix A.1 provides additional details.

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

[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.1.3 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) enrollment (or take-up) cost of enrolling in the program (k). For expositional clarity, we graph this enrollment cost as

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 ($\$F \geq 0$). This "bonus" F 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. 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 stamp program (blue budget set).

The figure abstracts from several features of the purchase requirement that will be relevant in some of the normative analyses in Section 5. First, enrollees could buy fractions of their total food stamp bonus F by paying a corresponding fraction of the purchase requirement R. In practice, however, the vast majority of enrollees paid the entire purchase requirement R and received all of the bonus F; this is consistent with the evidence that most enrollees spent more on food than they received in food stamps (Appendix A.2). Second, the food stamp bonus F varied with household size and the purchase requirement R varied with both household size and income in a progressive manner, and was capped at 30% of net income. Indeed, during our analysis period, the poorest households already had zero purchase requirement (although not many households had incomes low enough that they paid zero; see Appendix Table A4).

The 1979 reform changed the program from the black budget set to the blue budget set by eliminating the purchase requirement (R) without changing the amount of the food stamp transfer paid by the government (F). Under the free food stamp program, enrollees now 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 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.

Figure 1 illustrates two other important aspects of the program. First, it assumes that enrollment involves an ordeal cost k, which we model as a monetary enrollment cost.[12] Enrollment costs are 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, 2016), as well as with direct evidence that enrollment costs contribute to incomplete take-up (Schanzenbach, 2009; Finkelstein and Notowidigdo, 2019). Second, the figure's hypothetical enrolled households (households v and z) are depicted as *inframarginal*: 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 A.2). In particular, we estimate that most enrollees under the purchase requirement would have consumed more than $R+F$ in food even in the absence of the

additive, but in the analysis below we do not impose additivity (see Section 5).

[12] Enrollment costs may include the time cost of enrolling and (in the period we are studying) of retrieving food stamps from the local food stamp office each month (Congressional Budget Office, 1977). There may also be psychological enrollment costs like stigma (Moffitt, 1983; Anders and Rafkin, 2026).

program (as illustrated by person z). As a result, removing the purchase requirement had no effect on their budget constraint. We also estimate that most of the compliers who enroll in response to the elimination of the purchase requirement would have consumed more than F in food even absent the program (as illustrated by person v). As a result, their enrollment expanded their budget set but did not affect their marginal price of food.

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 food stamp program) rather than incurring the enrollment cost k.[13] 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 does not participate under the purchase requirement, but does 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. The key outcomes are food stamp enrollment and infant health. Why infant health? 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). In addition, high-quality administrative data make infant health one of the rare outcomes that can be measured consistently during this period and at scale. Finally, important work from the program rollout finds that food stamps improved infant health, and 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.

Food Stamp Enrollment: 1961–81. County-year counts of the number of participants in the food stamp program come from books published by the Food and Nutrition Service. Almond et al. (2011)'s replication package provides these data from 1961–1979; we digitized the data for 1980–81 from

[13] Not depicted for visual ease are never-takers who consume a lot more food but choose not to participate due to sufficiently high enrollment costs k.

books that we found in the Cornell libraries (U.S. Department of Agriculture, 1980–1981a).[14] As an alternative measure, we use data from the Bureau of Economic Analysis' Regional Economic Information System (REIS) (U.S. Department of Commerce Bureau of Economic Analysis, 1961–1981) to examine the total dollars transferred through the food stamp program in each county-year, deflating to $1970 using the CPI-U.

Infant Health: 1968–81. Birth-level data are available nationwide from the National Center for Health Statistics Vital Statistics starting in 1968 (National Center for Health Statistics, 1968–1987).[15] For live births, the data contain birth weight, gestational age, and an indicator for infant death in the first 12 months. They also contain the exact date of birth, state and county of birth, maternal race, and marital status.

Following Almond et al. (2011) and much of the literature on the impact of means-tested transfers on infant health outcomes, our primary outcomes are birth weight and the share of births that are low birth weight, defined as birth weight under 2,500 grams. However, for completeness, we also report supplemental analyses of impacts on infant mortality rate, gestational age, and birth counts. Infant mortality rate is a fortunately rare outcome. Gestational age may not be affected even if birth weight is, as increases in birth weight can come from increased intrauterine growth as well as increased gestational age; indeed, several studies of means-tested transfers find impacts on birth weight but not on gestational age (Bitler et al., 2025; Rossin-Slater, 2013; González and Trommlerová, 2022; Amarante et al., 2016). Gestational age is also missing for a large share of the births during our time period (Bitler et al., 2025).

1970 Census Data. We obtain county-level counts of total population and population in poverty from the 1970 Census.

2.1.2 Survey Data

Characteristics of Enrollees: 1973–81. We use annual data from the Panel Study of Income Dynamics (PSID) (Survey Research Center, 2026) from 1973–81 and data from the National Health and Nutrition Examination Survey (NHANES II) (National Center for Health Statistics, 1976–1980) from 1976–80 to study how demographic characteristics of enrollees change after the free food stamp reform. The PSID is a nationally representative, annual, longitudinal survey of about 4,800 households. The NHANES II is a one-time survey (conducted over five years) of approximately 20,000 individuals; it oversamples populations at elevated risk of malnutrition (including low-income individuals). Neither survey contains sub-state geographic identifiers in public-use data. Appendix B.8 provides more detail.

Food Purchases: 1977–80. To study how enrollees' food purchases compare to food stamp allotments, we use data from the USDA's Nationwide Food Consumption Survey (NFCS, 1977–1980).

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

[15] 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.

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 A.2 provides more detail.

2.2 Construction of Key Variables

The food stamp enrollment rate in county c and year t (FS_{ct}) is defined as:

$$\text{EnrollmentRate}_{ct} \equiv FS_{ct} = \frac{\text{Number of Food Stamp Enrollees in County } c \text{ and Year } t}{\text{Number of People in County } c \text{ in 1970}}. \quad (1)$$

We proxy for the share of households that are eligible for food stamps in a given county with the share of people in the county in poverty in the 1970 Census:[16]

$$\text{Poverty}_c = \frac{\text{Number of People in Poverty in County } c \text{ in 1970}}{\text{Number of People in County } c \text{ in 1970}} \quad (2)$$

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 in poverty (Equation 2) but not enrolled in food stamps (Equation 1):

$$\text{Exposure}_c \equiv \text{Poverty}_c - \text{EnrollmentRate}_{cz}. \quad (3)$$

Our empirical analysis of the free food stamp reform leverages county-level variation in Exposure_c, which captures the "slack" in county-level enrollment: the (approximate) share of people who are eligible for food stamps but not enrolled. This empirical strategy has contemporaneous historical backing, as the 1981 Report to Congress on the reform noted that enrollment in the program grew faster in areas with more slack: "a considerable portion of the observed differences among regions (and States) in participation increases is explained by differing potentials for growth" (U.S. Department of Agriculture, 1981b, p. 4). Intuitively, counties with a greater share of the population that is eligible for but not enrolled in food stamps have 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 our proxy for the county's eligibility rate (the poverty rate) and decreasing in its take-up rate (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 likely eligible households.

For our baseline definition of Exposure_c, we measure the county enrollment count in 1972, $z = 1972$. This choice allows us to examine trends across counties with differential exposure for a reasonably long period prior to the 1979 reform.[17] However, we show below that the choice of year is

[16] In general, eligibility is concentrated among people below the poverty line (Congressional Budget Office, 1977; Richardson, 1979). But poverty is a coarse proxy for eligibility and will include type I and type II errors. That may affect precision but need not bias results.

[17] 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 winsorize these estimates at zero, and for symmetry we winsorize the top 3.2% of the exposure distribution to the 96.8th percentile. Appendix Figure A1 shows the unwinsorized exposure measure and how we winsorized it. In everything else that follows we use the winsorized exposure measure.

inconsequential.

2.3 Sample Restrictions

We form a balanced panel of county-years for our analyses of administrative data. This excludes approximately 5% of (1970-population-weighted) counties that are missing census data, enrollment data, or infant health data in any relevant analysis year. For the analysis of the free food stamp reform, we limit the data to 1972–81 and to the approximately three-quarters of population-weighted counties that had the food stamp program in 1970 or earlier. This restriction ensures 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 use data from 1968 (the start of the birth weight data) through 1978 (the last year prior to the elimination of the purchase requirement). Following Almond et al. (2011)'s event-study sample, we restrict our analysis 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 free food stamp reform 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.

3 Impacts of the Free Food Stamp Reform

3.1 Empirical Framework

We estimate the following event studies:

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

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

where $\mathbf{1}\{t = \tau\}$ is an indicator equal to one if observation year t corresponds to calendar year τ. The parameters α_c^j are county fixed effects, γ_t^j are calendar-year fixed effects, and Exposure_c is our time-invariant, county-level exposure measure (defined in Equation 3). The outcomes FS_{ct} and y_{ct} denote the county-year food stamp enrollment rate (defined in Equation 1) and infant health outcome, respectively. 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 average impacts defined as the difference of these coefficients three years before and after the reform:

$$\beta_{\text{avg}} = \frac{1}{3}[(\beta_{1979} + \beta_{1980} + \beta_{1981}) - (\beta_{1976} + \beta_{1977} + \underbrace{\beta_{1978}}_{\equiv 0})] \tag{6}$$

and similarly for δ_{avg}.

The key identifying assumption for causal interpretation of the δ_τ's and β_τ's in Equations 4 and 5 as the impact of the free-stamps reform is that, absent the free-stamps reform, changes in food stamp enrollment and infant health outcomes between 1978 and each post-reform year would have been uncorrelated with county exposure. This is essentially a parallel trends assumption that counties with different levels of exposure would have experienced similar trends in outcomes in the absence of the reform.

Cross-sectional Patterns. There is considerable cross-sectional variation in Exposure_c across population-weighted counties, ranging from 0 to 0.2, with a mean of 0.06 and a 10 percentage point difference between the 10th and 90th percentile counties (Figure 2a).[18] Exposure_c is higher in the South, consistent with the historical evidence (Citizens' Board of Inquiry, 1968), as well as in the Great Plains region, but also exhibits substantial within-state variation (Figure 2b).

There is also considerable variation in the distribution of the 1978 share of births that are low birth weight across population-weighted counties, ranging from 0 to 0.2, with a mean share of 0.07, and a 4 percentage point difference between the 10th and 90th percentile counties (Figure 3a). Share low birth weight is higher in the Southeast and Southwest, but again exhibits meaningful within-state variation (Figure 3b). Counties with greater exposure to the free-stamps reform have a higher share of births with low birth weight (Figure 3c). This makes sense, since counties with higher poverty rates have mechanically higher exposure (see Equation 3). However, this also raises potential concerns about mean reversion when looking at impacts on birth weight, which we will assess by looking at the pattern of pre-period event-study coefficients.

3.2 Results

3.2.1 Enrollment Response

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.

[18] Appendix Figure A4 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.

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. It shows a sharp jump in enrollment after the elimination of the purchase requirement in more-exposed counties compared to less-exposed counties. The average increase in the enrollment rate three years after the reform relative to three years before in a county that was completely exposed relative to one that had no exposure was 0.325 (standard error = 0.027). Given average exposure prior to the reform of 0.06 (see Figure 2), this implies 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 1976–1978 of about 7.6 percent.[19] This interpretation assumes that the impact on a county with zero exposure would be zero (i.e., no "missing intercept"), which is plausible in our setting since zero exposure implies no "slack" in food stamps take-up that could respond to changes in the subsidy.

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. The increase in enrollment was concentrated at the bottom of the income distribution (Figure 5a); the average income of enrollees fell precipitously at the time of the reform, both in absolute terms and relative to those who were not enrolled (Figure 5b).

More broadly, Table 1 Panel (a) shows that across a variety of measures, the compliers with the free food stamp reform (column 3) look substantially poorer than always-takers (column 1).[20] For example, compliers' average equivalized household income[21] is about 60% of always-takers' income. Compliers are also more than five times more likely to be below 50 percent of the poverty line (39 percent, versus 7 percent for always-takers). These findings are particularly striking in light of the fact that the purchase requirement was progressive, so that lower-income households experienced smaller increases in the subsidy than higher-income households (Appendix Figure A7b). Appendix B.8 contains more detail on the data, variable definitions, and analyses, as well as additional results.

The high level of need among complier households raises the possibility of a health-capital investment channel via impacts on infant health. Indeed, the finding that compliers are lower-income than always-takers suggests the estimated impact of food stamp enrollment on birth weight from the free food stamp program might be larger than what was previously found for the program rollout. Panel (a) also shows that compliers were also more likely to be in households with young children: 80 percent of complier households compared to 55 percent of always-taker households. This serves as further motivation for studying impacts on birth weight.

[19] Appendix Figure A6 shows the estimated impact 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 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.

[20] Complier characteristics are derived from characteristics of enrollees before (column 1) and after (column 2) the free food stamp reform, an estimate of the share of enrollees (σ_c) who are compliers, and standard complier algebra. We estimate $\sigma_c = 1.85 \div (1.85 + 7.6) \approx 0.20$ where 1.85 is the 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.

[21] "Equivalized" denotes that we adjust for the number of adults and children in the household.

3.2.2 Impact on Infant Health

Birth Weight. Figure 6 shows no indication of an 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. The raw data by quartile show 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 average coefficients post-reform relative to pre-reform is -0.0014 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 0.0084 percent (standard error $= 0.024$ percent). For share low birth weight (Panel f), the difference in the estimated average coefficients post-reform relative to pre-reform is -0.0013 (standard error $= 0.0044$), suggesting that the impact on the average county was a statistically insignificant decline in the share low birth weight of 0.0078 percentage points (standard error $= 0.026$ percentage points). Indeed, we see no evidence of a statistically significant impact of the free food stamp reform anywhere in the birth weight distribution (Appendix Figure A8). The only marginally significant impact is a reduction in birth weight in the second-highest birth weight bin — between 4,000 and 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. We also see no impacts when we examine White and Black infants separately (Appendix Figure A9).[22]

Additional Outcomes. 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 preterm births (defined as the 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 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).

Appendix B.1 presents the analysis and results in detail. There is no evidence of an impact of the free food stamp reform on infant mortality rate. There is some evidence that the removal is associated with a decline in births (a somewhat counterintuitive result) and a decline in the share of births that are preterm. However, in both cases, the underlying event studies are not compelling. Moreover, as has been found in other work (e.g., Bitler et al., 2025; Rossin-Slater, 2013), gestational age is missing for a large share (over 20%) of our baseline analysis sample. 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.

[22] For the race-specific analysis, we further restrict the set of counties to the subset of baseline counties that have at least 25 births of that race in every county-year; this retains the vast majority of the race-specific populations represented by the baseline analysis samples (Appendix Table A3).

3.2.3 Identification and Robustness

Identification. The non-experimental nature of the variation in exposure naturally raises questions about the assumption that, absent the 1979 free food stamp reform, counties with different exposure would have experienced similar trends in enrollment and birth outcomes. However, reassuringly, trends 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). Moreover, given the large, immediate, and persistent increase in enrollment in 1979 in more-exposed relative to less-exposed counties, any confound would need to be a large shock in 1979 to food stamp enrollment that differentially affected high-exposure counties relative to low-exposure counties but did not affect infant health. That said, we explore in more detail three general types of threats to the identifying assumption and find little evidence for them.

Mean Reversion. Mean reversion is a particular concern for analyzing the impact of the reform on enrollment rates, since the exposure measure is itself a function of 1972 enrollment rates (see Equations 1 and 3). However, the results are not consistent with mean reversion. Because exposure is decreasing in the 1972 enrollment rate, mean reversion would produce a relative *increase* in enrollment in more-exposed counties starting after 1972, the year used to define our exposure variable, and this increase would persist. We do see a modest increase in enrollment rates in higher-exposure counties in 1972–75, but it is followed by a modest relative fall in 1975–78; mean reversion can explain neither this fall nor the sharp relative jump in enrollment in 1979.

Omitted Variables Bias. County-level exposure could be correlated with changes in confounding factors that correlate with enrollment or infant health. In particular, the fact that more-exposed counties are poorer (see Equation 3) 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.[23] It is unlikely that such confounds would coincide precisely with the free food stamp reform in 1979, but they still warrant consideration. In particular, variation in local business cycles across counties with different poverty rates (and hence different exposure) is a natural potential confound, since food stamp enrollment is highly countercyclical (Appendix Figure A11). In Appendix B.2 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 the slight pre-trends in food stamp enrollment but has a negligible impact on the magnitude of the estimated impact on enrollment and does not affect the null results for birth weight measures.

Exposure Designs. The exposure design assumes treatment effects are homogeneous across counties with different Exposure_c (Sun and Shapiro, 2022). However, poorer counties might be expected to have larger treatment effects of food stamp enrollment on birth weight. In this case, since poorer counties have higher exposure (and therefore higher impacts on enrollment), the homogeneity

[23] Indeed, Appendix Figure A10 suggests that the identifying variation in Exposure_c primarily comes from counties' poverty rates (our proxy for their share of the population eligible for food stamps) rather than take-up rates (i.e., the share of eligible households who enroll): adding controls for indicators for the decile of the county's baseline take-up rate $\times$ calendar-year indicators does not affect the estimated impact on enrollment (Panel a), but adding controls for indicators for the decile of the county's baseline poverty-rate $\times$ calendar-year indicators eliminates the impact of the reform (Panel b). This finding is consistent with the time-series evidence in Figure 5 and Table 1 that the reform drew in poorer enrollees.

assumption would bias *toward* finding an impact on birth weight, thus rendering our null effects conservative. Moreover, in Appendix B.3 we relax the homogeneity assumption and continue to find no evidence of an impact on birth weight or share low birth weight.

Robustness. We probed 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, to limiting our analysis to counties in the third and fourth quartiles of exposure (where most of the variation in exposure occurs), and to controlling for the county-by-county rollout of the WIC program in the mid-1970s (Bitler et al., 2025) and of community health centers (CHCs) in the late 1960s and 1970s (Köse et al., 2026). Appendix B.4 shows that our estimates are robust to these and several other alternative specifications that we discuss.

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

Since on average only 7.6 percent of the population was enrolled in food stamps prior to the free food stamp reform, most births will not be affected by the food stamp program. We should expect small intent-to-treat estimates of the free food stamp reform on birth weight for mechanical reasons.

To interpret magnitudes and compare estimates using different sources of variation, we therefore introduce an IV analysis of the impact of food stamp enrollment on birth weight. We report IV estimates using the free food stamp reform as an instrument for enrollment, and we also report estimates using the earlier county-by-county rollout of the food stamp program as an instrument for enrollment. The IV analyses scale the intent-to-treat estimates to account for the first stage, and the standard errors from the IV analyses reflect the resultant statistical uncertainty that comes from having a comparatively small first stage.

4.1 Estimating Equations

Section 3.1 presented our reduced-form equations for the free food stamp reform. Here we present our reduced-form equations for the rollout (previously analyzed by Almond et al., 2011), followed by the IV framework we use to analyze both reforms separately and together.

4.1.1 Reduced-Form Impact of the Food Stamp Rollout on Birth Weight

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

Because of the staggered implementation and the potential for heterogeneous treatment effects (Callaway and Sant'Anna, 2021; Sun and Abraham, 2021), we implement a stacked event-study analysis by constructing and "stacking" (i.e., vertically appending) datasets d where we compare

treatment cohorts to "clean controls." This approach follows Cengiz et al. (2019) and has recent econometric justification in Wing et al. (2024).

Each dataset d consists of the treatment counties that have the same rollout date and the control counties that did not roll out on that date. For treatment counties, we include up to 3 years before and after the rollout (i.e., relative years –3 to +2); for control counties, we include any of the corresponding 6 years that occurred before their county's rollout date. As all counties are eventually treated, the control group within each dataset consists of all counties in the sample, limited to the periods before they were treated. We then stack the datasets and estimate:

$$\text{FS}_{ctd} = \sum_{\tau \neq -1} \delta_\tau (\mathbf{1}\{r = \tau\}) + \alpha^1_{cd} + \gamma^1_{td} + \epsilon^1_{ctd} \tag{7}$$

$$y_{cmd} = \sum_{\tau \neq -1} \beta_\tau (\mathbf{1}\{r = \tau\}) + \alpha^2_{cd} + \gamma^2_{md} + \epsilon^2_{cmd} \tag{8}$$

where FS_{ctd} is the food stamp enrollment rate, y_{cmd} is a birth weight outcome, α^j_{cd} are county-by-dataset fixed effects and γ^1_{td} (γ^2_{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. We weight the analyses using 1970 population-weighted versions of the weights in Wing et al. (2024), so that the treatment-effect coefficients are interpretable as population-weighted average treatment effects on the treated across treatment cohorts; Appendix B.5 provides additional detail.

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. Given that all counties are eventually treated (i.e., eventually adopt the food stamp program), 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 lets 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 IV Analyses of the Impact of Food Stamp Enrollment on Birth Weight

We then analyze both reforms in the same IV framework. The second-stage estimating equation is:

$$y_{ctd} = \lambda \text{FS}_{ctd} + \alpha^2_{cd} + \gamma^2_{td} + e^2_{ctd} \tag{9}$$

where y_{ctd} is a birth weight outcome in county c, time period t, and dataset d, FS_{ctd} denotes the food stamp enrollment rate, α^2_{cd} are county-by-dataset fixed effects and γ^2_{td} are calendar time-by-dataset fixed effects. We instrument for FS_{ctd} with the first-stage equation:

$$\text{FS}_{ctd} = \delta Z_{ctd} + \alpha^1_{cd} + \gamma^1_{td} + e^1_{ctd} \tag{10}$$

where Z_{ctd} is the excluded instrument.[24]

[24]With a slight abuse of notation we denote calendar time with t for year or month depending on the reform (free stamps is always in years) and outcome (only infant-health outcomes for the rollout have monthly variation, while the first stage is always at the calendar-year level).

For the rollout, Z_{ctd} is an indicator that is 1 if relative time $r \geq 0$, where r depends on c, t and d; as with the preceding reduced-form analysis, we analyze up to 3 years before and after the rollout, clustering our analyses at the county level and weighting the analyses using 1970 population-weighted versions of the weights in Wing et al. (2024). For the free-stamps reform, Z_{ctd} is $\text{Post}_{td} \times \text{Exposure}_{cd}$ where Post is an indicator for years 1979–1981 and the dataset d does not vary. We analyze the three years before and after the reform (1976–1981). We cluster our analyses at the county level and weight our analyses by the 1970 population.

Finally, we use both instruments together by stacking the free-stamps and rollout datasets, normalizing the weights within each reform to sum to 1. This exercise imposes a homogeneous treatment effects assumption on λ across the instruments. This implicitly precision-weights estimates across the two reforms. We cluster standard errors by county also in the pooled analysis, so standard errors account for county-level autocorrelation across the reforms.

Exclusion Restriction. The IV analyses require the exclusion restriction that the reforms only affect birth outcomes via their impacts on food stamp enrollment. For the rollout, the IV does not require additional assumptions beyond what is needed for the reduced-form analysis of the rollout's impact on birth weight. For the free food stamp reform, however, the exclusion restriction does require additional assumptions beyond what was needed for the reduced form. Specifically, exclusion requires that the elimination of the purchase requirement has no impact on the always-takers, and that any impact on compliers operates only through the impact on enrollment. Appendix A.2 presents evidence on food stamp spending and benefits from before the free-stamps reform that is consistent with the exclusion restriction. Most enrolled individuals (always-takers) were already purchasing at least $R + F$ in food (i.e., person z in Figure 1). The reform set the purchase requirement R to 0 without changing the government benefit F. Therefore, the elimination of the purchase requirement had no effect on their budget constraint and should therefore not affect infant health among always-takers.

4.2 Results

4.2.1 Reduced-Form Impacts of the Rollout

Overall Enrollment and Birth Weight. The rollout led to a 5.5 percentage point (standard error = 0.2) increase in the enrollment rate, but no impacts on birth weight (Figure 7). The point estimates indicate a statistically insignificant impact of the rollout on log birth weight of –0.04% (standard error = 0.1) and a statistically insignificant impact on share low birth weight of –0.1 percentage points (standard error = 0.2). Appendix B.4 shows that these results are robust to a number of alternative specifications.

As with the free food stamp reform, we also examine the impact of the rollout on other measures of infant health. Once again, there is no evidence of an impact on infant mortality rate, but there is suggestive evidence in the baseline specification that the rollout increased gestational age and (counterintuitively) reduced births. While the event studies here are more compelling than those for the free food stamp reform, neither of these results is robust across the alternative specifications we explore. Appendix B.1 describes the analyses and results in more detail.

Birth Weight by Race. There is no impact of the rollout on birth weight for White infants, but for Black infants it is associated with a statistically significant increase in log birth weight and decline in the share low birth weight (Figure 8).[25] The average effect for Black infants is a 0.56% (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, while across these alternative specifications there is consistently no evidence of improvements in White birth weight (Appendix Figure A29).

Comparison to Almond et al. (2011). Almond et al. (2011) 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. In Appendix B.6 we discuss in more detail how our reduced-form estimates of the impact of the rollout on birth weight compare with Almond et al. (2011).

4.2.2 IV Estimates: Impact of Food Stamp Enrollment on Birth Weight

IV Results for Average Birth. For both birth weight measures and using either or both reforms as an instrument for enrollment, there is no evidence of a statistically significant impact of food stamps on birth weight (Table 2, Panel a). IV estimates using the free food stamp reform as an instrument for enrollment are considerably more precise than those using the rollout. For birth weight, the IV estimate using the free food stamp reform as the instrument (Column 3) indicates that enrollment reduces birth weight by 0.4% (standard error = 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%. The IV estimate using the rollout as the instrument (Column 2) indicates that enrollment increases birth weight by 1.0% (standard error = 2.9); the 95% confidence interval rules out a 7% increase in birth weight. For the share low birth weight, the IV estimate using the free-stamps reform (Column 6) is a statistically insignificant decline of 0.4 percentage points (standard error = 1.3); using the rollout instrument (Column 5), the IV point estimate is larger (a 4.5 percentage point reduction in the share low birth weight) but less precise (standard error = 3.9).

We pool both reforms to form a precision-weighted estimate of the impact of food stamp enrollment on birth weight. Column 1 shows that the first-stage estimate for the free food stamp reform (rescaled based on the average exposure in the sample) is about 37 percent that of the rollout (0.0185 ÷ 0.0506). However, the reduced-form estimates from the free-stamps reform are much more precise. Ultimately, the IV estimate from the free food stamp reform is more precise than the estimate from the rollout, and as a consequence, the precision-weighted estimates are more similar to those from the free food stamp 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

[25] As with the free food stamp reform analysis by race, we restrict the set of counties used to the subset of baseline counties that have at least 25 births of that race in every county-year analyzed. Once again, this retains the vast majority of the race-specific populations represented by the baseline analysis samples (Appendix Table A3).

or more (Column 4). Enrollment in food stamps decreases the probability of low birth weight by a statistically insignificant 0.7 pp, although power remains weaker (standard error = 1.3) (Column 7).

Of course, the relevant first stage when examining potential impacts on infant health is not the impact on overall food stamp enrollment, but the impact on food stamp enrollment among pregnant women. If the percentage point increase in enrollment from the reform were the same for pregnant women and the general population, no adjustment would be needed. As with the original work on the rollout, we cannot directly estimate the first stage among pregnant women; the county-year enrollment data are not disaggregated by demographics. However, the national time-series evidence from the PSID indicates that during our time period, the enrollment rate for households with a young (aged < 3) child (which we use as a proxy for the population of women at risk of giving birth) was about 85% of the overall enrollment rate (Appendix Figure A41). Therefore, if the enrollment response was proportional for pregnant women and other demographics, we should rescale our first stage to be 85% of the overall first stage, which would make the resulting IV estimates of the impact of food stamps on birth weight modestly larger in absolute value (Appendix Table A6). For example, using the free food stamp reform as an instrument for enrollment, we would estimate a decline in birth weight of –0.49 percent (rather than –0.41), and the 95 percent confidence interval would allow us to reject an increase in birth weight from enrollment of more than 2.3 percent. However, given that the time-series evidence suggests that, if anything, compliers with the free food stamp reform were more likely than always-takers to be families with young children (Table 1), it is also plausible that the appropriate scaling for the free food stamp instrument is greater than our 0.85 estimate, which would produce slightly smaller IV estimates.

IV Results by Race. IV analysis by race is further complicated by the lack of data on race-specific enrollment rates. To construct race-specific first-stage estimates, we impute county-specific, race-specific enrollment rates based on county-year enrollment data, national PSID data on race-specific take-up rates (i.e., enrollment among those eligible), and 1970 census data on race-specific poverty rates by state (Appendix B.7). The results suggest that the first-stage impact of each reform on enrollment rates is about five times larger for Black households than White households.

Using these race-specific first-stage estimates, we find no evidence of an impact of food stamps on White birth weight, whether we use each reform separately as an enrollment instrument or we pool them (Table 2, Panel c). For Black births, the IV estimates are also null for both birth weight outcomes when we use both reforms as precision-weighted instruments for food stamp enrollment; however, if we use the rollout alone as an instrument for enrollment we find statistically significant improvements (Table 2, Panel b). Specifically, for Black births, there is evidence from IV analysis of the rollout that food stamps produce a statistically significant 9.5 percentage point (standard error 3.9) decline in share low birth weight and a borderline statistically significant 5.1 percent increase in log birth weight (standard error = 2.6); however IV estimates for Black births using the free food stamp reform are insignificant for both outcomes. The pooled IV estimate of the impact of food stamps using both instruments is not statistically significant, for either log birth weight or share low birth weight. For instance, 95% confidence intervals rule out that food stamp enrollment causes more than a 3.6% increase in birth weight among Blacks.

Contextualizing the Findings. To help contextualize our findings, Appendix Table A7 compares our estimates to the existing literature that has studied the impact of WIC, and to a lesser extent food stamps, on birth weight. Appendix A.3 provides more details on the existing papers; it also describes how we convert their estimates to comparable IV estimates and construct precision-weighted estimates by program, outcome, and race group. The evidence indicates that WIC improves infant health but that the magnitudes are moderate; food-stamp-specific evidence of improvements in birth weight is weaker.

Impacts of food stamps and WIC for the average enrolled birth range from 0.3 to 2 percent impacts on birth weight and 0–3 percentage point impacts on share low birth weight. Looking separately by race, impacts for Blacks appear larger than for Whites, though they are naturally less precise.

For the average enrolled birth, the precision-weighted impacts for WIC are that enrollment raises birth weight by 0.97% (standard error = 0.31%) and reduces the share with low birth weight by 1.2 percentage points (standard error = 0.58). The precision-weighted estimates for food stamps are similar, but with standard errors that are twice as large: 1.1% on birth weight (standard error = 0.73%) and a 0.98 percentage point reduction in the share low birth weight (standard error = 0.82). As one way to benchmark the magnitudes, in our time period the Black–White gap in birth weight was about 8%, and the gap in share low birth weight was about 6 percentage points (Table 2). For the average household, the evidence is thus consistent with WIC increasing infant health by the equivalent of about a one-eighth to one-fifth reduction in the Black–White birth weight gap, depending on the outcome. Estimates from the food stamp program are similar in magnitude but statistically insignificant, and sometimes even wrong-signed.

5 Normative Analysis of Subsidies for Means-Tested Transfers

We 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 (see also Section 3.1 of Banerjee et al., 2024). Transfers' redistributive value depends on the private value of the transfer to enrollees and on enrollees' social marginal welfare weights. Whether recipients use these transferred resources in ways that improve their health outcomes is not directly relevant. We therefore sketch a stylized framework that illustrates the key empirical objects for determining the redistributive value of a transfer. We then calibrate these objects using our empirical evidence on the enrollment response to the increase in the food stamp subsidy.

5.1 Conceptual Framework

Setup. A government transfer program provides eligible enrollees a benefit B (a good whose market price is normalized to 1) at a subsidy $0 \leq s \leq 1$. That is, the household pays $(1-s)B$ in cash and receives B from the government; we refer to sB as the mechanical government spending on the transfer.

We define the average enrolled household's "normative net private value" (NNPV) per dollar

of mechanical government spending (sB) as $\text{NNPV}(s) = V(s) + D(s)$.[26] Enrollees' NNPV thus includes both their *perceived* net private value $V(s)$ and any distortions $D(s) \lesseqgtr 0$ that lead perceived private valuations to depart from how much a household "should" be willing to pay based on their normative utility. $V(s)$ corresponds to a willingness to pay in money-metric terms (see Appendix C.1.1 for formalization). Distortions $D(s)$ can reflect externalities or behavioral biases (internalities).[27]

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. We denote the social value of enrolling type θ — i.e., the welfare-weighted NNPV — by T_θ:

$$T_\theta(s) \equiv \lambda_\theta(V_\theta(s) + D_\theta(s)). \tag{11}$$

$T_\theta(s)$ captures the total social value of transferring B per dollar of mechanical government spending sB to type θ; conceptually, it corresponds to the numerator of a welfare-weighted MVPF (Hendren and Sprung-Keyser, 2020).

We deliberately leave the benefit B abstract. A useful benchmark is a frictionless cash transfer, in which a dollar of mechanical government spending generates a dollar of net private value ($V(1) = 1$). For most in-kind transfers, we expect $V(1) < 1$ because of enrollment costs and/or because the in-kind transfer distorts consumption. However, $V(1) > 1$ is also possible, for instance when market failures prevent households from purchasing the good privately at its cost to the government.

Analyzing an Increase in Subsidy. We use this framework to analyze the targeting and welfare properties of an increase in the subsidy from s_0 to $s_1 > s_0$. For simplicity, we assume that there are two types $\theta \in \{a, c\}$, where always-takers take up when $s \geq s_0$, and compliers 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 make a "constant proportional value" assumption that $V_a(s_0) = V_a(s_1)$; that is, for always-takers, the perceived net 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. 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.[28]

We say the increase in subsidy improves targeting if the social value of enrolling a complier is higher than that of enrolling an always-taker: $T_c(s_1) > T_a(s_1)$. Whether increasing the subsidy improves or worsens targeting (i.e., $T_c(s_1) \lesseqgtr T_a(s_1)$) is a priori ambiguous, depending on:

$$\lambda_a(V_a(s_1) + D_a(s_1)) \lesseqgtr \lambda_c(V_c(s_1) + D_c(s_1)). \tag{12}$$

On the one hand, simple economic logic favors using prices as a screen for enrollment: a higher

[26] For simplicity, we suppress the dependence on B because we analyze variation in s holding B constant.

[27] An externality could be a failure of efficient intra-household bargaining, in which parents do not internalize the positive health benefits that a transfer might have for their children. An internality could be if households have behavioral biases that drive a wedge between their valuations based on revealed preference and their normative willingness to pay (Bernheim and Taubinsky, 2018).

[28] This is a natural assumption to make with unit demand (e.g., Hendren, 2020; Finkelstein et al., 2019b). See Appendix C.1.2 for more discussion of the assumption's implications with continuous demand.

subsidy will lead only individuals with lower perceived net private value to enroll ($V_a(s_1) > V_c(s_1)$). Intuitively, compliers reveal themselves as having lower perceived net private value than always-takers since they were not willing to enroll at s_0 (therefore $V_c(s_0) \leq 0$) but always-takers were (therefore $V_a(s_0) \geq 0$).[29] On the other hand, a higher subsidy may also enroll individuals with higher social value, that is, $\lambda_c > \lambda_a$. Transfer programs are intended to reach households with little disposable income (and hence high λ_θ), and such households may have mechanically lower willingness to pay than richer households. That said, the existence of a tradeoff between private and social value is not guaranteed; welfare weights could be higher for always-takers if the poorest people always enroll because they are at subsistence and have high marginal utility of consumption. And there is no clear theoretical reason to suspect that always-takers or compliers have higher distortions ($D_a \lessgtr D_c$).

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 (sB), which are increasing in the subsidy, and fiscal externalities ($\text{FE}_\theta(s)$). 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

$$W(s_1) = \frac{\sigma_a \lambda_a (V_a(s_1) + D_a(s_1)) + \sigma_c \lambda_c (V_c(s_1) + D_c(s_1))}{1 + \overline{\text{FE}}} \tag{13}$$

and

$$W(s_0) = \frac{\lambda_a (V_a(s_0) + D_a(s_0))}{1 + \overline{\text{FE}}} = \frac{\lambda_a (V_a(s_1) + D_a(s_1))}{1 + \overline{\text{FE}}}, \tag{14}$$

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 Equation 14 to equate always-takers' 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:

$$W(s_1) > W(s_0) \iff \lambda_c (V_c(s_1) + D_c(s_1)) > \lambda_a (V_a(s_1) + D_a(s_1)). \tag{15}$$

That is, the sign of the welfare impact of the subsidy increase depends only on its targeting properties. Appendix C.2.1 presents a graphical version of this welfare comparison; it shows how we can think of the core welfare tradeoff in terms of the slope of a social value curve that adjusts the private value of the benefit by households' welfare weights.

Analyzing a Given Subsidy Relative to No Program. We can also ask whether a particular subsidy (s) raises welfare relative to not having the program, i.e., whether $W(s) \lessgtr 1$, where $W(s)$ is defined in equations 13 or 14, and the comparison with 1 comes from assuming the program is financed via a non-distortionary tax on the average person in society. It is possible that a higher subsidy can

[29] 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.

produce less social welfare per dollar of government spending but still raise social welfare relative to the program not existing; that is, $W(s_0) > W(s_1) > 1$. In fact, this is what we will find in our empirical application. Making the program free reduces welfare per dollar ($W(s_0) > W(s_1)$), but even the free program has $W(s_1 = 1) > 1$.

Analyzing Perceived Net Private Value. The same objects needed for the above welfare analyses also allow us to estimate $V(s_0) = V_a(s_0)$ and $V(s_1) = \sigma_c V_c(s_1) + \sigma_a V_a(s_1)$, the average perceived net private value per dollar of transfer for those who enroll, respectively, at the lower and higher subsidy. This parameter is relevant for, among other things, estimating how close the transfer comes to the value of a frictionless cash transfer and, relatedly, how much the program transfers are valued by enrollees.

5.2 Empirical Application: Free Food Stamps

Recall that the food stamp program is characterized by a purchase requirement $R \geq 0$ and a bundle of food stamps $B = R + F$, where F is the government transfer (Figure 1). The free food stamp reform closely mimics the conceptual experiment of increasing s, fixing mechanical government spending $sB \equiv F$. Eliminating the purchase requirement (i.e., setting $R \to 0$) increases the subsidy from $s_0 = \frac{F}{R+F} < 1$ to $s_1 = 1$ without changing F.[30]

To apply the conceptual framework in our empirical setting, we make four main assumptions. The first three simplify the inputs to welfare analysis. First, we assume an exclusion restriction that the free-stamps reform was a "pure" subsidy increase, without any additional policy changes; as we discuss below, this is not strictly true, but it seems to be a reasonable approximation for our purposes. Second, we assume that households had more disposable income than the purchase requirement, else enrollment responses to price variation could not reveal perceived private value. This assumption also appears plausible. As discussed in Section 1, the purchase requirement was capped at 30% of net income; moreover, households could buy small amounts of food stamps over the course of the month, so that enrollment did not require more liquidity than standard grocery purchases would. Notably, we do *not* rule out the presence of liquidity constraints; we incorporate them via the social welfare weights (λ_θ), as is standard in the MVPF literature (Hendren and Sprung-Keyser, 2020). Third, we make a revealed-preference assumption, i.e., that $D_\theta = 0$. Section 5.4.3 revisits these three assumptions.

Under these first three assumptions, the condition for the free food stamp reform (i.e., the increase in the subsidy from s_0 to $s_1 = 1$) to improve targeting (Equation 12) or to increase social welfare per dollar of government spending (Equation 15) collapses to:

$$W(1) > W(s_0) \iff \lambda_c V_c(1) > \lambda_a V_a(1). \tag{16}$$

[30] Notice that the policy therefore changes the total size of the food stamp bundle from $B = R + F$ to $B = F$, while leaving mechanical spending $sB = F$ unchanged. 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., $V_a(s_0, B) = V_a(1, B)$) to assume $V_a(s_0, R + F) = V_a(1, F)$. In other words, we assume that the per-dollar valuation of free stamps is constant even when the size of the food stamp bundle changes (from $R + F$ to F). See discussion in Appendix C.1.2.

Improved targeting on social welfare weights ($\lambda_c > \lambda_a$) is necessary but not sufficient for the move from the purchase requirement to free food stamps to improve targeting and increase social welfare per dollar of government spending. Meanwhile, the condition under which free food stamps ($s_1 = 1$) raises social welfare per dollar of government spending relative to no program (i.e., $W(s_1) > 1$) collapses to:

$$W(1) > 1 \iff \frac{\sigma_c \lambda_c V_c(1) + \sigma_a \lambda_a V_a(1)}{1 + \overline{\text{FE}}} > 1. \tag{17}$$

Our fourth assumption facilitates empirical estimation. We assume that a "Law of Demand" holds:

$$\text{household } \theta \text{ enrolls at subsidy } s_0 \iff V_\theta(1) \geq 1 - s_0. \tag{18}$$

This assumption says that consumer surplus per dollar of a free benefit B exceeds $1 - s_0$ if and only if the household would have bought the good at a per-dollar subsidy rate s_0. More intuitively, the assumption says that always-takers (who take up at $s_0 < 1$) value the program more than compliers who do not. This allows us to provide a demand-curve interpretation of the enrollment response we estimated in Section 3.2 and estimate $V_c(1)$ and $V_a(1)$. The assumption is natural but not guaranteed in settings with nonlinear subsidies, which are common for many in-kind transfers.[31] With nonlinear subsidies, enrollment under one subsidy level may not, without the additional assumption in Equation 18, reveal perceived value under a different subsidy level; Appendix C.1.3 specifies assumptions on primitives that would yield Equation 18.

We use the characteristics of compliers and always-takers to the free food stamp reform to estimate the social marginal welfare weights of each type (λ_a and λ_c). We use the level of s_0 under the purchase requirement and of $s_1 = 1$ under the free food stamp reform to estimate $V_a(1)$ and $V_c(1)$. These objects allow us to evaluate Equation 16. In addition, we use the estimated enrollment response to the free food stamp reform to estimate σ_c and σ_a, and recover $W(s_1)$ (Equation 17) and $V(1) \equiv \sigma_a V_a(1) + \sigma_c V_c(1)$.

Microfoundations. Our conceptual framework requires us to make a constant proportional value assumption in Section 5.1, which implies that for always-takers, their per-dollar value of the transfer sB does not change with s. Our empirical implementation requires the assumption of a Law of Demand, which implies that compliers value the free transfer less than always-takers. Appendices C.1.2–C.1.3 discuss microfoundations for these assumptions.

We can see an example of how these assumptions can be satisfied by referring back to the budget sets in Figure 1. Suppose always-takers are ***strongly inframarginal*** households: they would have consumed more than $R + F$ in the absence of the food stamp program (type z in Figure 1). Suppose compliers, meanwhile, are ***weakly inframarginal*** households: they would have consumed more than F but less than $R + F$ in the absence of the program (type v in Figure 1). In this case, for always-takers, the deadweight loss comes entirely from ordeal costs. For compliers, deadweight loss from (counterfactually) enrolling under the purchase requirement would come from both ordeal

[31] 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 (B), after which enrollees can buy 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.

costs as well as the fact that compliers do not view subsidized stamps as fungible with cash, since compliers are weakly — but not strongly — inframarginal. As a result, under the free food stamp reform, which shrinks the food stamp bundle B from $R + F$ to F, compliers' net value may increase per dollar transferred. If compliers are weakly inframarginal, that would also explain why we see a first-stage enrollment response: compliers choose to enroll only under the smaller food stamp bundle (F rather than $R + F$), which does not distort their food consumption.

The evidence is consistent with these assumptions. Using survey data on food expenditures among food stamp enrollees from both before and after the elimination of the purchase requirement, we separately estimate that most of the always-takers were strongly inframarginal, while none of the compliers were (Appendix A.2.2).

Figure 1 provides further intuition for the Law of Demand assumption. Households further to the left, such as v, have lower food demand and are screened out by the purchase requirement because unlike strongly inframarginal households such as z, purchasing the larger food stamp bundle distorts their consumption. When food stamps become free, this distortion disappears for v. Since both households are inframarginal under free stamps (weakly inframarginal), they value the food benefit dollar-for-dollar in consumption terms, and their net private valuations differ only through ordeal costs (k in Figure 1). The Law of Demand imposes an intuitive no-crossing-type condition: households screened out by the purchase requirement cannot "leapfrog" always-takers in their net private valuations when food stamps become free. This assumption rules out households like v having lower ordeal costs than households like z.

5.3 Estimating Perceived Net Private Value

Subsidy Variation from the Free Food Stamp Reform. Given the above assumptions, we can use the enrollment response to the increase in subsidy from the free food stamp reform to estimate perceived net private value for enrollees $V(1)$. If we only observed enrollment under free food stamps ($s = 1$), the perceived net private valuation for enrollees could be anywhere between 0 (if they are just indifferent between enrolling and not enrolling) and 1 (if they have no enrollment costs from enrolling and value each dollar of food stamps as a dollar of cash): $V(1) \in [0, 1]$.[32] However, if we observe enrollment both under a lower subsidy ($s_0 < 1$) and when free ($s_1 = 1$), we can identify distinct, and tighter, bounds on valuations *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

[32] Relative to other work using revealed preference to obtain the value of social *insurance* (e.g., Landais and Spinnewijn, 2021), being able to determine the upper bound of 1 is a feature of our setting. In the insurance case, willingness to pay has no upper bound. There, extrapolation into inframarginal demand involves parametric assumptions.

the increased subsidy.

In our empirical application we observe enrollment under both free stamps ($s = 1$) and purchase requirement ($s_0 = 0.65$) regimes. We therefore know $V_a \in [0.35, 1]$ and $V_c \in [0, 0.35]$. Moreover, we estimated in Section 3.2 that the higher subsidy increased enrollment by 1.85 percentage points off of a base of 7.6 percent; we therefore have $\sigma_a = \frac{7.6}{1.85+7.6} = 0.80$ and $\sigma_c = 1 - \sigma_a = 0.20$. The bounds for V_θ and our estimates of σ_θ thus allow us to tighten the bounds from $V(1) \in [0, 1]$ to $V(1) \in [0.28, 0.87]$.

Figure 9(a) illustrates the demand curve graphically. 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 ($s_1 = 1$) (also by definition), and $\sigma_a = 1 - \sigma_c = 0.80$ enrolled at a price of 0.35 ($s_0 = 0.65$). The nonparametric bounds on V_a and V_c are therefore given by the shaded blue rectangles. 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. 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, 1972).[33] We analyze the enrollment response to this earlier subsidy increase with the same empirical approach that we used for the free-stamps reform, but now exploiting variation in exposure (Equation 3) to this earlier reform based on the county-level enrollment rate in 1968, prior to the 1970 subsidy increase. Appendix C.3.1 presents a parallel set of empirical analyses. They show that higher-exposure counties were on similar enrollment trends as lower-exposure counties prior to 1970, and experienced a markedly larger increase in enrollment in 1970 relative to lower-exposure counties. Given the average exposure in 1968, the event-study estimates imply that the increase in subsidy from 0.40 to 0.65 increased the average enrollment rate by 2.4 percentage points (standard error = 0.22).

Figure 9(b) plots the demand curve for food stamps with the additional data point from the 1970 subsidy increase. 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 piecewise linear demand; it is now 59 cents ($V(1) = 0.59$) rather than the 58 cents above. However, the additional data point allows us to substantially tighten the nonparametric bounds: 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 ($V_a = 0.696$).

[33] It did so by increasing the amount of the "bonus" F, i.e., the amount paid by the government; an increase in F in turn increases the subsidy rate $\frac{F}{R+F}$. Appendix C.1.4 shows how to extend the empirical assumptions above to add this point to the demand curve. We do not use this reform to study the impact of food stamps on birth weight because the exclusion restriction would not hold: always-takers receive a benefit (that could in principle affect birth weight) in the form of higher government benefits F.

In Appendix C.3.2 we relax the assumption of piecewise linearity of the demand curve, which we used to obtain our baseline estimate of $V(1) = 0.59$. Following Kang and Vasserman (2025), we consider various shape restrictions and the bounds they imply. Under a benchmark curvature constraint and including the bonus reform, we obtain semiparametric bounds of the value of food stamps lying within [0.53, 0.65].

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 from 0.65 to 1 points to the existence of meaningful deadweight loss that dissipates some of the transfer's gross private value.

Why would households behave as if they perceive their value of food stamps to be worth only 59 cents per dollar? It is not because they lack any cash on hand to purchase food stamps; as discussed in Section 1, the purchase requirement was designed to be no greater than 30% of monthly household income, and was in fact zero for the very lowest-income households (see also Section 5.4.3). Rather, the dissipation of value reflects either that the food stamp program delivers more food than the household would choose to consume out of the equivalent cash, and/or that there are enrollment costs to receiving food stamps.

If all households were inframarginal at F and there were no distortions, food stamps would be fungible with cash and we could interpret the average value of a food stamp dollar being less than dollar-for-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 A.2.2). This suggests that much of the deadweight loss comes from ordeal costs that dissipate some of the transfer's gross private value. Of course, this discussion has maintained the assumption of the absence of distortions ($D = 0$), particularly zero behavioral biases. In the context of food stamps, a growing body of work has found evidence consistent with substantial mental accounting (e.g., Hastings and Shapiro, 2018; Song, 2022; Chorniy et al., 2025; Frisvold and Lozano Rojas, 2026; Allcott et al., 2026), and there is reason to believe that other behavioral biases like present focus could enter decisions. Such distortions could contribute to why the *as-if* demand curve is low. But they do not in general offer a clear sign for whether normative willingness to pay is higher or lower than perceived willingness to pay.

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 perceived by recipients as eroded relative to a frictionless cash transfer with $V(1) \equiv 1$. This estimate of deadweight loss exceeds what one gets from monetizing time costs associated with applying for food stamps — about 2 cents per dollar transferred (Finkelstein and Notowidigdo, 2019) — or from envelope-theorem-based estimates of ordeal costs from benefit take-up elasticities (17 cents per dollar in SNAP in Rafkin et al., 2025). It implies that the U.S. engages in substantially less redistribution than is commonly assumed under current accounting practices that treat the program as if it were equivalent to a frictionless cash

transfer valued dollar-for-dollar (e.g., Piketty et al., 2018; Blanchet et al., 2022; Congressional Budget Office, 2025).

5.4 Empirical Welfare Analysis

5.4.1 Welfare Tradeoff: Free Food Stamps vs. No Program and Partial Subsidy

The empirical evidence has shown that cost-sharing improves targeting on private value ($V_a = 0.696 > 0.175 = V_c$) but reduces targeting on social value (since we saw in Table 1 Panel (a) that compliers are substantially poorer than always-takers). To assess whether free stamps produce more welfare per dollar of mechanical government spending than the purchase requirement (i.e., whether $W(1) \lessgtr W(0.65)$ per Equation 16), we must convert the average income estimates for compliers and always-takers (Table 1 Panel a) into social marginal welfare weights λ_a and λ_c. To do so, we assume 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 to 1 for an average household in society from 1973–81. This assumption 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. The resultant average social marginal welfare weights are 6.2 for compliers, compared to 3.2 for always-takers (Table 1 Panel b). In other words, compliers have about twice as high social marginal utility from a dollar as always-takers.

Putting our estimates of V_c, V_a, λ_a, and λ_c together implies that social welfare per dollar of government spending is higher under cost-sharing than free food stamps: $\lambda_c V_c < \lambda_a V_a$. Compliers' social value $\lambda_c V_c = 6.2 \times 0.18 \approx 1.1$, while always-takers' social value $\lambda_a V_a = 3.2 \times 0.70 \approx 2.2$. In other words, even though compliers' social marginal welfare weights exceed always-takers' ($\lambda_c > \lambda_a$), empirically, we find this difference is not large enough to offset our estimates of compliers' lower private values ($V_c < V_a$). Appendix C.2.2 shows this empirical tradeoff graphically.

5.4.2 Social Welfare from Free Food Stamps Relative to No Program

We can also use the above estimates of V_c, V_a, λ_c, and λ_a together with our estimate of the complier share ($\sigma_c = 0.2$) to ask whether the free food stamp program increases social welfare relative to the program not existing. Given our estimates, Equation 13 indicates $W(1) = \frac{2.01}{1+\overline{\text{FE}}}$.

It is instructive to compare our finding of $W(1) = \frac{2.01}{1+\overline{\text{FE}}}$ to the redistributive value of providing cash frictionlessly to the same recipients as the free-stamps program (i.e., without ordeals, so that $V_a = V_c = 1$). This program would produce $W^{\text{cash}} = \frac{\lambda_a \sigma_a \cdot 1 + \lambda_c \sigma_c \cdot 1}{1+\overline{\text{FE}}} = \frac{\$3.8}{1+\overline{\text{FE}}}$ of redistributive value per mechanical dollar spent. Thus, relative to a frictionless cash benchmark, free food stamps waste about \$1.80 of social value per dollar spent ($= 3.80 - 2.01$). Of course, this counterfactual program may not be feasible, as ordeals may also be necessary to restrict enrollment to the same population that takes up under the free-stamps regime.

As another benchmark, we can ask whether the free-stamps program financed by a non-distortionary tax on the average person in society still improves social welfare (i.e., is $W(1) > 1$?). Despite the dissipation of about two-fifths of food stamps' perceived private value relative to a frictionless cash transfer ($V(1) = 0.59$), the answer is yes, unless the fiscal externality from the free

stamps program is greater than or equal to the mechanical cost. In practice, the net fiscal externalities from the program may in fact be negative due to improvement in lifetime earnings from individuals who receive food stamps as children (Hendren and Sprung-Keyser, 2020; Bailey et al., 2024), thus serving to further increase $W(1)$.[34] The key reason $W(1) > 1$ despite $V(1)$ that is much below 1 is simply that the welfare weights on food stamp enrollees are very high under standard social welfare functions. This captures the logic that food stamps is a redistributive program, but, of course, substantially less redistributive social preferences could imply $W(1) < 1$.

5.4.3 Assumptions and Limitations

Our normative analysis involves 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. In Appendix C.5, we discuss these assumptions and explore sensitivity to them in detail. Our baseline estimates imply that welfare-weighted valuations are twice as high for always-takers as for compliers ($\lambda_a V_a = 2.2$ versus $\lambda_c V_c = 1.1$), leading to our conclusion that $W(1) < W(0.65)$. We therefore focus on potentially unmodeled or unaccounted-for factors that could either double compliers' welfare-weighted valuations relative to always-takers', or halve our baseline estimate of $W(1) \approx 2$. Overall, our main conclusion that $W(0.65) > W(1) > 1$ appears relatively robust to relaxing these assumptions. We summarize some of the main considerations here.

Revealed Preference. We have assumed no distortions ($D = 0$), so that perceived net private value V reflects normative net private value NNPV. In practice, behavioral biases such as misperceptions or present focus may create nonzero distortions D. If welfare-weighted distortions are equal ($\lambda_a D_a = \lambda_c D_c$), they have no impact on the comparison of social welfare per dollar of government spending under free food stamps compared to the purchase requirement (see Equation 15); this would be the case, for example, if society values all children equally and food stamp enrollment benefits children in ways that households do not fully internalize. Our calibration exercises suggest that for nonzero distortions to alter our conclusions, they must be either implausibly large for all types, or compliers must be much more biased than always-takers; for instance, to reverse our results if no always-takers are present-focused, we would need that at least 62% of compliers are present-focused. Moreover, under these alternative calibrations $W(1)$ remains substantially above 1.

Liquidity Constraints. Since compliers are substantially poorer than always-takers, it is possible that some may not have enrolled at the positive purchase price ($s < 1$) due to liquidity constraints which create a high marginal utility of consumption out of current resources. This is not a *conceptual* problem for our analysis: our estimate of $V(1)$ correctly reflects their private value of food stamps,

[34] The likely source of positive fiscal externalities comes from individuals reducing their labor supply (and hence income tax revenue earned) to qualify for food stamps. In Appendix C.4, 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. Even if we had found larger impacts on labor supply, positive fiscal externalities from labor supply are likely low due to low tax-and-transfer rates on low-income households (e.g., Hendren and Sprung-Keyser 2020 use 12.9%). 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 estimated by Hoynes and Schanzenbach (2012) implies a positive fiscal externality of just 16 cents (standard error 70 cents) in tax revenue per dollar of food stamps transferred.

given their budget and borrowing constraints, and these constraints are reflected in higher social marginal welfare weights λ_θ.[35] Indeed, a core tradeoff captured in our framework is that using prices in the context of transfer programs may backfire, because low-income households may have mechanically lower willingness to pay due to tighter budget constraints.[36]

However, liquidity constraints may have *empirical* implications for our estimates of λ_θ, since they could create distributional concerns arising from differences in resources across enrollees. That is, there is scope for reasonable disagreement about whether standard welfare weights capture households' marginal utility of consumption when near subsistence. In practice, we find that even using extreme parameterizations of social welfare functions with CRRA utility (e.g., risk aversion parameter of 50), we cannot reverse our result that the increase in subsidy reduced targeting; at the other end of the spectrum, even with a risk aversion parameter of 1 (i.e., log utility), we still find $W(1) > 1$. Of course, because compliers are poorer than always-takers, if the social welfare function puts sufficiently high weight on only the very poorest members of society, we can get that the increase in subsidy improves targeting.

Another potential concern is that if liquidity constraints mean that always-takers had to incur borrowing costs to purchase stamps, making food stamps free could involve an (unaccounted-for) transfer to always-takers; this would suggest unaccounted-for gains from increasing the subsidy. In practice, this appears unlikely. Recall that most always-takers are strongly inframarginal. Therefore, food stamps do not distort their consumption and cause them to borrow more than otherwise; this is particularly the case since households could buy incremental amounts of stamps over the course of the month.

Finally, our approach to using the enrollment response to the subsidy change to estimate perceived net private value V does require that households had sufficient cash or liquid assets to pay the purchase requirement. Several pieces of evidence support this assumption: the purchase requirement was capped at 30% of monthly household income; most households bought more food than their food stamp benefits; households could buy increments of stamps (e.g., could buy 25% of stamps at once), so did not need savings beyond weekly grocery purchases; and the poorest households paid nothing even under the purchase requirement. One way of thinking about our results is that eligible non-enrolled households could have enrolled and purchased food stamps in small increments on the way to the grocery store; if they did not do so, but continued to buy groceries, that suggests low private valuations due to ordeals more than liquidity constraints. In addition to these institutional features, the evidence from the PSID suggests that most enrollees, although low-income, were not so low-income that they had zero cash on hand.[37]

[35] Our approach thus differs from envelope-theorem-based welfare analyses, which use causal effects on downstream outcomes (e.g., earnings) to infer willingness to pay V (Hendren and Sprung-Keyser, 2020). In those approaches, liquidity constraints can affect whether behavioral responses are enveloped out or represent first-order welfare gains. By contrast, we use price variation to recover willingness to pay directly, without needing to monetize downstream behavioral responses into first-order welfare gains.

[36] Of course, liquidity constraints can create a wedge between low-income households' willingness to accept cash for food stamps and their willingness to pay, a kind of income effect when the transfer F is large relative to income. Any demand curve estimated among low-income populations (for any good or means-tested program) is subject to this limitation (e.g., Finkelstein et al., 2019a), and this limitation is important to keep in mind when considering our estimates of average private value of food stamps $V(1)$ relative to providing cash frictionlessly.

[37] See Appendix Table A4. Only 1–2% of enrolled households had zero income net of housing costs while a larger share

Potential Policy Confounds. Our welfare analysis assumes that the policy can be modeled as a change in a uniform subsidy (s), with no other policy changes. Consistent with this, the 1979 reform did not change the government bonus (F). It also did not directly affect transaction costs of procuring food stamps (k); after the reform, most recipients still had to go to the local food stamp office to pick up food stamps.

The 1979 reform did, however, slightly reduce eligibility, particularly for higher-income households. If anything, this biases us against finding an enrollment response to the free-stamps reform and toward finding higher welfare gains from the free food stamp reform, since some of the greater poverty of compliers might stem from eligibility changes.

In addition, during the 1970s and 1980s, local jurisdictions gradually expanded the use of mail applications and issuance of food stamps. The elimination of the purchase requirement arguably made this switch easier; however, mail issuance was low during our analysis period, and there is no evidence that a shift to mail occurred in 1979, when we observe the sharp increase in enrollment. Moreover, where we can observe a shift to mail issuance in one state in 1981, we see little evidence of an enrollment response (Appendix C.5.2).

Finally, as discussed in Section 1, under the purchase requirement the subsidy was not in fact uniform; it decreased with income. Given that compliers faced a smaller subsidy change, when we incorporate person-specific subsidies, it only lowers V_c and thus pushes toward our conclusion that $V_a > V_c$.

6 Conclusion

As one of the largest means-tested transfer programs in the United States, the $100-billion-a-year food stamp program has been described both as a redistributive program and as an investment in child human capital (Hoynes and Schanzenbach, 2018). We leverage the major 1979 redesign of the food stamp program to shed new empirical light on both of these functions. The reform transformed the food stamp program from its original, partial subsidy design to the current, full subsidy (i.e., "free food stamps"). We estimate that the increase in subsidy raised enrollment by about 25%, and that marginal enrollees were substantially poorer than those already enrolled, with about 60% of their household equivalized income.

We do not find evidence that food stamps improve infant health when we use the free-stamps reform as an instrument for food stamp enrollment. 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%.

However, the redistributive value of the food stamp program does not require it to have causal impacts on infant health or other outcomes. Instead, its redistributive value depends on how poor the

(5% of households) had $0 purchase requirement. Just 2% of always-takers and 11% of compliers had post-tax income net of housing costs that was less than 10% of the Federal Poverty Level (FPL). One reason many households have positive income is that 30% were on AFDC, and thus had positive welfare income (nearly $3,000 on average in $1979). Mean income net of housing costs among compliers was about $5,200 while the average annualized 1978 purchase requirement was around $800–1,100 (authors' calculation based on PSID and benefit tables).

recipients are and how much private value the program delivers to them per dollar of government spending (Lieber and Lockwood, 2019). From this perspective, we estimate that although the food stamp program dissipated about 41 cents per dollar of government spending relative to a frictionless cash transfer, it is sufficiently well-targeted that it increases social welfare relative to having no program. We also document an empirical tradeoff in setting the level of the program's subsidy between targeting on private and on social value.

Beyond our specific findings, 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 and important question is whether enrollment caused by the move to free food stamps improved the long-run health and earnings of those who newly received food stamps in childhood — as has been found by studies examining the impact of the rollout of the initial, partially subsidized food stamp program (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 dollar-for-dollar that would be the case with a frictionless cash transfer. This dollar-per-dollar value is what is 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). Our findings suggest that the U.S. safety net may involve much more deadweight loss — and thus perform less redistribution — than is assumed under current, standard accounting practices. 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.

Agarwal, Sumit, H Luke Shaefer, Yasamean Zamani-Hank, Eric Finegood, Jenny LaChance, and Mona Hanna, "The Effects of the Rx Kids Unconditional Cash Prescription Programme During Pregnancy and Infancy on Birth Outcomes in the USA: A Population-Based, Quasi-Experimental Study," *Lancet Public Health*, 2026, *11* (6), e355–e363.

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

Allcott, Hunt, Amy Finkelstein, Anna Grummon, and Matthew J Notowidigdo, "The Effects of SNAP Sugary Drink Restrictions on Consumption and Welfare," Working Paper 35659, National Bureau of Economic Research 2026.

Allen, Joyce E., "Food Programs Grow Slowly from 1979 to 1983," *National Food Review*, 1983, *26*, 21–24. Food Situation and Review section.

Almond, Douglas, Hilary W Hoynes, and Diane Whitmore Schanzenbach, "Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes," *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.

Amarante, Verónica, Marco Manacorda, Edward Miguel, and Andrea Vigorito, "Do Cash Transfers Improve Birth Outcomes? Evidence from Matched Vital Statistics, and Program and Social Security Data," *American Economic Journal: Economic Policy*, 2016, *8* (2), 1–43.

Anders, Jenna and Charlie Rafkin, "The Welfare Effects of Eligibility Expansions: Theory and Evidence from SNAP," *American Economic Journal: Economic Policy*, 2026, *18* (3), 61–100.

Anderson, Patricia M and Bruce D Meyer, "Unemployment Insurance Takeup Rates and the After-Tax Value of Benefits," *Quarterly Journal of Economics*, 1997, *112* (3), 913–937.

Associated Press, "Newsom Wants to Pause Health Care Expansion to Low-Income Immigrants," AP News May 2025. Accessed January 31, 2026.

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.

Autor, David, Andreas Kostøl, Magne Mogstad, and Bradley Setzler, "Disability Benefits, Consumption Insurance, and Household Labor Supply," *American Economic Review*, 2019, *109* (7), 2613–2654.

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.

Barber, Sarah L and Paul J Gertler, "The Impact of Mexico's Conditional Cash Transfer Programme, Oportunidades, on Birthweight," *Tropical Medicine & International Health*, 2008, *13* (11), 1405–1414.

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," in B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds., *Handbook of Behavioral Economics: Applications and Foundations 1*, Vol. 1, Amsterdam: North-Holland, 2018, pp. 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," Working Paper 33182, National Bureau of Economic Research 2024.

_ , Danea Horn, Esra Kose, Maria Rosales-Rueda, and Arian Seifoddini, "Effects of WIC on Birth Outcomes: Evidence from Newly Digitized Data from the National Archives," Working Paper 34521, National Bureau of Economic Research November 2025.

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," *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," Technical Report, KFF, San Francisco, CA 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," Technical Report, Kaiser Family Foundation, San Francisco, CA 2019.

Brown, Gregory M, "End of Purchase Requirement Fails to Change Food Stamp Participation," *Monthly Labor Review*, 1988, *111* (7), 14–18.

California Department of Benefit Payments, "Food Stamp Program - Annual Food Stamp Questionnaire," February 1976.

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. SSRN Working Paper 11086.

Carlson, Steven and Joseph Llobrera, "SNAP Is Linked with Improved Health Outcomes and Lower Health Care Costs," Report, 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," Working Paper, 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," *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.

Cholli, Neil A and Derek Wu, "Multiple Program Participation in the Safety Net: Incidence, Impediments, and Implications," Working Paper 18636, Luxembourg Institute of Socio-Economic Research 2026.

Chorniy, Anna, Amy Finkelstein, and Matthew Notowidigdo, "Paternalistic Social Assistance: Evidence and Implications from Cash vs. In-Kind Transfers," 2025. Working Paper.

_ , Janet Currie, and Lyudmyla Sonchak, "Does Prenatal WIC Participation Improve Child Outcomes?," *American Journal of Health Economics*, 2020, *6* (2), 169–198.

Citizens' Board of Inquiry, *Hunger, U.S.A*, Boston: Beacon Press, 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, *43* (1), 864–878.

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, Congressional Budget Office, Washington, D.C. 1977. U.S. Government Printing Office.

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

Congressional Quarterly, "Food Stamp Legislation," in "CQ Almanac 1981," 37th ed., Washington, DC: Congressional Quarterly, 1982, pp. 466–471.

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 March 24, 2026.

Cook, Jason and Katherine Meckel, "Family Matters: Sibling Spillovers in Program Participation," 2026. Working Paper.

Currie, Janet and Nancy Cole, "Welfare and Child Health: The Link Between AFDC Participation and Birth Weight," *American Economic Review*, 1993, *83* (4), 971–985.

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.

Finkelstein, Amy and Matthew J Notowidigdo, "Take-Up and Targeting: Experimental Evidence from SNAP," *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*, 2019a, *127* (6), 2836–2874.

_ , _ , and Mark Shepard, "Subsidizing Health Insurance for Low-Income Adults: Evidence from Massachusetts," *American Economic Review*, 2019b, *109* (4), 1530–1567.

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.

Frisvold, David and Felipe Lozano Rojas, "Short-Run Impacts of SNAP Restriction Waivers on Beverage Purchases," Working Paper 35613, National Bureau of Economic Research 2026.

Ganong, Peter and Jeffrey B Liebman, "The Decline, Rebound, and Further Rise in SNAP Enrollment: Disentangling Business Cycle Fluctuations and Policy Changes," *American Economic Journal: Economic Policy*, 2018, *10* (4), 153–176.

_ , Avik Garg, and Maximilian Kasy, "Precision-Weighted Estimates and τ," 2026. Manuscript.

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.

González, Libertad and Sofia Trommlerová, "Cash Transfers Before Pregnancy and Infant Health," *Journal of Health Economics*, 2022, *83*, 102622.

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

_ , _ , and Josef Zweimüller, "Designing Disability Insurance Reforms: Tightening Eligibility Rules or Reducing Benefits?," *Econometrica*, 2024, *92* (1), 79–110.

Hastings, Justine and Jesse M Shapiro, "How Are SNAP Benefits Spent? Evidence from a Retail Panel," *American Economic Review*, 2018, *108* (12), 3493–3540.

Hawkins, Amelia, Christopher Hollrah, Sarah Miller, Laura R Wherry, Gloria Aldana, and Mitchell Wong, "The Long-Term Effects of Income for At-Risk Infants: Evidence from Supplemental Security Income," *American Economic Review*, 2025, *115* (9), 3081–3129.

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," *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, Dale M., "Current Issues in the Food-Aid Programs," *Congressional Record*, 1972.

Hoynes, Hilary and Diane Whitmore Schanzenbach, "US Food and Nutrition Programs," in Robert A. Moffitt, ed., *Economics of Means-Tested Transfer Programs in the United States, Volume 1*, Chicago: University of Chicago Press, 2016, pp. 219–301.

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

_ , **Marianne Page, and Ann Huff Stevens**, "Can Targeted Transfers Improve Birth Outcomes?: Evidence from the Introduction of the WIC Program," *Journal of Public Economics*, 2011, *95* (7-8), 813–827.

Hoynes, Hilary W. and Diane Schanzenbach, "Safety Net Investments in Children," *Brookings Papers on Economic Activity*, 2018, *2018* (Spring), 89–132.

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.

_ , **Diane Whitmore Schanzenbach, and Douglas Almond**, "Long-Run Impacts of Childhood Access to the Safety Net," *American Economic Review*, 2016, *106* (4), 903–934.

_ , **Leslie McGranahan, and Diane W Schanzenbach**, "SNAP and Food Consumption," in Judith Bartfeld, Craig Gundersen, Timothy M. Smeeding, and James P. Ziliak, eds., *SNAP Matters: How Food Stamps Affect Health and Well-Being*, Redwood City, CA: Stanford University Press, 2015, pp. 107–133.

Hoynes, Hilary Williamson and Diane Whitmore Schanzenbach, "Work Incentives and the Food Stamp Program," *Journal of Public Economics*, 2012, *96* (1-2), 151–162.

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," *Lancet Child & Adolescent Health*, 2021, *5* (5), 367–384.

Köse, Esra, Siobhan M. O'Keefe, and Maria Rosales-Rueda, "Does the Delivery of Primary Health Care Improve Birth Outcomes? Evidence from the Rollout of Community Health Centers," *Journal of Human Resources*, 2026, *61* (2), 1669–1705.

Landais, Camille and Johannes Spinnewijn, "The Value of Unemployment Insurance," *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.

Londoño-Vélez, Juliana, Luis Esteban Alvarez, Catherine Rodriguez, and Fabio Sánchez, "Equity and Efficiency in Financial Aid Targeting," 2025. Working Paper.

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," 2025. SSRN Working Paper 5692222.

Melvin, Shamar and Travis A Smith, "From Paper to Plastic: How the Transition to EBT Affected SNAP Enrollment," *Economics Letters*, 2022, *220*, 110852.

Moffitt, Robert, "An Economic Model of Welfare Stigma," *American Economic Review*, 1983, *73* (5), 1023–1035.

_ , "Estimating the Value of an In-Kind Transfer: The Case of Food Stamps," *Econometrica*, 1989, *57* (2), 385–409.

Montpetit, Sébastien, Luisa Carrer, and Pierre-Loup Beauregard, "A Welfare Analysis of Universal Childcare: Lessons from a Canadian Reform," 2025. Working Paper.

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," Research Bulletin 181447, Department of Applied Economics and Management, Cornell University 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," 2025. Working Paper.

National Academies of Sciences, Engineering, and Medicine, *A Roadmap to Reducing Child Poverty*, Washington, DC: The National Academies Press, 2019. Edited by Greg Duncan and Suzanne Le Menestrel.

National Cancer Institute, "Surveillance, Epidemiology, and End Results (SEER) Program Populations," https://seer.cancer.gov/popdata/download.html 2026. Surveillance Research Program, Division of Cancer Control and Population Sciences.

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

_ , "Second National Health and Nutrition Examination Survey (NHANES II)," https://wwwn.cdc.gov/nchs/nhanes/nhanes2/default.aspx 1976–1980. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

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

of Michigan Survey Research Center Institute for Social Research, University, "Panel Study of Income Dynamics, Public Use Dataset," Ann Arbor, MI 2026. Accessed June 29, 2026.

Office of the Federal Register, "Code of Federal Regulations, Title 7, Section 271.3," 1978. Revised as of January 1.

Ohio General Assembly, *Laws of Ohio: Legislative Acts Passed and Joint Resolutions Adopted by the 113th General Assembly, Regular Session, 1979–1980*, Vol. 138, Columbus, OH: State of Ohio, 1980. Accessed via HeinOnline Session Laws Library.

Ohls, James C and Harold Beebout, *The Food Stamp Program: Design Tradeoffs, Policy, and Impacts*, The Urban Institute, 1993.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman, "Distributional National Accounts: Methods and Estimates for the United States," *Quarterly Journal of Economics*, 2018, *133* (2), 553–609.

Price, David J and Jae Song, "The Long-Term Effects of Cash Assistance," *American Economic Journal: Applied Economics*, 2026, *18* (4), 154–199.

Pub. L. 91-671, "Food Stamp Act Amendments of 1970," 84 Stat. 2048 (Jan. 11, 1971) January 1971.

Pub. L. 92-603, "Social Security Amendments of 1972," 86 Stat. 1329 (Oct. 30, 1972) October 1972.

Pub. L. 93-86, "Agriculture and Consumer Protection Act of 1973," 87 Stat. 221 (Aug. 10, 1973) August 1973.

Pub. L. 94-585, "An Act to Amend the Social Security Act with Respect to Food Stamp Purchases by Welfare Recipients," 90 Stat. 2901 (Oct. 21, 1976) October 1976.

Rafkin, Charlie, Adam Solomon, and Evan Soltas, "Self-Targeting in US Transfer Programs," 2025. SSRN Working Paper 4495537.

Reader, Mary, "The Infant Health Effects of Starting Universal Child Benefits in Pregnancy: Evidence from England and Wales," *Journal of Health Economics*, 2023, *89*, 102751.

Reese, Robert B, J Gerald Feaster, and Garey B Perkins, "Bonus Food Stamps and Cash Income Supplements," Marketing Research Report, Economic Research Service, U.S. Department of Agriculture 1974.

Richardson, Joe, "A Concise History of the Food Stamp Program," CRS Report 79-244 EPW, Congressional Research Service Library of Congress, Washington, D.C. 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.

Rossin-Slater, Maya, "WIC in Your Neighborhood: New Evidence on the Impacts of Geographic Access to Clinics," *Journal of Public Economics*, 2013, *102*, 51–69.

Ruffini, Krista, "Does Unconditional Cash During Pregnancy Affect Infant Health?," Working Paper 72, Federal Reserve Bank of Minneapolis 2023.

Samuelson, Paul A, "The Pure Theory of Public Expenditure," *Review of Economics and Statistics*, 1954, *36* (4), 387–389.

Schanzenbach, Diane, "Experimental Estimates of the Barriers to Food Stamp Enrollment," Discussion Paper 1367-09, Institute for Research on Poverty, University of Wisconsin-Madison 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.

Sonchak, Lyudmyla, "The Impact of WIC on Birth Outcomes: New Evidence from South Carolina," *Maternal and Child Health Journal*, 2016, *20* (7), 1518–1525.

Song, Jintaek, "Propensity to Consume Food Out of SNAP and Its Welfare Implications," 2022. SSRN Working Paper 4293382.

Spinnewijn, Johannes, "Heterogeneity, Demand for Insurance, and Adverse Selection," *American Economic Journal: Economic Policy*, 2017, *9* (1), 308–343.

Strully, Kate W, David H Rehkopf, and Ziming Xuan, "Effects of Prenatal Poverty on Infant Health: State Earned Income Tax Credits and Birth Weight," *American Sociological Review*, 2010, *75* (4), 534–562.

Sun, Liyang and Jesse M Shapiro, "A Linear Panel Model with Heterogeneous Coefficients and Variation in Exposure," *Journal of Economic Perspectives*, 2022, *36* (4), 193–204.

__ **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," Technical Report, Mathematica Policy Research, Washington, DC 2007.

U.S. Bureau of Labor Statistics, "Employment-Population Ratio [EMRATIO]," Retrieved from FRED, Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org/series/EMRATIO 2026. Accessed June 10, 2026.

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

U.S. Department of Agriculture, *Food Stamp Statistical Summary of Project Area Operations Report*, Washington, DC: Food and Nutrition Service, U.S. Department of Agriculture, 1980–1981.

__ , "Report to Congress on the Food Stamp Act of 1977," Technical Report, U.S. Department of Agriculture January 1981.

U.S. Department of Agriculture, Economic Research Service, "Supplemental Nutrition Assistance Program (SNAP)–Key Statistics and Research," 2025. Accessed May 18,2026.

__ , "WIC Program," July 2026. Accessed October 6, 2026.

U.S. Department of Agriculture, Food and Nutrition Service, "Food Stamp Program: Allotments," *Federal Register*, 1971, *36* (145), 14118. July 29.

__ , "Food Stamp Program: Maximum Monthly Allowable Income Standards and Basis of Coupon Issuance, 48 States and District of Columbia," *Federal Register*, 1971, *36* (74), 7273. April 16.

__ , "Food Stamp Program: Allotments," *Federal Register*, 1972, *37* (97), 10012. May 18.

__ , "Food Stamp Program: Allotments," *Federal Register*, 1973, *38* (61), 8287. March 30.

__ , "Food Stamp Program: Allotments," *Federal Register*, 1974, *38* (210), 30119. November 1.

__ , *Food Stamp Facts: Allotments and Purchase Requirements*, Washington, DC: U.S. Government Printing Office, 1975.

__ , "Food Stamp Program: Allotments," *Federal Register*, 1976, *41* (128), 27366. July 2.

__ , "Food Stamp Program: Maximum Allowable Income Standards and Basis of Coupon Issuance, 48 States and District of Columbia," *Federal Register*, 1977, *42* (215), 58152–58153. November 8.

__ , "Food Stamp Program: Income Eligibility Standards and Coupon Allotments, 48 States and District of Columbia," *Federal Register*, 1978, *43* (225), 54199–54202. November 21.

__ , "Implementing the Food Stamp Act of 1977," *Federal Register*, 1978, *43* (201), 47846–47931. October 17.

__ , "Participation of State Agencies and Eligible Households," *Federal Register*, 1978, *43* (95), 21304. May 16; to be codified at 7 C.F.R. pt. 271.

U.S. Department of Commerce Bureau of Economic Analysis, "Regional Economic Information System (REIS)," https://apps.bea.gov/regional/downloadzip.html 1961–1981. Accessed May 22, 2025.

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.

U.S. General Accounting Office, "Millions Could Be Saved by Improving Integrity of the Food Stamp Program's Authorization-to-Participate System," Technical Report CED-82-34, U.S. General Accounting Office January 1982.

U.S. House of Representatives, Committee on Appropriations, *Department of Agriculture Appropriations for 1971: Hearings Before a Subcommittee of the Committee on Appropriations, House of Representatives, Ninety-First Congress, Second Session*, Washington, DC: U.S. Government Printing Office, 1970. Part 2.

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

U.S. Senate, Subcommittee on Nutrition, "Hunger in America: Ten Years Later," Hearing 96th Congress, 1st Session, Committee on Agriculture, Nutrition, and Forestry, Washington, DC April 1979.

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?," Working Paper 468, Princeton Industrial Relations Section 2002.

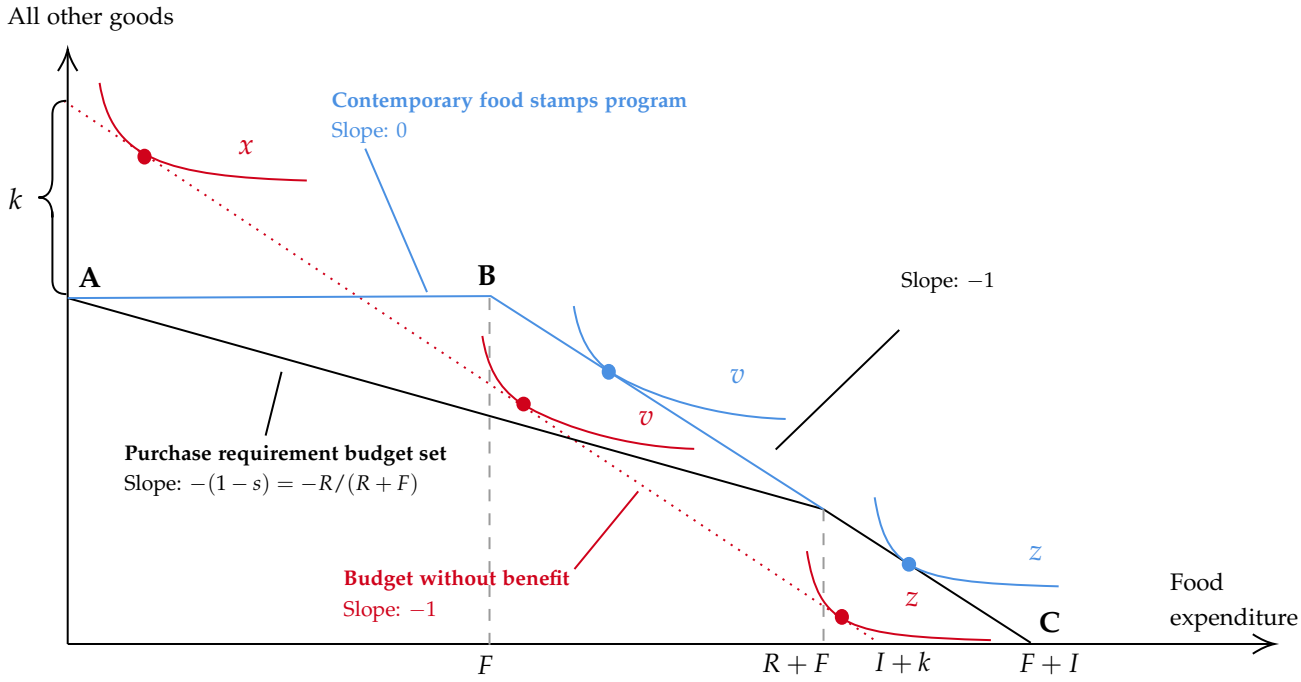
Wing, Coady, Seth M Freedman, and Alex Hollingsworth, "Stacked Difference-in-Differences," Working Paper 32054, National Bureau of Economic Research 2024.

Wu, Derek and Bruce D. Meyer, "Certification and Recertification in Welfare Programs: What Happens When Automation Goes Wrong?," *American Economic Journal: Economic Policy*, Forthcoming.

Yang, Frank, Piotr Dworczak, and Mohammad Akbarpour, "Comparison of Screening Devices," 2025. SSRN Working Paper 4456198.

7 Figures

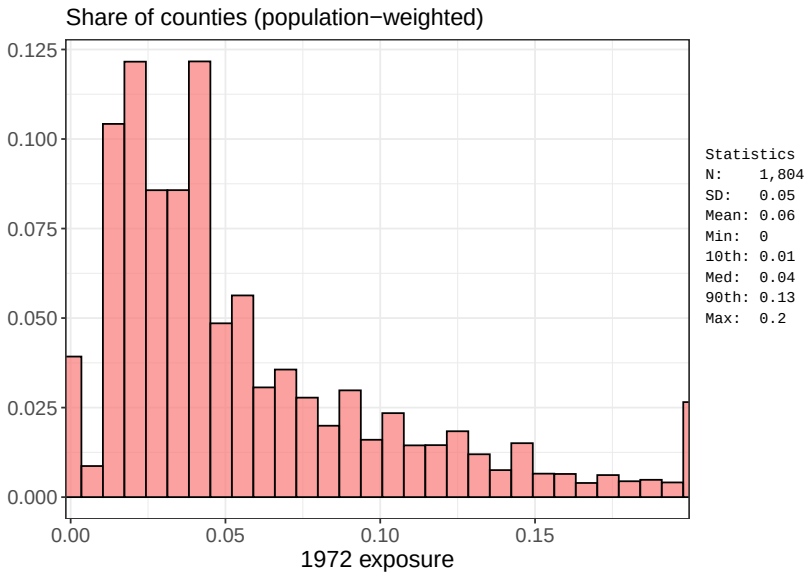
Figure 1: Impact of Eliminating the Purchase Requirement: Graphical Illustration



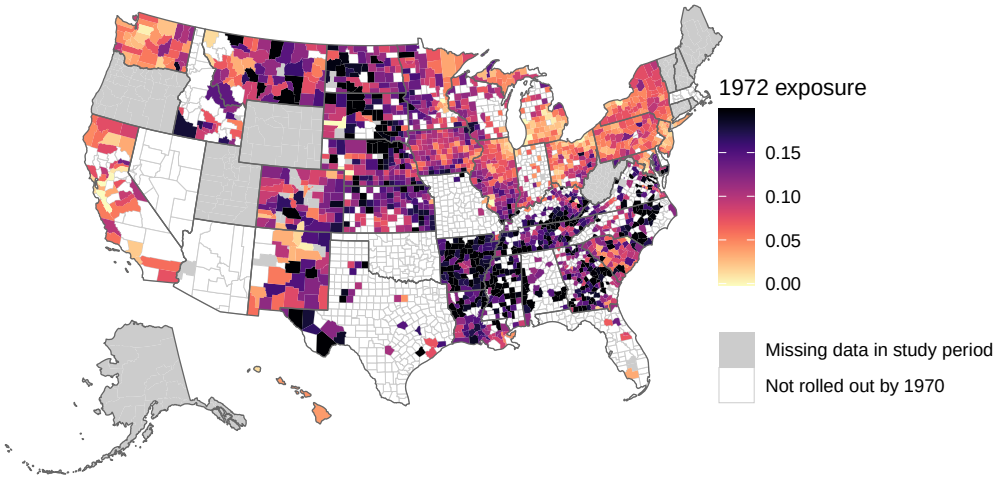
Notes: Figure shows budget sets and hypothetical choices for those not enrolled in the food stamp program (red line), those enrolled under the pre-1979 purchase requirement (black line) and those enrolled under the current, free stamps program (blue line). F denotes the food stamp payment by the government (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 the Free Food Stamp Reform

(a)

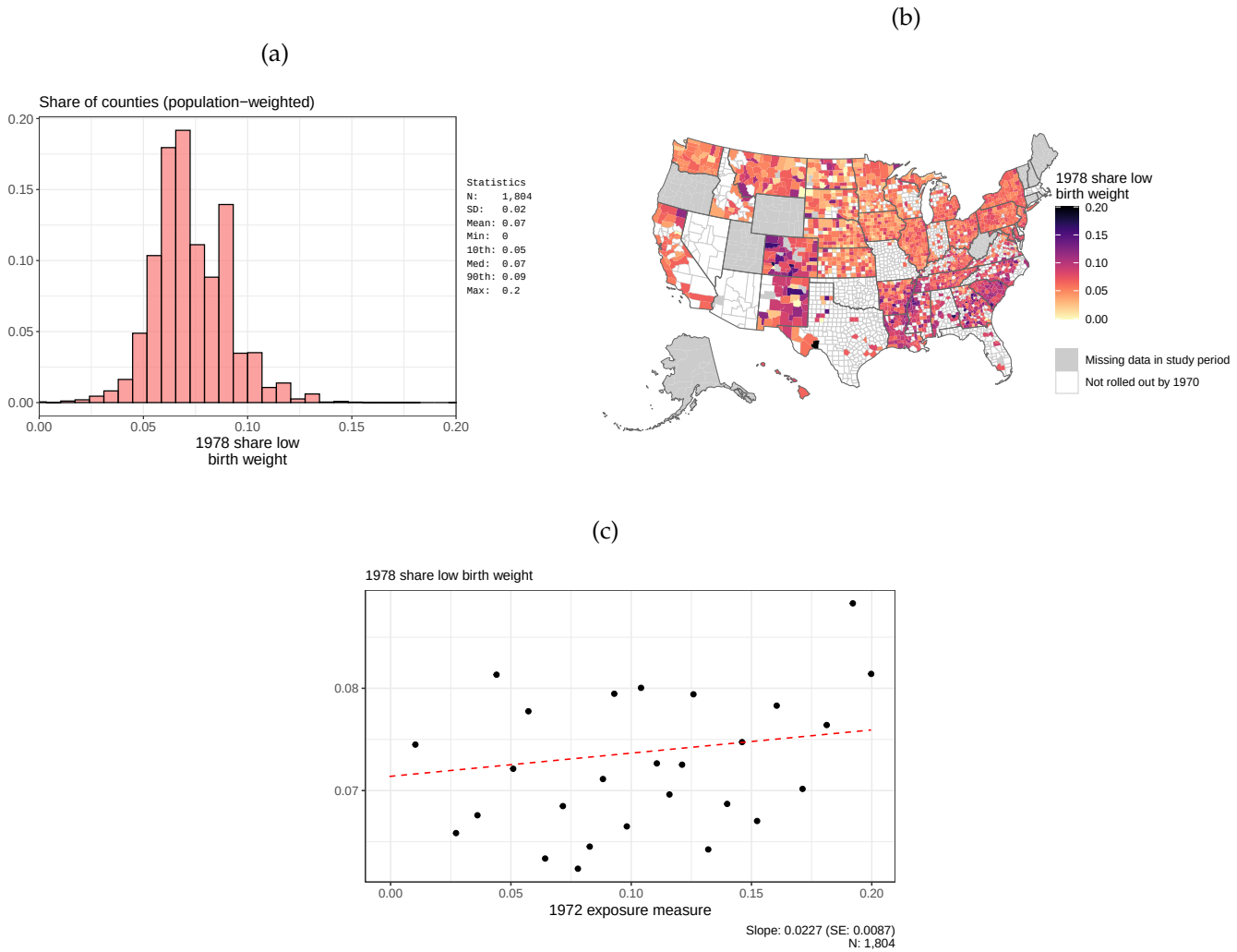


(b)



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 Appendix Table A1).

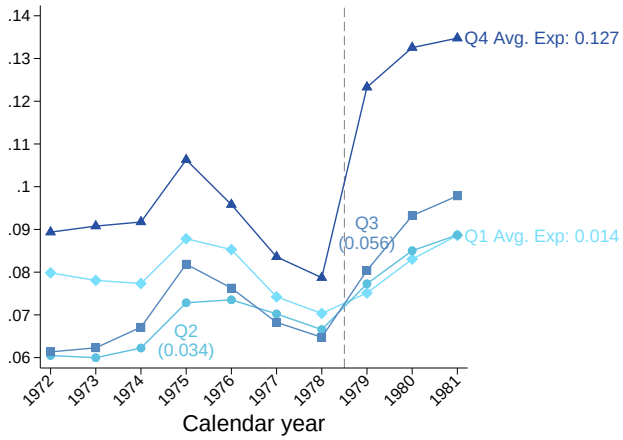
Figure 3: Variation in 1978 County-level Share Low Birth Weight



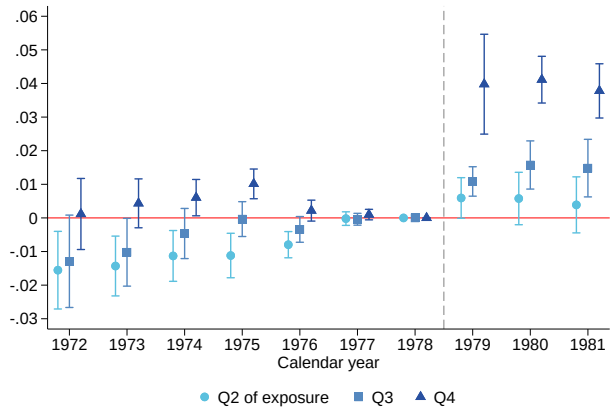
Notes: Figure describes the 1978 county-level share of births $< 2{,}500$ grams (low birth weight). Panel (a) shows a population-weighted histogram of share low birth weight, and Panel (b) shows a heat map of share low birth weight across counties. Panel (c) shows a binned scatterplot of 1978 share low 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 stamp reform (see Appendix Table A1).

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

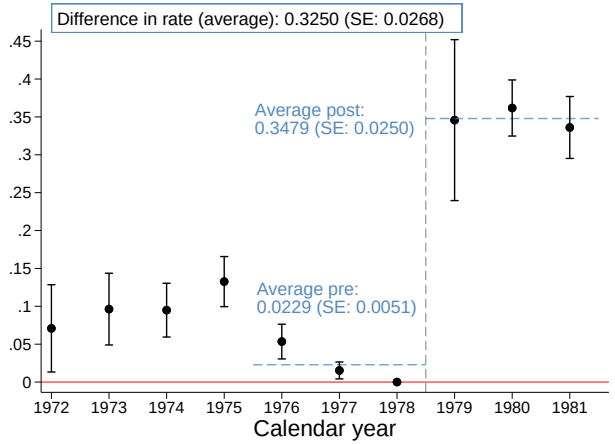
(a) Trends by Quartile of Exposure Measure



(b) Impacts by Quartile of Exposure Measure



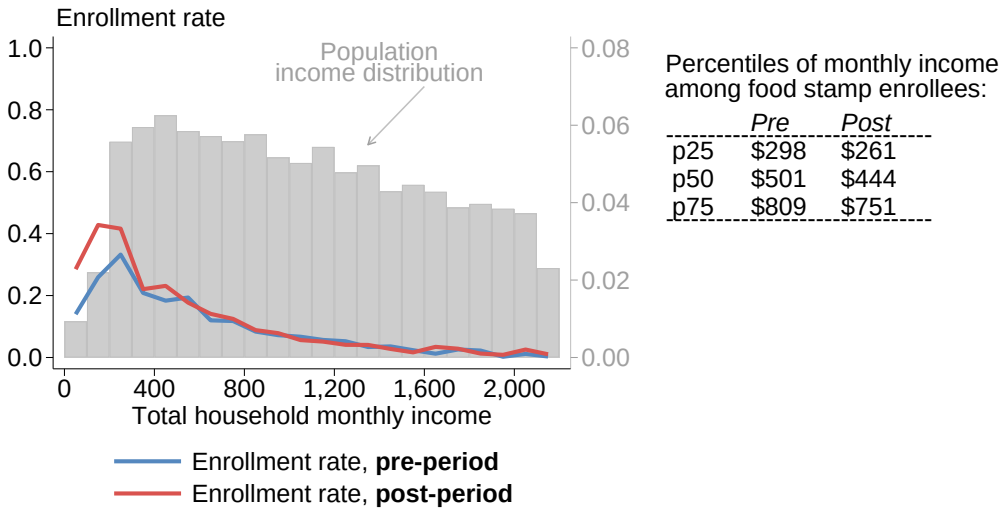
(c) Impact on Food Stamp Program Enrollment



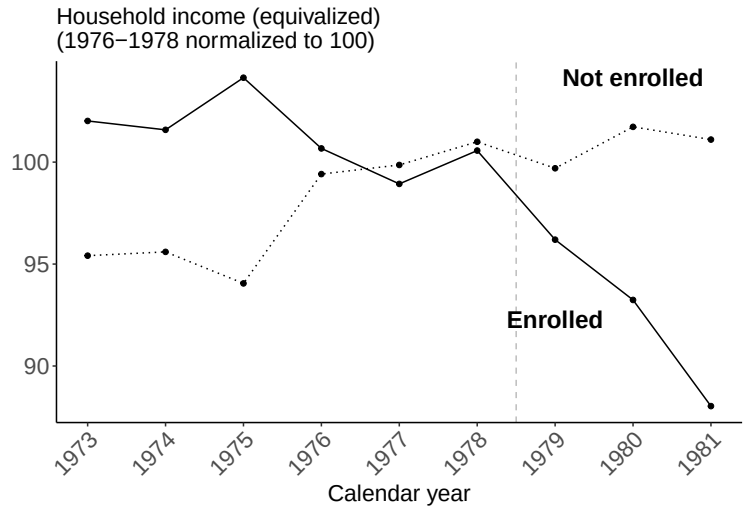
Notes: Panel (a) shows 1970 population-weighted average county-level enrollment rate over time by population-weighted quartile of the exposure measure. Panel (b) shows the results of estimating the first-stage Equation 4 using binary cuts of the 1972 exposure measure (where the lowest quartile of exposure comprises 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)

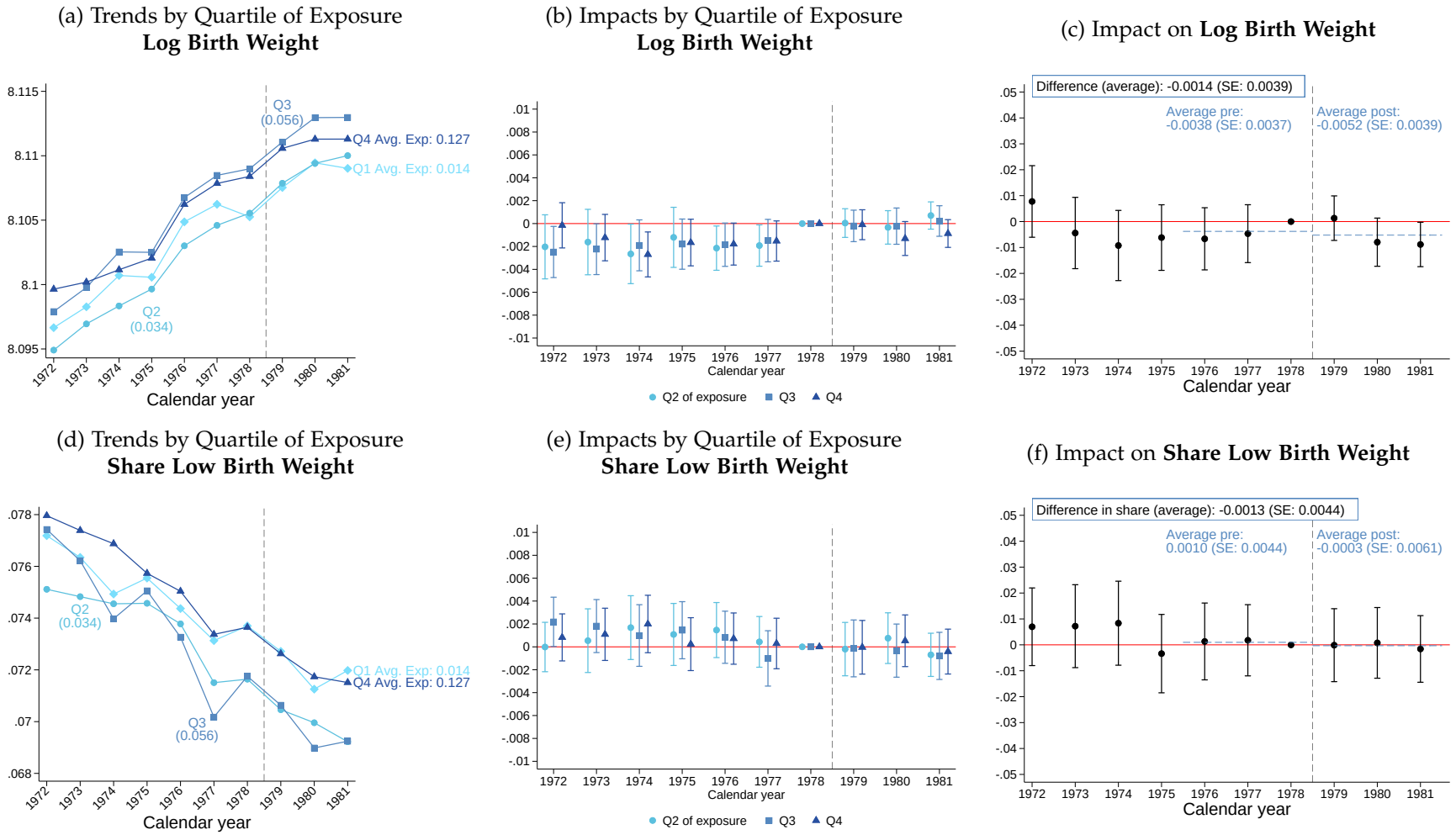


(b)



Notes: Panel (a) limits the PSID sample to households with income below the 99th percentile of 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-year observations 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 equivalized household income (i.e., adjusting for number of adults and children in the household) based on pooling observations from the PSID and NHANES; Panels (a) and (b) of Appendix Figure A5 show results separately for each survey. All income amounts are in $1979. See Appendix B.8 for details.

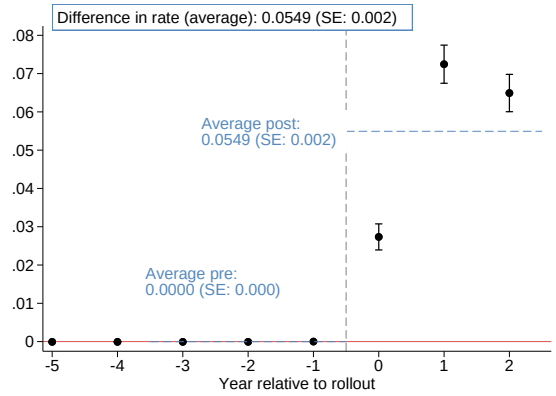
Figure 6: Free Food Stamp Reform: Impact on Infant Birth Weight



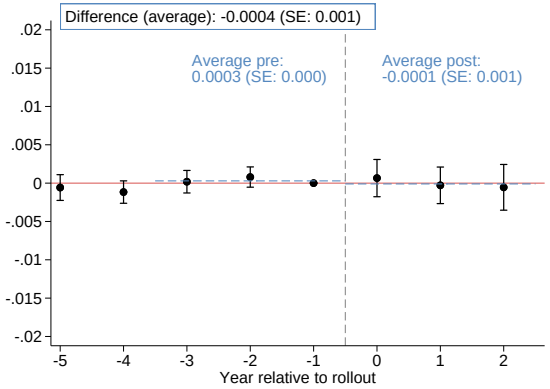
Notes: Panels (a) and (d) show 1970 population-weighted average county-level outcomes over time by population-weighted quartile of the exposure measure. Panels (b) and (e) show the results of estimating the reduced-form equation (Equation 5) using binary cuts of the 1972 exposure measure (where the lowest quartile of exposure comprises the comparison group) for log birth weight (Panel b) and share low birth weight (Panel e). In Panels (a), (b), (d), and (e), 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: Impact of Food Stamp Program Rollout: Event Study

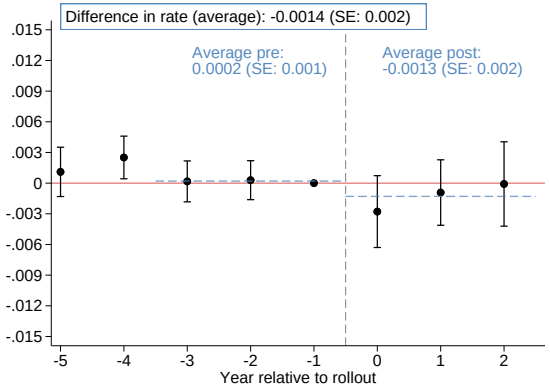
(a) Impact on Food Stamp Program Enrollment



(b) Impact on Log Birth Weight



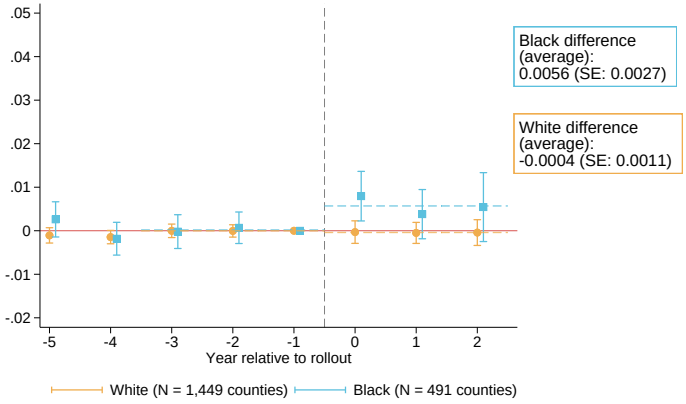
(c) Impact on Share Low Birth Weight



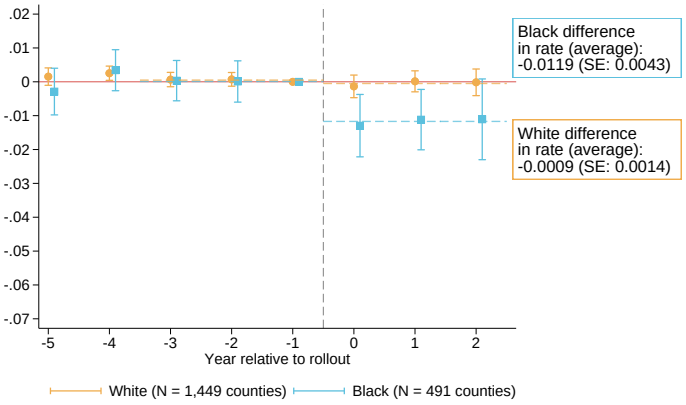
Notes: Figure shows event-study coefficients δ_τ from estimating equation 7 in Panel (a) and β_τ from estimating Equation 8 in Panels (b) and (c). The regressions are weighted by dataset-normalized weights, which are described in Appendix B.5. Vertical lines denote 95% confidence intervals computed using standard errors clustered at the county level. N = 1,520 counties.

Figure 8: Impacts on Infant Health Outcomes by Race, Rollout

(a) Rollout
Log Birth Weight



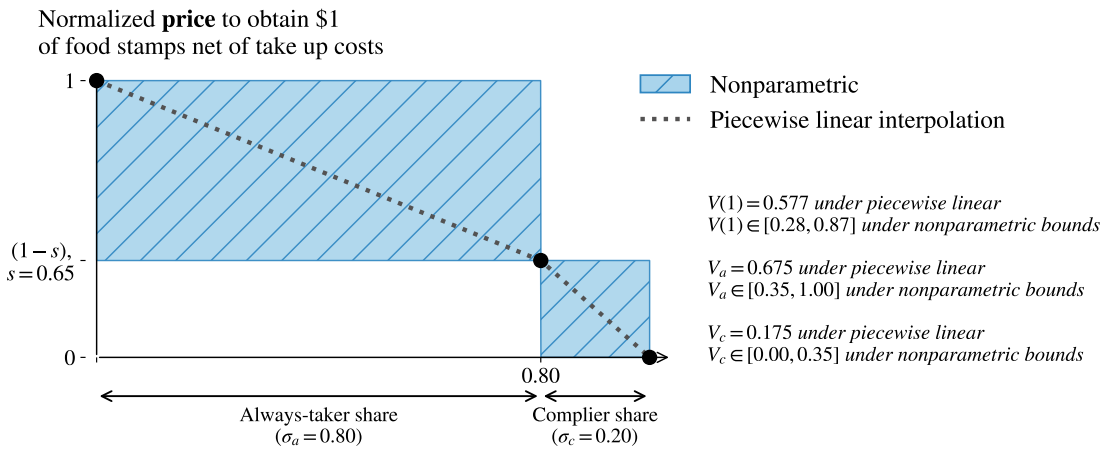
(b) Rollout
Share Low Birth Weight



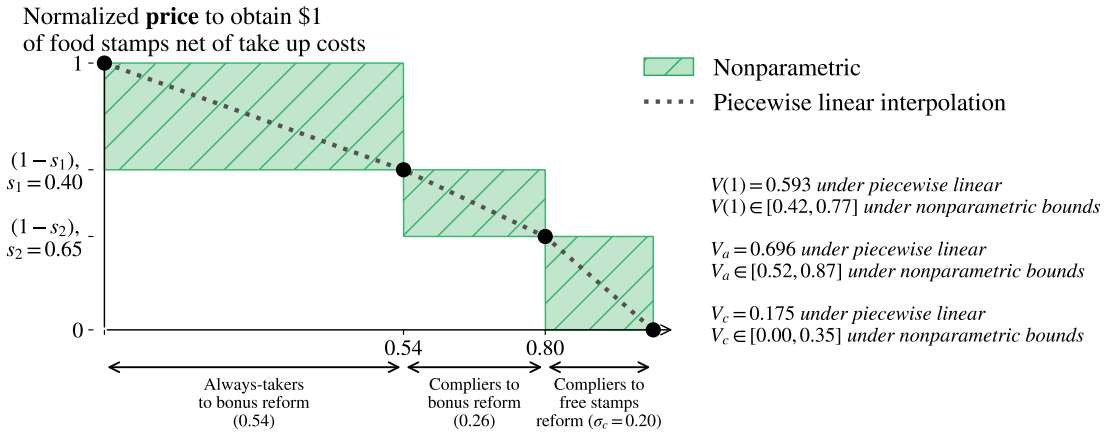
Notes: Panels (a) and (b) show β_τ from estimating equation 8 using birth weight outcomes for White and Black births, separately. The regressions are weighted by dataset-normalized weights, described in Appendix B.5. From the baseline sample for each analysis, the sample in this figure keeps only county-races with 25 or more births in each year relevant to the analyses (1972–1981 for the free-stamps reform, 1970–1975 for rollout). Vertical lines denote 95% confidence intervals computed using standard errors clustered at the county level. Sample size, in terms of counties, is denoted in each panel legend.

Figure 9: The Value of Food Stamps: Demand Curves

(a) With Purchase Requirement Only



(b) With Purchase Requirement and Bonus Reform



Notes: Panels (a) and (b) show points on the demand curve for food stamps. The points $(0,1)$ and $(1,0)$, written as (share enrolled, price), reflect the assumptions that no one would enroll in food stamps absent any subsidy, and that individuals observed enrolling must weakly prefer enrollment. The points with a subsidy less than 100% come from the take-up response to the elimination of the purchase requirement in 1979 (which increased the subsidy from 0.65 to 1) and the bonus reform (which increased the subsidy from 0.4 to 0.65). We linearly interpolate between these points to form an estimate of the average value of food stamps among all enrollees. The enrollment rate prior to the free-stamps reform was 7.6 percent; it increased by 1.85 percentage points in response to the free-stamps reform and increased by 2.4 percentage points from the bonus reform. This implies that 20% ($= 1.85 \div (7.6 + 1.85)$) of the post-reform enrollees were compliers with the free-stamps reform and approximately 26% ($= 2.4 \div (7.6 + 1.85)$) were compliers with the bonus reform. The free-stamps reform thus gives us the point $(0.8, 0.35)$ on the demand curve while the bonus reform gives us the point $(0.54, 0.6)$, where 0.54 ($= 0.8 - 0.26$) is the remaining share of households that are always-takers to the bonus reform.

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)
Panel A: Empirics			
Average annual household income ($1979)	8,187	7,285	3,605
	(189)	(163)	(1,132)
Equivalized income	4,108	3,797	2,528
	(76)	(78)	(486)
Share below 50% poverty level	0.066	0.129	0.388
	(0.009)	(0.008)	(0.066)
Share below 25% poverty level, given below 50%	0.160	0.233	0.534
	(0.038)	(0.030)	(0.217)
Share with child under three years old	0.555	0.604	0.804
	(0.015)	(0.013)	(0.095)
Panel B: Welfare weights			
Social marginal welfare weight	3.220	3.801	6.172
	(0.077)	(0.072)	(0.527)
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, accounting for the covariance between the pre-reform mean and the post-reform mean and assuming zero covariance between the mean estimates and the estimated 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 across all household-years, 1976–1981. See Appendix B.8 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. Similarly, "Share with child under three years old" is only estimated in the PSID, N = 2,553 households in the pre-period and N = 3,146 in the post-period.

Table 2: IV Analysis of Food Stamp Enrollment 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 births							
Post rollout	0.0506						
	(0.0016)						
Post free FS reform	0.0185						
	(0.0015)						
Enrollment rate		0.0098	-0.0041	-0.0032	-0.0448	-0.0041	-0.0067
		(0.0290)	(0.0120)	(0.0114)	(0.0387)	(0.0134)	(0.0128)
First-stage F		875.7	147.9	510.0	875.7	147.9	510.0
p-value (joint test of instr.)	<0.0001						
Hansen J p-value				0.657			0.318
Baseline Mean	0.0253	8.102	8.106	8.104	0.0732	0.073	0.073
Observations	3351298	3340293	10824	3351117	3340293	10824	3351117
Counties	2853	1520	1804	2853	1520	1804	2853
Panel B: Black births							
Post rollout	0.1353						
	(0.0046)						
Post free FS reform	0.0421						
	(0.0050)						
Enrollment 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 C: White births							
Post rollout	0.0250						
	(0.0006)						
Post free FS reform	0.0075						
	(0.0010)						
Enrollment 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: Column 1 shows the first-stage estimates δ from estimating Equation 10 and the remaining columns show the IV estimates λ of the impact of enrollment on birth weight from estimating Equation 9 using either the free-stamps reform (columns 3 and 6), the rollout (columns 2 and 5) or both (columns 4 and 7) as instruments for enrollment. The sample includes relative years -3 to +2 for food stamp program rollout and years 1976–1981 for the 1979 reform. The instrument for the free-stamps reform is a three-year post-period indicator interacted with exposure; for the rollout, it is an indicator for relative year ≥ 0. First-stage estimates for the free food stamp reform are scaled to reflect the average change in exposure; likewise "Post free FS reform" impact is the exposure $\times$ post coefficient rescaled 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. The regressions are weighted by county population (county-race population for race-specific estimates) in 1970 for free food stamp reform county-years and dataset-normalized weights (Appendix B.5) for rollout county-years. The weights within each sample (rollout and free food stamp reform) have been rescaled to sum to 1. Standard errors are clustered at the county level.

Appendices for Online Publication

Contents

A Setting and Background 53

B Empirical Appendix for Reduced-Form and IV Analyses 63

C Appendix for Welfare Analysis 75

D Appendix Figures and Tables 91

A Setting and Background

A.1 Food Stamp Program: History of Major Reforms

A.1.1 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 about 40% (Hoover, 1972).

The so-called "bonus reform" — effective January 1970 — increased and standardized benefit levels. It was fueled by concerns about low program enrollment, and greatly increased the bonus. In addition, one nationalized benefits schedule took the place of the region-specific ones. The policy also made slight changes to the purchase requirement (slightly decreasing it at the bottom of the income distribution, and carrying out minor changes at the top that seem to account for inflation, shown in Appendix Figure A44). As a result of the reform, the mean subsidy rate increased to about 65% (63.2%) (Hoover, 1972; Richardson, 1979). We will use this "bonus reform" as an additional source of variation in the subsidy rate in Section 5.3.

These changes to the benefits schedule 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. This legislation 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.1.2 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 toward monthly food consumption. The government bonus (F in the notation in Figure 1) was intended to 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. 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, participating households essentially all chose the maximum allotment $R + F$ (see Appendix A.2.2).

As noted, the food subsidy the household received varied with household income. The subsidy was 100 percent for the poorest households who had no purchase requirement ($R = 0$), and the purchase requirement increased by about \$1 with every \$3 increase in income (Congressional Budget Office, 1977).[38] 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 A43).

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" (ATP) with their purchase

[38] Transfers from other government programs such as Aid to 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).

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.

A.1.3 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 A7, Panel (a) shows that the bonus increased annually (reflecting 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%. Given that the pre-reform subsidy was decreasing in income, the reform entailed a larger increase in the subsidy for higher-income households, as shown in Appendix Figure A7, Panel (b).[39]

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.

The reform 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). However, data from the PSID suggest that any impact on enrollment from elimination of welfare-program-based categorical eligibility was likely quantitatively small. The share of food stamp households that are also participating in these programs is flat over time (Appendix Figure A42).

Ordeals. The 1979 reform did not change the basic "ordeals" associated with receiving food stamps. After the reform, households generally still had to go to the local post office or other distributor to pick up their food stamps. However, two aspects of the reform are worth noting.

First, the 1979 reform also required local food stamp offices to provide outreach and nutrition education materials, as well as access to bilingual personnel and materials for applicants (U.S. Department of Agriculture, Food and Nutrition Service, 1978b). Additionally, the legislation set standards for program offices in terms of location and hours of operation, aiming to ensure that households were able to apply and purchase stamps, especially in rural areas (U.S. Senate, Subcommittee on Nutrition, 1979). However, unlike the timing and requirements of the elimination of the purchase requirement, which were uniformly applied at the federal level, the implementation of these non-benefit-related provisions was left to the states, and, in most cases, specific service areas and certification offices (U.S. Department of Agriculture, Food and Nutrition Service, 1978b). Contemporaneous evidence suggests that implementation was resisted and delayed; the National Association of Farmworker Organizations reported delays in implementing bilingual services, while

[39] Appendix Figure A43 shows bonuses and purchase requirements before and after the reform.

other witnesses documented continued administrative barriers and resistance from local food stamp offices in implementing accessibility reforms (U.S. Senate, Subcommittee on Nutrition, 1979).

Second, the detailed implementation guide of the Food Stamp Act of 1977 formalized that states were allowed to issue coupons via mail to any household (U.S. Department of Agriculture, Food and Nutrition Service, 1978b). Prior to this, when households were on AFDC or other cash welfare, Public Assistance Withholding (PAW) rules allowed the government to deduct the purchase requirement directly (PAW was disallowed and then re-imposed several times during the 1970s) (Pub. L. 91-671, 1971; Pub. L. 92-603, 1972; Pub. L. 93-86, 1973; Pub. L. 94-585, 1976). Some jurisdictions (e.g., in California) already allowed mail issuance of coupons to any household (including those who were not cash welfare recipients) before 1979 (California Department of Benefit Payments, 1976). The 1979 reform did not require mail issuance, and we found no historical evidence that suggests systematic expansion of mail issuance exactly in 1979, in contrast to the large jump in enrollment when food stamps became free that year.

In Appendix C.5.2 we present several pieces of quantitative evidence that indicate that the reform did not meaningfully change ordeals associated with application or take-up in 1979.

A.2 Food Purchases Relative to Food Stamp Benefits

Most food stamp enrollees are inframarginal; 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 1977–80 Surveys of Food Consumption in Low Income Households.

A.2.1 Existing Evidence of Inframarginality

Contemporary Free Food Stamp 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, Inc., 2012; Hoynes et al., 2015b; 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, Inc., 2012; Hoynes et al., 2015b) while transaction-level data from a large U.S. 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 inframarginal (Hastings and Shapiro, 2018).

Estimates from the 1970s Purchase Requirement. We were only able to locate one estimate of food spending relative to food stamp receipt from the time period of the purchase requirement. Moffitt (1989) analyzed a cross-sectional survey of the Puerto Rican population in 1977, estimating that about 8 percent had food expenditures below F, and that "a bit more, 13 percent, have expenditures below [F] plus $2, possibly a more relevant figure if there is measurement error." Below, we conduct our own, similar analyses for the national population using the Surveys of Food Consumption and find similar estimates for the share of enrollees under the purchase requirement who were inframarginal.

However, there is additional indirect evidence that suggests that most enrollees under the purchase requirement were inframarginal. Under the purchase requirement, food stamp participants had the option to purchase either the full amount or a fraction of the food stamps available to them in a given month. Estimates from the period suggest that only about 6% or less of food stamp enrollees under the purchase requirement received less than their full potential allotment of food stamps (Reese et al., 1974; Clarkson, 1976; MacDonald, 1977). We produce similar findings below using the Survey of Food Consumption, which is a different dataset from what these prior papers analyzed.

A.2.2 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 the benefit schedule specified by the Federal Register. We directly analyze food expenditures relative to food stamp 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. The two surveys (henceforth collectively referred to as NFCS) contain information for 7,625 low-income households. Household-level survey weights are included. We convert all monetary amounts (e.g., income and expenditures) to $1978.

We form monthly food expenditure measures from NFCS reports of the typical amount that households 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 food stamp 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 household expenses (e.g., rent and utilities).

Data on Food Stamp Benefit Schedules. Appendix Figure A43 provides examples of benefit schedules before and after the free-stamps reform.

Measuring R and F in 1978. For our analyses of inframarginality, we assign all NFCS households (*regardless of the year they were observed in*) the benefits they would have been prescribed under the 1978 rules. Specifically, we use the benefits schedule published in the Federal Register in May 1978 (U.S. Department of Agriculture, Food and Nutrition Service, 1978c). This allows us to compare food purchases in all years to a single, fixed reference point.

As shown in Appendix Figure A43, benefits were defined such that a household of size h with net income i would buy a coupon allotment of food stamps worth $\overline{F}_h$ from the government at a price of R_{ih}. This resulted in a bonus $F_{ih} = \overline{F}_h - R_{ih}$. The value of the coupon allotment $\overline{F}_h$ was determined by the Thrifty Food Plan (TFP), a minimum monthly food budget needed for a household of size h to afford a nutritionally adequate diet. The value of R_{ih} did not exceed 30% of the household's net income. Thus, the transfer let the household afford the TFP without spending more than 30% of its net income.

Net income, defined in the Code of Federal Regulations (Office of the Federal Register, 1978), was the household's total gross income minus a set of deductions. 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). The NFCS contains both pre- and post-tax income, as well as information on dependent care expenses and shelter costs. Thus, we are able to partially implement these deductions to approximate net income in all years.

By inflation-adjusting this estimated net income to 1978 dollars, we are able to assign each NFCS household the coupon allotment $\overline{F}_h$, purchase requirement R_{ih}, and bonus $F_{ih} = \overline{F}_h - R_{ih}$ that it would have had in 1978.

Defining and Measuring Inframarginality. We measure inframarginality relative to the 1978 benefit schedule. We define two kinds of inframarginality. A household is labeled as *strongly inframarginal* if it spends at least $\overline{F}_h = R_{ih} + F_{ih}$ on food. It is labeled as *weakly inframarginal* if it spends at least F_{ih} on food. Strong inframarginality thus implies weak inframarginality but the converse is not true.[40]

We suspect more measurement error in our measurement of weak inframarginality than strong inframarginality. This is because our estimate of net income for determining R_{ih} and F_{ih} likely comes with unavoidable measurement error inherent to the NFCS data, and this error propagates into our estimation of "weak inframarginality" (i.e., share with food spending more than F_{ih}). However, it does not affect our assessment of strong inframarginality because $\overline{F}_h = R_{ih} + F_{ih}$ depended only on household size. Of course, estimates of both weak and strong inframarginality will be affected by any measurement error in the food spending variable.

Results. Several pieces of evidence suggest that households are inframarginal. First, under the purchase requirement, food stamp 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 A45 shows that the vast majority (94%) of food stamp enrollees purchased their entire coupon allotment. This is consistent with pre-1979 food stamp enrollees all being strongly inframarginal.

Second, we consider the pattern of monthly food expenditures. Appendix Figure A46 shows the distribution of monthly food expenditures relative to the household-specific 1978 $R + F$ for enrollees

[40] This classification differs from the definition of strongly and weakly inframarginal in Section 5.2 in that we of course do not observe counterfactual food consumption in the absence of the program, and must classify households by actual consumption.

in the pre-reform period (top panel) and the post-reform period (bottom panel). Appendix Table A8 combines the pre-period and post-period mean food spending relative to F and to $R+F$ with standard complier algebra to infer characteristics of always-takers (Column 1) and compliers (Column 3). The first row of Panel (a) reports the share who are weakly inframarginal (food spending at least F) and the first row of Panel (b) 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 estimates of food expenditures and of food stamp bonuses F (which depend on income) naturally come with (potentially substantial) measurement error. Intuitively, if all households are strongly inframarginal but the survey measures both food expenditures and income with noise, then some households may be misclassified as not strongly inframarginal. Second, if there is bunching at the kink of the nonlinear 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 that prior to the elimination of the purchase requirement, always-takers were at least *weakly* inframarginal: 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, they suggest that most of these households were also *strongly* inframarginal: 63% of households on food stamps had monthly food purchases in excess of their total allotment $R+F$. Consistent with some measurement error, if we slightly relax the definition of strongly inframarginal to account for potential mismeasurement, then the estimated share of strongly inframarginal households increases to 83 percent. Combined with the evidence in Appendix Figure A45 that most (94%) 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 inframarginal.

Column 3 shows similar calculations for compliers, using standard complier algebra together with our estimated complier share of 0.20 (see Section 3.2). Panel (a) indicates that 31% of compliers were weakly inframarginal, but as discussed above, we suspect our estimate of weak inframarginality is biased downward due to measurement error. If we similarly relax the definition of weakly inframarginal, then the estimated share increases from 31 percent to 51 percent. Panel (b) indicates that the share of compliers who are strongly inframarginal is not distinguishable from 0 (our estimate is actually negative at -0.37, but has a standard error of 0.21). The implication is that had compliers (counterfactually) enrolled under the purchase requirement, they would have valued food stamps below dollar-for-dollar transferred, in part because food stamps distort their purchases toward more food consumption than they would have chosen unconstrained.

Altogether, our assessment is that most of the deadweight loss in the free-stamps program reflects ordeal costs. However, some of the less than dollar-for-dollar valuation for free food stamps may reflect the fact that just under 20% of enrollees have less than dollar-for-dollar preferences for food spending: 69% of compliers and 7% of always-takers, noting that compliers are 20% of the overall enrollees.

A.3 Existing Literature on Food Subsidies and Birth Weight

Appendix Table A7 summarizes literature estimating the causal impact of the WIC program and the food stamp program on birth weight. We made several adjustments to the estimates reported in the original papers in order to make them comparable across papers. If the authors do not directly report an average treatment effect (ATE) of program enrollment on birth weight, but do report a reduced-form impact on birth weight and a statistically significant first-stage effect on enrollment, we scale the reduced-form intent-to-treat (ITT) estimates by the first stage to report an ATE of program enrollment on birth weight. We also convert estimates reported in grams to percentage terms.

Below we provide more detail on the exact estimates in each paper and any conversions that we undertook. Specifically, we report the in-text location for each estimate, the conversion of estimates reported in grams to percent terms, and the nature of the first-stage measure reported in each paper.

1. Sonchak (2016) employs maternal fixed effects to directly report the ATE of WIC receipt in their Table 2. Race-specific estimates are contained in this same table. Average birth weight and share low birth weight are found in Table 1. Birth weight estimates are reported in grams, and the estimate transformations are as follows:

   $$\% \text{ impact on average birth weight} = 100 \times \frac{51.638g}{3{,}133.75g}$$

   $$\% \text{ impact on Black birth weight} = 100 \times \frac{62.951g}{3{,}036.29g}$$

   $$\% \text{ impact on White birth weight} = 100 \times \frac{38.630g}{3{,}250.92g}$$

2. Chorniy et al. (2020) employ maternal fixed effects to directly report the ATE of WIC receipt on birth weight, as well as sample averages, in Table 3. Race-specific estimates are contained in this same table. Birth weight estimates are reported in grams, and the estimate transformations are as follows:

   $$\% \text{ impact on average birth weight} = 100 \times \frac{9.164g}{3{,}154g}$$

   $$\% \text{ impact on Black birth weight} = 100 \times \frac{16.084g}{3{,}051g}$$

   $$\% \text{ impact on White birth weight} = 100 \times \frac{3.252g}{3{,}268g}$$

3. Rossin-Slater (2013) uses variation induced by WIC clinic openings and closings (a proxy for clinic access) and maternal fixed effects to report an ITT estimate on birth weight, as well as sample averages in Table 6. Note that in Appendix Table 2, this ITT estimate is scaled by a first-stage estimate of the effect of any clinic access in the zip code in 39 weeks post-conception on clinic access during the actual pregnancy timeframe. Appendix Table 2 also reports an estimate of the effect of clinic access on WIC food receipt during pregnancy, which is the relevant first-stage estimate for converting the reported and scaled ITT into a comparable ATE of the impact of WIC receipt on birth weight. This estimate has a t-statistic < 2, so we do not report the first-stage or ATE. Race-specific estimates are not reported in the paper. Birth weight ITT estimates are reported in grams, and the estimate transformation is as follows:

   $$\% \text{ impact on average birth weight} = 100 \times \frac{27.302g}{3{,}275.41g}$$

4. Hoynes et al. (2011) use a staggered rollout design and report an ITT estimate of the impact of WIC rollout on birth weight in Table 3. Note that from Table 3, we back out the ITT estimate on share low birth weight using the reported mean and percent change: $\frac{-0.1}{100} \times 0.072$. Race-specific estimates are not reported in the paper. On page 819, they report that in 1988, 8% of pregnant women participated in WIC according to the National Survey of WIC Participants. This is the first-stage measure that we use to estimate the ATE of WIC receipt on birth outcomes. Hoynes et al. (2011) use this measure to directly report a treatment-on-the-treated estimate from a birth weight estimate in a model incorporating state×year fixed effects, also in Table 3, though we opt for the baseline estimate without additional fixed effects. Birth weight ITT estimates are reported in grams, and the estimate transformation is as follows:

   $$\% \text{ impact on average birth weight} = 100 \times \frac{2.7g}{3{,}316g}$$

5. East (2020) exploits variation in immigrant eligibility for the food stamp program in a difference-in-differences (DiD) design. While the paper focuses on child health outcomes, she reports impacts on birth weight, ITT and sample averages, as a part of the explanation of possible mechanisms for improved child health in Table 6. Race-specific estimates are not reported in the paper. Also contained in Table 6 is an estimate of the impact of eligibility on food stamp enrollment for a CPS sample of immigrant mothers with a high school education or less. This is the first-stage estimate we use to calculate an ATE of food stamp receipt on birth weight. Birth weight ITT estimates are reported in grams, and the estimate transformation is as follows:

   $$\% \text{ impact on average birth weight} = 100 \times \frac{6.52g}{3{,}295g}$$

6. Bitler et al. (2025) use a staggered rollout design and report an ITT estimate of the impact of the rollout of WIC on birth weight and sample means in Table 1. Note that share low birth weight estimates are reported in percentage point terms, so we multiply reported estimates for share low birth weight by 0.01 to get the estimate in comparable probability terms. Figure 2 provides an enrollment rate for pregnant women by race, constructed using data on WIC enrollment and natality data on all births. This is the first-stage estimate we use to estimate ATE, which Bitler et al. (2025) similarly do in Table 4. Pooled-race estimates are not reported in the paper. Birth weight ITT estimates are reported in grams, and the estimate transformations are as follows:

   $$\% \text{ impact on Black birth weight} = 100 \times \frac{11.709g}{3{,}080.08g}$$

   $$\% \text{ impact on White birth weight} = 100 \times \frac{5.123g}{3{,}339.085g}$$

7. Estimates for the Free Food Stamp Reform and Food Stamps Rollout come from Table 2 in this paper. The first-stage and ATE estimates are exactly as reported in Table 2, except in the case of the ATE for birth weight, where we exponentiate the log birth weight estimate to get it in percent terms. The ITT estimates are the reported reduced-form estimates from Figures 6, 7, and 8 and Appendix Figure A9, and we similarly exponentiate the log birth weight estimates. Note that all free-stamps reform ITT estimates are additionally scaled by the sample average exposure prior to exponentiation. Average exposure for the pooled, Black, and White samples is 0.057, 0.050, and 0.057, respectively. Below, we detail the exponentiation of all relevant log birth weight estimates.

- From Figure 6 and Appendix Figure A9, Free-Stamps Reform ITT estimates:
  - % impact on average birth weight $= 100 \times (e^{-0.0014 \times 0.057} - 1)$
  - % impact on Black birth weight $= 100 \times (e^{0.0007 \times 0.050} - 1)$
  - % impact on White birth weight $= 100 \times (e^{-0.0069 \times 0.057} - 1)$
- From Figures 7 and 8, Food Stamps Rollout ITT estimates:
  - % impact on average birth weight $= 100 \times (e^{-0.0004} - 1)$
  - % impact on Black birth weight $= 100 \times (e^{0.0056} - 1)$
  - % impact on White birth weight $= 100 \times (e^{-0.0004} - 1)$
- From Table 2, Free-Stamps Reform ATE estimates:
  - % impact on average birth weight $= 100 \times (e^{-0.0041} - 1)$
  - % impact on Black birth weight $= 100 \times (e^{0.0123} - 1)$
  - % impact on White birth weight $= 100 \times (e^{-0.0518} - 1)$
- From Table 2, Food Stamps Rollout ATE estimates:
  - % impact on average birth weight $= 100 \times (e^{0.0098} - 1)$
  - % impact on Black birth weight $= 100 \times (e^{0.0512} - 1)$
  - % impact on White birth weight $= 100 \times (e^{-0.0260} - 1)$

We calculate precision-weighted estimates from this literature by program and by race. Following Ganong et al. (2026), let $\widehat{\theta}_i$ denote estimate i and let $\widehat{\sigma}_i$ denote its standard error. We calculate the between-study variance as:

$$\widehat{\tau}^2 = \max\left\{0, \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{\theta}_i - \overline{\widehat{\theta}}\right)^2 - \frac{1}{n}\sum_{i=1}^{n}\widehat{\sigma}_i^2\right\}. \tag{19}$$

We then calculate the precision-weighted mean as:

$$\widehat{\overline{\theta}}' = \frac{\sum_{i=1}^{n}\left(\widehat{\sigma}_i^2 + \widehat{\tau}^2\right)^{-1}\widehat{\theta}_i}{\sum_{i=1}^{n}\left(\widehat{\sigma}_i^2 + \widehat{\tau}^2\right)^{-1}}. \tag{20}$$

Here, $\overline{\widehat{\theta}}$ is the simple mean of the n estimates. Equation 19 estimates the component of dispersion across estimates that remains after accounting for sampling variation. Equation 20 places greater weight on more precisely estimated effects, while allowing for between-study heterogeneity through $\widehat{\tau}^2$. When we include only estimates from this paper in the precision-weighted estimates, we use the precision-weighted race-specific estimates for the food stamp program in Appendix Table A7 that are directly reported from Table 2 of this paper instead of being estimated via Equations 19 and 20.

B Empirical Appendix for Reduced-Form and IV Analyses

B.1 Additional Infant Health Outcomes

We present reduced-form effects of the free food stamp reform (using Equation 5) and of the rollout (using Equation 8) on three other available birth measures: infant mortality rate, gestational age, and births themselves.

B.1.1 Data

Gestational Age. We make three sample restrictions to handle gestation. As in Chen et al. (2016), we limit gestation observations to births with at least 22 weeks of gestation. We keep only observations with less than or equal to 48 weeks of gestation to deal with outliers that are more likely errors. We set all gestational ages reported outside that range to missing. This censors only the highest half percentile of gestation values.

Second, as shown in Appendix Figure A3, 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 a mother's last menstrual period (LMP). This consistent reporting ensures more accurate gestation calculations. For state-years that do not report LMP, we treat all gestation values as missing.

Third, in 1981, NCHS began imputing gestation for births missing the day of the mother's 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 to missing, as well.

Infant Mortality Rate. The 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., 1972–81 for the free-stamps reform analysis, and 1970–75 for the rollout); for log infant mortality rate and log birth counts, we drop the county if the non-logged outcome is zero in any of the relevant years. Consistent with prior work (Bitler et al., 2025), we find that gestational age is missing for a non-trivial share of our sample: over 20% of the baseline analysis sample for the free-stamps reform and one-third of the sample for the rollout analysis.

As a result, for the free-stamps reform analysis, our analysis samples are 78.44% (1,359 counties) of the (population-weighted) 1,804-county baseline sample (77.99% of births in baseline sample) for gestation outcomes, 95.72% (1,145 counties; 95.72% of births) for infant mortality rate, and 100% for birth counts. For the rollout analysis, the analysis samples constitute 66.68% (932 counties) of the (population-weighted) 1,520-county baseline sample (60.94% of births in baseline sample) for the gestation outcomes, 95.45% (1,067 counties; 96.94% of births) for infant mortality rate, and 100% for birth counts.

B.1.2 Results

Free Food Stamp Reform. Appendix Figure A19 summarizes the results for each outcome across specifications; Appendix Figures A23 through A26 present the underlying event studies for each outcome, with Panel (a) reporting our baseline specification and Panels (b) through (f) reporting results from the alternative specifications that we also explored in our analyses of enrollment and birth weight (see Appendix Figure A18).

Appendix Figure A19 shows some evidence of a (puzzling) decline in births and of an increase in gestational age (and decrease in the share preterm). However, corresponding event studies show pre-trends in both births and gestational age in some specifications (though these are purged in others), making these results hard to interpret with confidence. There is little evidence of an impact on infant mortality rate.

Rollout. Appendix Figure A28 summarizes the findings for each outcome across specifications; Appendix Figures A33 through A36 present the underlying event studies for each outcome, again with Panel (a) reporting our baseline specification and Panels (b) through (f) reporting results from the alternative specifications that we also explored in our analyses of enrollment and birth weight (see Appendix Figure A27). Here the event studies are more compelling, but the evidence of increases in gestational age and declines in births is not robust across specifications.

B.2 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 countercyclical, 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 Proxy_{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 free food stamp reform. The Freyaldenhoven et al. (2019) procedure adjusts the focal variable — such as the food stamp enrollment rate or infant health — using leads of the policy variable as instruments for the

proxy. 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 as 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).[41] 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.[42]

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

$$y_{ct} = \lambda \text{Proxy}_{ct} + \sum_{\tau \neq 1978,1975} \beta_\tau (\mathbf{1}\{t = \tau\} \times \text{Exposure}_c) + \alpha_c^2 + \gamma_t^2 + u_{ct}^2 \tag{21}$$

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

$$\text{Proxy}_{ct} = \pi \left(\text{Post}_{t+1} \times \text{Exposure}_c\right) + \sum_{\tau \neq 1978,1975} \phi_\tau \left(\mathbf{1}\{t = \tau\} \times \text{Exposure}_c\right) + \alpha_c^1 + \gamma_t^1 + u_{ct}^1. \tag{22}$$

The excluded instrument is the lead of the policy variable defined as $\text{Post}_{t+1} \times \text{Exposure}_c$, where Post_{t+1} is an indicator that is 1 if calendar year t is 1978 or later. We normalize both 1978 and 1975 (relative years –1, –4) to be equal to each other and zero.[43] For both of the above equations, we cluster at the county level and weight by 1970 county population.

Appendix Figure A14 presents the results, following the template in Freyaldenhoven et al. (2019). Panel (a) plots the event study for 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

[41] 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.

[42] 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 be treated as zero in 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 A13 and A39 where we require employment be observed across all low-wage sectors.

[43] 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.

in the years leading up to 1976. Because food stamp enrollment is countercyclical, this force could plausibly decrease 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 β_τ event-study coefficients shown in Panel (a) (based on estimating Equation 5 with Proxy_{ct} as the outcome), and compares them (in green) to the δ_τ enrollment coefficients from Equation 4 shown in Figure 4c (in blue). This process involves a normalization that rescales the proxy coefficients so that β_{1975} equals δ_{1975}, the enrollment coefficient for 1975. 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. 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 (green series, Panel (b)) from the event-study coefficients 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 estimate in Panel (c) is 0.3141 (standard error = 0.050), which is economically and statistically indistinguishable from the unadjusted first-stage estimate of 0.3250 (standard error = 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 A21 through A26 and find minimal change to the reduced-form estimates.

B.3 Heterogeneity in Treatment Effects Across Deciles of Exposure Measure

Our exposure design assumes treatment effects are homogeneous across counties with different Exposure_c. With an untreated control group, it is possible to relax and test the homogeneity assumption in the exposure design (Sun and Shapiro, 2022). In our context, we approximate this by using the bottom decile of exposure, which is close to untreated (exposure = 0.007). Specifically, we form "stacks" of counties by exposure decile and compare each decile stack to the bottom decile; this design thus only imposes homogeneity within relatively homogeneous deciles.

More specifically, we construct a stacked dataset in which counties in each population-weighted exposure decile $s \in \{2, 3, \ldots, 10\}$ (i.e., each decile other than the lowest) are separately paired with the lowest-decile comparison group, forming a distinct "stack" s. The population-weighted average county-level exposure measure in the lowest decile is 0.007, nearly zero. Each county appears once within each stack for which it is eligible: as a treated observation in stack s if it belongs to decile s, or as a comparison observation if it belongs to the lowest decile.

We then estimate the following specification jointly across the stacked dataset:

$$\text{FS}_{cst} = \sum_{\tau \neq 1978} \delta^s_\tau \left(\mathbf{1}\{t = \tau\} \times \text{Exposure}_c \right) + \alpha^1_{cs} + \gamma^1_{ts} + u^1_{cst} \tag{23}$$

$$y_{cst} = \sum_{\tau \neq 1978} \beta^s_\tau \left(\mathbf{1}\{t = \tau\} \times \text{Exposure}_c \right) + \alpha^2_{cs} + \gamma^2_{ts} + u^2_{cst} \tag{24}$$

where c indexes county, s indexes the stack (i.e., the exposure decile being compared against the

lowest decile), and t indexes calendar year. Exposure$_c$ is the continuous exposure measure, and $\mathbf{1}\{t = \tau\}$ is an indicator equal to one if observation year t corresponds to calendar year τ, with 1978 omitted as the reference year. α_{cs} are county-by-stack fixed effects and γ_{ts} are calendar-year-by-stack fixed effects. We weight by the 1970 population (or stacked weights, as defined by Equation 25 and also used in the rollout analysis) and cluster standard errors by county.

Appendix Figure A15 reports estimates of δ_τ^s and β_τ^s, showing decile-specific impacts on enrollment and birth weight. We find large and precisely estimated increases in enrollment in deciles 3–10 of exposure, consistent with our baseline design. We see no increase in birth weight or share low birth weight in these deciles; if anything, across deciles, impacts on birth weight appear to be decreasing in the impact on enrollment.

B.4 Robustness Analysis

Free Food Stamp Reform. We probe the robustness of our estimates of the increase in enrollment from the free food stamp reform (Equation 4) and the null effects on both birth weight measures from the free food stamp reform (Equation 5). We explore a number of alternative specifications, which we describe here.

First, 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. Second, there is also a well-known, large decrease in births over our study period due to an "echo" from the baby boom (Appendix Figure A12). Since exposure may correlate with county-level differences in cohort effects, we also try controlling for the size of the 1970 county-level female population of child-bearing age, interacted with calendar-year fixed effects. Third, 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. Finally, because the WIC program was rolled out county-by-county in the mid-1970s (Bitler et al., 2025) and community health centers (CHC) were rolled out county-by-county in the late 1960s and 1970s (Köse et al., 2026), both could affect infant health outcomes. We tried controlling for indicator variables for county-years for which each program was in effect.[44]

Appendix Figure A18 shows that our estimates survive these tests; Appendix Figures A20 to A22 present the underlying event studies. 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.[45]

Finally, we note that our analysis stops in 1981 because in that year Congress passed three major pieces of food legislation, which restricted food stamp eligibility and reduced benefit inflation-indexing (Congressional Quarterly, 1982; Allen, 1983). Nevertheless, Appendix Figure A16 shows results through 1984. We find a *decrease* in birth weight (Panel b) and an increase in share low birth

[44]We use year of first non-planning CHC grant to a county from Köse et al. (2026) replication package, and we pull WIC rollout years from Figure 1 in Bitler et al. (2025).

[45]We consider the implications of the smaller estimate with state-by-year fixed effects for the normative analysis in sensitivity checks (Appendix C.5).

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.

Rollout. We probe the robustness of our estimates of the increase in enrollment from the rollout (Equation 7) and the null effects on both birth weight measures from the rollout (Equation 8). We find that results 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-by-time fixed effects, include the 1970 female population of child-bearing age interacted with time fixed effects, or include controls for the county-by-county rollout of WIC and of CHCs.

Appendix Figure A27 summarizes the comparison across these alternative specifications, while Appendix Figures A30 to A32 present the underlying event studies. We also show robustness to these alternative specifications by race (see Appendix Figure A29, and underlying event studies in Appendix Figures A37 and A38).

B.5 Reduced-Form Analysis of Rollout: Additional Details

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

Treatment and control weights, at the dataset d and relative year r level, are constructed as:

$$w_{dr} = \begin{cases} \dfrac{\text{Pop}_{d}^{\text{treat}} / \text{Pop}_{\text{tot}}^{\text{treat}}}{\text{Pop}_{dr}^{\text{treat}} / \text{Pop}_{r}^{\text{treat}}}, & \text{if county–time is treated,} \\ \dfrac{\text{Pop}_{d}^{\text{treat}} / \text{Pop}_{\text{tot}}^{\text{treat}}}{\text{Pop}_{dr}^{\text{ctrl}} / \text{Pop}_{r}^{\text{ctrl}}}, & \text{if county–time is control.} \end{cases} \tag{25}$$

The subscripts denote level of population summation where tot is the total across relative years and datasets, d is the total within datasets, r is the total within relative year, and dr is the total within dataset and relative year.

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:

$$x_{ct} = \alpha_c^1 + \gamma_t^1 + \sum_{\tau \neq -1} \delta_\tau (\mathbf{1}\{r = \tau\}) + \epsilon_{ct}^1 \tag{26}$$

$$y_{cm} = \alpha_c^2 + \gamma_m^2 + \sum_{\tau \neq -1} \beta_\tau (\mathbf{1}\{r = \tau\}) + \epsilon_{cm}^2 \tag{27}$$

where x_{ct} is a measure of the food stamp program (e.g., enrollment rate); y_{cm} is an outcome (e.g., birth weight); α_c^j is a county fixed effect (or county-by-race) and γ_t^1 or γ_m^2 is a year- or year-month fixed effect.

Since all counties are eventually treated, there is a standard time-period-cohort multicollinearity problem. To identify relative time indicators β_τ, 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 rollout.

B.6 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 the 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.[46] They do not report average 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 begin in 1968; AHS make a similar restriction for their event study analysis (their Figure 5).
- 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.

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

- 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 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 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 Appendix Tables A9 and A10 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:

$$Y_{ct} = \alpha + \delta \text{FSP}_{ct} + \beta \left(\text{CB60}_c \times t\right) + \phi X_{ct} + \eta_c + \gamma_t + \varepsilon_{ct} \tag{28}$$

where Y_{ct} is an outcome (e.g., birth weight), η_c is a county fixed effect, and γ_t is a time fixed effect. CB60_c are 1960 county characteristics (specifically share Black population, share urban, share farmland, and log population), which they interact with a linear time trend. The vector X_{ct} 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. Following AHS, the key variable of interest, 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 Appendix Tables A9 and A10 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 28 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 Appendix Tables A11 and A12 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 Appendix Tables A9 and A10. 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 28) 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 28 to a stacked dataset specification (with only county and time fixed effects):

$$y_{cmd} = \alpha_{cd} + \gamma_{md} + \delta \text{FSP}_{cmd} + \epsilon_{cmd} \tag{29}$$

where y_{cmd} is an outcome (e.g., birth weight), α_{cd} is a county-by-dataset fixed effect, and γ_{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 25).[47]

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):

$$y_{cmd} = \alpha_{cd} + \gamma_{md} + \sum_{\tau \neq -1} \beta_\tau (\mathbf{1}\{r = \tau\}) + \epsilon_{cmd} \tag{30}$$

where the terms are identical to those of Equation 29 except for the introduction of relative year 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:

$$\beta_{\text{avg}} = \frac{1}{3}\left[(\beta_0 + \beta_1 + \beta_2) - (\beta_{-3} + \beta_{-2} + \underbrace{\beta_{-1}}_{=0})\right]. \tag{31}$$

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 30 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.[48] Finally,

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

[48] These limitations include ensuring enrollment and birth weight are observed in each year 1970–1975 and dropping counties with any county-race-year cell with fewer than 25 births. Finally, we also ensure that after these limitations,

Columns 13–14 define relative year as relative to the 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).

Assessment. 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 the rollout of the food stamp program shows a statistically significant decrease in share of low birth weight among Black births and an insignificant and economically small impact on White births (Appendix Table A12, 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 4.2.1, in some specifications but not others, there is evidence of a statistically significant improvement in Black birth weight but there is no evidence in any specification of improvements in White birth weight (Appendix Figure A29).

Finally, since AHS do not estimate effects for a log birth weight outcome, Appendix Table A13 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 A13 corresponds to Column 5 in Appendix Table A9, and Columns 2–11 correspond to the earlier Columns 5–14.

B.7 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 Black to White households applies uniformly across counties.

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

$$\text{FS}_{ct} = \frac{e_{ct}}{\text{pop}_c}, \tag{32}$$

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

$$\text{FS}_{ct} = \text{FST}_{ct} \times \text{pov}_c, \tag{33}$$

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

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.

Our goal is to construct race-specific enrollment rates FS_{cgt} for groups $g \in \{w, b\}$ (White and Black). 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 A40, Panel a). These are defined as the number of PSID households of race g enrolled in food stamps divided by the number of PSID households of race g with income below the census needs standard.[49] We define the relative take-up ratio as

$$\delta_t = \frac{\text{FST}_{bt}}{\text{FST}_{wt}}. \tag{34}$$

Since δ_t is approximately constant over time (Appendix Figure A40, 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

$$p_{cg} = \text{pov}_{s(c)g} \times \text{pop}_{cg}, \tag{35}$$

where $\text{pov}_{s(c)g}$ is the state-level race-specific poverty rate from the census and pop_{cg} 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:

$$\text{FST}_{ct} = \frac{\text{FST}_{cbt}\, p_{cb} + \text{FST}_{cwt}\, p_{cw}}{p_{cb} + p_{cw}}. \tag{36}$$

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

$$\text{FST}_{cwt} = \text{FST}_{ct} \cdot \frac{p_{cb} + p_{cw}}{\delta\, p_{cb} + p_{cw}}, \tag{37}$$

$$\text{FST}_{cbt} = \text{FST}_{ct} \cdot \frac{p_{cb} + p_{cw}}{p_{cb} + p_{cw}/\delta}. \tag{38}$$

Finally, race-specific county-year enrollment rates are

$$\text{FS}_{cgt} = \text{FST}_{cgt} \times \text{pov}_{s(c)g}. \tag{39}$$

[49] 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.

B.8 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 five 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, preschool, 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 A5 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 monthly 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). We calculate social marginal welfare weights as the inverse cube of equivalized income (where we winsorize equivalized income at the 5th percentile), divided by the weighted mean of this inverse cube for all household-years 1976–1981. We compare the 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:

$$y_{htd} = \sum_{p \neq 1978} \alpha_p \cdot \mathbf{1}\{p = t\} + \gamma \text{NHANES}_d + \varepsilon_{htd} \tag{40}$$

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 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:

$$y_{htd} = \alpha + \beta \cdot \text{NHANES}_d + \delta \cdot \text{Post}_t + \varepsilon_{htd} \tag{41}$$

where α is the pre-period mean (normalized to PSID levels via the NHANES indicator), δ is the average change 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 41 together with our estimate of the complier share of $\sigma_c = 0.2$ (see Section 3.2) to estimate complier means as:

$$\bar{y}_{\text{complier}} = \alpha + \frac{\delta}{\sigma_c} \tag{42}$$

where α is the pre-period mean from Equation 41 and δ is the post-period coefficient (also in Equation 41). 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 41. This accounts for estimation uncertainty in both the first-stage estimates and the period-specific survey means. Appendix Table A4 also estimates enrollee household characteristics using this method for PSID household-years only.

Results. Appendix Figure A5 summarizes the results for equivalized income. Panels (a) and (b) show the results separately for the PSID and NHANES so that within-dataset trends are visible; the remaining panels pool them for increased power. The data indicate that after the free food stamp reform, mean equivalized household income falls among enrollees; trends among those not enrolled in food stamps are less pronounced.

C Appendix for Welfare Analysis

C.1 Microfoundations and Discussion of Welfare Assumptions

C.1.1 Primitives and Money-Metric Private Value

Let household $\theta \sim G$ have utility $U_\theta(c, f)$ over a numeraire good c and food expenditure f, with both prices normalized to one and income y. Define indirect utility without food stamps as

$$v_\theta(y) = \max_{c,f \geq 0} \{U_\theta(c, f) : c + f \leq y\},$$

and indirect utility from receiving B in food stamps as

$$v_\theta^F(y; B) = \max_{c,f \geq 0} \{U_\theta(c, f) : c + (f - B)_+ \leq y\}.$$

The household's money-metric willingness to pay for B, denoted $\widetilde{X}_\theta(B)$, is defined by:

$$v_\theta^F\left(y - \widetilde{X}_\theta(B); B\right) = v_\theta(y).$$

Thus, $\widetilde{X}_\theta(B)$ is the amount of income the household would be willing to give up to receive B in food stamps.

Let $\kappa_\theta(s, B)$ denote the utility cost of enrolling at subsidy s. Define utility from enrolling as

$$v_\theta^E(y; s, B) = \max_{c,f \geq 0} \{U_\theta(c, f) - \kappa_\theta(s, B) : c + (f - B)_+ \leq y - (1 - s)B\}.$$

We define the corresponding money-metric enrollment cost $\widetilde{k}_\theta(s, B)$ implicitly by

$$v_\theta^F\left(y - (1 - s)B - \widetilde{k}_\theta(s, B); B\right) = v_\theta^E(y; s, B).$$

Thus, $\widetilde{k}_\theta(s, B)$ is the amount of income loss that would leave the household as well off as incurring the non-price cost of enrollment.

Household θ's net private value for the food stamp bundle B (including both the purchase requirement R and the bonus F) per dollar of mechanical government spending sB is

$$V_\theta(s, B) = \frac{\widetilde{X}_\theta(B) - \widetilde{k}_\theta(s, B) - (1 - s)B}{sB},$$

and the average value among enrolled households is

$$V(s, B) = \mathbb{E}[V_\theta(s, B) \mid \text{Enroll}_\theta(s, B) = 1].$$

C.1.2 Constant Proportional Value

Section 5.1 imposed the "constant proportional value" (CPV) assumption:

$$\text{(A1, CPV for always-takers)} \qquad V_a(s_0, B) = V_a(s_1, B), \tag{43}$$

which we modify for our empirical setting to take into account that the subsidy bundle changes size (see Section 5.2 and Footnote 30) to:

$$\text{(A1}') \qquad V_a(s_0, R + F) = V_a(s_1, F). \tag{44}$$

Assumption A1 says that the welfare impact of the free food stamp reform on always-takers operates exclusively through a change in the subsidy. The subsidy may mechanically change how much of the bundle the government finances, but it does not independently affect households' valuations of the benefit B or their non-price costs of enrollment. For intuition, consider a standard setting with unit demand and no enrollment costs. Always-takers purchase under either subsidy rate (s_0 or $s_1 > s_0$). By envelope logic, a change in the subsidy affects their welfare only through the direct change in their budget constraint; their perceived private value of a dollar of mechanical government spending is the same under the higher and lower subsidy rate.

Assumption 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 government spending unchanged. The assumption therefore rules out additional impacts of the free food stamp reform on always-takers' valuations or non-price enrollment costs, such as stigma.

To see the restriction on primitives, let tildes $\widetilde{\cdot}$ denote levels rather than per-F units. Accordingly, $\widetilde{k}_\theta(s, R+F) \equiv k_\theta(s, R+F) \cdot s(R+F) = k_\theta(s, R+F) \cdot F$ is the dollar-level counterpart of the per-dollar transaction or ordeal cost $k_\theta(s, R+F)$ used in the main text. Then, $\widetilde{V}_\theta \equiv V_\theta \cdot s(R+F) = V_\theta \cdot F$. We define $\widetilde{X}_\theta(R+F)$ as gross willingness to pay for $R+F$ in food stamps, and $\widetilde{k}_\theta(s, R+F)$ as the transaction or ordeal costs associated with purchasing $R+F$ in food stamps (see Appendix C.1.1 for formal definitions), with

$$\widetilde{V}_\theta(s, R+F) \equiv \widetilde{X}_\theta(R+F) - \widetilde{k}_\theta(s, R+F) - R. \tag{45}$$

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

$$\Delta_a(s, R+F) := \widetilde{V}_a(s, R+F) - \widetilde{V}_a(1, F) \tag{46}$$

so that

$$\Delta_a(s_0, R+F) = \left[\widetilde{X}_a(R+F) - \widetilde{X}_a(F) - R\right] - \left[\widetilde{k}_a(s_0, R+F) - \widetilde{k}_a(1, F)\right]. \tag{47}$$

Then

$$V_a(s_0, R+F) = V_a(1, F) + \frac{\Delta_a(s_0, R+F)}{F}, \tag{48}$$

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

$$\widetilde{k}_a(s_0, R+F) = \widetilde{k}_a(1, F). \tag{49}$$

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.

C.1.3 Law of Demand

To estimate perceived net private value we impose what we call a "Law of Demand" (Equation 18). Here we elaborate on why we need to make this assumption and provide microfoundations for it.

Why Do We Need the Assumption? The key issue is that observed enrollment under the *purchase requirement* may not reveal 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 are a function of the subsidy (e.g., if they are multiplicative in bundle size), 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 B and vary only the subsidy rate s. Purchases up to B are subsidized, while purchases above B face the market price. For households who consume less than B 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 enrollees' value per dollar of government spending under the free program. 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 C.1.2, define net value per dollar of the benefit bundle (but gross of the purchase price) as

$$Y_\theta(s, B) := \frac{\widetilde{X}_\theta(B) - \widetilde{k}_\theta(s, B)}{B}. \tag{50}$$

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):

$$\text{(A2, } Y\text{-CPV for compliers)} \qquad Y_c(s_0, R + F) = Y_c(1, F). \tag{51}$$

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

$$0 \leq V_c(1, F) < 1 - s_0. \tag{52}$$

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:

$$\text{(A3, monotonicity)} \qquad \inf_{\omega_a \in \Omega_a} V_a(1, F; \omega_a) \geq \sup_{\omega_c \in \Omega_c} V_c(1, F; \omega_c), \tag{53}$$

where Ω_θ is the set of preference primitives consistent with type θ's observed behavior.[50] A3 is an intuitive sorting condition that rules out always-takers having lower values of free stamps than compliers.

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

$$V_a(1, F) \geq 1 - s_0 > V_c(1, F). \tag{54}$$

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

[50] 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.

on the demand curve (Appendix C.1.4).[51]

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

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

Equivalently,

$$\frac{\widetilde{X}_c(F)}{F} - \frac{\widetilde{X}_c(R+F)}{R+F} = \frac{\widetilde{k}_c(1,F)}{F} - \frac{\widetilde{k}_c(s_0,R+F)}{R+F}. \tag{56}$$

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 (defined in Appendix A.2.2):

$$\frac{\widetilde{X}_c(R+F)}{R+F} = \frac{\widetilde{X}_c(F)}{F} = 1. \tag{57}$$

Equation 56 then requires ordeal costs to be proportional to bundle size:

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

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 A.2.2 finds evidence of):

$$\widetilde{X}_c(F) = F, \qquad \widetilde{X}_c(R+F) < R+F. \tag{59}$$

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

$$\frac{\widetilde{k}_c(1,F)}{F} - \frac{\widetilde{k}_c(s_0,R+F)}{R+F} = 1 - \frac{\widetilde{X}_c(R+F)}{R+F} > 0. \tag{60}$$

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 smaller 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 56 in this case is that *both* gross valuation and ordeal costs scale proportionally with bundle size:

$$\frac{\widetilde{X}_c(R+F)}{R+F} = \frac{\widetilde{X}_c(F)}{F}, \qquad \frac{\widetilde{k}_c(s_0,R+F)}{R+F} = \frac{\widetilde{k}_c(1,F)}{F}, \tag{61}$$

[51] A3 is sufficient but not necessary for our empirical exercise in Section 5.3. There, we only require that always-takers' *average* value of free stamps be at least $1 - s_0$, the value of the marginal complier, so there could exist always-takers with lower private values than some compliers.

which then implies that

$$\frac{\widetilde{X}_c(R+F)}{\widetilde{X}_c(F)} = \frac{\widetilde{k}_c(s_0, R+F)}{\widetilde{k}_c(1,F)} = \frac{R+F}{F}. \tag{62}$$

Intuitively, Equation 62 says that in this case both preferences and ordeal costs satisfy the same proportional valuation in bundle size.

C.1.4 Bonus Reform

Under the bonus reform, let the purchase requirement R be fixed, while the government bonus rises from F_L to F_H. Thus, in regime $j \in \{L, H\}$, the household pays R and receives a bundle $R + F_j$. The subsidy rate is therefore

$$s_j = \frac{F_j}{R+F_j}. \tag{63}$$

We assume that F_H is the same magnitude as in the free-stamps program. We want to know the household's value per dollar of the free benefit, $V_\theta(1, F_H)$. Index households who are *always-takers* to the bonus reform by a' and households who are compliers to the bonus reform by c'. We assume an analog to the CPV assumption:

$$\text{(Bonus CPV)} \qquad Y_{c'}(s_L, R+F_L) = Y_{c'}(1, F_H), \tag{64}$$

where $Y(s,B)$ is defined in Equation 50 and noting that $Y_{c'}(1,F_H) = V_{c'}(1,F_H)$. Bonus CPV therefore assumes that gross value, net of ordeal costs, per dollar of the low-bonus bundle is equal to value per dollar of the eventual free benefit.

Because bonus-reform compliers do not enroll in the low-bonus regime, Equation 64 implies

$$V_{c'}(1,F_H) < 1 - s_L. \tag{65}$$

We also maintain the monotonicity assumption A3, but now for bonus-reform always-takers with respect to bonus-reform compliers:

$$\text{(A3}', \text{ monotonicity)} \qquad \inf_{\omega_{a'} \in \Omega_{a'}} V_{a'}(1, F_H; \omega_{a'}) \geq \sup_{\omega_{c'} \in \Omega_{c'}} V_{c'}(1, F_H; \omega_{c'}). \tag{66}$$

As with the free-stamps reform, absent further information, bonus-reform complier values can approach $1 - s_L$ from below, so

$$\sup_{\omega_{c'} \in \Omega_{c'}} V_{c'}(1, F_H; \omega_{c'}) = 1 - s_L. \tag{67}$$

A3′ therefore implies

$$V_{a'}(1,F_H) \geq 1 - s_L > V_{c'}(1,F_H). \tag{68}$$

Moreover, compliers to the free-stamps reform satisfy $V_c(1,F_H) < 1 - s_H < 1 - s_L$ by A2. Thus,

among households who eventually enroll when stamps are free,

$$\text{enroll under the low bonus} \iff V_\theta(1, F_H) \geq 1 - s_L. \tag{69}$$

The bonus reform therefore contributes the point on the demand curve at the observed price.[52]

To see primitives that would produce the Bonus CPV assumption for bonus-reform compliers, suppose

$$\widetilde{X}_{c'}(B) = B, \tag{70}$$

so compliers to the bonus reform are strongly inframarginal. Then Bonus CPV requires

$$\frac{\widetilde{k}_{c'}(s_L, R + F_L)}{R + F_L} = \frac{\widetilde{k}_{c'}(1, F_H)}{F_H}. \tag{71}$$

Thus, ordeal costs must scale so that ordeal costs per dollar of the bundle are constant across the two regimes. For example, this would hold if ordeal costs scaled proportionally with bundle size.

C.2 Graphical Analysis of the Welfare Consequences of Increasing the Subsidy

C.2.1 Conceptual Graphical Analysis

Appendix Figure A47 depicts graphically the welfare consequences of raising the subsidy. The graphs assume $D_\theta = 0$ and $\overline{\text{FE}} = 0$, but we now 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; in our context this corresponds to all those enrolled at the higher subsidy $s_1 = 1$. 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 in the main text, this need not be the case.

In Panel (a), the welfare weights are assumed constant across enrollees. Since demand is downward-sloping, the social value of always-takers therefore exceeds the social value of compliers

[52] Instead of Bonus CPV, we could make an assumption that compliers to the bonus reform have additive ordeal costs and value each extra dollar of the bonus dollar-for-dollar. This assumption would produce a similar conclusion.

($\lambda_a V_a(1) > \lambda_c V_c(1)$), and a lower subsidy produces more social value per dollar than a higher subsidy.[53] 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 and deliver $\lambda_c V_c(1) > \lambda_a V_a(1)$. Finally, Panel (d) starts with the situation in (b) but flattens the demand curve; this can also deliver the conclusion that $\lambda_c V_c(1) > \lambda_a V_a(1)$. These graphs thus 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.

C.2.2 Empirical Version of Welfare Graph

Appendix Figure A48 shows the empirical version of the above conceptual graph. The blue line in the figure 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).

C.3 Estimating Perceived Net Private Value: Extensions and Robustness

C.3.1 Additional Subsidy Variation: 1970 Bonus Reform

We provide more detail on our empirical analysis of the 1970 "bonus reform" used in Section 5.3. In 1969, Congress modified the program to increase the subsidy starting in January 1970; although this subsidy increase was 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, 1972).

Specifically, the bonus reform raised the "bonus" amount F, thus increasing the subsidy rate $\frac{F}{R+F}$. The reform also made small changes to the purchase requirement (see Appendix Figure A44). We use both of these sources of variation as a shock to enrollment, and combine them when measuring the subsidy changes for our demand curve. Because (see Appendix A.2) 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, 1972).

The bonus reform occurred at a national level. Therefore, to study its impact on enrollment, we use an exposure design similar to Section 3.1. Because the bonus reform occurred in a period during the food stamp 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., $\text{Exposure}_c = \text{Poverty}_c - \text{EnrollmentRate}_{cz}$) as before (see Equation 3) but now we use $z = 1968$ so that we measure county enrollment in 1968,

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

prior to the bonus reform. Appendix Figure A49 shows the distribution of our county-level exposure measure, winsorized at the population-weighted 1st (0.04) and 99th percentiles (0.33). The average 1968 exposure measure among these counties is 0.10 (standard deviation 0.06).

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

$$\text{FS}_{ct} = \sum_{\tau \neq 1969} \delta_\tau \left(\mathbf{1}\{t = \tau\} \times \text{Exposure}_c \right) + \alpha_c + \gamma_t + \mu_{o(c)t} + u_{ct} \tag{72}$$

where $\mu_{o(c)t}$ are rollout-cohort-by-calendar-year fixed effects, with $o(c)$ denoting the year in which county c initially rolled out the food stamp program. The extra fixed effects are required because of the concurrent program expansion.

Results. We first present visual evidence in Appendix Figure A50 Panel (a), akin to Figure 4 (Panel a) showing trends in enrollment rates by quartile of exposure, from 1967 to 1972. 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.[54]

Appendix Figure A50 Panel (b) shows the results from estimating the event study in Equation 72. 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%.

C.3.2 Relaxing Piecewise-Linearity

Our baseline estimate of $V(1) = 0.59$ assumed piecewise linearity of the demand curve, which is a strong assumption. We therefore also consider shape restrictions to relax piecewise-linearity and make fewer 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 flat" or "too steep" relative to the observed variation.[55] When $r \to 0$, we recover the piecewise-linear benchmark, under which admissible

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

[55] 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

$$\underline{\beta} = \frac{\hat{\beta}}{1 - r} \qquad \text{and} \qquad \overline{\beta} = \hat{\beta}(1 - r). \tag{73}$$

demand curves are the piecewise-linear functions interpolating the observed points. When $r \to 1$, we recover the nonparametric bounds above.

Appendix Figure A51 shows the results. The light blue and green regions in Panels (a) and (b) show the bounds on willingness to pay when we allow for any monotone demand curve consistent with the observed data from the free food stamp reform; the dark blue and green regions in these panels show how the bounds are tightened if we impose the shape restriction $r = 0.5$. Panel (c) 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 bonus 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 in [0.53, 0.65] if we include the bonus reform and [0.48, 0.68] if we exclude it.

C.4 Fiscal Externalities: Employment

Our main estimates assumed $\overline{\text{FE}} = 0$. As discussed in the main text (see especially Footnote 34), the main potential source of positive fiscal externalities comes from individuals reducing their labor supply (and hence income tax revenue earned) in order to qualify for food stamps; Hendren and Sprung-Keyser (2020) calibrate that these fiscal externalities are likely to be small. We complement these estimates by examining the impact of the free food stamp reform 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 (CBP). 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 infer CBP county-industry employment or payroll 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 B.2 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 that are in the primary rollout analysis sample and have employment outcome data in relative years –3 to +2. 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 700 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, which covers 78% of the baseline rollout sample (by population) and 29% of the 1970 U.S. population.

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

For the free food stamp reform, 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 A13). 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 average exposure is about 0.06, food stamp exposure reduces employment by 0.13% (standard error = 0.6%) and increases employment in exposed industries by 0.8% (standard error = 0.7%). We can therefore rule out employment declines of 1.3% overall and 0.66% in exposed industries. Next, we examine impacts of the food stamp rollout using the stacked specification (Equation 8). In Appendix Figure A39, we find tight null results; for employment, we find that the rollout reduced employment by 0.8% (standard error = 0.7%) and raised employment in exposed industries by 0.4% (standard error = 1%).

C.5 Empirical Welfare Analysis: Sensitivity Analysis

We provide more detail behind the sensitivity analyses of our welfare results discussed in Section 5.4.

C.5.1 Normative Assumptions

Our finding that free stamps produce less welfare per dollar of mechanical expenditure than cost-sharing (i.e., $W(1) < W(0.65)$) stems from the fact that welfare-weighted valuations are more than twice as high for always-takers as for compliers (i.e., $\lambda_a V_a = 2.2$ versus $\lambda_c V_c = 1.1$). We therefore 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 A14 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, assuming $\text{FE}_c = \text{FE}_a = 0$; we discuss relaxing this below.

Distortions (D_θ). Rows 2 through 4 consider distortions (D_c and D_a), which our baseline analysis assumed to be zero. Distortions can include behavioral biases or within-household bargaining constraints that would lead enrollment behavior not to reflect normative private value. 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 $W(1) < W(0.65)$. 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 39 cents per dollar. This would imply severe distortions for compliers in percent terms (over 200% 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 $D_c < 0.188$, 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 a stylized model of a particular behavioral bias would imply about distortions. 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-focused 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 $D_c = 0.188$ (what we need to overturn results if only compliers experience distortions), we would need that always-takers are all neoclassical ($\pi_a = 0$) while 62% of compliers are present-focused with $\beta = 0.7$.

Welfare Weights (λ_θ). Rows 5 through 9 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 welfare weights capture liquidity constraints, insofar as they reflect the social value of transferring \$1 to households. Under liquidity constraints, welfare weights based on annual data may only imperfectly reflect the instantaneous social value of transferring funds (e.g., because liquidity or savings constraints make it difficult to smooth consumption).

Reassuringly, our results are robust to a range of alternative welfare weights. For example, even if we double the complier welfare weight (Row 5), or assume different redistributive preferences by allowing a coefficient of relative risk aversion of 1 or 50 (Rows 6 and 7), we still get that targeting is worse under free food stamps.

Of course, since compliers are needier than always-takers, there exist welfare weight schedules that can allow us to conclude that raising the subsidy improves targeting. Trivially, if the social planner *only* values the compliers and not the always-takers, then making food stamps free must mechanically increase welfare per dollar relative to cost-sharing. Generalizing this idea, we explore

what step-function-type welfare function would be required to reverse our results. We imagine 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 8, we show that if the social planner places positive and equal weight only on households whose annual income is less than 50% of the 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 9). An implication is that if the social planner has a step-function welfare weight schedule with a cutoff at 100% of the FPL, that could not reverse our result. The extreme redistributive preferences implied by the step-function welfare weights with cutoff $\kappa = 63\%$ of FPL 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).

Likewise, flattening the welfare weight schedule, while reinforcing the conclusion that targeting is worse under the free food stamp reform, reduces our calibration of $W(1)$. However, even if we use log utility (i.e., CRRA with risk aversion of 1), we still obtain that $W(1) = 1.1$ (Row 6). The result of $W(1) > 1$ is hard to reverse, even with these relatively flat welfare weights, because even log utility places very high welfare weights on food stamp enrollees relative to the average person in society. Even so, our finding that $W(1) > 1$ is not mechanical, since enough compliers with low private valuations and insufficiently high welfare weights could have instead produced $W(1) < 1$.

Variation in Subsidy Across Households. Since we have abstracted from the fact that under the purchase requirement, R was not constant but increased with income, we show in Rows 13 and 14 that this abstraction is actually conservative for our targeting result.[56] Intuitively, compliers were poorer than always-takers. Assuming a constant subsidy rate thus overstates compliers' maximum willingness to pay for a dollar of food stamps, relative to using their actual (higher) subsidy rate that accounts for their lower income. Using the PSID, we form estimates of complier and always-taker subsidies using their income. We then estimate V_c and V_a using their PSID subsidy rates, rather than assuming homogeneity.[57] As compliers face higher subsidies, this makes V_c lower than our main estimate, so V remains low and $V_a > V_c$.[58]

Net Private Value (V_θ). Our estimates of V_θ come from interpolating bounds on demand. In Rows 10 through 14, we therefore consider how alternative assumptions on these demand curves would

[56] A related point 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 5% of 4-person enrolled households would have had a \$0 purchase requirement (Appendix Table A4).

[57] Specifically, we simulate a household-year's net income using the 1978 food stamp earned income and shelter deduction schemes. Additionally, we apply the 1979 standard deduction which replaced several of the pre-reform deductions which require variables we do not observe across all PSID years 1976–1981 (e.g., child care spending). This standard deduction is \$60 in 1979, and we CPI-adjust the deduction in all other years. We then match household-years to 1978 benefits using their net income. We calculate subsidy rate to be the ratio of bonus to bonus plus purchase requirement, and we obtain complier means according to Equation 42 for the subset of enrollees that receive a positive bonus amount in 1978 benefits and are below the FPL. From these benefit estimates, we similarly obtain the share of all enrolled households, without applying additional eligibility limitations, with \$0 purchase requirement in 1978 benefits in Appendix Table A4.

[58] Our imputed estimate of V_a rises slightly, because the PSID-estimated subsidy is slightly lower than the administrative subsidy estimates that we use in the primary analysis.

affect results. We consider changing the valuations to be at the upper bound (Row 10) or lower bound (Row 11) of the nonparametric bounds in Figure 9; these shift $V(1)$ to either 0.77 or 0.42 but cannot reverse the comparison of free stamps to cost-sharing. And in all of these cases, $W(1)$ remains above 1.

However, we can reverse the conclusion that free food stamps reduce targeting if we assume compliers have valuations V_c at their upper bound, and always-takers' valuation at their lower bound (Row 12). This would say that all compliers and all always-takers are just indifferent, but in the opposite direction.[59]

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 $V(1)$ and $W(1)$. In Rows 15 and 16, we consider the smallest and largest first stages that we estimate in Appendix Figure A18 (either including state-by-year fixed effects or interacting the log of the female population age 15–44 with calendar year), which moves $V(1)$ from our baseline estimate of 0.59 to 0.63 or 0.59; naturally, $W(1)$ is therefore largely unchanged.

Fiscal Externalities (FE). Our baseline analysis assumed zero fiscal externalities. Positive fiscal externalities from earnings responses must be small because taxes on low-income households were only about 12.9% (at least, under the assumptions in Hendren and Sprung-Keyser, 2020). Empirically, we found little evidence that food stamps affect employment in Appendix C.4, which would imply small fiscal externalities. If we assign both parties a conservatively large fiscal externality of 16 cents (which instead comes from the earnings impact in Hendren and Sprung-Keyser, 2020), that still cannot flip the comparison of W_a to W_c (Row 17).[60] Row 18 shows that if we assign *only* always-takers the fiscal externality of 16 cents, we still get that the increase in subsidy reduces targeting. Row 19 shows that, if compliers have zero FE, the always-taker FE would need to be at least \$1.07 per dollar of spending to overturn our targeting result. Moreover, even in this case, we continue to find $W(1) > 1$.

C.5.2 Potential Policy Confounds

Did the Policy Coincide with Reductions in Ordeals? A key assumption for the normative analysis is that the free-stamps reform did not change the hassles associated with the food stamp program. Note that this assumption is not needed for the positive analysis of the impact of food stamps on birth weight, since we can think of any coincident changes as part of the policy bundle that triggered an increase in enrollment (the first stage).

[59] 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.

[60] 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 65 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, or \$916). Moreover, the final fiscal externality in Hendren and Sprung-Keyser (2020) incorporating other longer-run forces is smaller than this, at 5 cents.

The reform did not meaningfully change ordeals associated with application or take-up. As discussed in Appendix A.1.3, the most concerning threat to our normative interpretation is that the free stamps legislation clarified that local jurisdictions could offer mail issuance of coupons to all enrollees. That said, elimination of the purchase requirement may have made it easier for enrollees to do mail issuance of coupons, since now enrollees need not mail in their purchase requirements.

In practice, however, several pieces of evidence suggest little quantitative impact of any confounds from this. First, by 1981, less than one quarter of food stamp recipients could have received their stamps through direct mail issuance.[61] Second, a 1981 Report to Congress on the free food stamp reform concluded, based on qualitative inferences, that perhaps 10% of the enrollment increase from the reform owed to changes in ordeals (U.S. Department of Agriculture, 1981b). Third, at the state level, our detailed historical searches did not find evidence that a large share of states expanded mail applications or issuance exactly in 1979. This suggests that the stark increase in enrollment exactly in 1979 in higher- relative to lower-exposure counties reflects the removal of the purchase requirement rather than changes in ordeals.

In fact, searching state-by-state, we could only find two state-level reforms that occurred in 1979 on mail issuance of food stamp coupons, and one reform that occurred in 1981: in Maine in 1979 (this state was already dropped due to lack of county-level enrollment data), Iowa in 1979 (just 1.94% of the sample), and Ohio in 1981 (7.19% of the sample, see discussion below). Moreover, any reforms that happened at the state level would be captured by state-by-year fixed effects. Our first-stage estimates attenuate but remain large and significant in this specification (Appendix Figure A20).

Fourth, we implemented two case studies to gauge the potential importance of changes in mail issuance. First, at the county level, mail issuance should be more important in rural counties where potential recipients lived farther from food stamp offices than those in urban counties. However, when we restrict the analysis of enrollment to the 71% of the population that lived in urban counties, we see, if anything, a larger first stage (Appendix Figure A17a).

Second, we could identify the exact timing of implementation of mail issuance of food stamp coupons in Ohio (it occurred in 1981) (Ohio General Assembly, 1980). We therefore estimate our first stage within Ohio (Appendix Figure A17b). Here again, we see a large jump in 1979, coincident with the (known to be uncontaminated) free-stamps reform. The impact of mail issuance is not clearly visually distinguishable in 1981, although enrollment increases slightly in 1981 versus 1980. Looking only at the average impact through 1980, and rescaling by the average exposure in Ohio, the average enrollment impact is 0.71 percentage points (approximately $0.2305 \times 0.031 \times 100$).

Changes in Eligibility. Although the elimination of the purchase requirement was the centerpiece of the free food stamp reform, the reform also changed some eligibility rules, which may have reduced the share eligible, particularly for higher incomes (see Appendix A.1.3). However, accounting for potential unmodeled eligibility reductions would actually strengthen the normative case for cost-

[61] Specifically, 75% of benefits were delivered via ATP mail cards, which still required turning that card in for stamps at an office (U.S. General Accounting Office, 1982). Moreover, some project areas still used strictly in-person delivery of stamps but not ATP mail cards, so the 75% is a weak lower bound on the share that used in-person delivery of stamps.

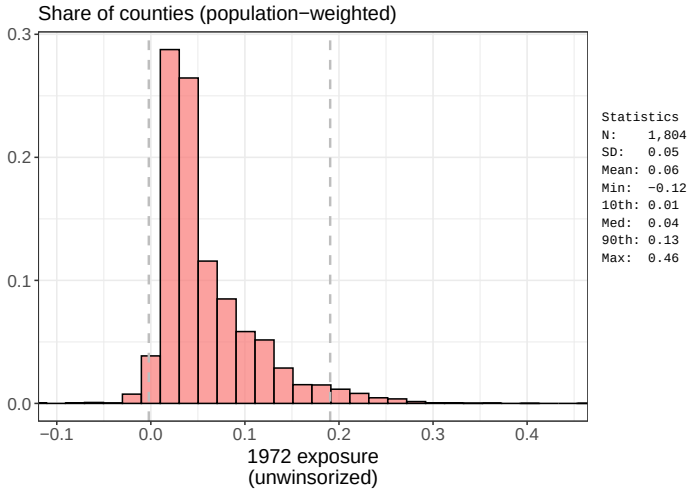
sharing.[62] If eligibility restrictions reduced the share of eligible enrollees but did not affect the characteristics of eligible individuals, this would imply that our estimated first-stage enrollment response underestimates the enrollment response to an increase in subsidy. In other words, the increase 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 13 and 14). In practice, the eligibility restrictions actually had more bite at higher incomes and thus reduced the average income of eligible households; 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 $V(1)$.

[62] 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.

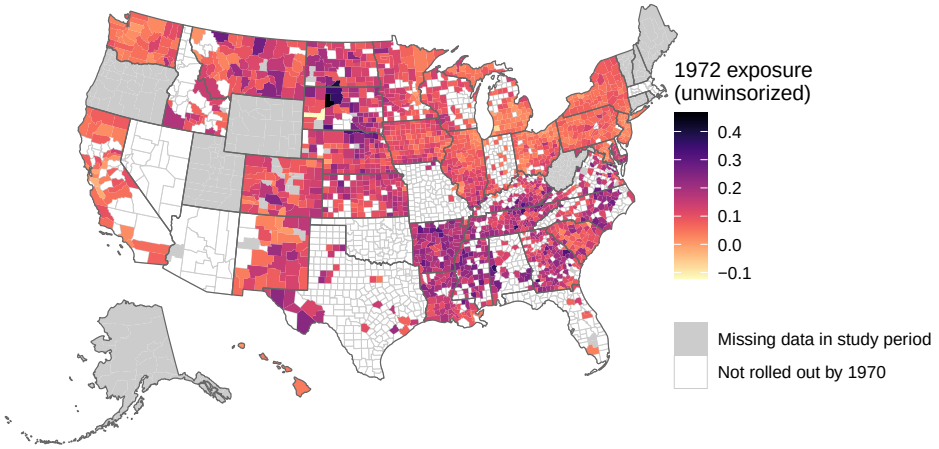
D Appendix Figures and Tables

Figure A1: Variation in Exposure (Unwinsorized)

(a)

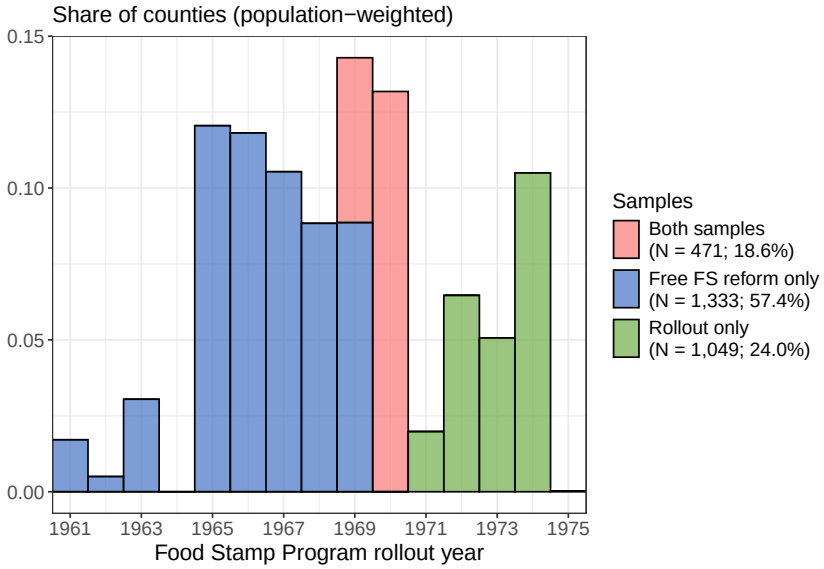


(b)



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., unwinsorized) exposure measure. Our baseline exposure measure (e.g., Figure 2) is winsorized 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 is shown in the figure legend. The analysis sample for the rollout analysis includes counties that rolled out in July 1969 or later.

Figure A3: NCHS Table of Last Menstrual Period (LMP) Reporting

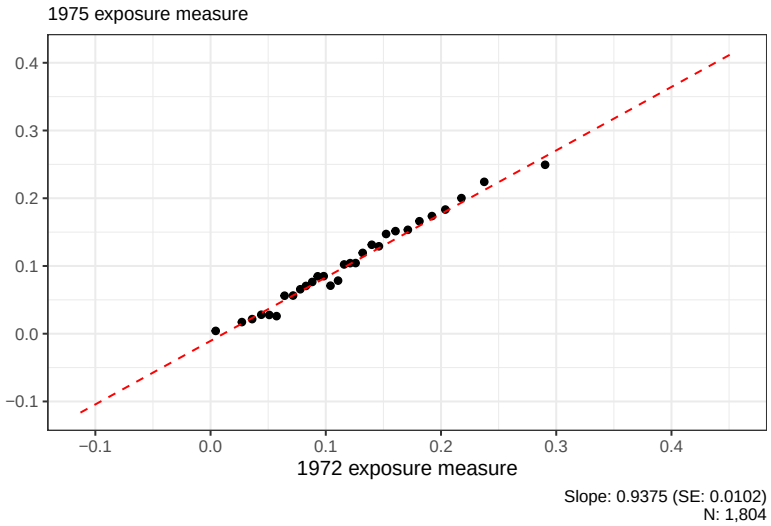
Appendix table. States reporting date last normal menstrual period began: United States, 1968-78

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

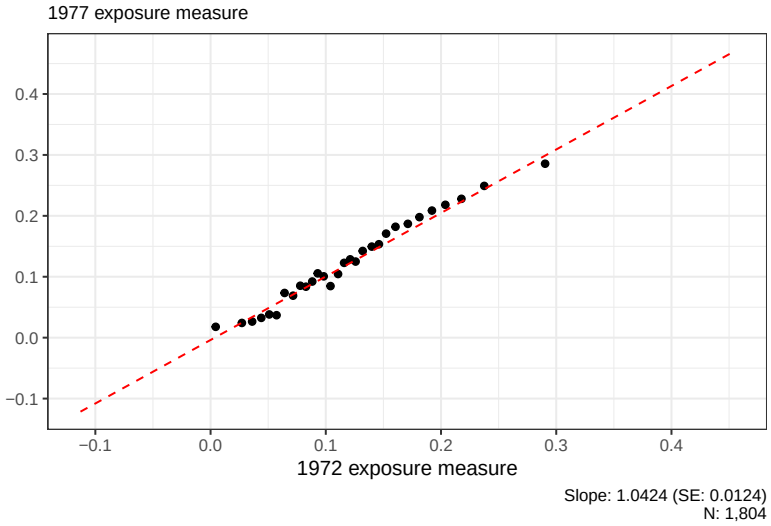
Notes: Table denotes, for each state, which years report LMP for gestation period calculation. Note that (not reported in this table) Connecticut began reporting LMP in 1981, Texas by 1980, and New Mexico did not report LMP in the years included in the analysis (U.S. Department of Health and Human Services,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

Figure A4: Correlation between Year-specific Exposure Measures

(a)

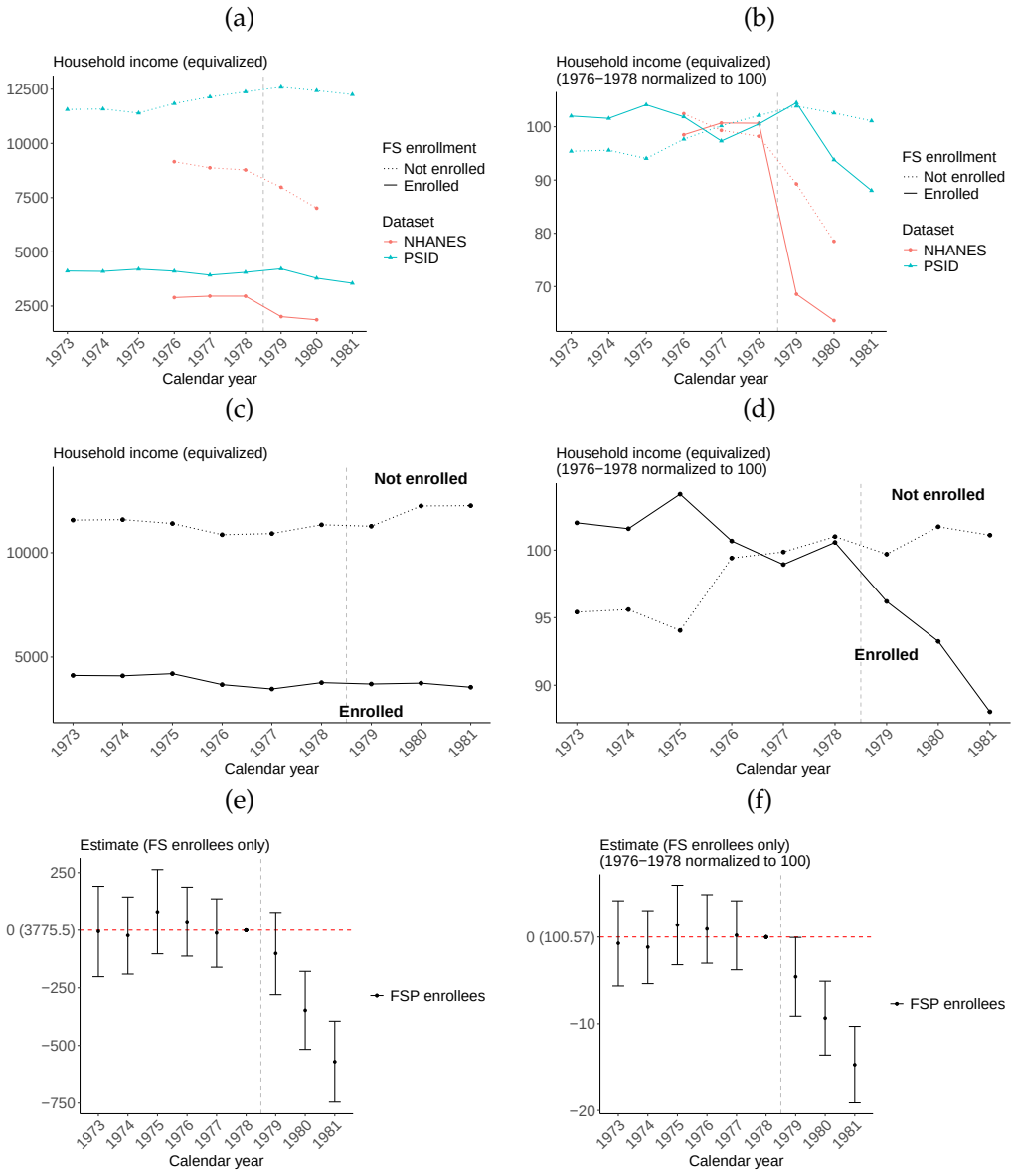


(b)



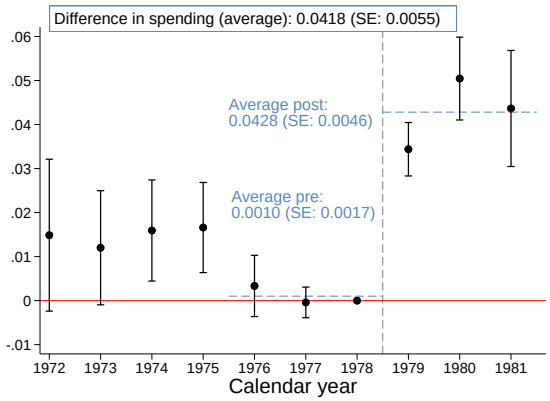
Notes: Figure shows a binned scatterplot of alternative year-specific exposure measures against the 1972 exposure measure, as described in Equation 3 and depicted in Appendix Figure A1, weighted by 1970 county population. Panel (a) describes the correlation of the 1972 exposure measure with the 1975 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 A5: 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 40 with the error bars indicating the 95% confidence interval (from standard errors clustered at the 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 A6: Impact of Free-Stamps Reform on Food Stamp Spending per Capita

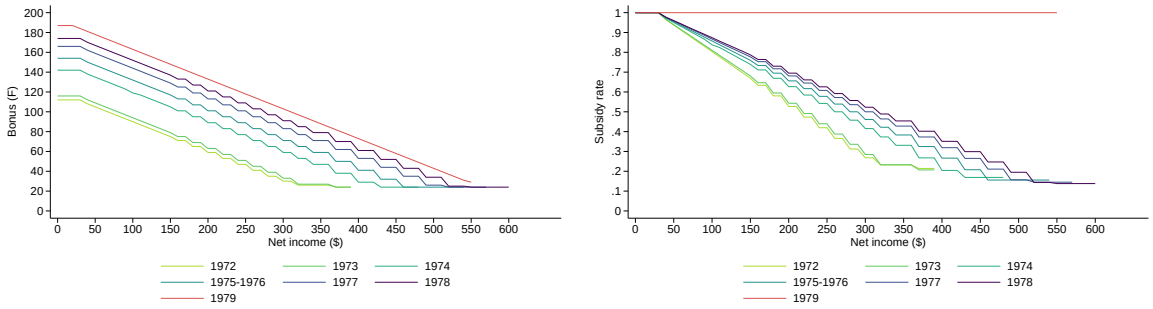


Notes: Figure shows the results of estimating Equation 4. Dependent variable is county-year food stamp spending per capita (in thousands of 1970 dollars). N = 1,774 counties. Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level.

Figure A7: Benefit Schedules: Transfers and Subsidy Rate

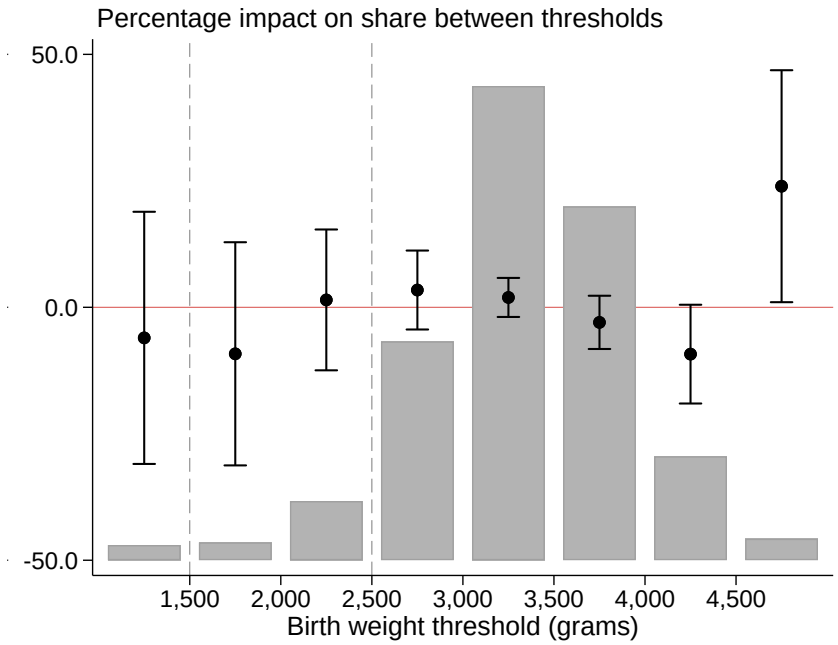
(a) Food Stamp Bonus (F)

(b) Food Stamp Subsidy Rate ($F/(R+F)$)



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 and $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 U.S. Department of Agriculture, Food and Nutrition Service (1971a, 1972, 1973, 1974, 1976, 1977, 1978a)), supplemented by U.S. Department of Agriculture, 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 schedule specifies benefits in $3-wide monthly net income bins (instead of $10-wide).

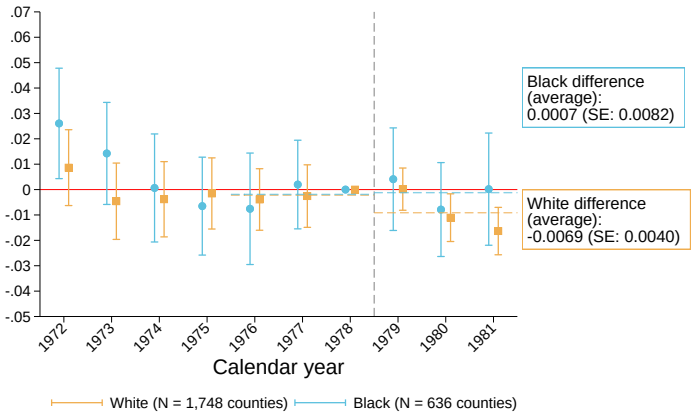
Figure A8: Free Food Stamp Reform: Impacts on Share of Births Within Birth Weight Ranges



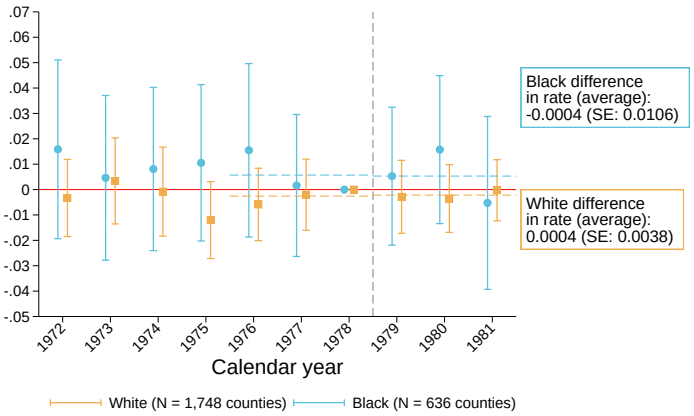
Notes: Figure shows scaled estimates of β_{avg} (as defined in Equation 6) from estimating Equation 5 where the outcome is the share of births that fall within a specified number of grams. Because the share of births in 1976–1978 in various birth weight ranges varies greatly (as shown in the light gray 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 ($<$ 1,500 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 A9: Impacts on Infant Health Outcomes by Race, Free-Stamps Reform

(a) Free Food Stamp Reform
Log Birth Weight



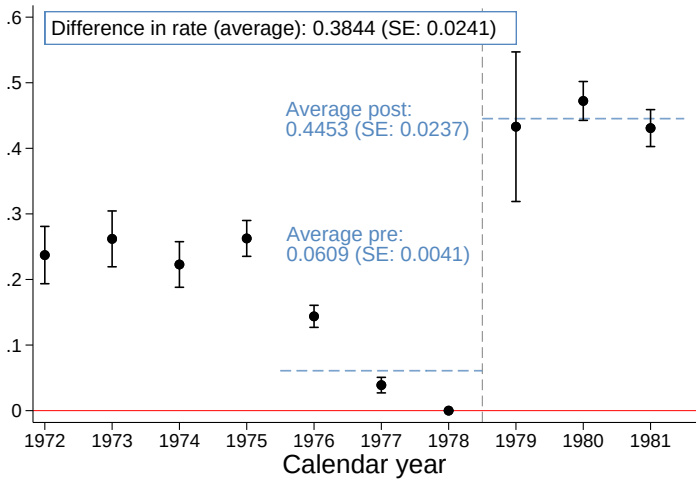
(b) Free Food Stamp Reform
Share Low Birth Weight



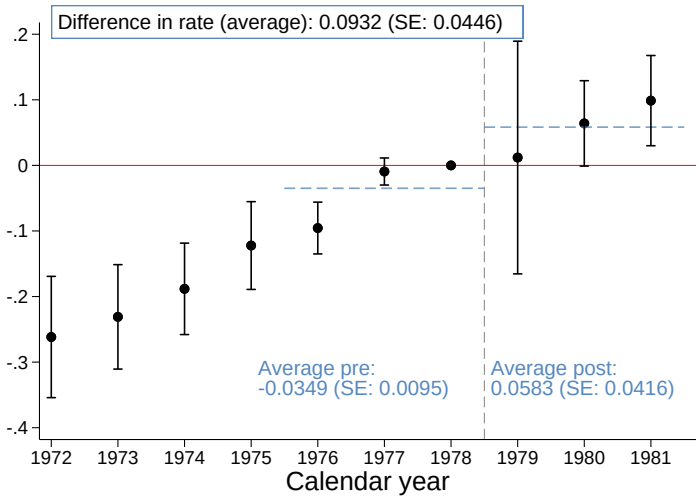
Notes: Panels (a) and (b) show the results of estimating the reduced-form Equation 5 using a continuous 1972 exposure measure on infant birth weight outcomes, for White and Black births separately. Regressions for Panels (a) and (b) are weighted by 1970 county-race population, and standard errors are clustered at the county level. 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-stamps reform, 1970–1975 for rollout). Vertical lines denote 95% confidence intervals computed using standard errors clustered at the county level. Sample size, in terms of counties, is denoted in each panel legend.

Figure A10: Impact of Free Food Stamp Reform on Enrollment, with Additional Covariates

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

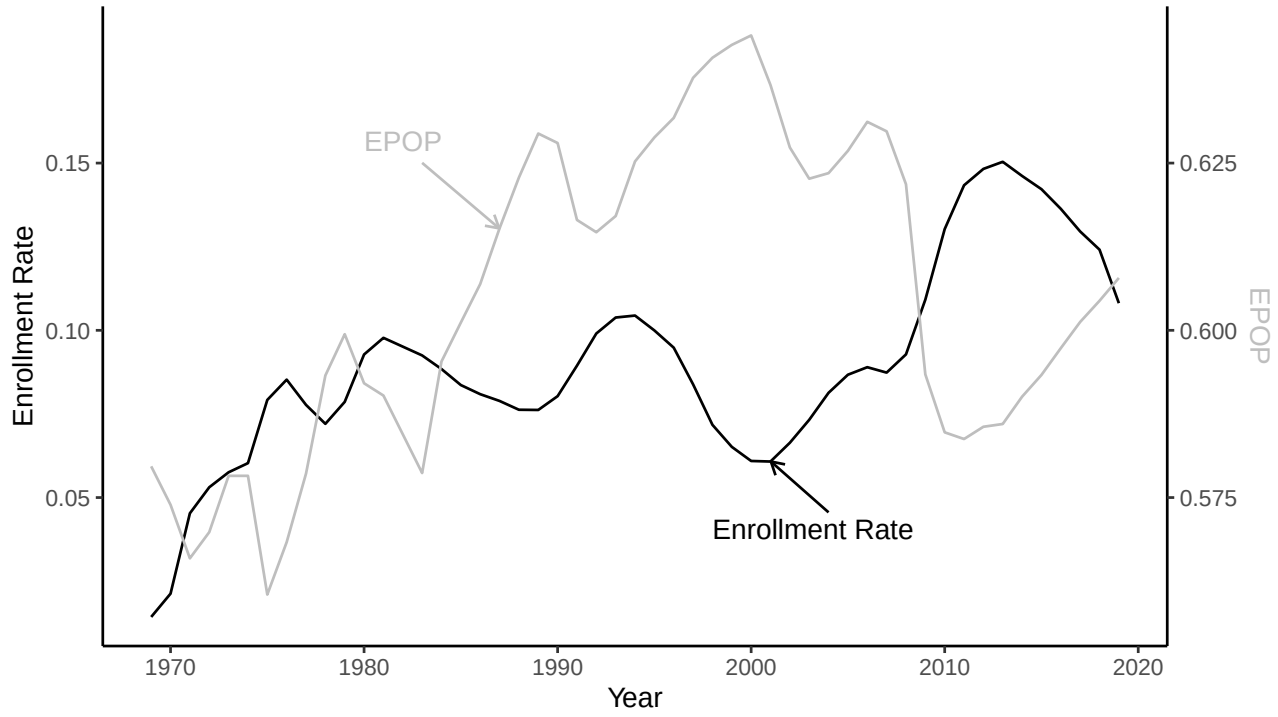


(b) County's Poverty Rate Decile × Year Fixed Effects



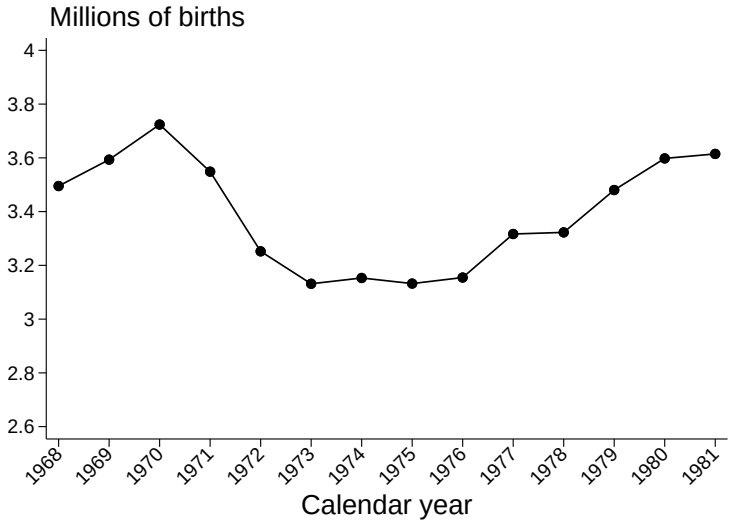
Notes: Figure shows the results of estimating Equation 4 with additional controls. Panel (a) adds controls for the county's take-up rate decile interacted with year fixed effects. The take-up rate is defined as the number of individuals enrolled as a share of the number of individuals eligible to enroll; 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 A11: 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 come from the USDA (USDA Food and Nutrition Service, 2025), data on U.S. population come from the SEER database (National Cancer Institute, 2026), and data on EPOP come 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.

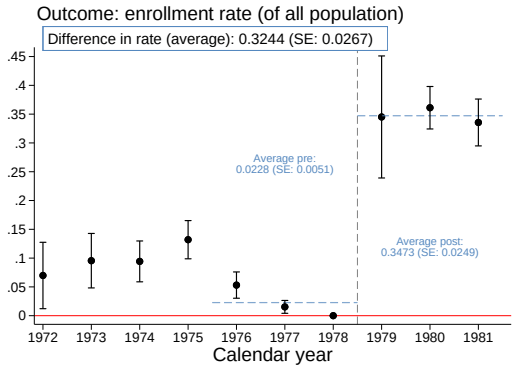
Figure A12: National Trend in Total Births



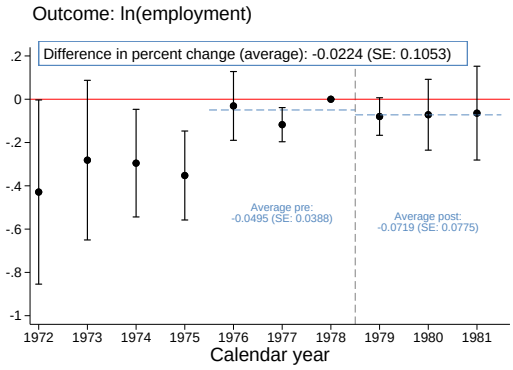
Notes: Figure shows the number of births nationally (across all states except Alaska).

Figure A13: Free-Stamps Reform Impacts on Imputed Employment

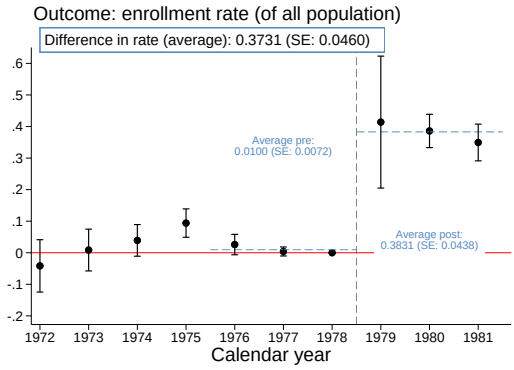
(a) First Stage: Impact on Food Stamp Program Enrollment (All Industries)



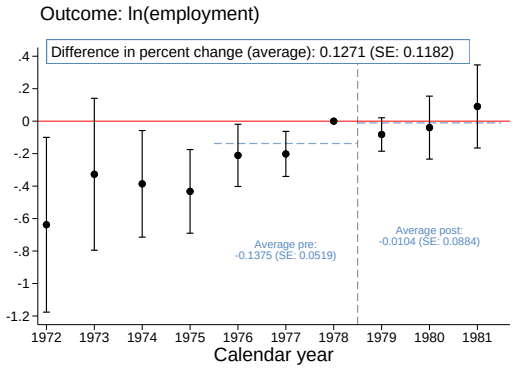
(b) Reduced Form: Impact on Log Employment (All Industries)



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



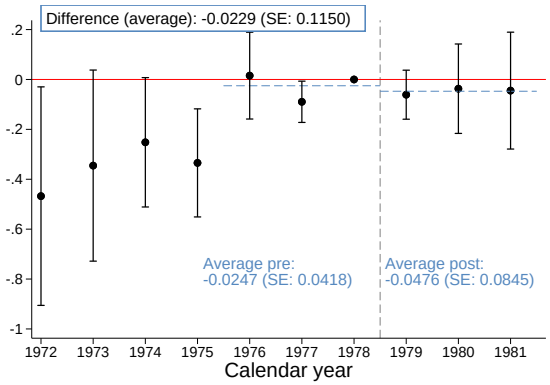
(d) Reduced Form: Impact on Log Employment (Low-wage Industry Subset)



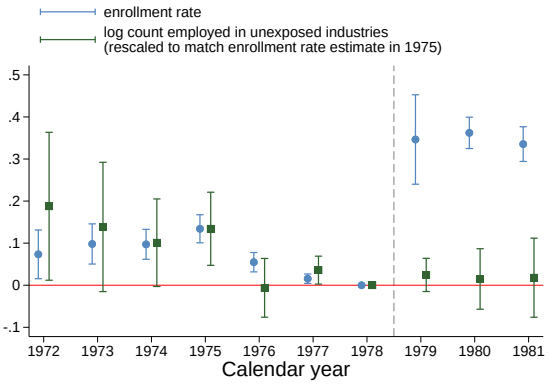
Notes: Panels (a) and (c) show the results of estimating the first-stage Equation 4. Panels (b) and (d) show the results of estimating the reduced-form Equation 5. Both estimations use a continuous 1972 exposure measure. The log employment outcome comes from the imputed County Business Patterns datasets 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 the food stamp program by 1970 and that have both food stamp program enrollment data and nonzero 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 A14: Impact of Free Food Stamp Reform on Enrollment, Adjusting for Economic Confounds

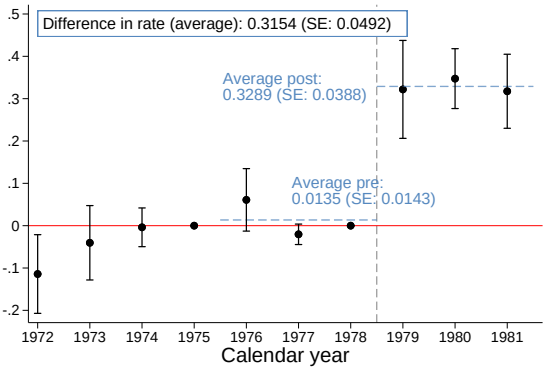
(a) Log Employment Count in Unexposed Industries



(b) Overlaying Enrollment Rate and Rescaled Employment

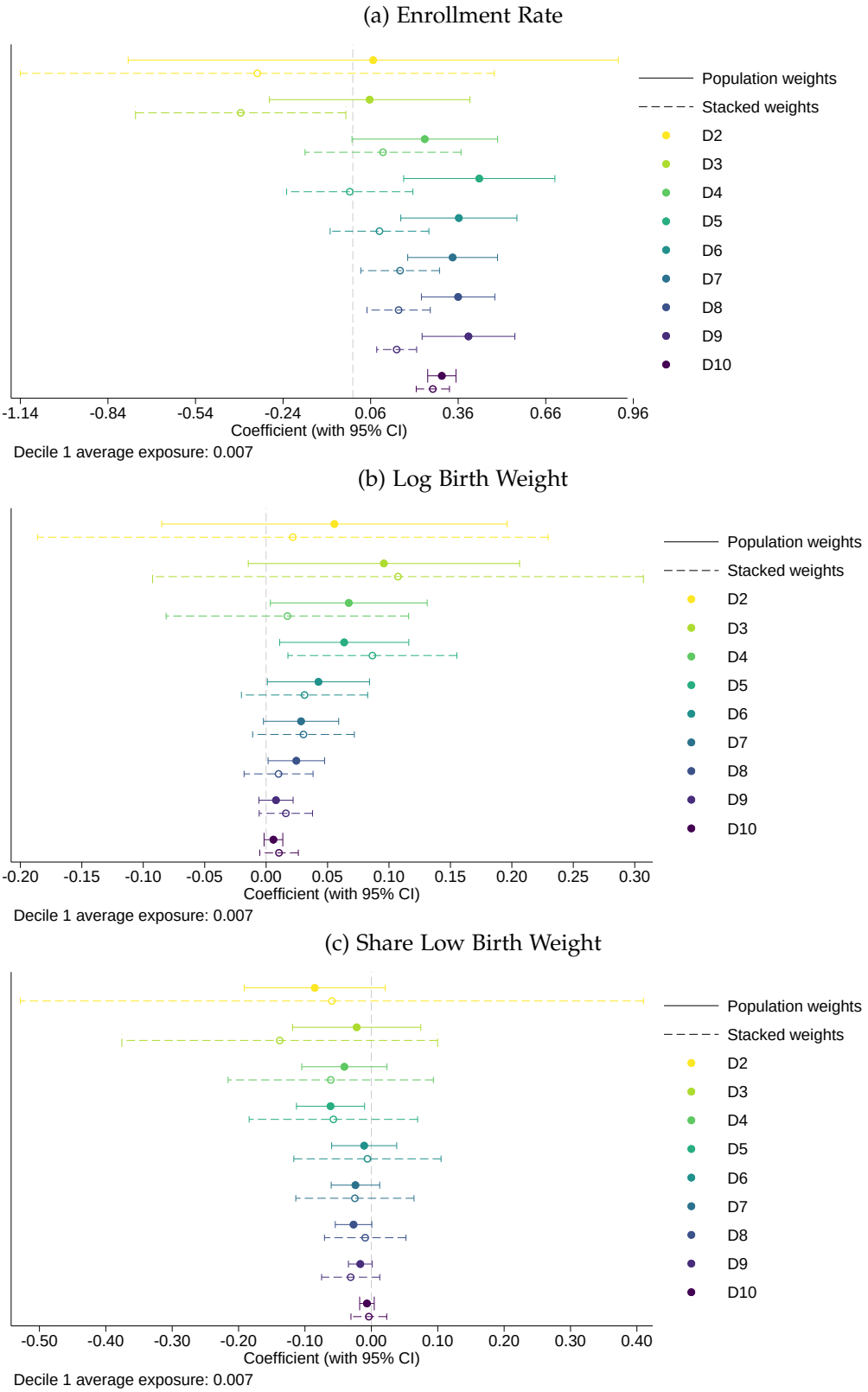


(c) Enrollment Rate



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 1975 for the first stage. Panel (c) shows the results of estimating Equation 21 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 22. 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.

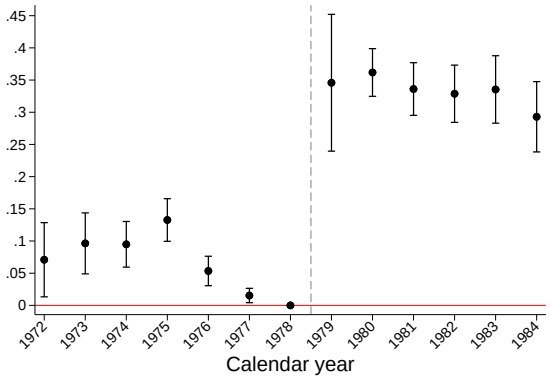
Figure A15: Impact of Free-Stamps Reform Across Deciles of Exposure



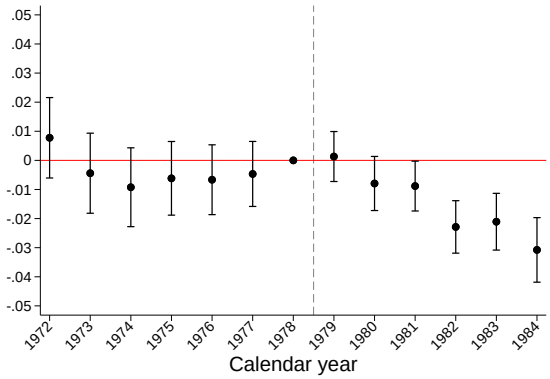
Notes: Figure compiles the 3-year difference coefficients (as in Equation 6) and the 95% CIs from the stacked-decile specification for each decile individually, where the comparison group is the lowest decile of exposure, Decile 1. The results display estimates from Equations 23 and 24 on a dataset stacking comparisons between each Decile 2 through 10 and Decile 1, using 1970 county population weights or dataset weights. Standard errors are clustered at the county level.

Figure A16: Percentage Impacts of Later Post-reform Period on Share Within Birth Weight Ranges

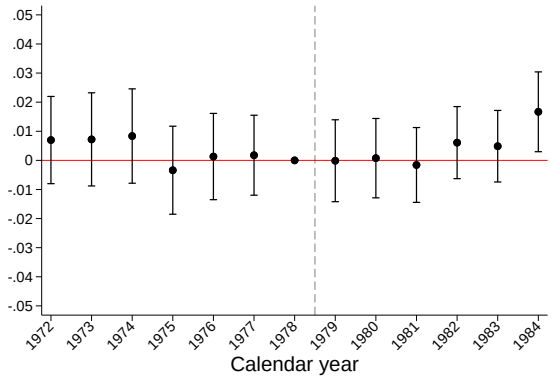
(a) Food Stamp Program Enrollment



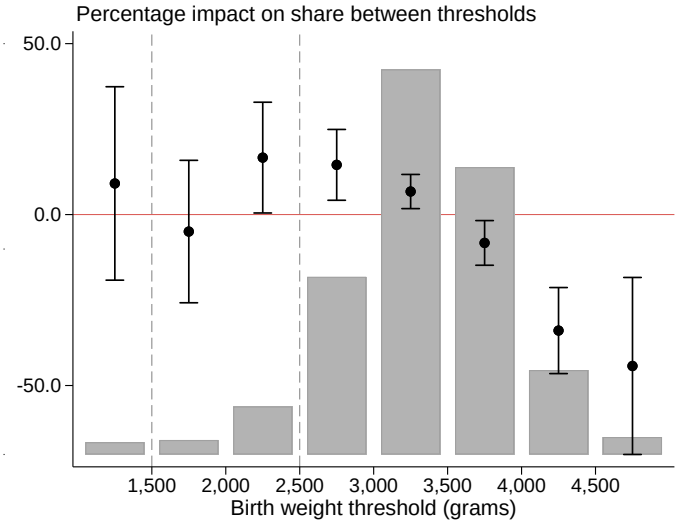
(b) Log Birth Weight



(c) Share Low Birth Weight



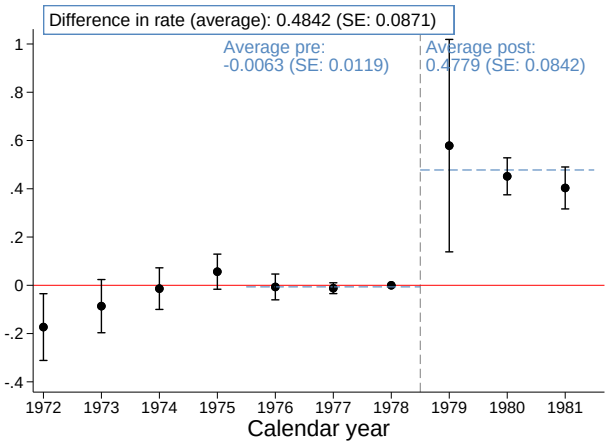
(d) Percentage Impacts of Later Post-reform Period on Share Within Birth Weight Ranges



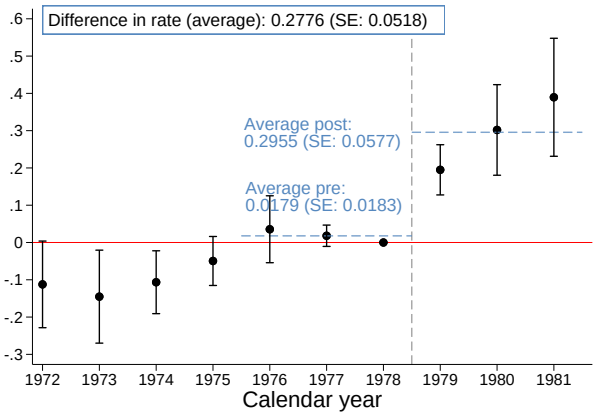
Notes: Panels (a), (b), and (c) show the results of estimating Equations 4 and 5 on enrollment and birth weight outcomes using a continuous 1972 exposure measure for an extended post-reform period. Panel (d) shows the scaled results of estimating the reduced-form Equation 5 using a continuous 1972 exposure 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 "$\geq$ 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 ($<$ 1,500 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 A17: Impact of Free-Stamps Reform: Case Studies

(a) Urban Counties

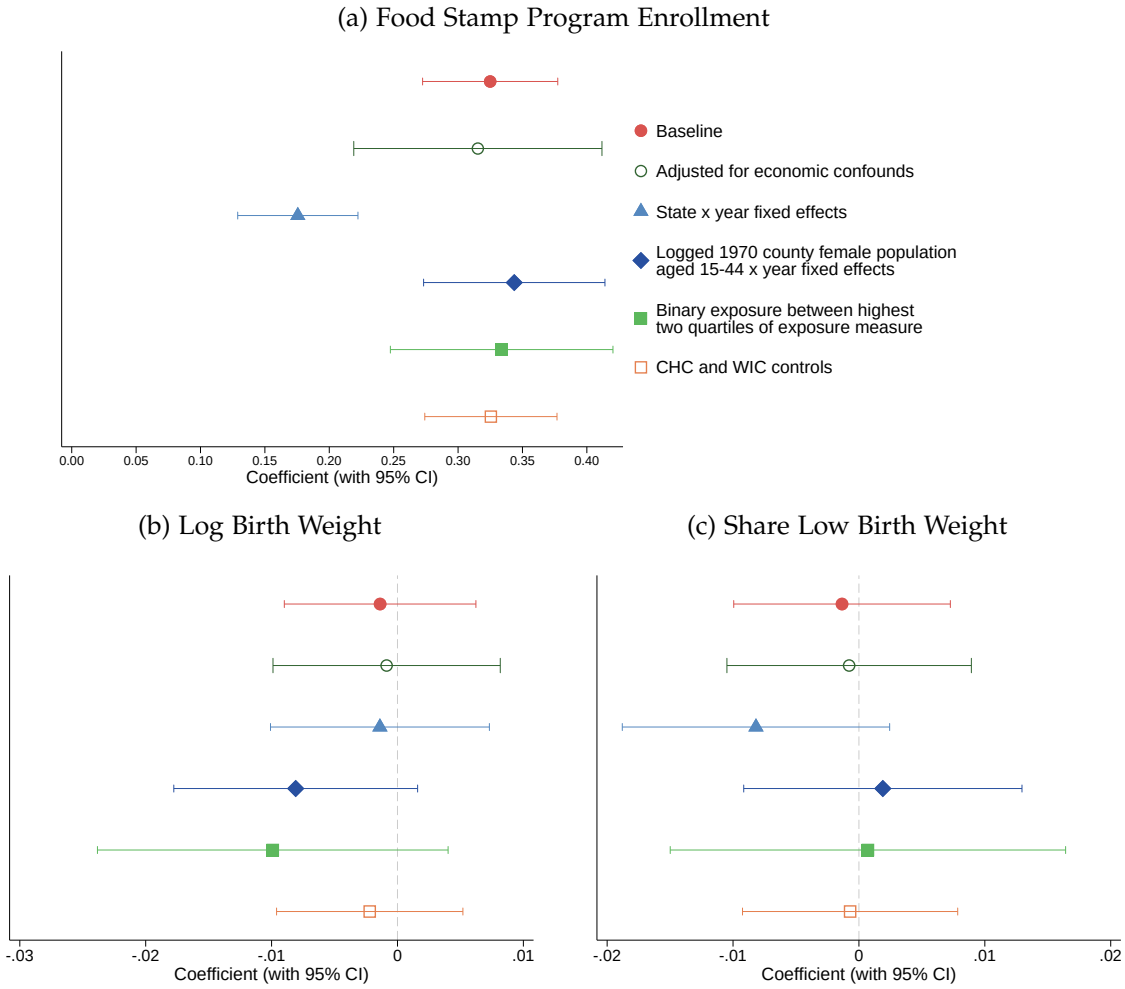


(b) Ohio



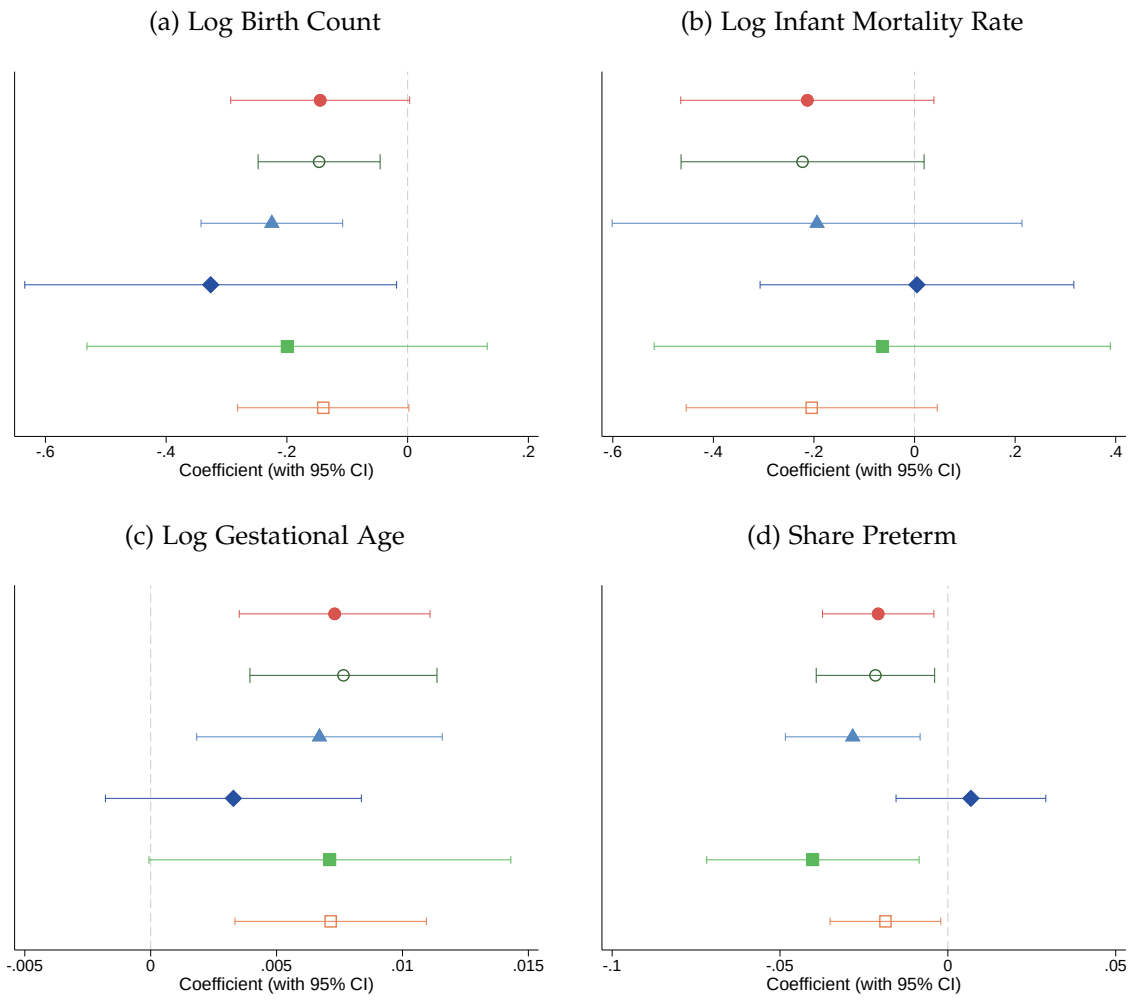
Notes: Figure shows the results of estimating Equation 4 for urban counties in Panel (a) where N = 289 counties, and for Ohio counties in Panel (b) where N = 80 counties. Regressions are all weighted by 1970 county-level population. Standard errors are clustered at the county level.

Figure A18: Free Food Stamp Reform: Impacts Across Specifications



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 five 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 on food stamps as a proxy for the local business cycle (see Appendix B.2). The "state × year" and "Logged 1970 county female population aged 15–44 × year" specifications add those respective fixed effects, and "CHC and WIC controls" add two separate simple controls for county-years with each program rolled out. 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 A20 through A22 present the underlying event studies for all of these specifications.

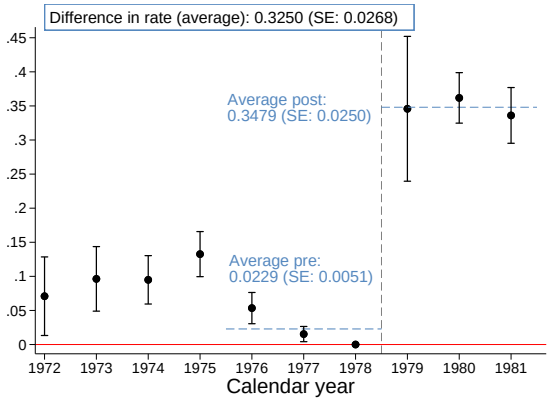
Figure A19: Free Food Stamp Reform: Impacts on Additional Infant Health Outcomes



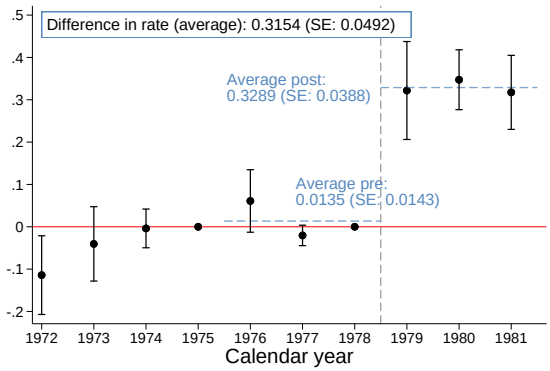
Notes: Figure presents estimates of β_{avg} (as defined in Equation 6) from estimating Equation 5 for the baseline specification and five alternative specifications. The "adjusted for economic confounds" specification implements the proxy approach for economic confounds from Freyaldenhoven et al. (2019), using county-year data on log employment counts in industries unlikely to have many workers on food stamps as a proxy for the local business cycle (see Appendix B.2). The "state × year" and "Logged 1970 county female population aged 15–44 × year" specifications add those respective fixed effects, and "CHC and WIC controls" add two separate simple controls for county-years with each program rolled out. 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 A23 through A26 present the underlying event studies.

Figure A20: Impact of Free-Stamps Reform on Enrollment Rate: Alternative Specifications

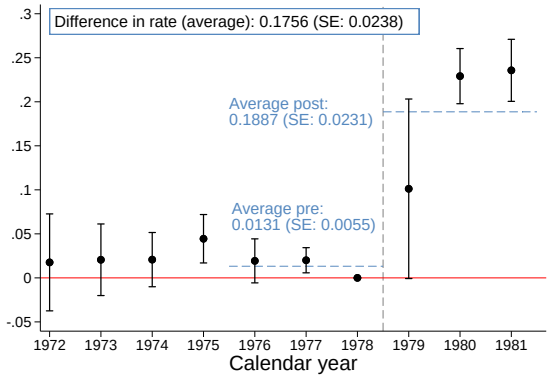
(a) Baseline



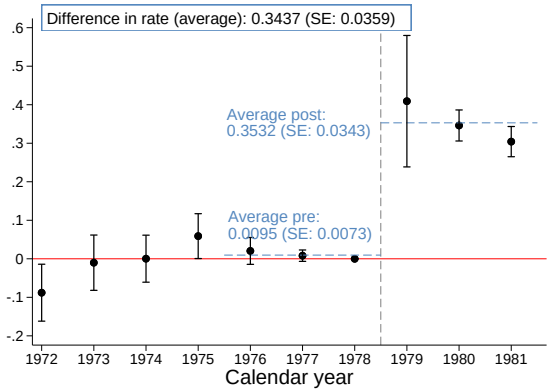
(b) Adjusted for Confounds Using Unexposed Employment



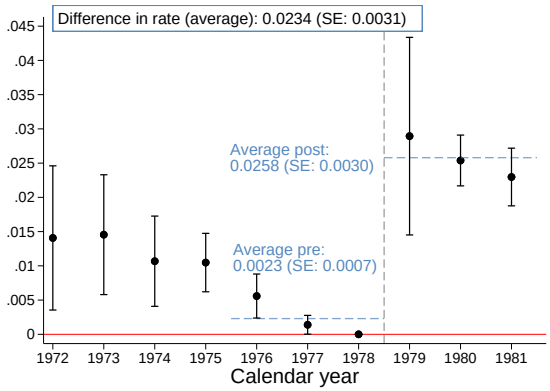
(c) State × Year Fixed Effects



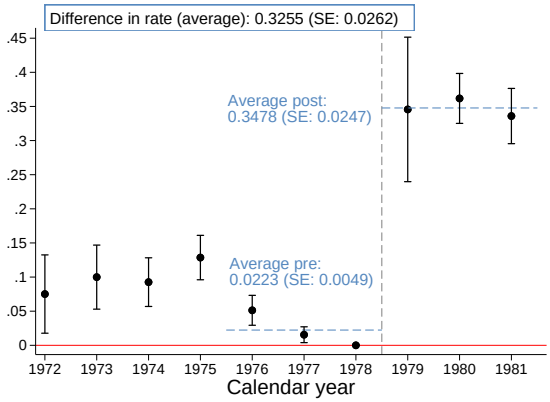
(d) Log Population Women Aged 15–44 × Year Fixed Effects



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



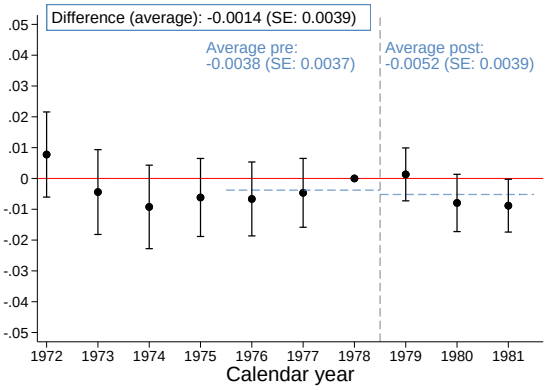
(f) CHC and WIC Controls



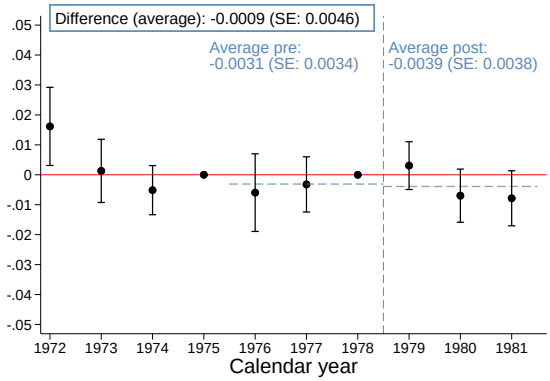
Notes: Figure shows the results of estimating Equation 4 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A18 for more details on each alternative specification. N = 1,804 counties for the baseline sample in Panel (a) and Panel (f). 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 A21: Impact of Free-Stamps Reform on Log Birth Weight: Alternative Specifications

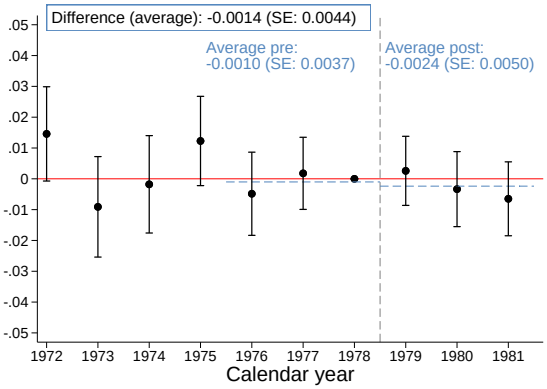
(a) Baseline



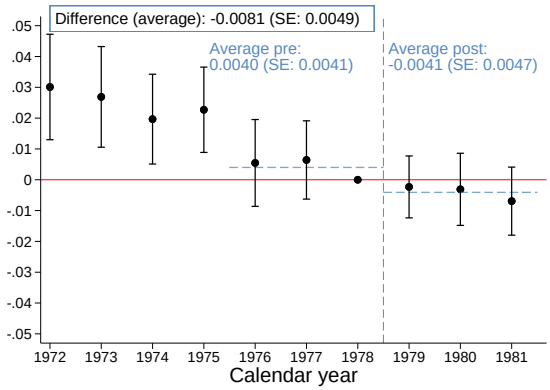
(b) Adjusted for Confounds Using Unexposed Employment



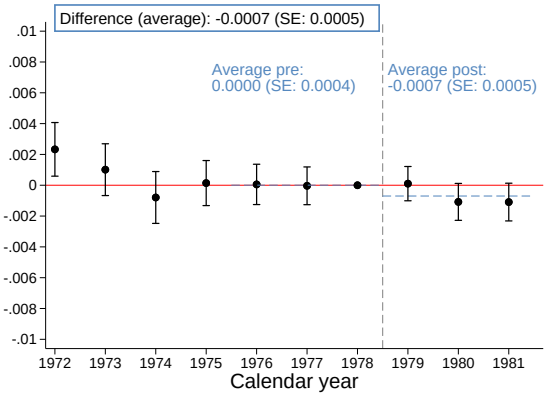
(c) State × Year Fixed Effects



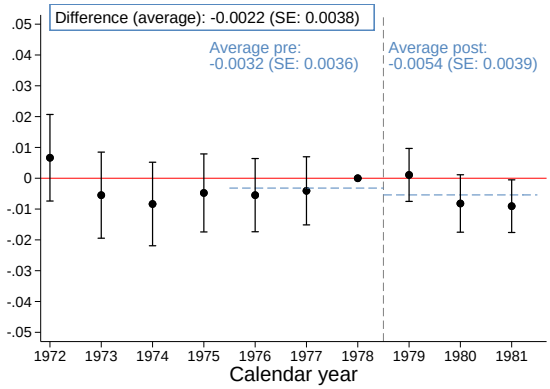
(d) Log Population Women Aged 15–44 × Year Fixed Effects



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



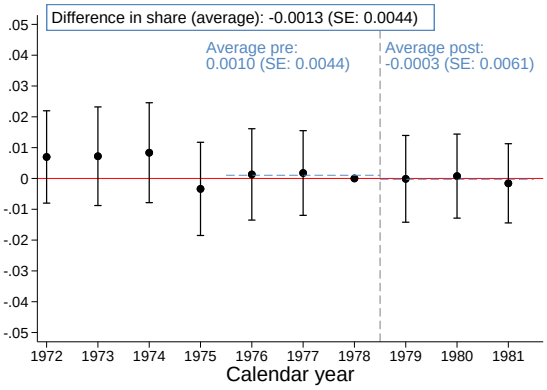
(f) CHC and WIC Controls



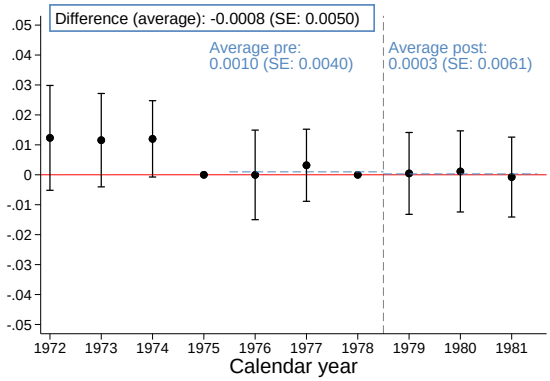
Notes: Figure shows the results of estimating Equation 5 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A18 for more details on each alternative specification. N = 1,804 counties for the baseline sample in Panel (a) and Panel (f). 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 A22: Impact of Free-Stamps Reform on Share Low Birth Weight: Alternative Specifications

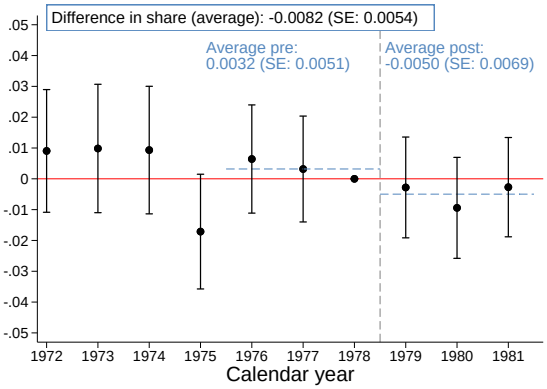
(a) Baseline



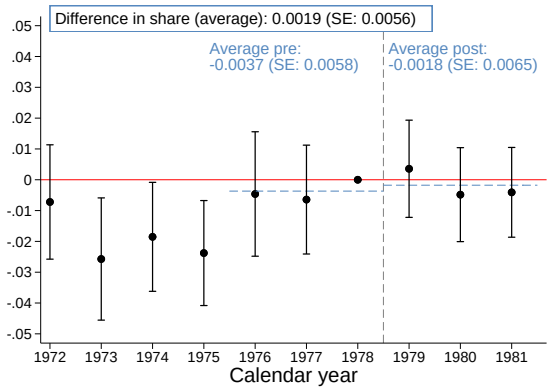
(b) Adjusted for Confounds Using Unexposed Employment



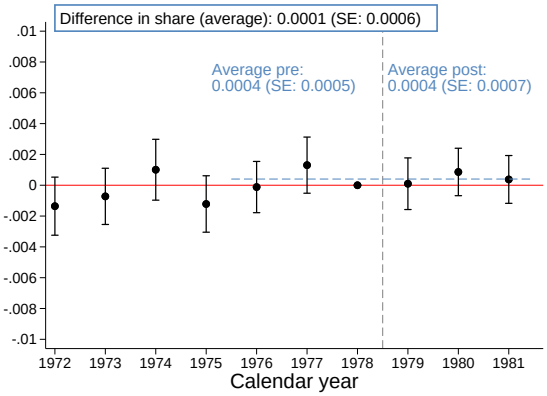
(c) State × Year Fixed Effects



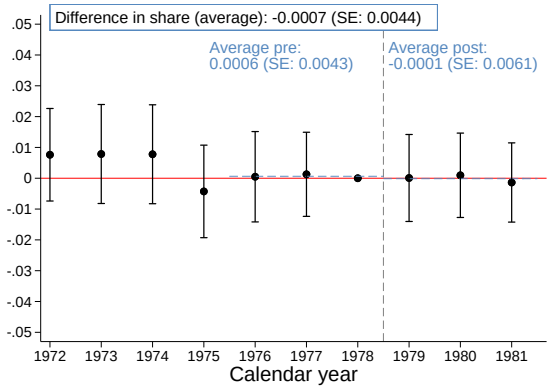
(d) Log Population Women Aged 15–44 × Year Fixed Effects



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



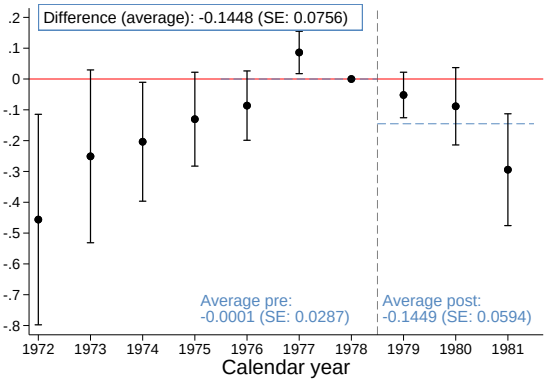
(f) CHC and WIC Controls



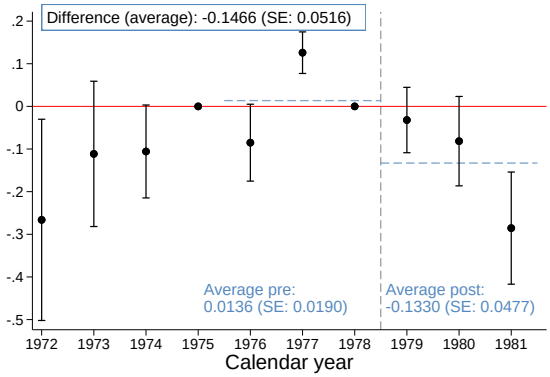
Notes: Figure shows the results of estimating Equation 5 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A18 for more details on each alternative specification. N = 1,804 counties for the baseline sample in Panel (a) and Panel (f). 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 A23: Impact of Free-Stamps Reform on Log Birth Count: Alternative Specifications

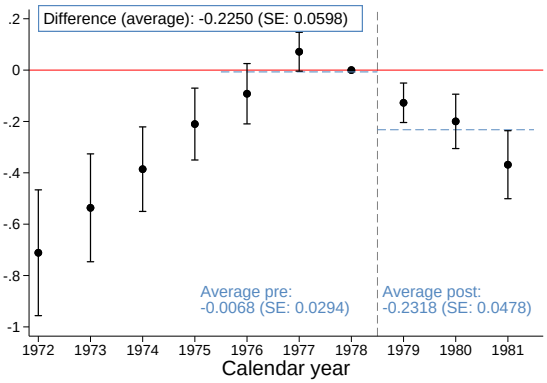
(a) Baseline



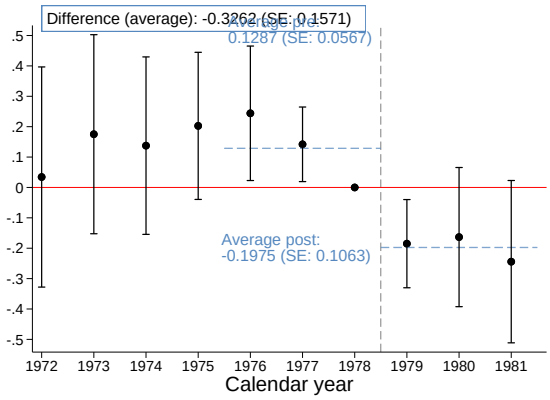
(b) Adjusted for Confounds Using Unexposed Employment



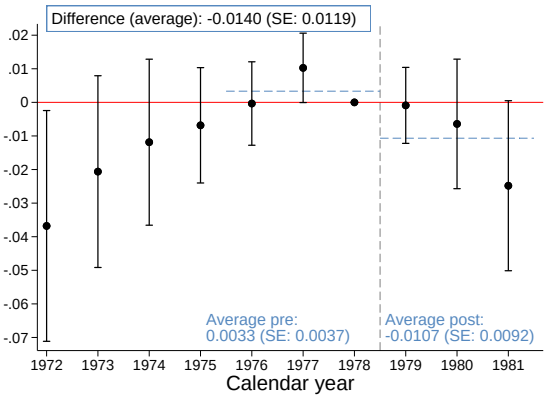
(c) State × Year Fixed Effects



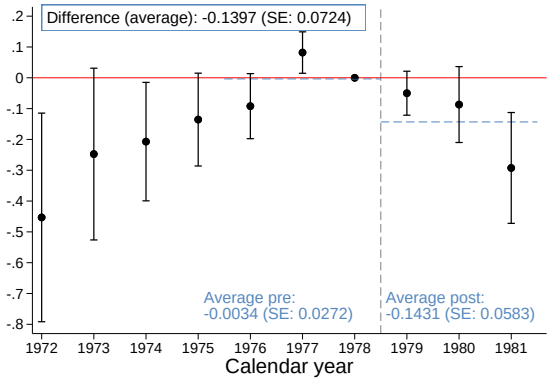
(d) Log Population Women Aged 15–44 × Year Fixed Effects



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



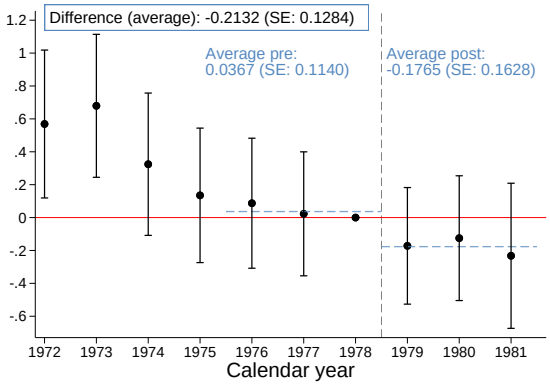
(f) CHC and WIC Controls



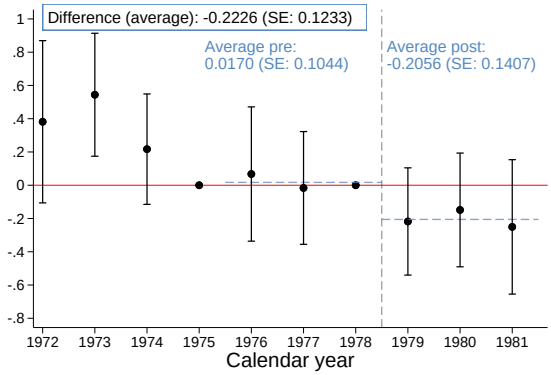
Notes: Figure shows the results of estimating Equation 5 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A18 for more details on each alternative specification. N = 1,804 counties for the baseline sample in Panel (a) and Panel (f). 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 A24: Impact of Free-Stamps Reform on Log Infant Mortality Rate (Deaths per 1,000 Births): Alternative Specifications

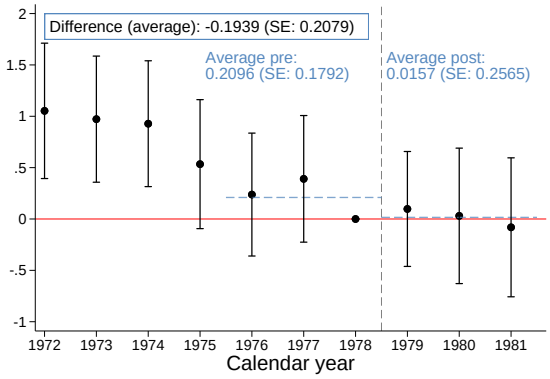
(a) Baseline



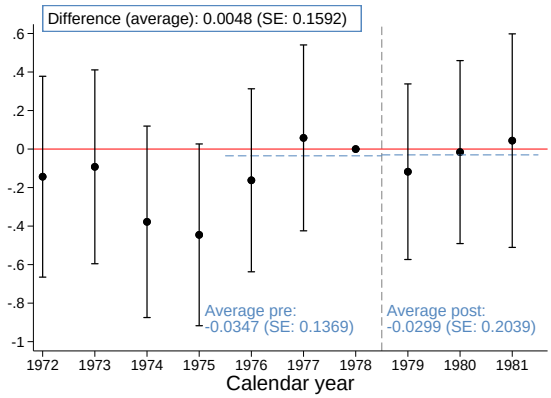
(b) Adjusted for Confounds Using Unexposed Employment



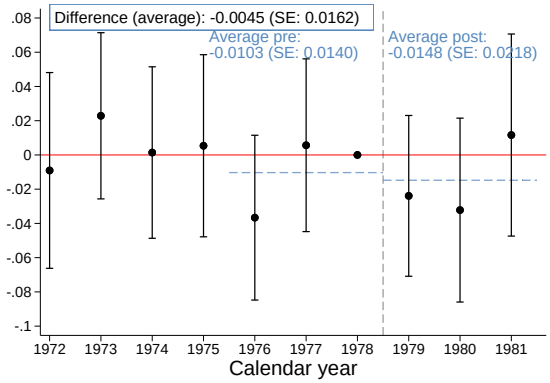
(c) State × Year Fixed Effects



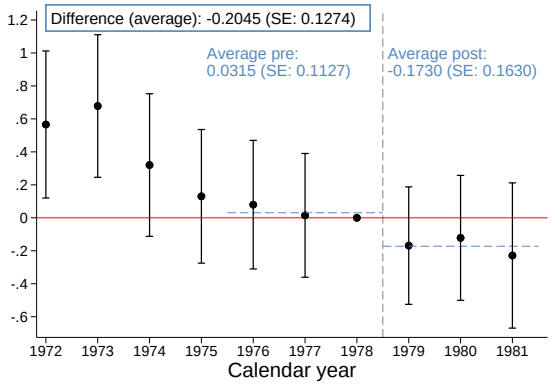
(d) Log Population Women Aged 15–44 × Year Fixed Effects



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



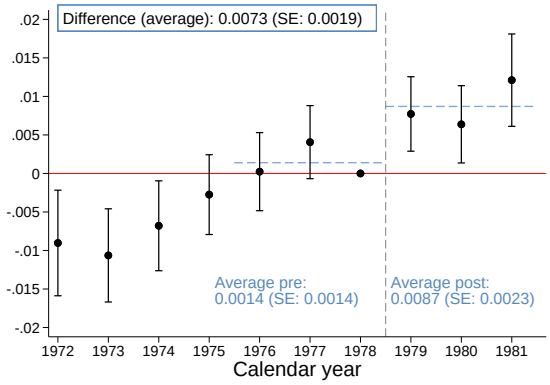
(f) CHC and WIC Controls



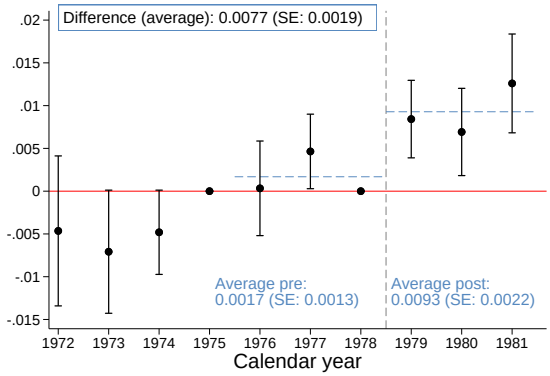
Notes: Figure shows the results of estimating Equation 5 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A18 for more details on each alternative specification. See Appendix B.1 for outcome-specific limitations applied to the free food stamp reform baseline sample. N = 1,145 counties for the baseline estimates in Panel (a) and Panel (f). 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 A25: Impact of Free-Stamps Reform on Log Gestational Age: Alternative Specifications

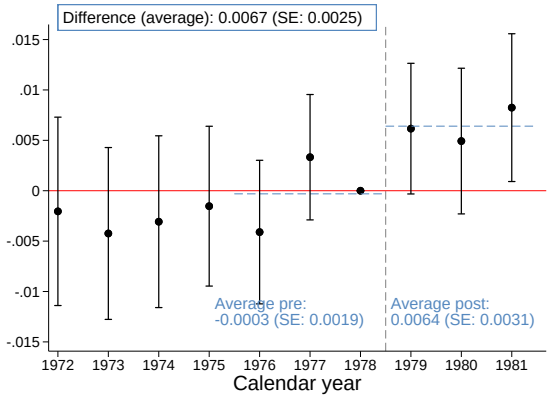
(a) Baseline



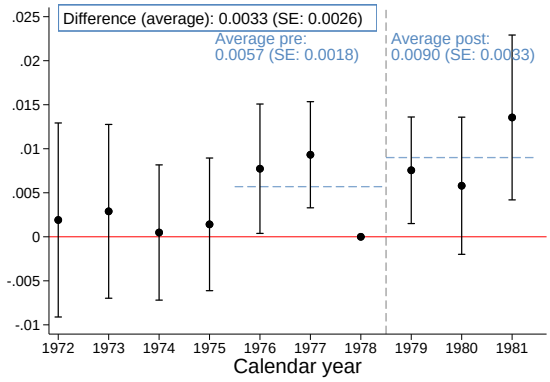
(b) Adjusted for Confounds Using Unexposed Employment



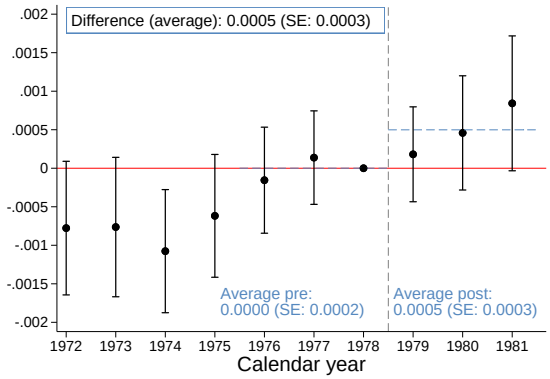
(c) State × Year Fixed Effects



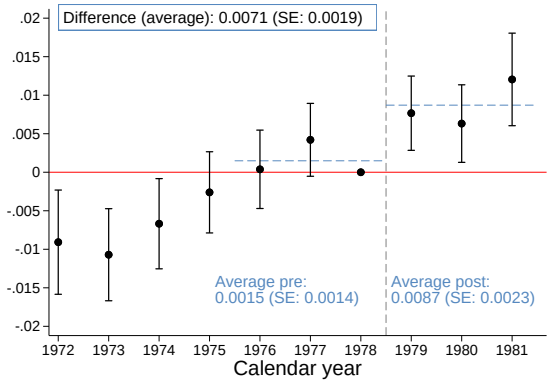
(d) Log Population Women Aged 15–44 × Year Fixed Effects



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



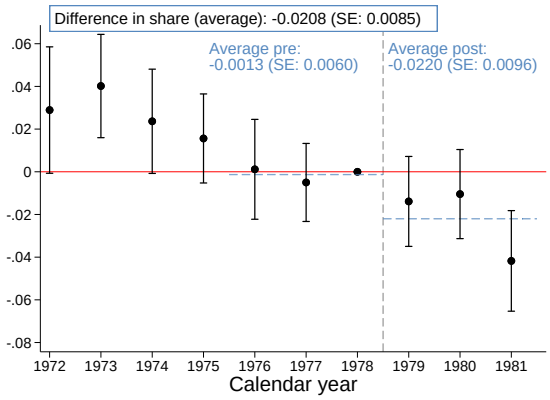
(f) CHC and WIC Controls



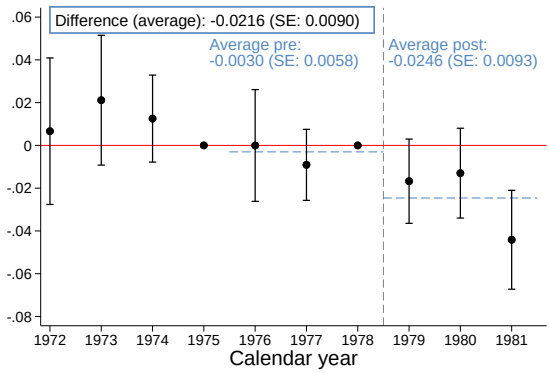
Notes: Figure shows the results of estimating Equation 5 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A18 for more details on each alternative specification. See Appendix B.1 for outcome-specific limitations applied to the free food stamp reform baseline sample. N = 1,359 counties for the baseline estimates in Panel (a) and Panel (f). 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 A26: Impact of Free-Stamps Reform on Share Preterm (< 37 weeks): Alternative Specifications

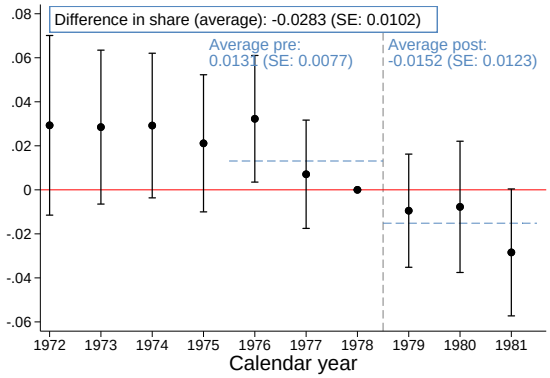
(a) Baseline



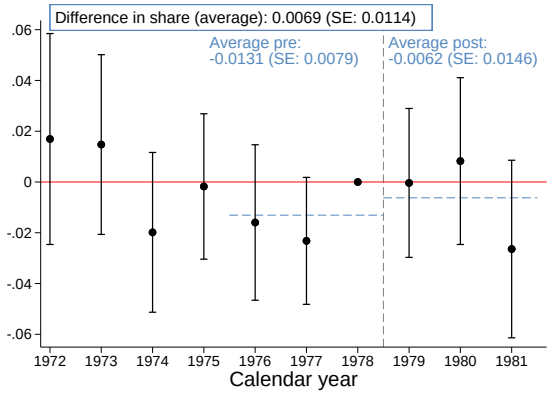
(b) Adjusted for Confounds Using Unexposed Employment



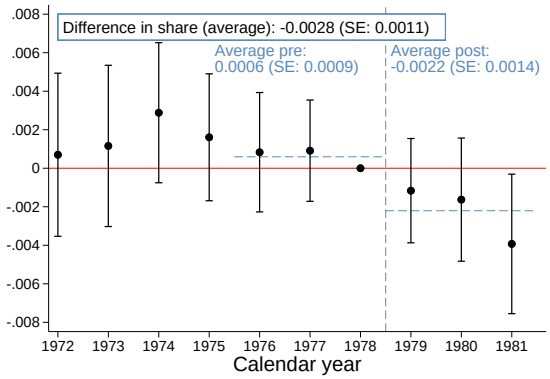
(c) State × Year Fixed Effects



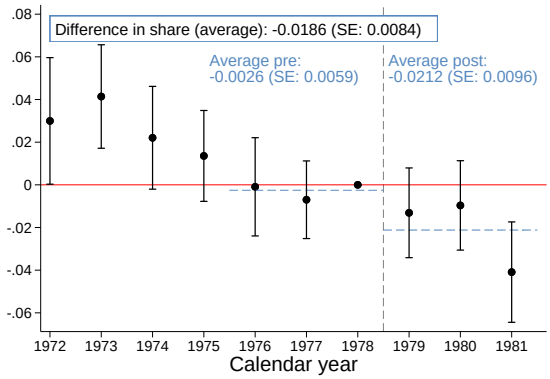
(d) Log Population Women Aged 15–44 × Year Fixed Effects



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



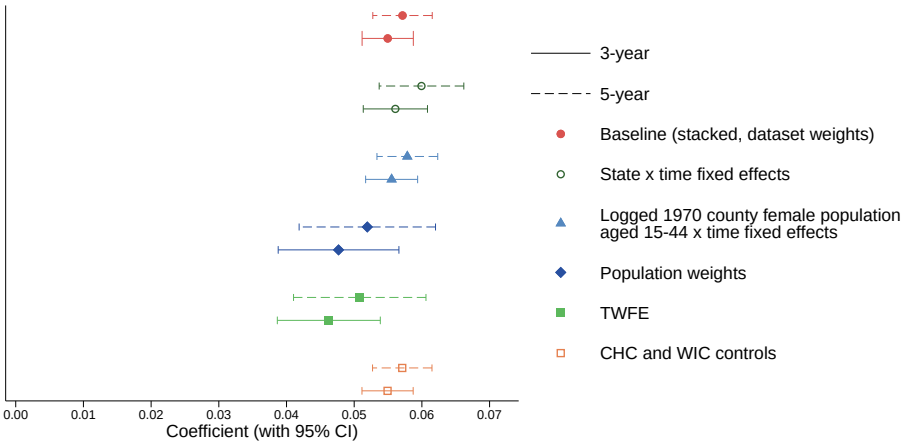
(f) CHC and WIC Controls



Notes: Figure shows the results of estimating Equation 5 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A18 for more details on each alternative specification. See Appendix B.1 for outcome-specific limitations applied to the free food stamp reform baseline sample. N = 1,359 counties for the baseline estimates in Panel (a) and Panel (f). 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.

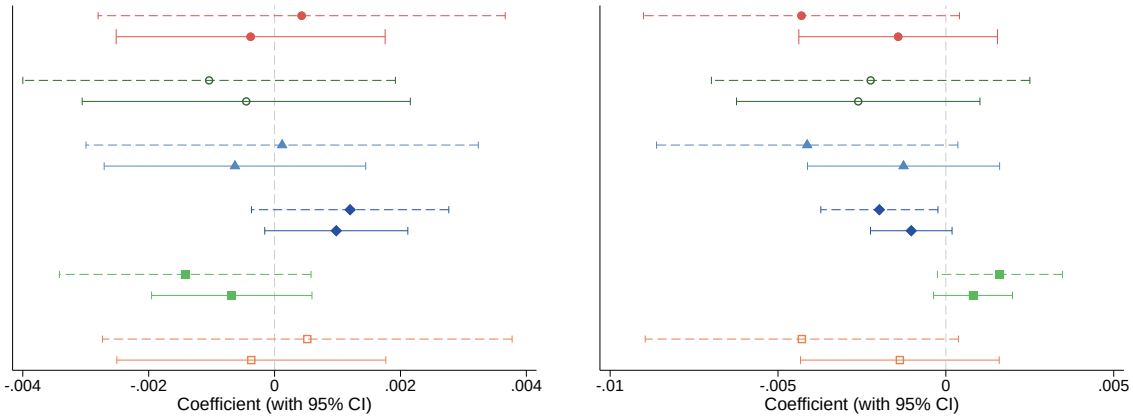
Figure A27: Rollout Impacts Across Specifications
Pooled Births

(a) Food Stamp Program Enrollment



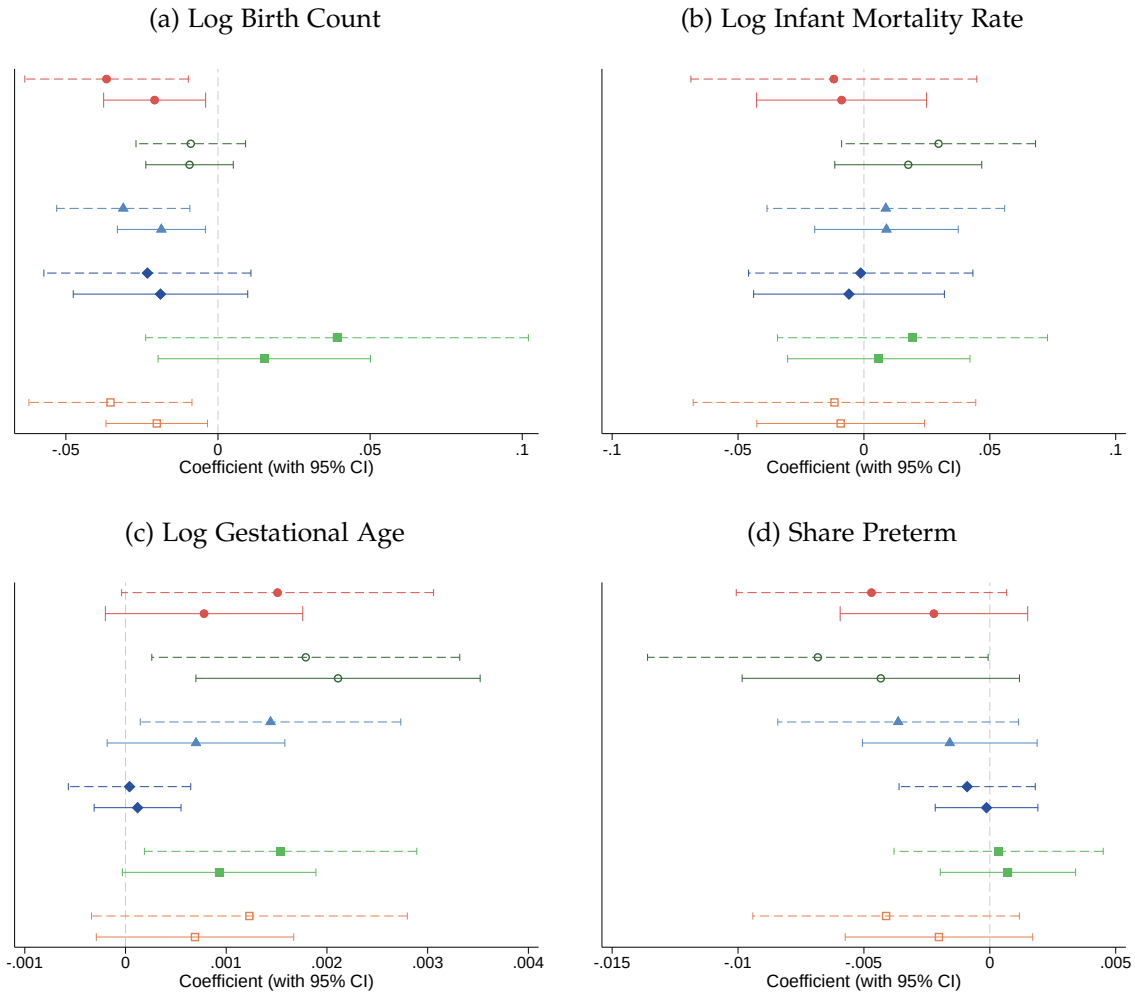
(b) Log Birth Weight

(c) Share Low Birth Weight



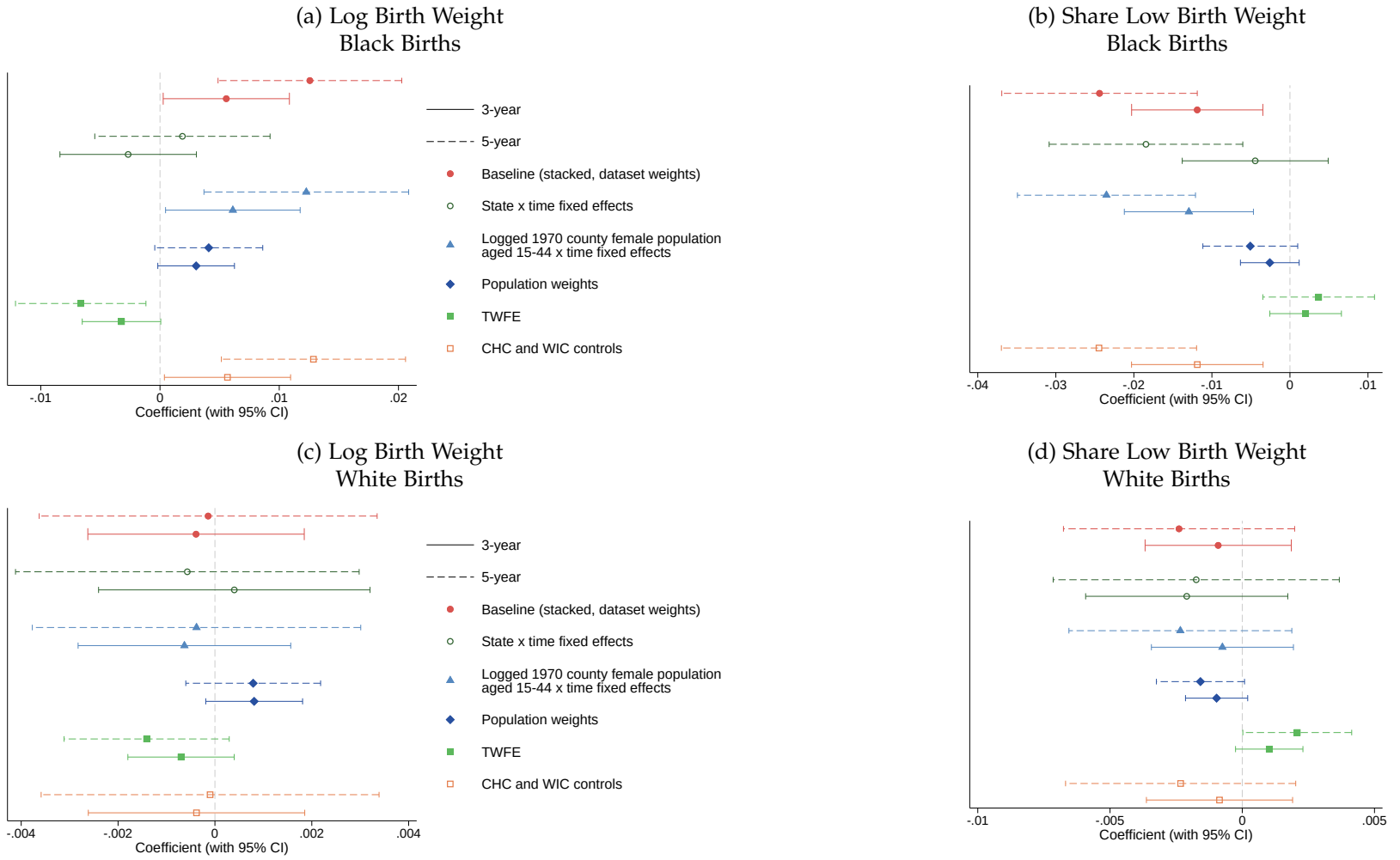
Notes: Figure compiles the 3-year and 5-year difference coefficients (as in Equation 31) and 95% CIs from the baseline specification and five alternative specifications for pooled races. Across specifications, we cluster standard errors at the county level. The baseline specification estimates the event-study 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 26 and 27 using a flat dataset, population weights, and binning at relative years beyond -5 and +5. "CHC and WIC Controls" add two separate simple controls for county-years with each program rolled out, interacting the indicator with dataset indicators. Appendix Figures A30 through A32 present the underlying event study estimates.

Figure A28: Rollout Impacts Across Specifications
Pooled Births, Additional Infant Health Outcomes



Notes: Figure compiles the 3-year and 5-year difference coefficients (as in Equation 31) and 95% CIs from the baseline specification and five 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 26 and 27 using a flat dataset, population weights, and binning at relative years beyond -5 and +5. "CHC and WIC Controls" add two separate simple controls for county-years with each program rolled out, interacting the indicator with dataset indicators. Appendix Figures A33 through A36 present the underlying event study estimates.

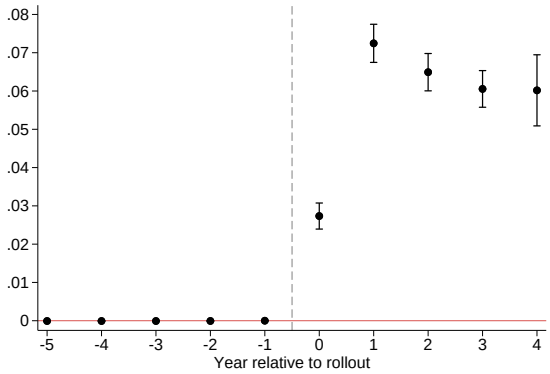
Figure A29: Rollout Impacts Across Specifications
Heterogeneity by Race



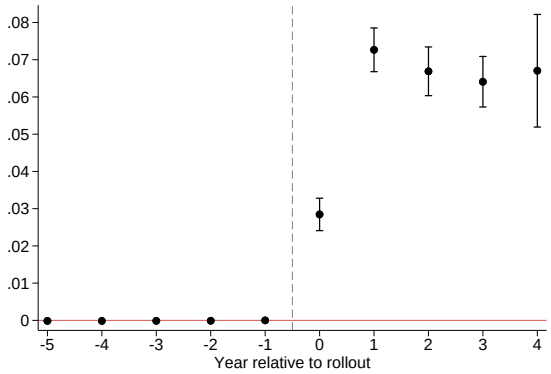
Notes: Figure compiles the 3-year and 5-year difference coefficients (as in Equation 31) and 95% CIs from the baseline specification and five alternative specifications for a sample of Black and White births, separately. Across specifications, we cluster standard errors at the county level. The baseline specification estimates the event-study specification from Equations 7 and 8 using a stacked dataset and dataset weights 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 26 and 27 using a flat dataset, race-county population weights, and binning at relative years beyond -5 and +5. "CHC and WIC Controls" add two separate simple controls for county-years with each program rolled out, interacting the indicator with dataset indicators. Appendix Figures A37 through A38 present the underlying event study estimates.

Figure A30: Rollout Impacts on Food Stamp Program Enrollment Across Specifications

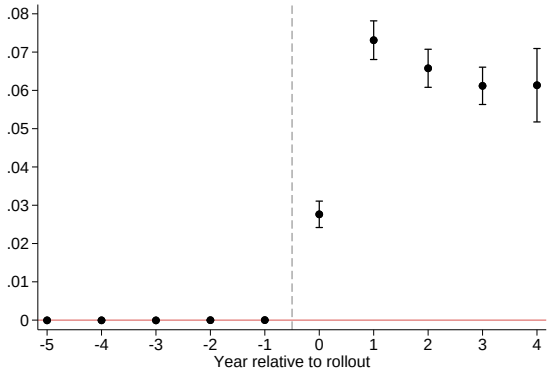
(a) Baseline



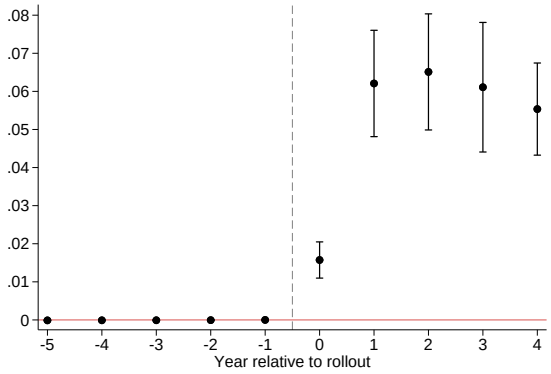
(b) State × Year Fixed Effects



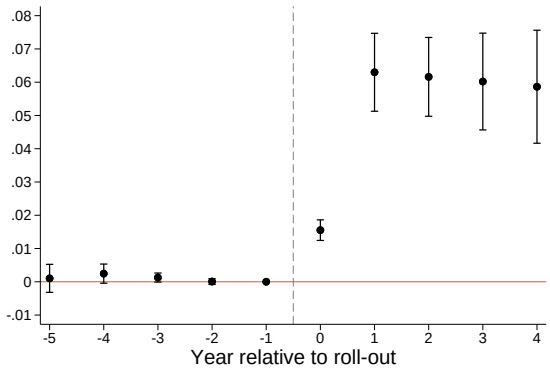
(c) Log Population Women Aged 15–44 × Year Fixed Effects



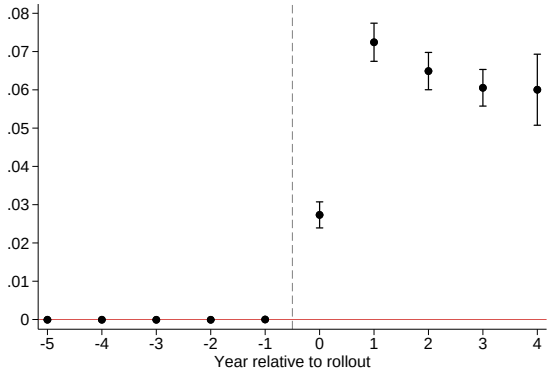
(d) 1970 County Population Weights



(e) TWFE



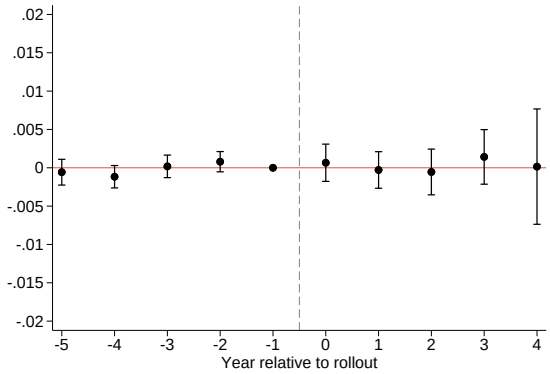
(f) CHC and WIC Controls



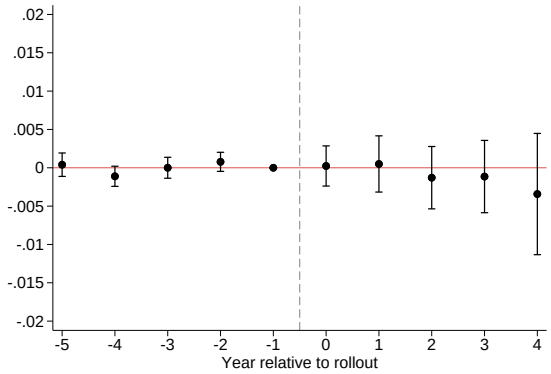
Notes: Figure shows the results of estimating δ_τ from Equation 7 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A27 for more details on each alternative specification. N = 1,520 counties for the baseline estimates in Panel (a), as well as for population-weighted estimates in Panel (d), TWFE estimates in Panel (e), and CHC/WIC controls in Panel (f). 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 A31: Rollout Impacts on Log Birth Weight
Across Specifications

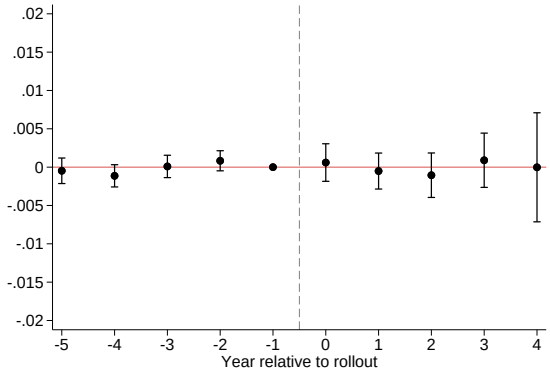
(a) Baseline



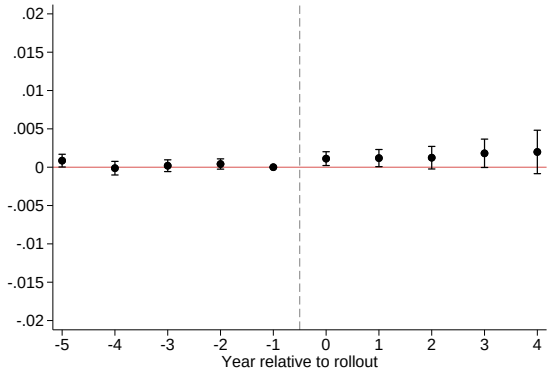
(b) State × Year Fixed Effects



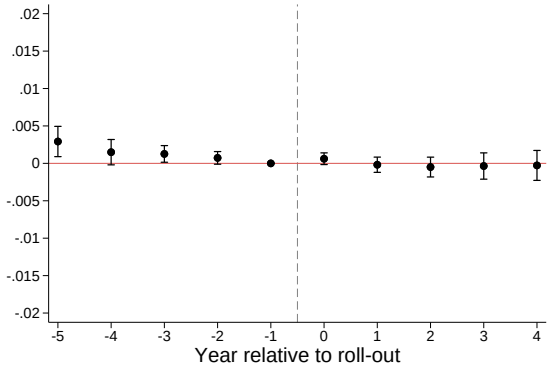
(c) Log Population Women Aged 15–44
× Year Fixed Effects



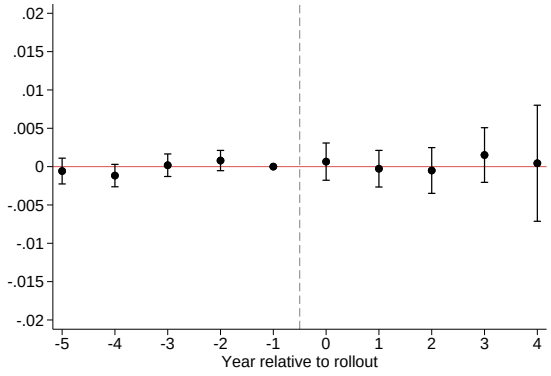
(d) 1970 County Population Weights



(e) TWFE



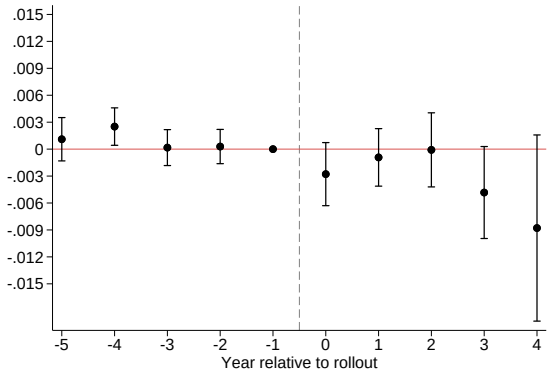
(f) CHC and WIC Controls



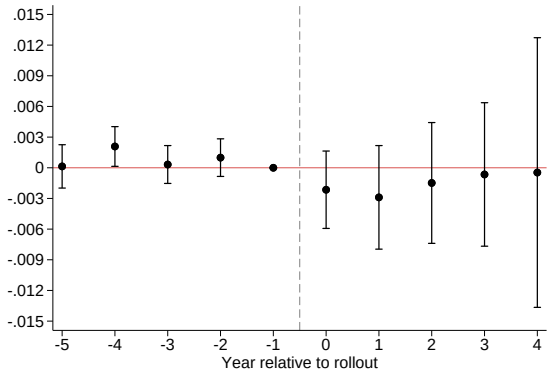
Notes: Figure shows the results of estimating β_τ from Equation 8 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A27 for more details on each alternative specification. N = 1,520 counties for the baseline estimates in Panel (a), as well as for population-weighted estimates in Panel (d), TWFE estimates in Panel (e), and CHC/WIC controls in Panel (f). 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 A32: Rollout Impacts on Share Low Birth Weight Across Specifications

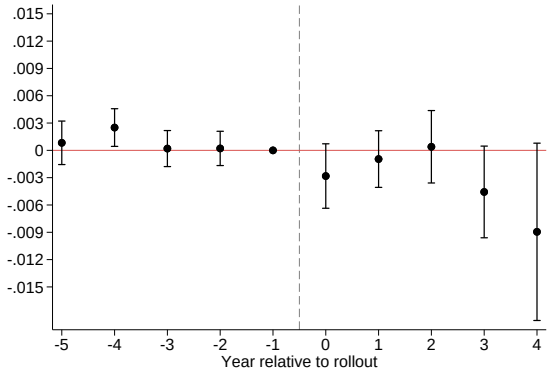
(a) Baseline



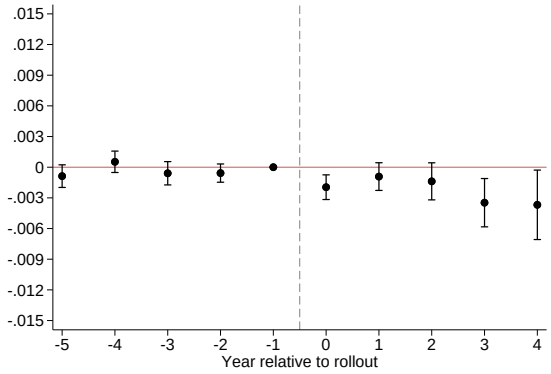
(b) State × Year Fixed Effects



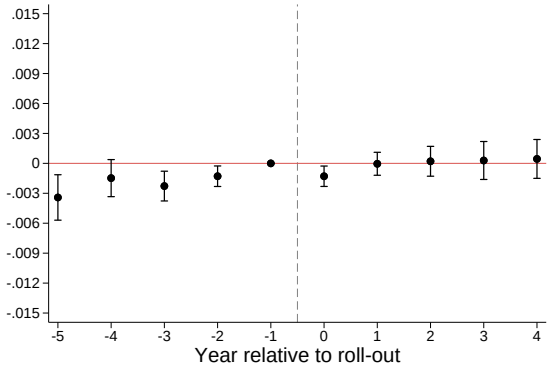
(c) Log Population Women Aged 15–44 × Year Fixed Effects



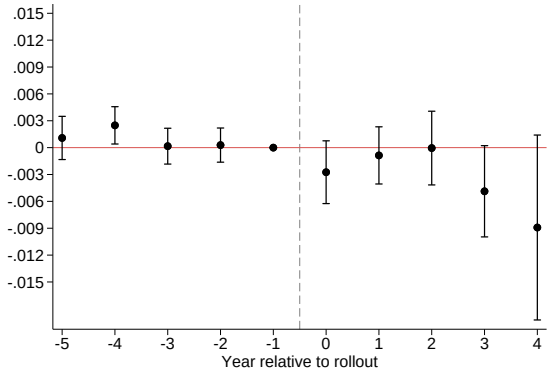
(d) 1970 County Population Weights



(e) TWFE



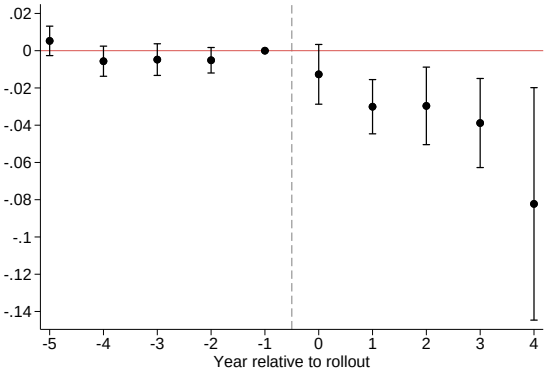
(f) CHC and WIC Controls



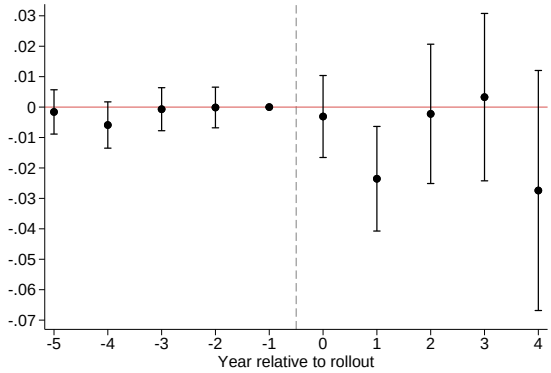
Notes: Figure shows the results of estimating β_τ from Equation 8 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A27 for more details on each alternative specification. N = 1,520 counties for the baseline estimates in Panel (a), as well as for population-weighted estimates in Panel (d), TWFE estimates in Panel (e), and CHC/WIC controls in Panel (f). 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 A33: Rollout Impacts on Log Birth Count
Across Specifications

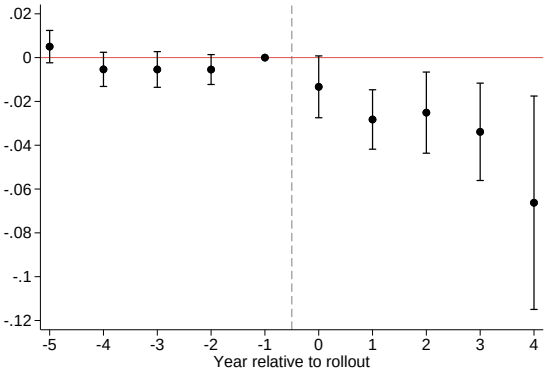
(a) Baseline



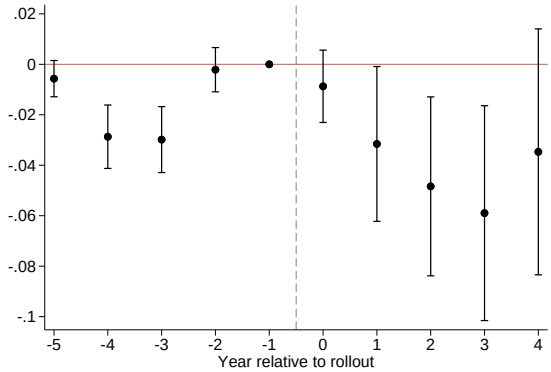
(b) State × Year Fixed Effects



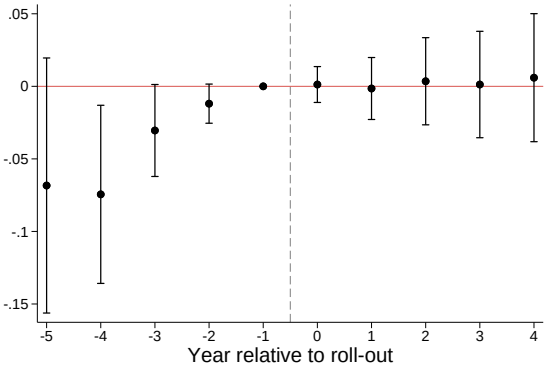
(c) Log Population Women Aged 15–44
× Year Fixed Effects



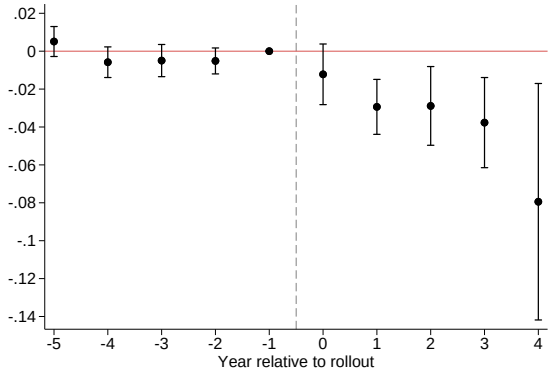
(d) 1970 County Population Weights



(e) TWFE



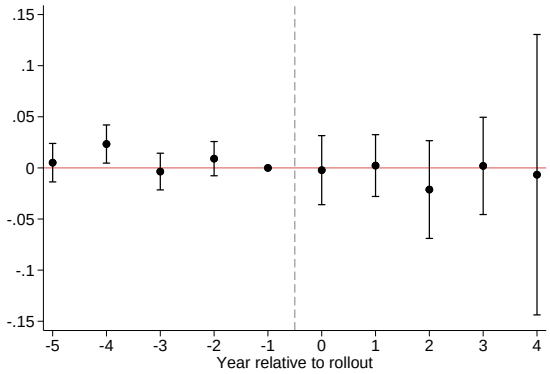
(f) CHC and WIC Controls



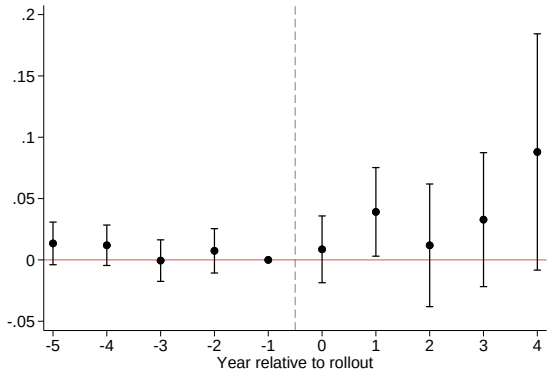
Notes: Figure shows the results of estimating β_τ from Equation 8 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A27 for more details on each alternative specification. See Appendix B.1 for outcome-specific limitations applied 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), TWFE estimates in Panel (e), and CHC/WIC controls in Panel (f). 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 A34: Rollout Impacts on Log Infant Mortality Rate
(Deaths < 12 Months per 1,000 Births)
Across Specifications

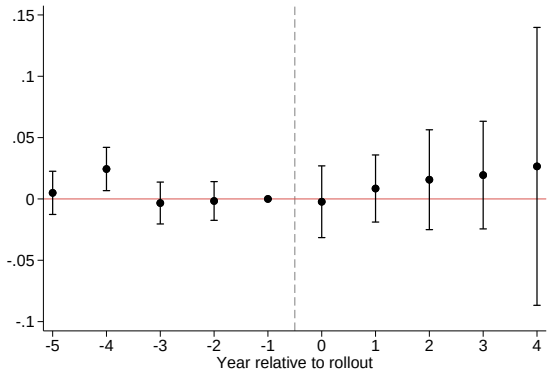
(a) Baseline



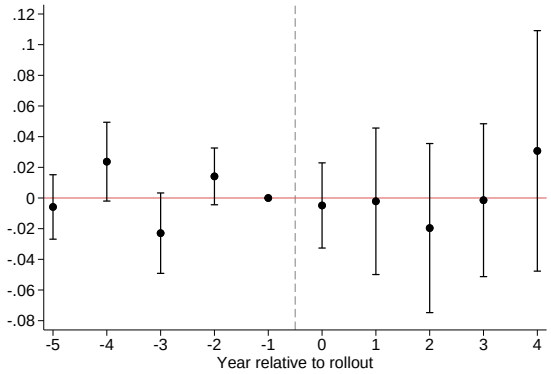
(b) State × Year Fixed Effects



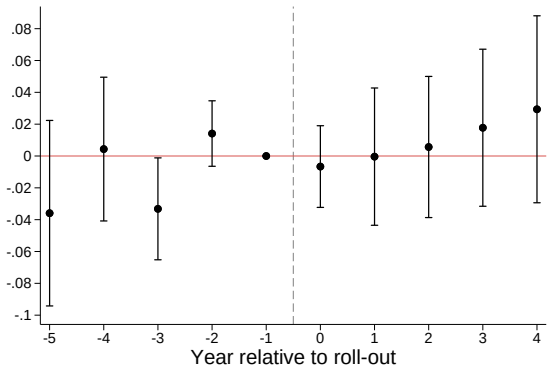
(c) Log Population Women Aged 15–44 × Year Fixed Effects



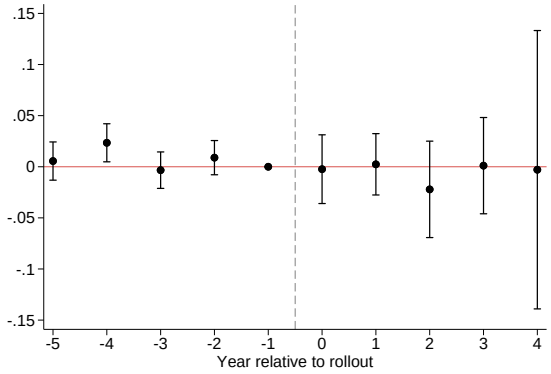
(d) 1970 County Population Weights



(e) TWFE



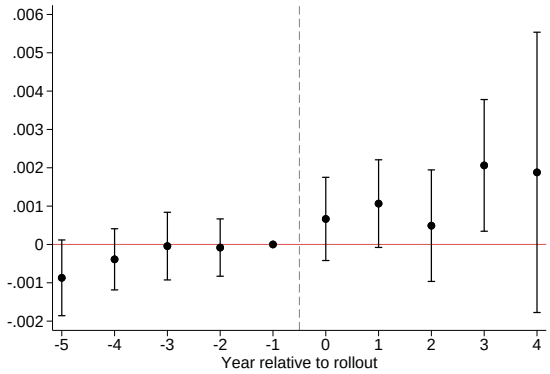
(f) CHC and WIC Controls



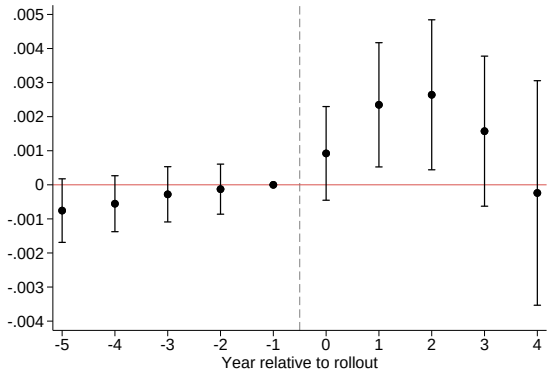
Notes: Figure shows the results of estimating β_τ from Equation 8 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A27 for more details on each alternative specification. See Appendix B.1 for outcome-specific limitations applied 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), TWFE estimates in Panel (e), and CHC/WIC controls in Panel (f). 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 A35: Rollout Impacts on Log Gestational Age Across Specifications

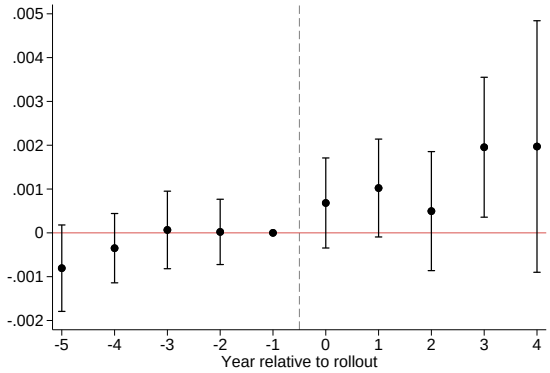
(a) Baseline



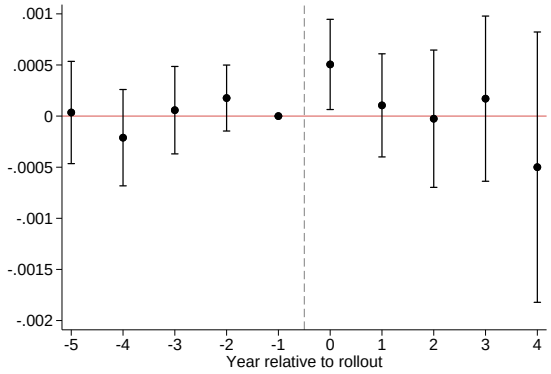
(b) State × Year Fixed Effects



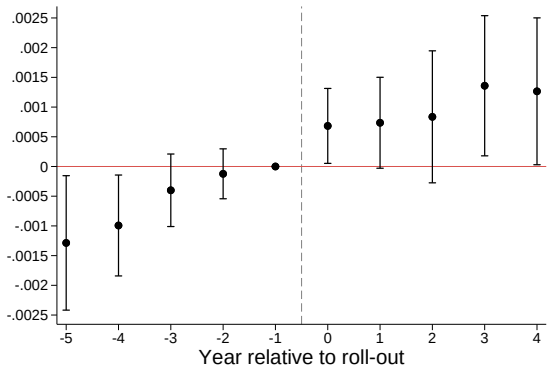
(c) Log Population Women Aged 15–44 × Year Fixed Effects



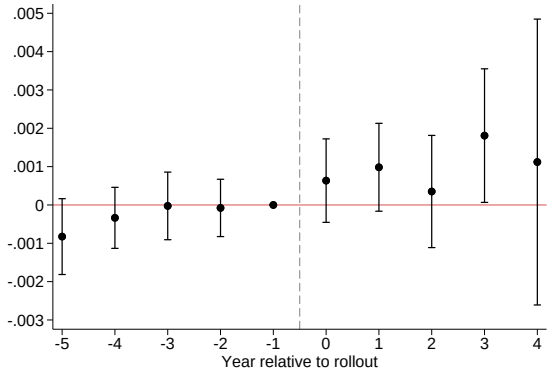
(d) 1970 County Population Weights



(e) TWFE



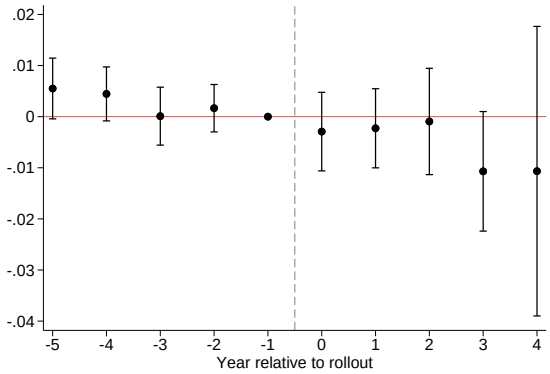
(f) CHC and WIC Controls



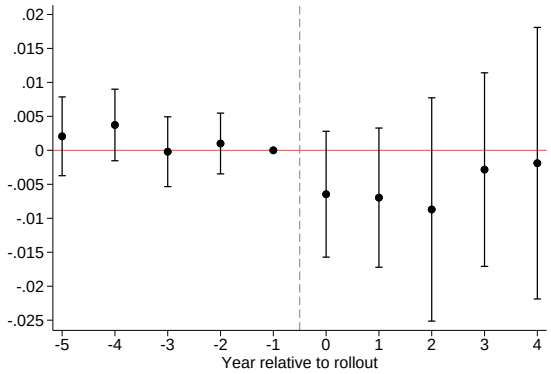
Notes: Figure shows the results of estimating β_τ from Equation 8 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A27 for more details on each alternative specification. See Appendix B.1 for outcome-specific limitations applied 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), TWFE estimates in Panel (e), and CHC/WIC controls in Panel (f). 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 A36: Rollout Impacts on Share Preterm (< 37 weeks gestation) Across Specifications

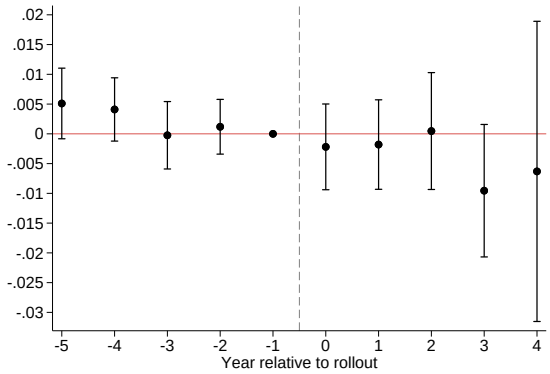
(a) Baseline



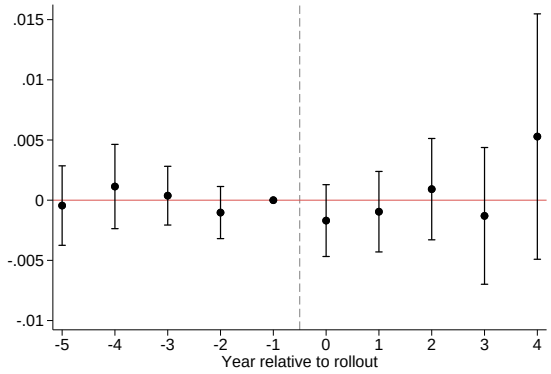
(b) State × Year Fixed Effects



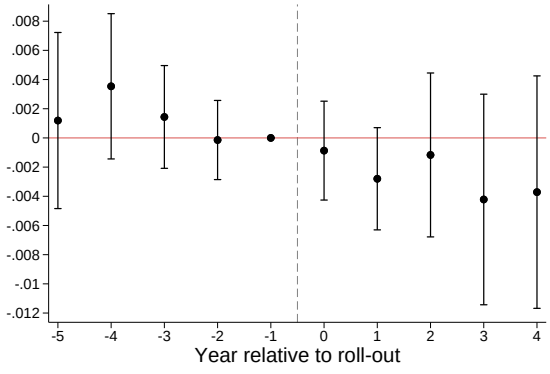
(c) Log Population Women Aged 15–44 × Year Fixed Effects



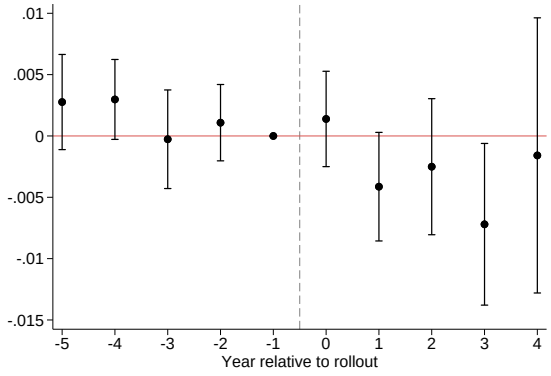
(d) 1970 County Population Weights



(e) TWFE



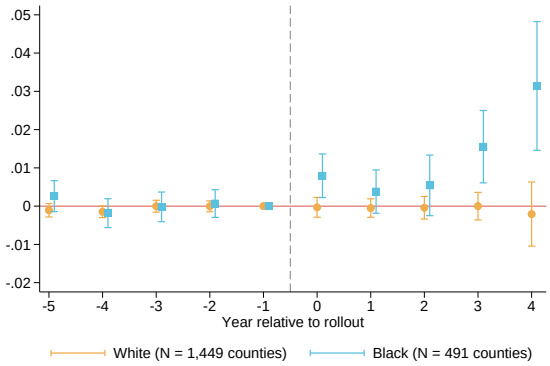
(f) CHC and WIC Controls



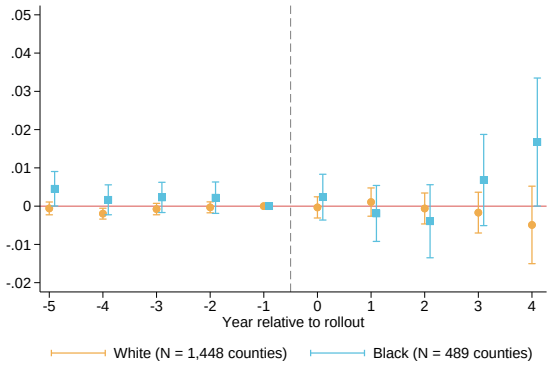
Notes: Figure shows the results of estimating β_τ from Equation 8 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A27 for more details on each alternative specification. See Appendix B.1 for outcome-specific limitations applied 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), TWFE estimates in Panel (e), and CHC/WIC controls in Panel (f). 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 A37: Rollout Impacts on Log Birth Weight
Across Specifications by Race

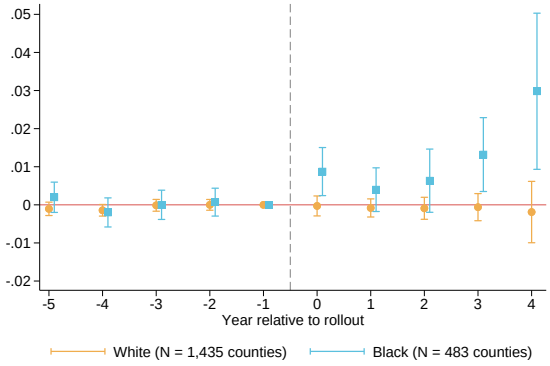
(a) Baseline



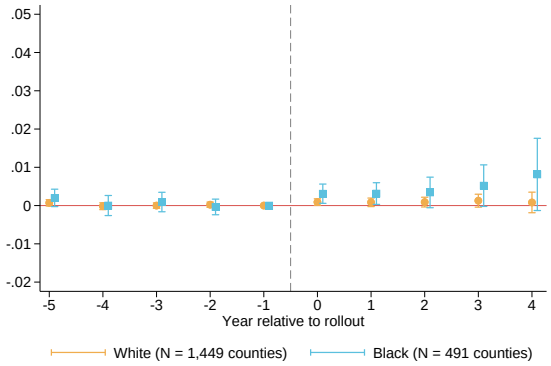
(b) State × Year Fixed Effects



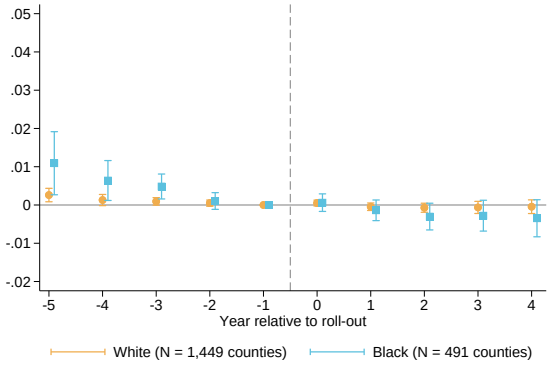
(c) Log Population Women Aged 15–44
× Year Fixed Effects



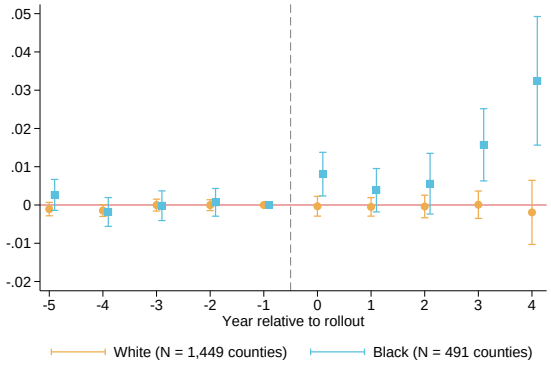
(d) 1970 County Population Weights



(e) TWFE



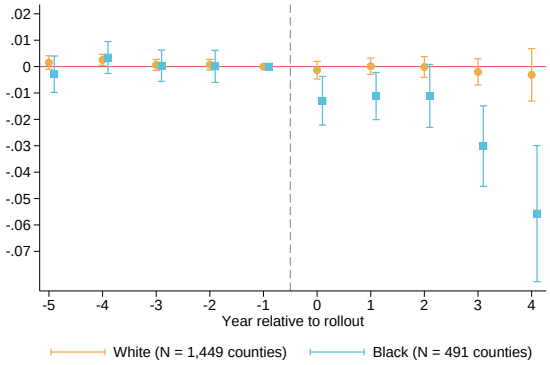
(f) CHC and WIC Controls



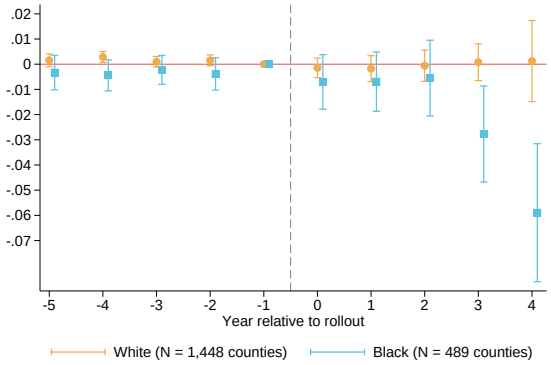
Notes: Figure shows the results of estimating β_τ from Equation 8 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A27 for more details on each alternative specification. The baseline rollout sample is additionally limited to counties that have at least 25 race-specific births observed across years 1970–1975. Standard errors are clustered at the county level.

Figure A38: Rollout Impacts on Share Low Birth Weight Across Specifications by Race

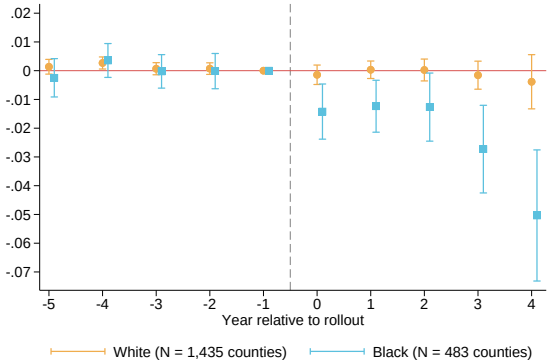
(a) Baseline



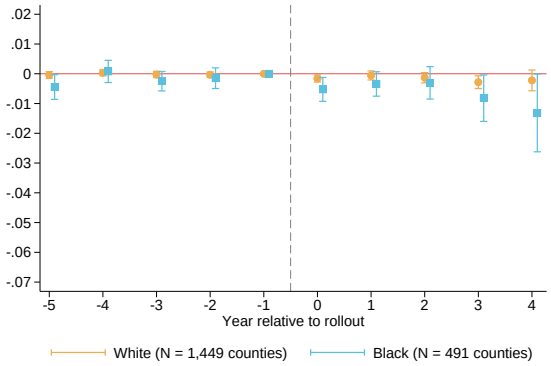
(b) State × Year Fixed Effects



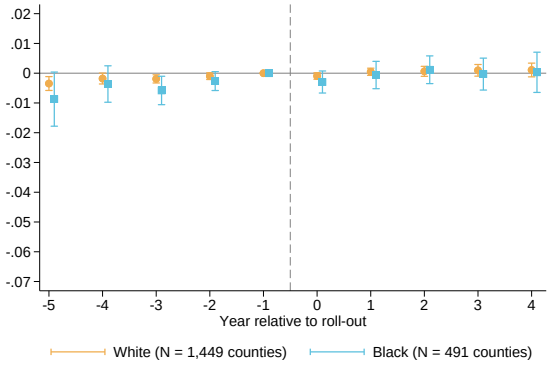
(c) Log Population Women Aged 15–44 × Year Fixed Effects



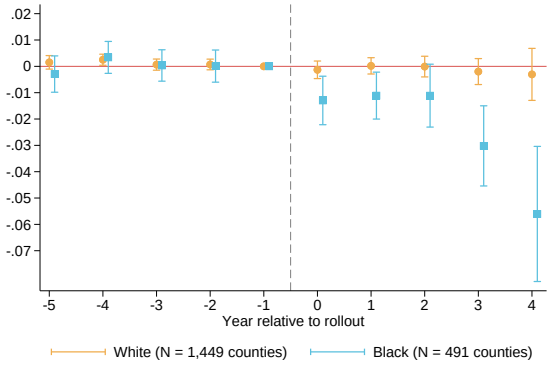
(d) 1970 County Population Weights



(e) TWFE



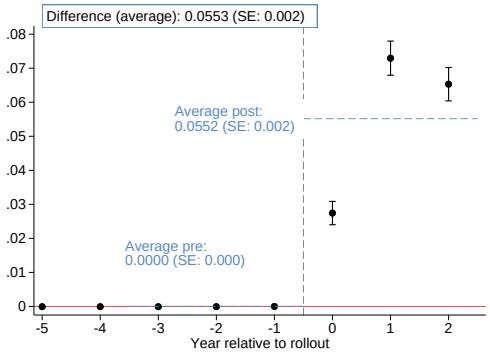
(f) CHC and WIC Controls



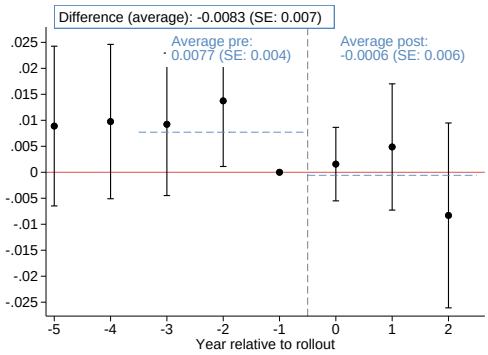
Notes: Figure shows the results of estimating β_τ from Equation 8 for the baseline specification (Panel a) and five alternative specifications (Panels b through f). See notes to Appendix Figure A27 for more details on each alternative specification. The baseline rollout sample is additionally limited to counties that have at least 25 race-specific births observed across years 1970–1975. Standard errors are clustered at the county level.

Figure A39: Impact of Food Stamp Program Rollout on Employment

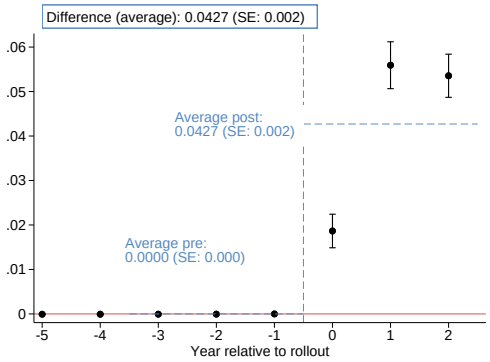
(a) First Stage: Food Stamp Program Enrollment



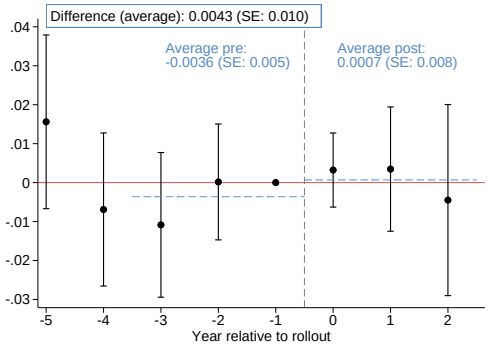
(b) Log Employment (All Industries)



(c) First Stage: Food Stamp Program Enrollment



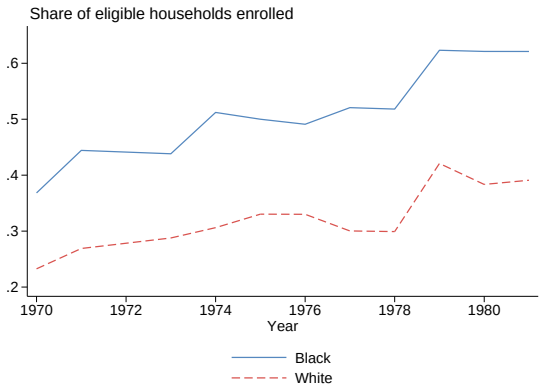
(d) Log Employment (Low-wage Industries)



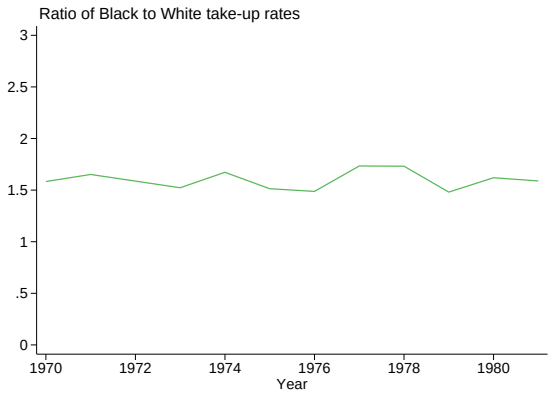
Notes: Figure shows estimates of δ_τ from Equation 7 in Panels (a) and (c) and from Equation 8 in Panels (b) and (d). The regressions are weighted by dataset-normalized weights, described in Equation 25. Since food stamp program enrollment and CBP data are both constructed as annual measures, relative 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 log 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 A40: PSID: Food Stamp Program Take-up

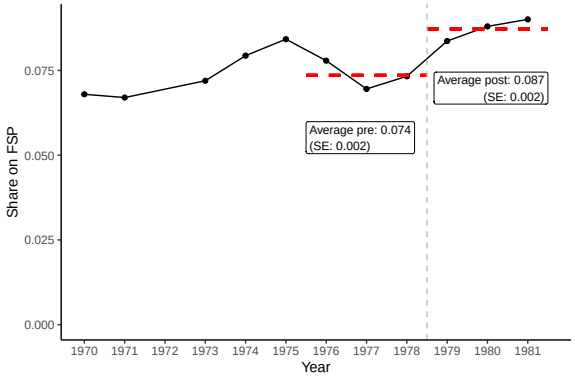
(a) Take-up Rate by Race



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

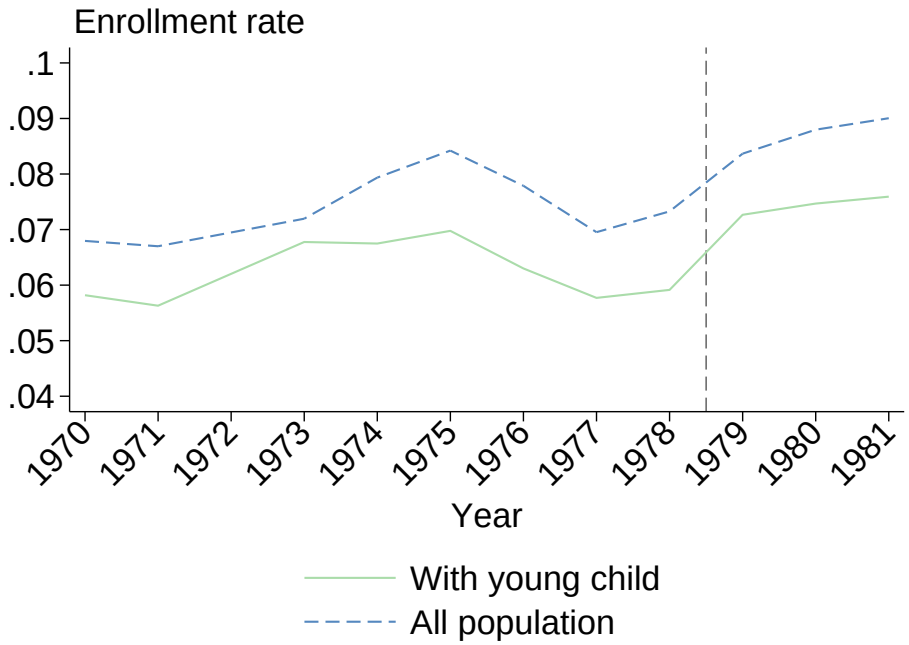


(c) Food Stamp Program Enrollment over Time



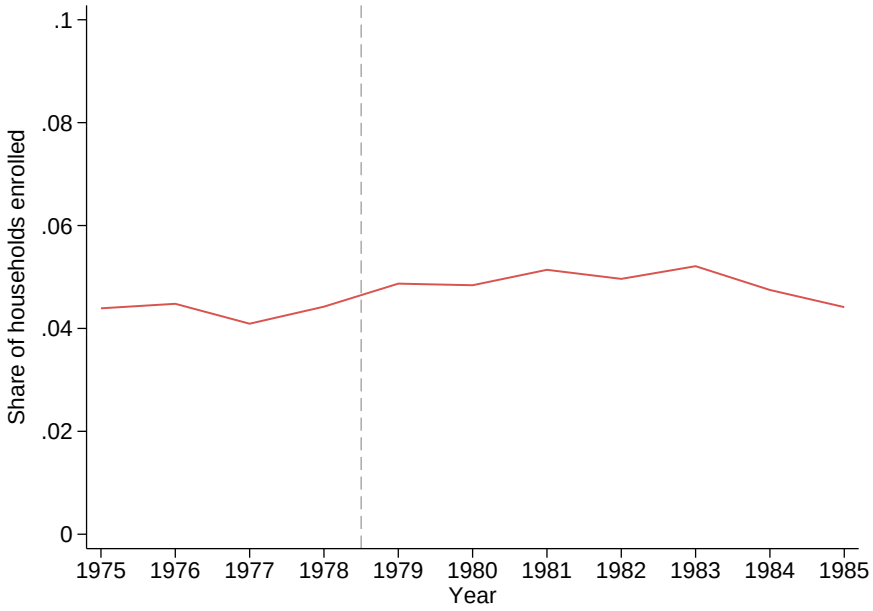
Notes: Panel (a) plots the national, annual take-up rate of food stamps for White and Black households, separately, with take-up defined as the share of the eligible population that is enrolled. Panel (b) plots the ratio of those race-specific take-up rates (Black:White) in each year between 1970 and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

Figure A41: Food Stamp Enrollment Rates in the PSID
Households with a Young Child



Notes: Figure shows PSID-weighted time series of food stamp program enrollment (food stamp transfers $>$ \$0) for households with a young child (child under age 3) and the entire sample population, separately. On average, the enrollment rate is slightly lower among households with a young child (average of 6.63 percent) than in the overall population (average of 7.82 percent).

Figure A42: Share of Households Enrolled in the Food Stamp Program and AFDC, SSI, or General Assistance



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

Figure A43: Benefits Table, January 1978 and January 1979

(a) January 1978

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

(b) January 1979

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

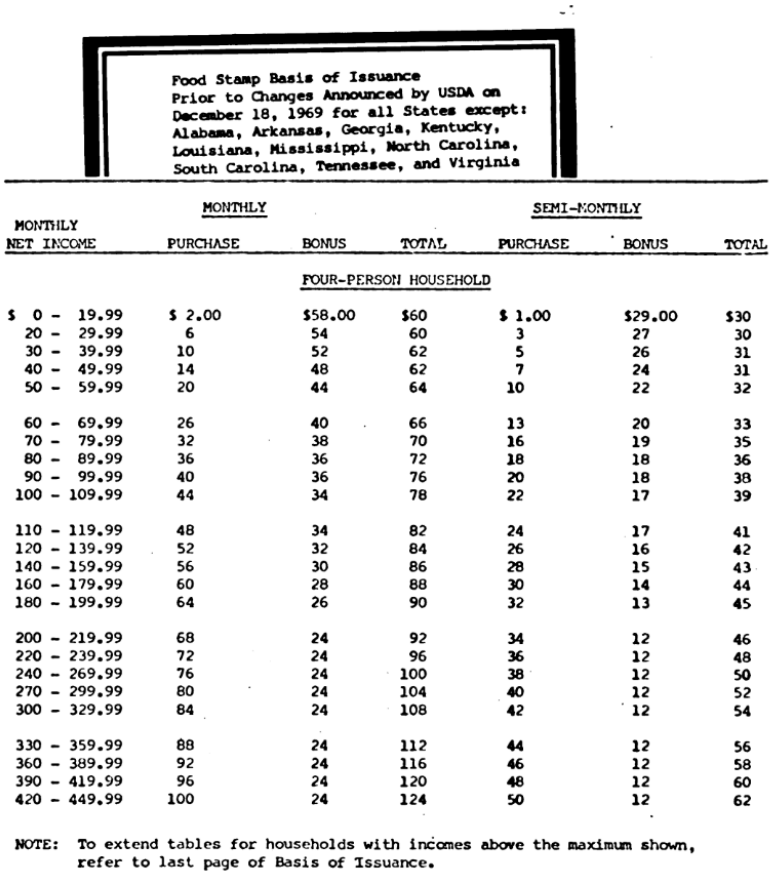
TABLE I

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. The schedule in Panel (a) took effect in January 1978 and is found in Federal Register Vol 42 Iss 215 (U.S. Department of Agriculture, Food and Nutrition Service, 1977). The schedule in Panel (b) took effect in January 1979 and is found in Federal Register Vol 43 Iss 225 (U.S. Department of Agriculture, Food and Nutrition Service, 1978a). Every household of the same size received the same-value stamps allotment, and in Panel (a), 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.

Figure A44: Benefits Table, December 1969 and April 1971

(a) December 1969

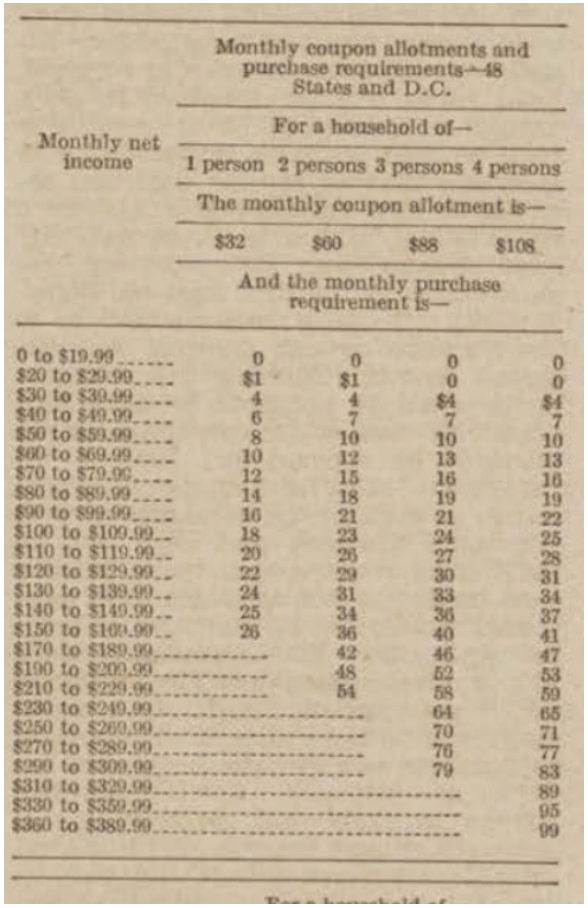


Food Stamp Basis of Issuance
Prior to Changes Announced by USDA on December 18, 1969 for all States except: Alabama, Arkansas, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, and Virginia

MONTHLY NET INCOME	MONTHLY PURCHASE	MONTHLY BONUS	MONTHLY TOTAL	SEMI-MONTHLY PURCHASE	SEMI-MONTHLY BONUS	SEMI-MONTHLY TOTAL
			FOUR-PERSON HOUSEHOLD			
$ 0 - 19.99	$ 2.00	$58.00	$60	$ 1.00	$29.00	$30
20 - 29.99	6	54	60	3	27	30
30 - 39.99	10	52	62	5	26	31
40 - 49.99	14	48	62	7	24	31
50 - 59.99	20	44	64	10	22	32
60 - 69.99	26	40	66	13	20	33
70 - 79.99	32	38	70	16	19	35
80 - 89.99	36	36	72	18	18	36
90 - 99.99	40	36	76	20	18	38
100 - 109.99	44	34	78	22	17	39
110 - 119.99	48	34	82	24	17	41
120 - 139.99	52	32	84	26	16	42
140 - 159.99	56	30	86	28	15	43
160 - 179.99	60	28	88	30	14	44
180 - 199.99	64	26	90	32	13	45
200 - 219.99	68	24	92	34	12	46
220 - 239.99	72	24	96	36	12	48
240 - 269.99	76	24	100	38	12	50
270 - 299.99	80	24	104	40	12	52
300 - 329.99	84	24	108	42	12	54
330 - 359.99	88	24	112	44	12	56
360 - 389.99	92	24	116	46	12	58
390 - 419.99	96	24	120	48	12	60
420 - 449.99	100	24	124	50	12	62

NOTE: To extend tables for households with incomes above the maximum shown, refer to last page of Basis of Issuance.

(b) April 1971

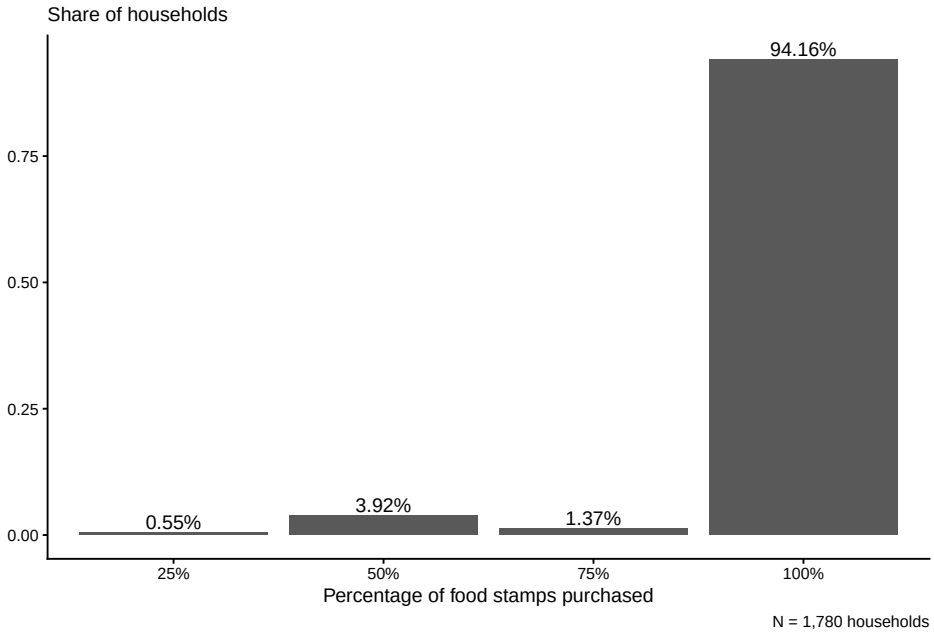


Monthly coupon allotments and purchase requirements—48 States and D.C.

Monthly net income	For a household of— 1 person	2 persons	3 persons	4 persons
	The monthly coupon allotment is— $32	$60	$88	$108
	And the monthly purchase requirement is—			
0 to $19.99	0	0	0	0
$20 to $29.99	$1	$1	0	0
$30 to $39.99	4	4	$4	$4
$40 to $49.99	6	7	7	7
$50 to $59.99	8	10	10	10
$60 to $69.99	10	12	13	13
$70 to $79.99	12	15	16	16
$80 to $89.99	14	18	19	19
$90 to $99.99	16	21	21	22
$100 to $109.99	18	23	24	25
$110 to $119.99	20	26	27	28
$120 to $129.99	22	29	30	31
$130 to $139.99	24	31	33	34
$140 to $149.99	25	34	36	37
$150 to $169.99	26	36	40	41
$170 to $189.99		42	46	47
$190 to $209.99		48	52	53
$210 to $229.99		54	58	59
$230 to $249.99			64	65
$250 to $269.99			70	71
$270 to $289.99			76	77
$290 to $309.99			79	83
$310 to $329.99				89
$330 to $359.99				95
$360 to $389.99				99

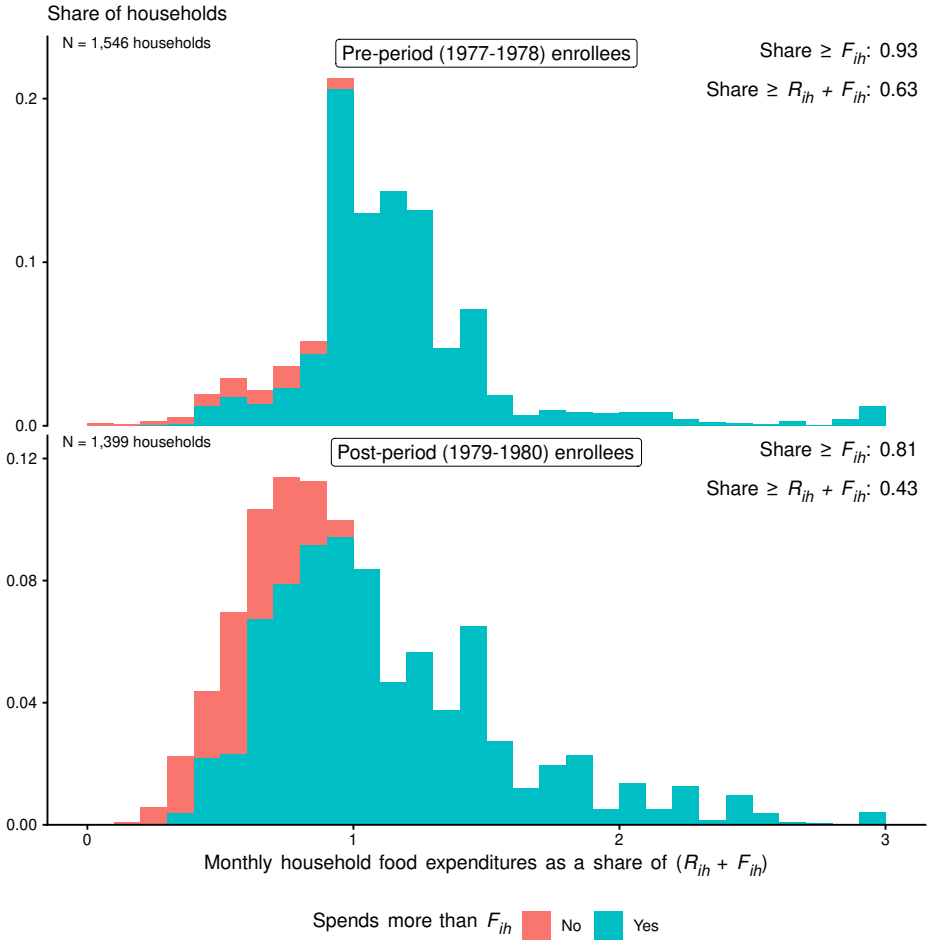
Notes: The schedule in Panel (a) was in effect in December 1969 and is found in the USDA appropriations hearings for fiscal year 1971 (U.S. House of Representatives, Committee on Appropriations, 1970). The schedule in Panel (b) took effect in April 1971 and is found in Federal Register Vol 36 Iss 74 (U.S. Department of Agriculture, Food and Nutrition Service, 1971b). 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.

Figure A45: NFCS: Distribution of the Percentage of Food Stamps Purchased



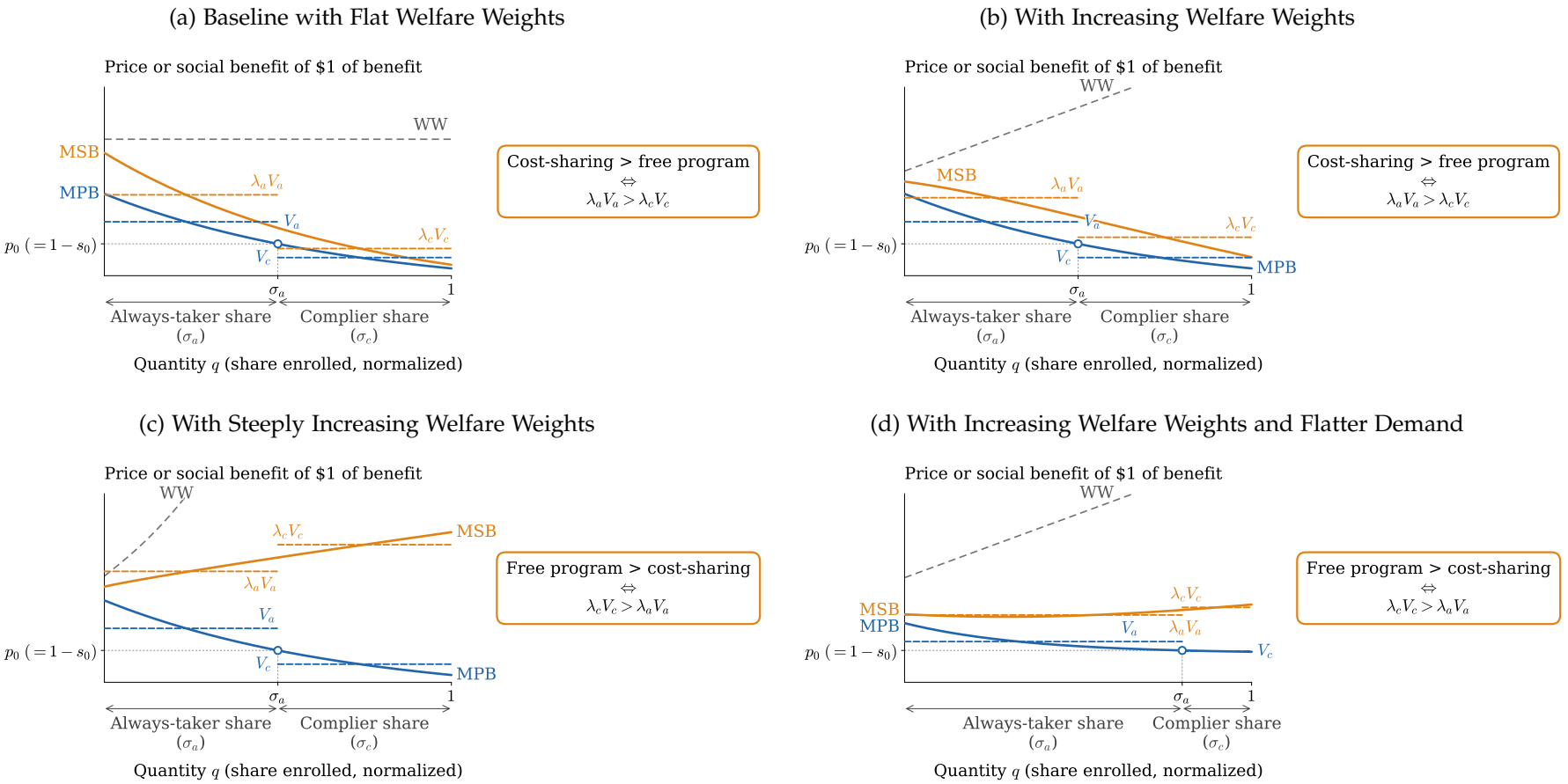
Notes: Figure shows the percentage of the food stamp benefit that households purchased in the month prior to responding to the 1977–1978 Nationwide Food Consumption Survey. Shares are weighted by the NFCS weights. The sample is restricted to households that were surveyed in 1977–1978 and were enrolled in food stamps at the time of their response. See Appendix A.2.2 for more details on the data.

Figure A46: NFCS: Food Spending Normalized to 1978 $R + F$



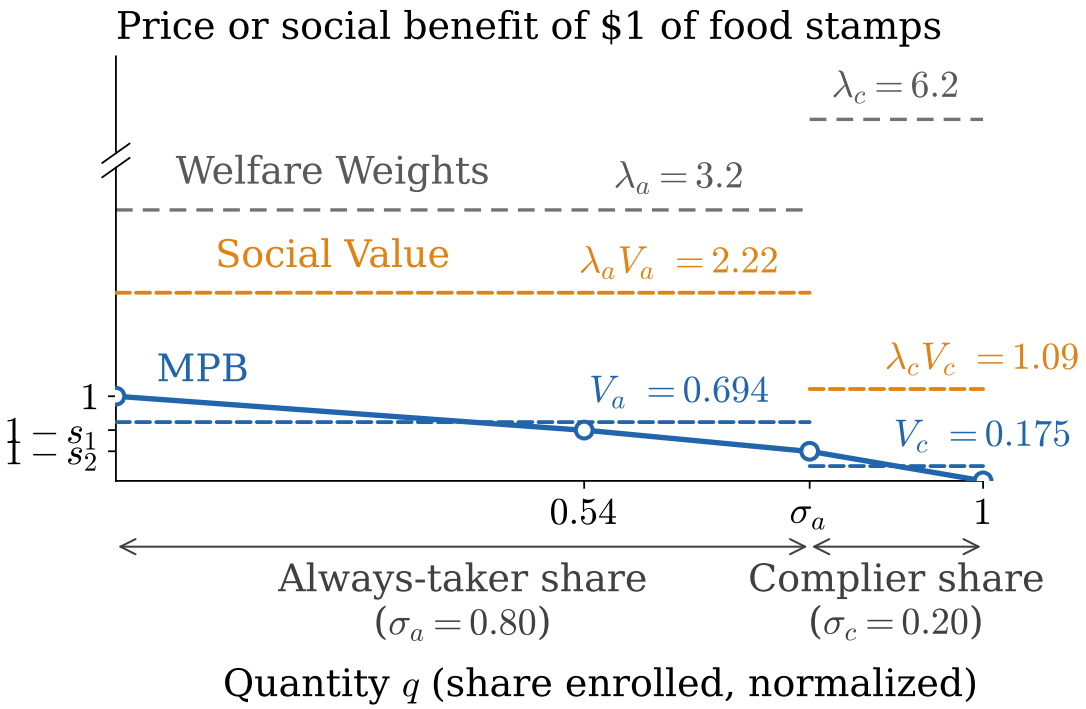
Notes: Figure shows a histogram of monthly food spending among food stamp enrollees expressed as a share of their 1978 $R_{ih} + F_{ih}$, winsorized to 3. Colors indicate whether monthly food spending exceeded the bonus F_{ih}. Sub-panels are split into the time period prior to the elimination of the 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 all years, net income and benefit amounts are determined using the regulations in effect in 1978, as described in Appendix A.2.2. 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. All dollar amounts are adjusted to $1978.

Figure A47: Illustration of Conceptual Framework



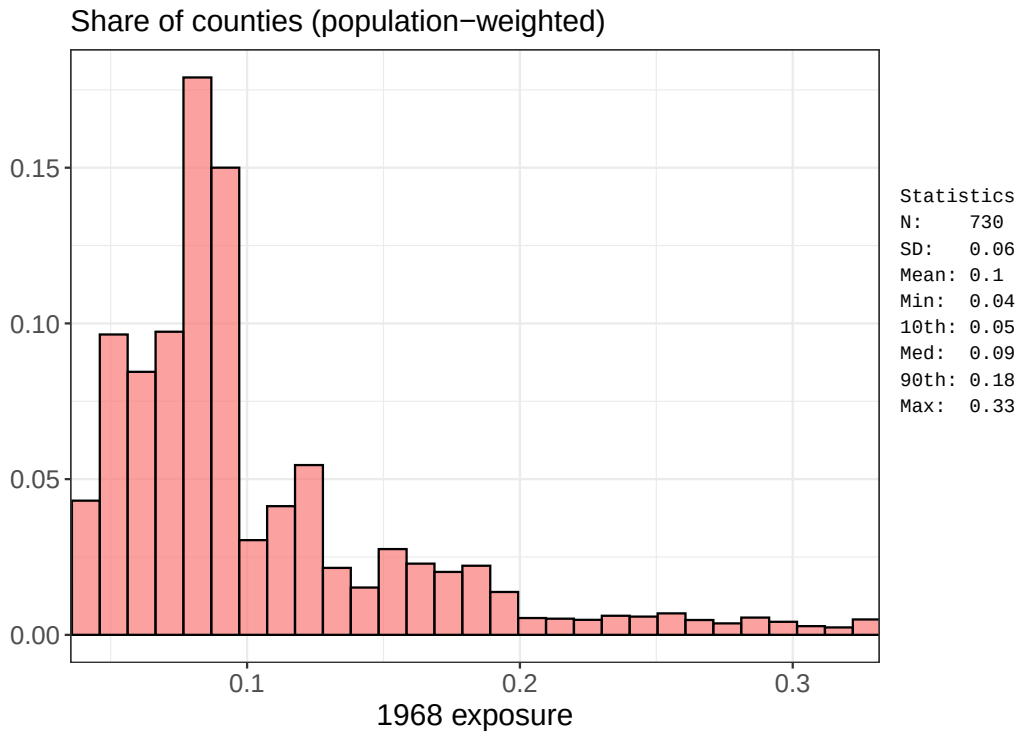
Notes: Figures present the Marginal Private Benefit curve in solid blue (demand, corresponding to perceived private values of a fully subsidized transfer $V_\theta(1)$), Welfare Weights in gray (λ_θ), and Marginal Social Benefit curves in orange ($\lambda_\theta V_\theta(1)$). Types θ are sorted by demand 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 a free transfer is desirable ($\lambda_a V_a \lessgtr \lambda_c V_c$). 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 A48: The Welfare Value of Free Stamps vs. Cost-Sharing: Empirical Version of Appendix Figure A47



Notes: Figure presents an empirical version of Appendix Figure A47. 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.

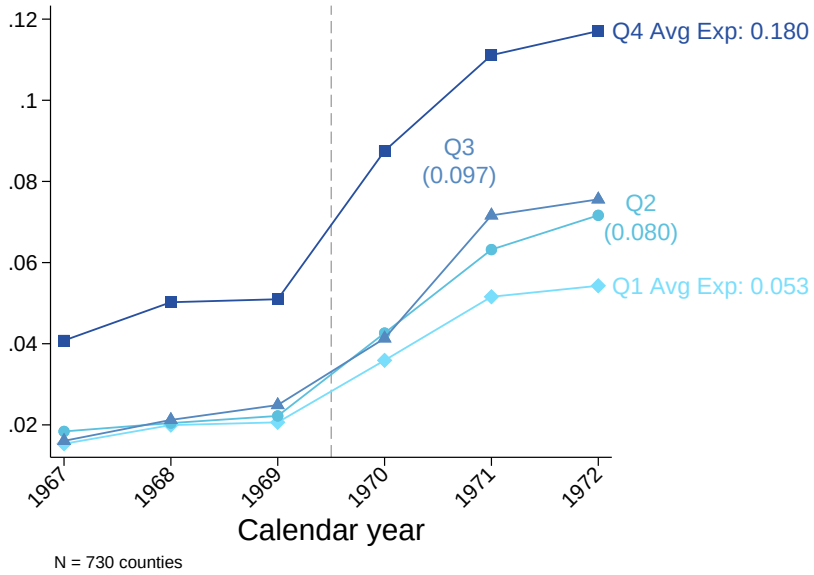
Figure A49: 1968 Exposure Measure



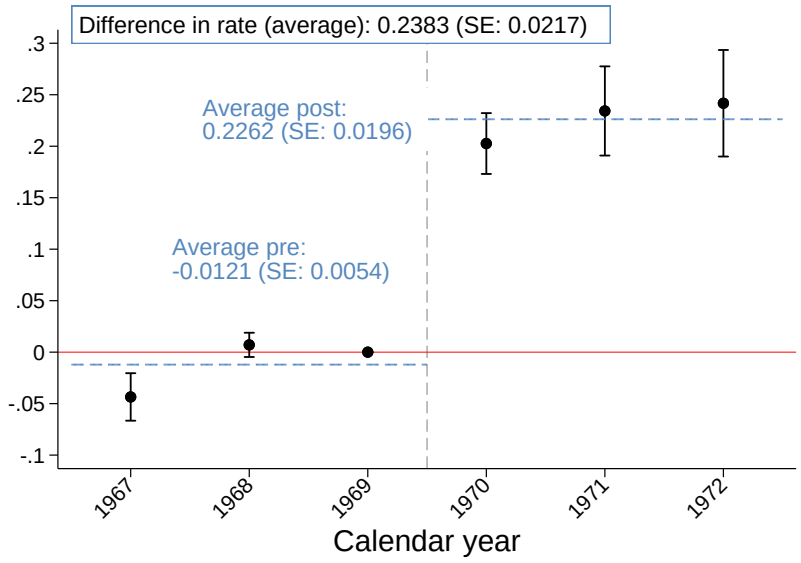
Notes: Figure shows 1970 population-weighted distribution of the county-level 1968 exposure measure, according to Equation 3. The exposure measure is winsorized at the 1st and 99th percentiles, 0.035 and 0.332, respectively. Sample is limited to the subset of the free food stamp reform analysis baseline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

Figure A50: 1970 Bonus Reform: Enrollment Rate

(a) Trends by Quartile



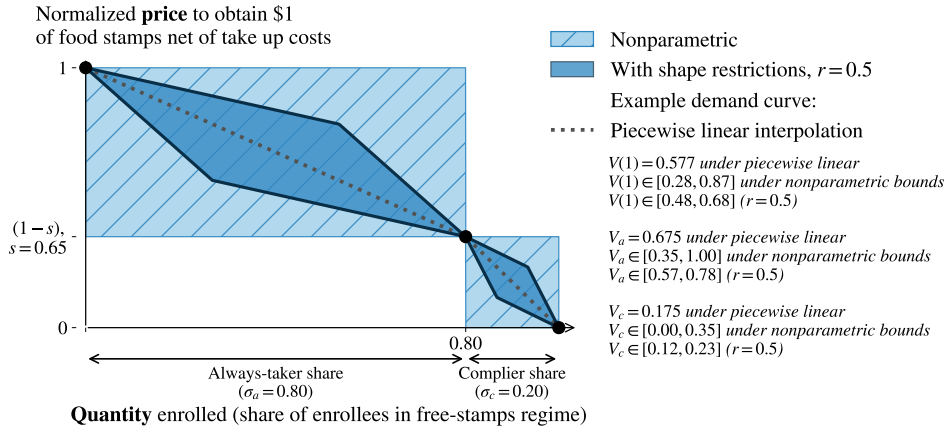
(b) Impact on Enrollment



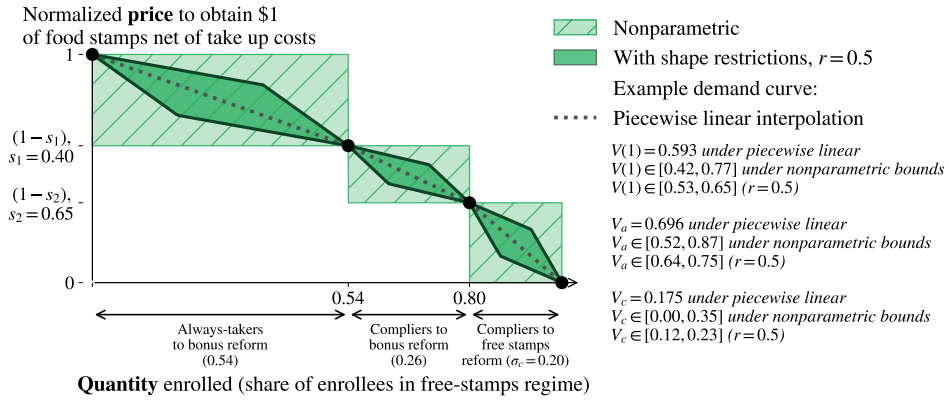
Notes: Panel (a) shows 1970 population-weighted average county-level food stamp enrollment over time by population-weighted quartile of the 1968 exposure measure. Darker shades in the series color scale correspond with increasing quartile of exposure. The average exposure measure within each quartile is displayed within series labels. Panel (b) shows the results of estimating Equation 72, which looks at the impact of the bonus reform on the food stamp enrollment rate using the 1968 exposure measure, and controls for rollout-cohort-by-calendar year fixed effects. Sample is limited to the subset of the free food stamp reform analysis baseline sample counties where the food stamp program was rolled out in 1967 or earlier. Standard errors are clustered at the county level (N = 730 counties).

Figure A51: Bounds on Value of Food Stamps

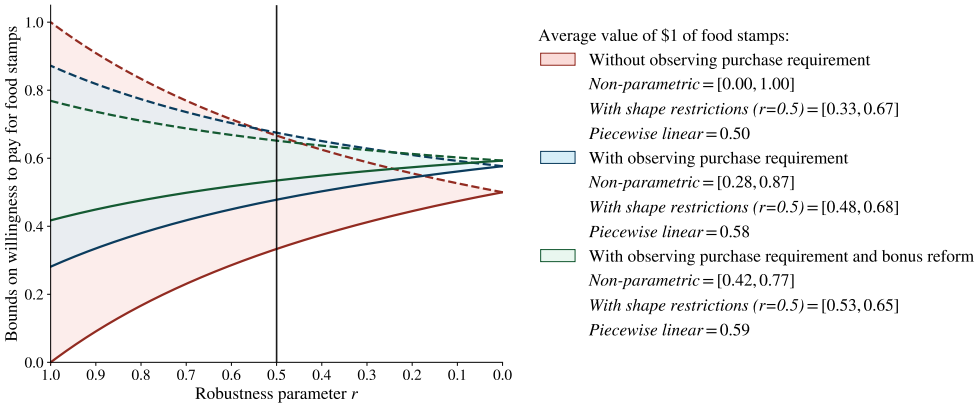
(a) With Purchase Requirement Only



(b) With Purchase Requirement and Bonus Reform



(c) Robustness to Parametric Restrictions



Notes: See notes to Figure 9. The dark shaded rhombuses correspond to semiparametric demand-curve regions using $r = 0.5$ from Kang and Vasserman (2025) (see Footnote 55 for the definition of r). 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 $r \in (0, 1)$, we 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 (Equation 73).

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 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, enrollment, 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 enrollment data aggregate 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: Race-specific Sample Restrictions

Restriction	White sample			Black sample		
	Counties (N)	**Share of 1970 population (%)**	**Baseline white pop retained (%)**	**Counties (N)**	**Share of 1970 population (%)**	**Baseline Black pop retained (%)**
(a) Free FS reform race-specific sample						
Baseline free food stamp reform sample	1,804	71.25%		1,804	71.25%	
Exclude county-races missing birth weight or race population data in any year 1972–1981	1,800	71.15%		908	63.19%	
Exclude county-races with fewer than 25 births in any year 1972–1981	1,748	70.81%	99.92%	636	56.71%	95.84%
(b) Rollout race-specific sample						
Baseline rollout sample	1,520	39.99%		1,520	39.99%	
Exclude county-races missing birth weight or race population data in any year 1970–1975	1,520	39.99%		825	34.43%	
Exclude county-races with fewer than 25 births in any year 1968–1978	1,449	39.90%	99.65%	491	29.69%	91.00%

Notes: Table outlines the sample restrictions imposed upon the baseline samples described in Appendix Table A1 for race-specific analyses.

Table A4: Household Characteristics of Food Stamp Enrollees

Variable	(1) Pre-period (always takers) (1976–1978)	(2) Post-period average (1979–1980)	(3) Complier average (Estimated from Agg FS)
Share AFDC households	0.320	0.298	0.206
	(0.014)	(0.015)	(0.097)
Annual AFDC transfers, given AFDC ($1979)	2,790	2,879	3,244
	(88)	(111)	(671)
Annual income net of housing costs			
Mean income net of housing costs ($1979)	6,338	6,107	5,167
	(179)	(194)	(1,223)
Share with $0	0.011	0.021	0.064
	(0.003)	(0.005)	(0.027)
Share 0–10% of FPL	0.025	0.041	0.107
	(0.004)	(0.006)	(0.038)
Share 11–50% of FPL	0.196	0.214	0.286
	(0.012)	(0.014)	(0.084)
Share 51–75% of FPL	0.222	0.221	0.220
	(0.013)	(0.014)	(0.089)
Share 76–100% of FPL	0.184	0.170	0.116
	(0.012)	(0.013)	(0.083)
1978 Benefits			
Share with $0 purchase requirement	0.027	0.045	0.118
	(0.005)	(0.007)	(0.040)
Households (N)	2,552	2,080	

Total income net of housing costs is defined as post-tax income minus annual rent and mortgage payments, winsorized at the 99th percentile. Table shows pre- and post-period household characteristics for household-years enrolled in food stamps, measured three years before and two years after the free-stamps reform (1976–1980). Column 3 reports mean outcomes among compliers, estimated using the aggregate first stage scaled by baseline average 1972 exposure, which generates a measure of the share of post-reform enrollees who join due to the reform. Standard errors are shown in parentheses and account for estimation uncertainty in both the first-stage estimate (clustered at the county level) and the pre/post means (clustered at the household level). Estimates use PSID weights which sum to the count of enrolled households in each period. See Appendix B.8 for additional details.

Table A5: 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 A6: IV Analysis Scaling by Relative Share of Enrolled Households with Young Children

		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: Baseline							
Post rollout	0.0506						
	(0.0016)						
Post free FS reform	0.0185						
	(0.0015)						
Participation rate		0.0097	-0.0041	-0.0033	-0.0446	-0.0041	-0.0067
		(0.0290)	(0.0120)	(0.0114)	(0.0387)	(0.0134)	(0.0128)
First-stage F		876.2	147.9	510.3	876.2	147.9	510.3
p-value (joint test of instr.)	<0.0001						
Hansen J p-value				0.659			0.321
Baseline Mean	0.0253	8.102	8.106	8.104	0.0732	0.073	0.073
Observations	3351298	3340293	10824	3351117	3340293	10824	3351117
Counties	2853	1520	1804	2853	1520	1804	2853
Panel B: scale by enrollment ratio of families with young child to all population in PSID							
Post rollout	0.0429						
	(0.0013)						
Post free FS reform	0.0157						
	(0.0013)						
Participation rate (imputed)		0.0114	-0.0049	-0.0038	-0.0526	-0.0049	-0.0079
		(0.0342)	(0.0142)	(0.0134)	(0.0456)	(0.0159)	(0.0151)
First-stage F		876.2	147.9	510.3	876.2	147.9	510.3
p-value (joint test of instr.)	<0.0001						
Hansen J p-value				0.659			0.321
Baseline Mean	0.0215	8.102	8.106	8.104	0.0732	0.073	0.073
Observations	3351298	3340293	10824	3351117	3340293	10824	3351117
Counties	2853	1520	1804	2853	1520	1804	2853

Notes: Panel (a) is as in Table 2; see notes for specification and sample definition. Panel (b) rescales the first-stage estimates by the ratio of enrollment rate of families with a young child to the general sample population in the PSID in 1976–1978 (= 0.848).

Table A7: Effects of the Women, Infants, and Children (WIC) and Food Stamp Programs on Birth Weight: Literature Review

			Birth Weight (%)		*Share Low Birth Weight*			
Study	Program	First Stage	ITT	ATE	ITT	ATE	Mean	Methods
Panel A: Pooled races								
Sonchak (2016)	WIC			1.648[a,b] (0.170)		-0.025[a] (0.003)	0.1033	Mother Fixed Effects
Chorniy et al. (2020)	WIC			0.291[a,b] (0.155)		-0.008[a] (0.004)	0.0956	Mother Fixed Effects
Rossin-Slater (2013)[d]	WIC	0.0234 (0.0123)	0.834[b] (0.244)		-0.005 (0.006)		0.06	Mother Fixed Effects IV clinic access
Hoynes et al. (2011)	WIC	0.08	0.081[b] (0.037)	1.018 (0.456)	-0.000072 (0.0004)	-0.0009 (0.005)	0.072	Staggered rollout
Precision-weighted estimate	WIC			0.9752 (0.3096)		−0.0119 (0.0058)		
East (2020)	Food Stamp Program	0.093 (0.051)	0.198[b] (0.090)	2.135 (0.969)	-0.001 (0.001)	-0.011 (0.011)	0.07	DiD, eligibility change
Free Stamps Reform (this paper)	Food Stamp Program	0.0185 (0.0015)	-0.00812[c] (0.0226)	-0.409[c] (1.195)	-0.0013 (0.0044)	-0.0041 (0.0134)	0.0730	Exposure design
Food Stamps Rollout (this paper, AHS)	Food Stamp Program	0.0506 (0.0016)	-0.0400[c] (0.0100)	0.9848[c] (2.9286)	-0.0014 (0.0020)	-0.0448 (0.0387)	0.0732	Staggered rollout
Precision-weighted estimate	Food Stamp Program			1.1168 (0.7290)		−0.0098 (0.0082)		
Panel B: Black								
Sonchak (2016)	WIC			2.074[a,b] (0.250)		-0.034[a] (0.005)	0.1273	Mother Fixed Effects
Chorniy et al. (2020)	WIC			0.527[a,b] (0.221)		-0.016[a] (0.006)	0.123	Mother Fixed Effects
Bitler et al. (2025)	WIC	0.36	0.3801[b] (0.1304)	1.056 (0.3622)	-0.0031[d] (0.0011)	-0.0086 (0.0030)	0.1343	Staggered rollout
Precision-weighted estimate	WIC			1.2188 (0.3693)		−0.0191 (0.0061)		
Free Stamps Reform (this paper)	Food Stamp Program	0.0421 (0.0050)	0.0041[c] (0.0476)	1.2376[c] (1.2149)	-0.0004 (0.0106)	-0.0103 (0.0175)	0.128	Exposure design
Food Stamps Rollout (this paper, AHS)	Food Stamp Program	0.1353 (0.0046)	0.5616[c] (0.2715)	5.2533[c] (2.7682)	-0.0119 (0.0043)	-0.0949 (0.0512)	0.1324	Staggered rollout
Precision-weighted estimate	Food Stamp Program			1.5519 (1.1272)		−0.0171 (0.0163)		
Panel C: White								
Sonchak (2016)	WIC			1.188[a,b] (0.227)		-0.014[a] (0.004)	0.0744	Mother Fixed Effects
Chorniy et al. (2020)	WIC			0.0995[a,b] (0.2217)		-0.0004[a] (0.0045)	0.0653	Mother Fixed Effects
Bitler et al. (2025)	WIC	0.15	0.1534[b] (0.0564)	1.0229 (0.3762)	-0.0006[d] (0.0004)	-0.0039 (0.0028)	0.0643	Staggered rollout
Precision-weighted estimate	WIC			0.7381 (0.2722)		−0.0061 (0.0033)		
Free Stamps Reform (this paper)	Food Stamp Program	0.0075 (0.0010)	-0.0400[c] (0.0232)	-5.0481[c] (2.9340)	0.0004 (0.0038)	0.0055 (0.0285)	0.060	Exposure design
Food Stamps Rollout (this paper, AHS)	Food Stamp Program	0.0250 (0.0006)	-0.0400[c] (0.1010)	-2.5665[c] (6.2552)	-0.0009 (0.0014)	-0.0286 (0.0824)	0.0664	Staggered rollout
Precision-weighted estimate	Food Stamp Program			−4.8295 (2.7219)		0.0024 (0.0270)		

Notes: [a] ATE directly reported. [b] Reported birth weight effects are in levels, so percentage effects are calculated as 100 times the estimated gram effect divided by the reported baseline mean. [c] For the Free-Stamps Reform and Food Stamps Rollout (this paper), percentage effects are calculated by exponentiating the coefficient on log birth weight. [d] The t-statistic for the relevant first-stage estimate is < 2, so we do not calculate an ATE. Precision-weighted estimates follow Ganong et al. (2026), and the food stamp program race-specific precision-weighted estimates reflect the pooled analysis estimates in Table 2. See Appendix A.3 for additional details and sources.

Table A8: Estimates of Inframarginality Relative to 1978 Benefit Schedule for R and F

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	0.81	0.31
	(0.01)	(0.02)	(0.12)
Food spending $\geq 0.9 \cdot F$	0.95	0.86	0.51
	(0.01)	(0.02)	(0.10)
Food spending $\geq 1.1 \cdot F$	0.88	0.76	0.26
	(0.02)	(0.02)	(0.14)
Panel B: Share strongly inframarginal			
Food spending $\geq R + F$	0.63	0.43	-0.37
	(0.04)	(0.03)	(0.21)
Food spending $\geq 0.9 \cdot (R + F)$	0.83	0.53	-0.68
	(0.02)	(0.03)	(0.19)
Food spending $\geq 1.1 \cdot (R + F)$	0.49	0.35	-0.24
	(0.04)	(0.03)	(0.22)

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 1977–1980 Nationwide Food Consumption Survey. In all years, net income and benefit amounts are determined using the regulations in effect in 1978, as described in Appendix A.2.2. 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. 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). All amounts are in $1978. See Appendix A.2.2 for more details.

Table A9: 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 28, 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 by county. Standard errors are in parentheses. Note that race group is determined by the race of the mother in each birth observation.

Table A10: 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 × 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 5–8). 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 28, with the additional controls as indicated. See Appendix Table A9 for more details on sample and included fixed effects.

Table A11: 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	28	28	28	28	28	28	29	29	30	30	30	30	30	30
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 28, 29, and 30. 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 Appendix Table A9. 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" means we use the weights recommended by Wing et al. (2024); see Appendix B.5. 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 enrollment 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 counties with any county-race-years with fewer than 25 births, instead of dropping race-county-quarter cells with fewer than 25 births, at Column 11. Across all columns, standard errors are all clustered by 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 A12: 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	28	28	28	28	28	28	29	29	30	30	30	30	30	30
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 28, 29, and 30. 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 Appendix Table A10. 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" means we use the weights recommended by Wing et al. (2024); see Appendix B.5. 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 enrollment 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 counties with any county-race-years with fewer than 25 births, instead of dropping race-county-quarter cells with fewer than 25 births, at Column 11. Across all columns, standard errors are all clustered by 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 A13: 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	28	28	28	29	29	30	30	30	30	30	30
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 28, 29, and 30. 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 a specification analogous to Column 5 in the replication of AHS in Appendix Table A9. 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" means we use the weights recommended by Wing et al. (2024); see Appendix B.5. 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 enrollment 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 counties with any county-race-years with fewer than 25 births, instead of dropping race-county-quarter cells with fewer than 25 births, at Column 8. Across all columns, standard errors are all clustered by 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 A14: 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)	W_1 (14)
1. **Baseline estimate**	**0.803**	**0.197**	**3.22**	**6.17**	**0.696**	**0.175**	**0**	**0**	**0**	**0**	**0.593**	**2.24**	**1.08**	**2.01**
Distortions (D_θ)														
2. Distortion equal to valuation	0.803	0.197	3.22	6.17	0.696	0.175	0.696	0.175	0	0	0.593	4.48	2.16	4.02
3. Minimal distortion to overturn results	0.803	0.197	3.22	6.17	0.696	0.175	0.393	0.393	0	0	0.593	3.50	3.50	3.50
4. Minimal complier distortion to overturn results	0.803	0.197	3.22	6.17	0.696	0.175	0	0.188	0	0	0.593	2.24	2.24	2.24
Welfare Weights (λ_θ)														
5. Double complier welfare weights	0.803	0.197	3.22	12.34	0.696	0.175	0	0	0	0	0.593	2.24	2.16	2.22
6. Log utility	0.803	0.197	1.87	2.51	0.696	0.175	0	0	0	0	0.593	1.30	0.44	1.13
7. CRRA parameter = 50	0.803	0.197	3.30	10.86	0.696	0.175	0	0	0	0	0.593	2.30	1.90	2.22
8. Welfare weight = 1 iff income $<$ 50% FPL	0.803	0.197	2.31	13.64	0.696	0.175	0	0	0	0	0.593	1.61	2.39	1.76
9. Welfare weight = 1 iff income $<$ 63.1% FPL	0.803	0.197	2.84	11.38	0.696	0.175	0	0	0	0	0.593	1.97	1.99	1.98
Net Private Value (V_θ)														
10. Upper bound for V_a, V_c (with bonus reform)	0.803	0.197	3.22	6.17	0.871	0.350	0	0	0	0	0.769	2.81	2.16	2.68
11. Lower bound for V_a, V_c (with bonus reform)	0.803	0.197	3.22	6.17	0.520	0	0	0	0	0	0.417	1.67	0.00	1.34
12. Lower bound for V_a (with bonus), upper bound for V_c	0.803	0.197	3.22	6.17	0.520	0.350	0	0	0	0	0.486	1.67	2.16	1.77
13. V_a, V_c midpoint of bounds from subsidy rate	0.803	0.197	3.22	6.17	0.718	0.140	0	0	0	0	0.604	2.31	0.86	2.03
14. Midpoint of bounds from subsidy rate (with bonus)	0.803	0.197	3.22	6.17	0.570	0.140	0	0	0	0	0.485	1.84	0.86	1.64
First Stage (σ)														
15. Smallest first stage (state$\times$year fixed effects)	0.883	0.117	3.22	6.17	0.696	0.175	0	0	0	0	0.635	2.24	1.08	2.10
16. Largest first stage (log female population 15–44$\times$year)	0.794	0.206	3.22	6.17	0.696	0.175	0	0	0	0	0.588	2.24	1.08	2.00
Fiscal Externalities (FE)														
17. Conservative FE from Hoynes and Schanzenbach (2012) and Hendren and Sprung-Keyser (2020)	0.803	0.197	3.22	6.17	0.696	0.175	0	0	0.160	0.160	0.593	1.93	0.93	1.73
18. Conservative FE, always-takers only	0.803	0.197	3.22	6.17	0.696	0.175	0	0	0.160	0	0.593	1.93	1.08	1.78
19. Necessary FE to overturn result, always-takers only	0.803	0.197	3.22	6.17	0.696	0.175	0	0	1.073	0	0.593	1.08	1.08	1.08

Notes: 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 are assumed to be zero for both compliers and always-takers. We present sensitivity analyses varying the first-stage estimate, welfare weights, and distortion assumptions.

- $W_\theta = \frac{\lambda_\theta(V_\theta + D_\theta)}{1+FE_\theta}$ corresponds to the welfare ratio for type θ. W_a is equivalent to $W(0.65)$ from Equation 14.
- Rows 2–4 make assumptions about the distortions. Row 2 assumes distortion equals private valuations. Rows 3 and 4 assume the minimal distortion necessary to overturn the result.
- Rows 5–9 make different assumptions about welfare weights. Row 5 doubles compliers' welfare weights λ_c. Rows 6 and 7 re-estimate λ_c if we use CRRA utility with a risk aversion parameter of 1 or 50, respectively. Rows 8–9 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 8) or below 63.1% FPL (Row 9) and 0 otherwise. Row 9 shows the highest cutoff that would suffice to overturn our result under this class of welfare weights; W_c and W_a are not exactly equal because these results come from simulations using the PSID data, where the finite sample prevents us from choosing a cutoff that equalizes the two values exactly.
- Rows 10–14 assume different estimates of V_a and V_c that are consistent with revealed preference in Figure 9.
- Rows 15 and 16 make different assumptions about σ_c, based on robustness specifications for the first stage (Appendix Figure A18).
- Our primary estimate for fiscal externalities is 0. In Rows 17 and 18, we use a fiscal externality of 0.16 (Hoynes and Schanzenbach, 2012; Hendren and Sprung-Keyser, 2020). Row 19 presents the always-taker fiscal externality sufficient to overturn the result if compliers have zero FE.