The Welfare Effects of Eligibility Expansions: Theory and Evidence from SNAP *

Jenna Anders†
Charlie Rafkin‡

January 12, 2021

Please click here for the latest version.

Abstract

We study the U.S. rollout of eligibility expansions in the Supplemental Nutrition Assistance Program and show that expanding eligibility raises enrollment among the inframarginal (always-eligible) population. This evidence motivates a general model of the optimal eligibility threshold for welfare programs with incomplete take-up, where inframarginal take-up responds to that threshold. The optimal threshold depends on the size of the inframarginal response and the extent to which information frictions and stigma influence that response. An online experiment provides evidence that the eligibility threshold affects stigma. Given our empirical results and certain modeling assumptions, the SNAP eligibility threshold is lower than optimal.

*We thank Hunt Allcott, Judi Bartfeld, Leo Bursztyn, Clément de Chaisemartin, Raj Chetty, Jon Cohen, John Conlon, Zoë Cullen, Esther Duflo, Amy Finkelstein, Jon Gruber, Benny Goldman, Craig Gundersen, Basil Halperin, Emma Harrington, Nathan Hendren, Lisa Ho, Jeffrey Liebman, Stephen Morris, Whitney Newey, Ben Olken, Emily Oster, Amanda Pallais, Dev Patel, Indira Puri, Frank Schulbach, Amy Ellen Schwartz, Jesse Shapiro, Evan Soltas, Dmitry Taubinsky, and participants at workshops at Harvard and MIT for helpful discussions. We thank Rian Flynn and Amber Zheng for excellent research assistance. We are grateful to Thalia Rubio for editing. We thank Timothy Harris for sharing data on SNAP work requirement waivers. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1122374 and Grant No. 1745303 and by Harvard’s Foundations of Human Behavior Initiative. The online experiment was pre-registered at the AEA RCT Registry under AEARCTR-0005566, and survey instruments are available on the authors’ websites. The experiment received exempt status from MIT’s Committee on the Use of Humans as Experimental Subjects (#E1962) and Harvard’s Institutional Review Board (#IRB20-0326).

†Department of Economics, Harvard University.
‡Department of Economics, MIT.
1 Introduction

Social programs in the United States are characterized by incomplete take-up, and there is substantial heterogeneity in take-up across programs. Meanwhile, there is also large heterogeneity in eligibility criteria across programs. In fact, in some social programs, such as the Supplemental Nutrition Assistance Program (SNAP, also known as food stamps), the eligibility threshold even varies across states. When viewed in tandem, the relationship between eligibility criteria and take-up is striking: programs with less stringent income eligibility thresholds have higher take-up rates (Figure 1).

This relationship between programs’ take-up rates and eligibility is consequential. In the simplest model of how to set eligibility thresholds, policymakers trade off giving larger benefits to only the poorest people or spreading the benefit more thinly to a larger number of people. But if eligibility thresholds affect take-up within the eligible population, the policymaker no longer faces this basic trade-off alone. Targeting benefits only to the poorest households could decrease take-up for these groups, relative to what would occur under a broader eligibility rule. Of course, the across-program correlation between eligibility rules and take-up among the eligible population could be driven by the unique contexts, enrollment rules, and histories of these programs. As a result, it is important to determine whether there is a causal relationship between the eligibility threshold and take-up of the already-eligible population.

Does the eligibility threshold affect take-up of social programs? If so, how does this phenomenon affect programs’ optimal eligibility? In this paper, we provide novel evidence that the eligibility threshold affects take-up among low-income individuals who are always eligible for SNAP, regardless of the threshold. We then propose a general model of welfare program participation that allows us to study optimal policy when the eligibility threshold endogenously affects take-up.

We focus on SNAP for several reasons. First, it is a large program (with an annual budget of about $70 billion) that forms an important part of the U.S. public assistance system. Second, SNAP eligibility rules are at the center of an ongoing public discussion. Third, SNAP publishes anonymized public-use administrative data (via the Department of Agriculture’s Quality Control files), which we use to form our main outcome of log enrollment counts. The administrative data alleviates concerns that the results could reflect the mismeasurement of individuals’ eligibility status or program participation reporting biases.

We begin by providing evidence that eligibility expansions in SNAP raised enrollment among the lowest-income individuals that are always SNAP-eligible. States can choose to expand SNAP eligibility standards beyond the federal minimum of 130% of the Federal Poverty Level (FPL). We focus our analysis on individuals at 50–115% of the FPL, a group of individuals eligible for SNAP in every state because they are poorer than the federal minimum eligibility. Leveraging an event-study design (using variation across

---

1We limit to programs with a budget of $5 billion or more; Appendix B gives details about constructing this figure.

2Moreover, if relatively less poor households enroll at higher rates, this could itself explain the relationship.

3For example, the Trump administration proposed eliminating state discretion in setting eligibility thresholds. The proposed change is in Federal Register 2019.
states and years), we find that raising the eligibility threshold by 10 percent of the FPL (e.g., from 130% to 140%) boosts enrollment by over 1 percent among the inframarginal group that was always eligible for SNAP. This take-up response among the inframarginal population is consistent with a small literature documenting a similar phenomenon for public health insurance programs, where it is called a “welcome-mat effect” or a “woodwork effect” (because already-eligible individuals appear “out of the woodwork” to take up the health program). We use the term “inframarginal effects” to avoid negative or positive connotations. As another way of benchmarking the magnitude, we find that for every person who joins SNAP because she becomes newly eligible, 0.9 inframarginal people join the program.\footnote{By contrast, in Sacarny et al. (2020)’s analysis of inframarginal effects in the Oregon Health Insurance Experiment, about 0.1 inframarginal children join the program for every new person who joins. Section 2 discusses several reasons why the inframarginal effects may be much larger in SNAP.}

Our setting also yields a clean placebo test: the policy change that permits states to change their SNAP eligibility threshold also gave other bureaucratic benefits to states, and we show that states which adopted the policy \textit{without} expanding SNAP eligibility saw no increases in SNAP enrollment.

Having provided evidence of quantitatively large inframarginal effects in an important public assistance program, we propose a general economic framework for analyzing optimal eligibility in the presence of inframarginal effects. Individuals who are eligible for a welfare program take up the program benefit as long the benefit exceeds a private take-up cost (e.g., stigma) and they are aware of the program.\footnote{While we focus on stigma costs, our framework permits any cost that depends on the eligibility threshold. Another possible cost embedded within our framework is uncertainty about eligibility, as in e.g. Kleven and Kopczuk (2011), which we discuss in Section 2.} Both the cost and information (awareness) can depend on the eligibility threshold. The planner trades off (i) a standard redistributive motive in which she values giving a bigger benefit to people with higher welfare weights against (ii) a new motive, inframarginal effects, in which relaxing eligibility thresholds raises take-up.\footnote{In our benchmark model, we hold labor supply constant, but we show in an appendix that similar intuitions apply in a more elaborate environment.} We derive an optimality condition for the eligibility threshold in which our key empirical fact, the inframarginal effect, enters as an observable elasticity (Chetty, 2009; Kleven, Forthcoming). Moreover, we prove that a “naïve” planner who ignores inframarginal effects but otherwise behaves optimally will often (although not always) set the eligibility threshold too low (or equivalently, the benefit size too high) relative to a “sophisticated” planner who is aware of inframarginal effects.\footnote{We discuss the precise conditions for this statement in Section 3. Intuitively, we require that if stigma is an important mechanism, people who would have signed up for SNAP regardless enjoy a reduction in stigma costs when the eligibility threshold increases. We also require that people whose behavior is governed by stigma are not much more elastic to the eligibility change than people whose behavior is governed by lack of information.}

Similar to in a Baily (1978)-Chetty (2006) framework, the optimality condition features a fiscal externality of the inframarginal effects and recipients’ willingness to pay (WTP) for a higher eligibility threshold. Recipients’ WTP depends on why inframarginal effects exist. If inframarginal effects are mostly driven by individuals responding to changing costs (e.g., eligibility expansions reduce stigma), the new enrollees behind the inframarginal effects are just indifferent between taking up and not, so they do not value the eligibility expansion; however, those who would have enrolled regardless now pay lower stigma costs to take up. If inframarginal effects reflect improved optimization (e.g., eligibility expansions increase information...
mation, or awareness of the program), the new enrollees capture the entire benefit of the program, while the previously enrolled do not gain. Either stigma or information can serve as a force to raise the eligibility threshold. However, if the inframarginal effects arise mostly due to increased awareness, the model’s normative conclusion that the planner should raise the eligibility threshold is generally strengthened.\(^8\)

To shed light on whether stigma could contribute to inframarginal effects, we conducted an online experiment with a nationally representative sample of more than 2,000 participants in March 2020. In particular, we provide information about the eligibility threshold in one state to influence beliefs about the mean eligibility threshold across states.\(^9,10\) The treatment increases participants’ beliefs about the share of individuals who are eligible for SNAP in the entire U.S. by 9 percentage points on average (standard error: 0.8 pp). We elicit beliefs about whether others stigmatize SNAP (“second-order” stigma) and whether the participant herself stigmatizes SNAP (“first-order” stigma). Our treatment decreases an index of second-order stigma by \(-0.050\) standard deviations (SE: 0.027, \(p = 0.061\)), consistent with a stigma mechanism. On the other hand, we find no effect on first-order stigma, perhaps because one’s own views about SNAP are difficult to move in a light-touch information treatment.\(^11\) While of course subject to concerns about external validity, we view the results from the online experiment as providing evidence that stigma costs could play a role in inframarginal effects.

Combining our model with the empirical estimates of the inframarginal effect, we find that the social planner makes a quantitatively large mistake if she ignores the inframarginal population’s utility gains from a higher eligibility threshold. As noted, under natural assumptions, our propositions developed in the model delivers that the social planner will set the eligibility threshold too low if she ignores inframarginal effects.\(^12\) But how large will the planner’s mistake be?

In our welfare analysis, we employ three distinct approaches, each offering advantages and disadvantages. First, using only the local optimality condition, we provide a numerical benchmark of the importance of the inframarginal effect for welfare. Under our benchmark assumptions, the naïve planner who ignores inframarginal effects acts “as if” the agents in the model are 30% less risk averse than they really are. Consequently, the naïve planner over-redistributes because she substantially underestimates the value of expanding eligibility. Second, we generalize this conclusion with the Marginal Value of Public Funds...
framework (Hendren, 2016; Hendren and Sprung-Keyser, 2020): neglecting inframarginal effects can cause the naïve planner to substantially underestimate the welfare impact of an eligibility expansion, although results depend on the calibration. Finally, we impose more structure to solve for a globally optimal eligibility threshold and show that the optimal threshold will be as much as 10% too low if the planner ignores inframarginal effects.

Throughout the welfare analysis, we assume that stigma and information both play a role in inframarginal effects. Importantly, we present robustness to both mechanisms and show that the general conclusions are insensitive to the mechanism. While we view the experiment as a useful complement that enriches our understanding of inframarginal effects, we emphasize that these normative conclusions by no means hinge on the experiment’s results. Our normative conclusions generally hold under weaker conditions if there is no program stigma.\textsuperscript{13} Thus, as we show in robustness exercises, if one is not persuaded by the experiment and wishes instead to suppose that information frictions are entirely responsible for inframarginal effects, that leads to an even stronger conclusion that the planner should raise the eligibility threshold given our modeling assumptions.

**Contributions and related literature.** Every social program makes some determination about program eligibility (even if the program is universal). Yet much of the vast literature on program design focuses on other policy instruments besides the eligibility threshold. Figure\textsuperscript{2} quantifies this statement. We collected all 278 papers published in the *American Economic Review* between 2010–2018 and the *Quarterly Journal of Economics* between 2010–2019 that met one of 33 search terms about social welfare programs (see Appendix\textsuperscript{B} for details). After reading the abstracts and/or introductions of these papers, we concluded that 76 of them were primarily about effects or design of social welfare programs, 49 of which involved the study of a specific policy instrument. Yet only 7 (14% of the 49) examined the eligibility criteria as a policy tool (i.e., as an instrument the planner could use). On the other hand, 25 of the 76 papers about welfare programs consider eligibility thresholds as a source of variation for estimating the program’s treatment effect. Put another way, while economists regularly exploit eligibility thresholds for causal inference, they are often neglected as an aspect of optimal program design. Our paper is among the first to combine empirical estimates of endogenous take-up from eligibility thresholds with a theoretical model that permits welfare analyses of current program rules.

Our work therefore advances several literatures. First, we add to the large body of research in public economics that deals with the optimal design of social programs. Much of this work considers the optimal benefit level when take-up is distorted by moral hazard; for example, an influential series of papers considers the optimal unemployment insurance (UI) benefit when the benefit size distorts UI duration (Baily, 1978; Gruber, 1997; Chetty, 2006, 2008). Analyses of optimal eligibility are rarer.\textsuperscript{14,15} For instance, we depart from

---
\textsuperscript{13} An exception to this statement appears in the analysis of the MVPF, which we describe in more detail in Section 5.

\textsuperscript{14} Fetter and Lockwood (2018) is a recent example that studies optimal eligibility for old-age insurance. Other papers, e.g. Diamond and Sheshenski (1995), Golosov and Tsyvinski (2005), and Low and Pistaferri (2015), study optimal eligibility in the context of disability insurance.

\textsuperscript{15} Our discussion of restricted eligibility departs the most common prior motivation for restricting eligibility, from Nichols and
Kroft (2008), who considers social spillovers in UI, by considering the instrument of optimal eligibility as a means of using inframarginal effects to achieve socially desirable outcomes.

Two related papers, Kleven and Kopczuk (2011) and Hanna and Olken (2018), model the means test as an instrument for optimal program targeting in the presence of exclusion (Type I) and inclusion (Type II) errors. Hanna and Olken (2018) consider a screening problem when the government cannot measure income well, as is more common in the developing world. Kleven and Kopczuk (2011) provide a theoretical model of a social planner who can use eligibility rules, in addition to program complexity and benefit size, to screen recipients. Our model differs from Kleven and Kopczuk (2011) by emphasizing how the eligibility threshold might directly affect stigma and information.

Other related work considers three categories of reasons for incomplete program take-up: hassle costs, incomplete information, and stigma (Moffitt, 1983; Currie, 2004; Heckman and Smith, 2004; Aizer, 2003; Bhargava and Manoli, 2015; Friedrichsen et al., 2018; Finkelstein and Notowidigdo, 2019). Our experiment contributes to this literature by providing clean evidence that aspects of program design may affect stigma costs.

Second, we link research on optimal program design to the growing literature in behavioral public economics (Bernheim and Taubinsky, 2018). More generally, this paper suggests that individuals’ utility depends on social norms, and government policy plays an important role in shaping these norms. Economists have only begun to explore how policy may influence psychological forces like shame or guilt, which may in turn may have important consequences for social welfare. For instance, we provide empirical support for the claim, promulgated in the sociology and historical literature, that programs like Social Security are not stigmatized precisely because they are not means tested. We also conduct welfare analysis in the presence of non-classical behavior, as in, for example, Finkelstein and Notowidigdo (2019) or Butera et al. (2020).

Third, we contribute to the study of the Supplemental Nutrition Assistance Program, the subject of a wide-ranging literature within economics. We draw on the data used in Ganong and Liebman (2018), who study how changes in the economic environment, coupled with changes in SNAP program design, affected SNAP enrollment through 2012. Other work studies how SNAP receipt affects household members’ nutrition, health or other outcomes (Almond et al., 2011; Hoynes et al., 2016; Bronchetti et al., 2019; Bailey et al., 2018).

Zeckhauser (1982), which suggests that limiting program participation can induce self-targeting; later work has examined whether introducing ordeals can improve the targeting properties of social programs (Alatas et al., 2016; Deshpande and Li, 2019; Gazze, 2019; Gray et al., 2019).

This literature includes several papers studying stigma surrounding SNAP take-up and the rollout of the Electronic Benefits Transfer (Daponte et al., 1999; Currie and Grogger, 2001; Atasoy, 2009; Klerman and Danielson, 2011; Manchester and Mumford, 2012; Eck, 2018). Alsan and Yang (2019) study how fear of deportation affects SNAP and Supplemental Security Income take-up in Hispanic communities.

For instance, Lindbeck et al. (1999) also provides a theoretical treatment of why program design may shape norms.

In a history of the U.S. welfare state, Katz (1986) writes, “With a strong, articulate middle-class constituency, social insurance, especially social security, carries no stigma… public assistance, which has become synonymous with welfare, is, of course, restricted to the very poor. Its recipients carry the historic stigma of the unworthy poor, and, as a consequence, they are treated meanly” (p. ix). Supporters of Universal Basic Income advance similar arguments.

Currie (2003) provides a review of the U.S. food assistance programs and Bartfeld et al., eds (2016) gives extensive coverage to additional research on SNAP.
No prior work has considered optimal eligibility in SNAP. Finally, we contribute to the small literature on inframarginal effects. Despite their normative implications, these effects have received little attention in public economics. The health literature on Medicaid expansions finds evidence of inframarginal effects (Aizer and Grogger, 2003; Sommers and Epstein, 2011; Frean et al., 2017; Sacarny et al., 2020), but it has not considered their implications for optimal program design.21,22

Roadmap. Section 2 establishes evidence of an inframarginal effect in SNAP. Section 3 proposes a model of optimal eligibility thresholds with inframarginal effects. Section 4 presents our online experiment. Section 5 presents welfare analysis. Section 6 concludes.

2 Infra marginal Effects in SNAP

This section documents the empirical relationship that motivates this paper. States that have less stringent eligibility standards tend to have higher take-up in SNAP among inframarginal people — people whose incomes are low enough that they are eligible everywhere, regardless of the state’s eligibility threshold. Appendix B provides more information about the dataset construction.

2.1 SNAP Data

We obtain the total number of people who participate in SNAP from the SNAP Quality Control (QC) files, which are administrative data from the U.S. Department of Agriculture on a random sample of SNAP participants (United States Department of Agriculture Food and Nutrition Service, 2019). The data record granular information about household characteristics, benefit size, and incomes of SNAP participants. Using these files, we construct the total counts of program participants and are below a given income threshold, for each state and year from 1996 to 2016.23,24 The Quality Control files are administrative data, so they

---

20 Ratcliffe et al. (2008) studies the effect of categorical eligibility on SNAP take-up but does not examine the effect of eligibility thresholds.

21 There is little evidence that inframarginal effects would generalize outside the Medicaid setting. Much of the inframarginal effects documented in the health literature pertain to within-household take-up of the already-eligible population — for instance, new Medicaid take-up among children who are already eligible because children face less stringent Medicaid requirements than adults, as in Sacarny et al. (2020). By contrast, we show that entire households that were already eligible may sign up when eligibility requirements are relaxed. Moreover, our experiment is the first evidence that stigma may contribute to inframarginal effects.

22 Outside of the health literature, Leos-Urbel et al. (2013) find that eligibility expansions boost take-up among inframarginal recipients of a free school breakfast program. They study a reform that granted universal eligibility for school breakfast in New York City; programs with universal eligibility may be very different than programs like SNAP where eligibility remains restricted. Moreover, program take-up among children may be subject to very different social dynamics and information frictions than among adults.

23 We construct our dataset by modifying the publicly available replication code for Ganong and Liebman (2018).

24 We end the sample in 2016 because the SNAP Policy Database (which provides our measures of SNAP policy variation) stops in that year.
record people’s incomes and household size accurately, thereby addressing concerns about measurement error from [Meyer et al. 2015] and others.

**Sample and outcomes.** Our main sample comprises individuals with household income between 50%–115% of the FPL. We focus on individuals above 50% of the FPL because take-up is very high among individuals below 50% of the FPL, regardless of the state’s eligibility threshold. As a result, there is little scope for increased take-up among this group. We also exclude individuals between 115%–130% of the FPL to address concerns about measurement error: we might consider an individual near the threshold as “inframarginal” when in fact she would be ineligible under a different eligibility regime because of additional restrictions such as asset tests. We also construct a sample of the entire inframarginal population, i.e. individuals at 0–130% of the FPL.

Using this sample, our main outcome is a measure of take-up *counts*. In particular, we use log total enrollment within specific income groups — in our main regression, among people earning 50–115% of the FPL. Almost all individuals in this range are eligible for SNAP in every state. This allows us to study inframarginal recipients; we are not counting increased enrollment among people who are newly eligible. Compared to take-up rates, this outcome has the advantage of not involving imperfect measures of the share of people who are eligible for the program (e.g., the total number of people between 50–115% of the FPL) as the outcome variable.

We form take-up *rates* as a secondary outcome. Following [Ganong and Liebman 2018], we divide the number enrolled (from the QC data) by the number of people within a given band of the income distribution in the state from the Current Population Survey’s Annual Social and Economic Supplement (CPS ASEC) ([Ruggles et al. 2020]). For instance, in our main specification, the denominator is the number of people in the CPS who are between 50–115% of the FPL. Crucially, the denominator does not exclude people who are otherwise ineligible for SNAP due to work requirements or asset histories. Thus the take-up rates are likely underestimates. There also may be measurement error in reported incomes in the CPS. We show that measurement error in the CPS data cannot explain our results in Section 2.6.

Just as in other work estimating SNAP take-up, a possible limitation to the analysis is that we do not always observe household assets or work histories, which can affect SNAP eligibility. First, federal rules restrict households with sufficient assets from participating in SNAP. In practice, only a small fraction of households are ineligible for SNAP under these asset histories. Second, under the Personal Responsibility and Work Opportunity Act (PRWORA), single households must meet certain work requirements to participate. However, the changes in these requirements do not coincide with changes in the eligibility threshold, and we also show that the results are similar among households with dependents. We also control for the requirements in robustness checks.

**Policy changes.** Federal rules require that households below 130% of the FPL are eligible to participate in SNAP. Beginning in 2001, states had the option to expand eligibility to additional households up to 200%
of the FPL under Broad-Based Categorical Eligibility (BBCE). Not every state that adopted the BBCE took the option to expand the eligibility threshold. In Section 2.6, we note that adopting the BBCE did entail additional changes to state welfare programs, but we reject that these changes can explain the inframarginal effects we document here. Ultimately, 30 states expanded SNAP eligibility through the BBCE through 2016, four of which adjusted eligibility twice during this period (Figure A.1). Expansions occurred throughout the period, but they were especially likely to occur in 2001–2002 and 2010–2011. The states that do roll out an eligibility expansion are generally distributed across the country, although there are no states in the Great Plains region that implement an expansion (Figure A.2).

2.2 Econometric Strategy

We estimate an event-study regression that leverages the variation in eligibility provided through the BBCE. We index each event by event-time $t$, where $t = 0$ represents the year of the event. We set $t = -1$ in all years for untreated states. We define the “event eligibility rate” in each state $s$ as the eligibility rate as a percent of the FPL after the BBCE expansion in treated states and the federal minimum (130%) in untreated states. We use a balanced panel: we limit the sample to the five years before and after treatment for treated states, and include all years in control states. We normalize our coefficients relative to the year before the event and estimate:

$$y_{s,t,r} = \sum_{r \in R} \eta' (1(t = r)_{s,t,r} \times \text{event eligibility rate}_s) + \delta_s + \gamma_t + X'_{s,t,r} \phi + \epsilon_{s,t,r}$$

(1)

where $R$ is the set of event periods, $s$ indexes states, $t$ indexes years, event eligibility rate $s$ measures the eligibility rate as a ratio of the FPL, $\delta_s$ is state fixed effects, and $\gamma_t$ is year fixed effects. We include $X$, a vector of additional linear controls for the state unemployment rate, the log of the number of people in a given income group in the state (measured in the CPS), SNAP outreach spending per person earning under 130% FPL in the states (transformed with $\sinh^{-1}$), and an index of other SNAP policies implemented around the same time (as in Ganong and Liebman [2018], henceforth the “Ganong-Liebman index”). In our primary estimates, we use $\ln(\text{enrollment}_{s,t,r})$ as the dependent variable ($y_{s,t,r}$). The coefficient of interest $\eta'$ represents the marginal effect of 1 pp increase in the eligibility rate (expressed in terms of the FPL) on enrollment in event-time $r$. This specification encodes a standard pre-trends test for whether $\eta' = 0$ when $r < 0$. Our primary specifications are unweighted. We present standard errors clustered at the state level in this and all subsequent analyses that use state-year variation.

---

25We provide more context about the BBCE in Appendix B. These thresholds correspond to gross income tests. Households must also pass a net income test: net of allowable deductions (e.g., an earnings deduction amounting to 20% of their earned income), their income must be below 100% of the Federal Poverty Level. This is true regardless of the gross income test set by the state.

26In cases in which the BBCE is implemented partway through a year, $t = 0$ represents the first fully treated year.

27We drop the four states with two events in the event-study analysis, as well as the two states that have events too recently to have sufficient post-period data. This leaves 45 states (including the District of Columbia).

28For instance, event eligibility rate $s = 1.3$ represents that the state has the minimum threshold of 130% of the FPL.

29See Appendix B for details on the variables that enter the index.
We also pool the data in this sample to estimate:

$$y_{s,t} = \eta \text{ eligibility rate}_{s,t} + \delta_s + \gamma_t + X_{s,t} \phi + \epsilon_{s,t}. \tag{2}$$

The variable eligibility rate ${s,t}$ represents the eligibility as a percent of the FPL in a given state-year, so $\eta$ is the average effect on inframarginal people after an eligibility expansion.

**Discussion of controls.** Given that our state and year fixed effects remove fixed differences in outcomes across states and across years, the identifying assumption is that there are no time-varying within-state trends in enrollment (not absorbed by our time-varying state controls). One concern is that states that impose the eligibility increase have faster population growth in the inframarginal sample. To address this concern, we control for the log count of the people within the inframarginal income group from the CPS. The economic environment and the policy environment are also relevant for SNAP take-up.\(^{30}\) We control for the state unemployment rate to address the concern that states with eligibility increases may simply have more financial distress, and we include the Ganong-Liebman index to address the concern that states that expand eligibility may also impose other policies relevant to enrollment. We present robustness to additional threats to identification later in this section. Altogether, these controls do not have a dramatic effect on our results. The most important control is for the count of people who are eligible, which we show eliminates a modest (and insignificant) pre-trend in our event study.

### 2.3 Results

**Descriptive evidence.** Before presenting the formal estimates, we begin by visualizing inframarginal effects in the raw data. In Figure 3A, we present total SNAP enrollment per 1,000 people (population-wide) in state-years with eligibility thresholds equal to 130% FPL versus above 130% FPL. We normalize the enrollment by the total population in all states with the relevant income rule to aggregate enrollment counts across states.

First, without the eligibility expansion, very few individuals with household income above 130% FPL take up the program, while with the eligibility expansion, mass appears above 130% FPL where individuals are newly eligible. This confirms that the QC data give sensible estimates of the enrollment counts, and that the eligibility changes relax a binding constraint for some individuals. Second, individuals below the threshold also enroll at higher rates with looser eligibility restrictions. These inframarginal effects — the increased enrollment below the threshold — are the subject of our attention. From here, we analyze the sample of SNAP recipients earning 50–115% of the FPL.\(^{31}\)

Figure 3B presents a binscatter of the cross-sectional relationship between SNAP take-up among these inframarginal individuals (i.e., earning 50–115% FPL) and the state’s eligibility threshold at the state-year

\(^{30}\)For additional discussion of the importance of these factors for SNAP take-up, see Mabli et al. (2014) and Ganong and Liebman (2018).

\(^{31}\)We note a slight excess mass around 75% of the FPL, which may be an artifact of the QC data; however, inframarginal effects appear throughout the income distribution.
level. We observe five different eligibility thresholds chosen by states between 1996–2016. Mean take-up is roughly 10 pp lower in states with eligibility at 130% of the FPL, the most stringent eligibility standard permitted under federal law.

**Event-study specifications.** For confidence that the raw data reflect inframarginal effects and are not driven by confounds, we turn to our event study (Equation (1)). We plot log enrollment among our inframarginal sample by event period, relative to event period -1 (Figure 4A). We find no evidence of pre-trends leading up to the policy change. After the policy change, enrollment increases steadily. Our benchmark estimates suggest that increasing the eligibility level by 10 pp of the FPL boosts the number enrolled by 1-2 percent in the five years following the policy change. The enrollment increase among the 0–130% FPL sample looks very similar (Figure 4B).

To show the effect of controls on our empirical estimates, we present in Figure 4 the event study with state and year fixed effects only (Panel A) and then add the control for the log of the total number of people between 50 to 115% of the FPL (Panel B). Overall the results are similar without controls. With no controls at all (Panel A), we see some visual evidence of a pre-trend prior to treatment, although the trend is small in magnitude and vanishes 3 years before treatment. Once we control for the log of the CPS population totals (Panel B), any pre-trend vanishes, and the results in Panel B are very close to those in Figure 4.

**Placebo.** We conduct a placebo test that offers a useful validation of the above results. We observe nine states implement the BBCE without expanding eligibility beyond 130% of the FPL. Most of these states adopted the BBCE around the same time as did the states in the main event study (2009–2011). This suggests a natural placebo test: we study the effect of the BBCE in the states that did not expand eligibility but did implement the BBCE. To implement the placebo test, we show an event study (as in Equation (1)), where treatment represents states that implemented the BBCE but did not expand eligibility (Figure 5). We use log enrollment among the 50–115% of FPL sample as the dependent variable. This event study yields no effect, confirming that there are not other aspects of the BBCE besides the changing eligibility threshold which drive increased enrollment among the inframarginal population.

**Combined estimates.** The event-study specification and placebo test confirm the existence of inframarginal effects. To obtain the pooled effect over all periods, and parsimoniously present robustness to different specifications, Table 1 estimates Equation (2). Our preferred specification (Column 1) uses the sample used in the event study and includes state and year fixed effects, and controls for the state unemployment rate, outreach spending, and the Ganong-Liebman controls. The independent variable is the eligibility threshold as a ratio of the Federal Poverty Level, so that increasing it by 1 corresponds to increasing the threshold by 100% of the FPL. We find that $\eta = 0.107$ and reject $\eta = 0$ at $p < 0.05$. These estimates can implement the BBCE for bureaucratic reasons, as the policy can simplify program administration, or to relax the SNAP assets test. See Appendix B.

33We exclude states that did increase eligibility from this test, so the regression includes 19 states. A handful of states which adopted BBCE without changing their eligibility thresholds at that point did expand eligibility at a later date. Here, we exclude these states, but the results are similar when they are included and we add a control for the eligibility threshold.

34We control for the inverse hyperbolic sine of outreach spending to address state-years with zero outreach spending. [Burbidge et al. 1998].
suggest that raising the eligibility rate by 10 pp of the FPL (e.g., from 130% to 140%) boosts take up by 1.07 percent. The modal eligibility increase in our sample is from 130% to 200% of the FPL, which delivers a 7.5 percent increase in take-up among this sample ($0.7 \times 1.07 \approx 0.75$). The results in Column 1 are consistent with the event study plot.

The rest of Table 1 shows that our estimate of inframarginal effects is robust to the particular choice of the specification. Column 2 separates the Ganong-Liebman index into separate indicators for each component variable. Column 3 reverts to the index form of these controls but adds new controls for lagged unemployment and the prevalence of waivers relaxing the SNAP work requirements for able-bodied adults without dependents (ABAWDs), beginning in 2010. Column 4 excludes the years 2008–2011 (the Great Recession). Column 5 weights by state-year population. Column 6 computes the treatment effect as the difference between the average of the event study coefficients in the post period and the average of the coefficients in the pre period, weighting all post periods equally. Finally, Column 7 uses all years of data we have (a balanced panel of 50 states and D.C. from 1996–2016), instead of only the event study sample of a 5-year window around the eligibility increase. It also includes states that change eligibility several times or reduce eligibility. Throughout the table, the results are stable: estimates of $\eta$ range from 0.10 to 0.12.

We also repeat the exercise for two different samples in Table A.1 and find similar results. Panel A shows enrollment responses in the 0-130% FPL sample. The estimates are consistent with the main results but generally lower. This attenuation reflects that our dependent variable (SNAP enrollment) has less scope to rise when almost all people from 0–50% of the FPL already take up SNAP. Panel B assesses enrollment among households with children, as these people are likely not subject to the additional ABAWD work requirements that were relaxed and reimposed during the sample period. Here, we see similar sized, though noisier, effects in this sample. Together, these results and Table 1 provide strong evidence of inframarginal effects from the BBCE, as the effect persists across specifications and samples.

**Distributional effects.** We explore treatment effect heterogeneity by household income (Figure 6). We estimate a version of Equation (2), using take-up rates instead of log enrollment counts so that the values are more directly comparable across income groups with different base rates. Take-up rates increase most among those earning 130–160% FPL, who are barely ineligible before an expansion. The effect in this group is larger than the largest effect in the inframarginal population, among those earning 100–130% FPL. However, even after the expansion, take-up in the newly eligible group is still much lower than any other group. We also see that the treatment effect size is increasing with household income within the inframarginal sample; however, this may partially reflect that the base take-up rate is much lower among households with relatively more income.

---

35 We use data on ABAWD waivers from data generously shared by [Harris Forthcoming].
36 Figure A.5B also shows the event study where the sample includes only SNAP recipients in households with children.
2.4 Interpreting the Magnitude of the Results

We now provide three ways of interpreting the magnitude of the results.

Take-up elasticity. We estimate the elasticity of take-up with respect to the share of the population who is eligible. The elasticity will also play a critical role in the theoretical model.

We employ an instrumental variables approach to estimate this elasticity. The share of the population eligible for SNAP is affected by confounding conditions which also affect the number of people below a certain income level. The eligibility expansions provide plausibly exogenous shocks to the share eligible. Thus we instrument for the log share eligible for SNAP using the state-and-year-specific income cutoff as a ratio of the Federal Poverty Level. The exclusion restriction is that eligibility expansions are not associated with take-up of inframarginal people except through changes in the share eligible.

We return to Equation (2) from Section 2. We use a log-log specification, with \( \ln(\text{take-up}) \) and \( \ln(\text{share eligible}) \) as the dependent and independent variables, respectively. The estimating equation is:

\[
\ln(\text{take-up})_{s,t} = \eta \ln(\text{share eligible})_{s,t} + \chi'_{s,t} \phi + \delta_t + \gamma_t + \varepsilon_{s,t},
\]

where we instrument for \( \ln(\text{share eligible}) \) using the state eligibility threshold as a ratio of the FPL. Here \( \eta \) represents an elasticity rather than a level effect.

We present the IV estimates using all the data (Table 2, Panel A) and the event-study sample (Panel B). We document a strong first stage: in the full sample, increasing the eligibility threshold by 10% of the FPL increases the share of a state population that is eligible by 7.28% (\( t \)-stat = 21.37), with similar results for the event-study sample. Our 2SLS estimate in the full sample is \( \eta_m = 0.161 \) (SE: 0.061); the estimate in the event-study sample is \( \eta_m = 0.097 \) (SE: 0.070). We also document that simple OLS regressions of log take-up on the log share eligible have the opposite sign, likely due to the omitted variables bias we described above.

The exact coefficient we estimate for \( \eta_m \) depends on whether we use the event study sample from Section 2 or all the data at our disposal. The limited sample is noisier, since we employ an instrument and rely on imprecise denominators (i.e., number of people eligible and number of people in a given income group) from the CPS. The full sample is more precise and larger, although the estimates are not statistically distinguishable.\(^{37}\)

Comparison to inframarginal effects in Medicaid. We now convert our inframarginal effect estimate to the same units as Sacarny et al. (2020) to compare magnitudes. Sacarny et al. (2020) find that about 0.1 previously-eligible children enter Medicaid for every adult who entered Medicaid from the Oregon Health Insurance experiment. To compare to this point estimate, we employ the magnitude of the workwood effect among the entire inframarginal population (Table A.1A).\(^{38}\) We find that .9 inframarginal people between

\(^{37}\)The coefficient in the full sample may be larger due to the following. Very early in the sample period, all states had low eligibility thresholds, lower shares of the population eligible, and lower take-up. Thus the elasticity is larger since the average take-up rate is lower.

\(^{38}\)Let the point estimate for the entire inframarginal population from Table A.1A, Column 1 be \( \hat{\eta} \). We then estimate a version of Equation (2), using the log of the total number of people on the program as the dependent variable (and controlling for the log of the
0–130% of the FPL are induced to take up the program for every newly eligible person who takes up the program.

We propose several reasons why our treatment effects may be as much as 9 times larger than Sacarny et al. (2020). First, we have no reason to expect that inframarginal effects will be of the same magnitude across programs and over time. Second, Sacarny et al. (2020) examine take-up of health insurance for already-eligible children in households that are not already eligible. Across-household spillovers may be more important than within-household spillovers. Third, expanding the SNAP eligibility threshold for gross income does not loosen other eligibility criteria (e.g., the net income threshold). These criteria may bind for people with higher incomes. In support of this point, note that not many people above 130% of the FPL actually enroll in SNAP upon an eligibility expansion (Figure 3A). As a result, an eligibility expansion can lead to higher take-up among the inframarginal population without many newly eligible people joining the program.

**Comparison to outreach spending.** A final way of benchmarking our effects is to compare the take-up from inframarginal effects to the take-up from direct SNAP spending on information and outreach. We use state × time variation in our measure of spending on outreach to SNAP-eligible households from the SNAP Policy Database and estimate a version of Equation (2):

\[ y_{s,t} = \beta \text{spending}_{s,t} + \delta s + \gamma t + \chi X_{s,t} + \varepsilon_{s,t}, \]

where spending\(_{s,t}\) represents different spending measures. We include the same controls as in our main estimates of \(\eta\) and do not limit to a window around the eligibility expansion. As in Table [1], we use log enrollment for people between 50 and 115% of the FPL (ln(enrollment)\(_{s,t}\)) as our outcome. We use two spending measures: (i) the annual dollars spent on outreach, and (ii) the inverse hyperbolic sine transform of the annual outreach dollars (as in Burbidge et al. (1988)), to address outliers in the raw outreach data.

We find that money spent on outreach is generally ineffective at boosting take-up (Table [3]). Doubling outreach spending yields an insignificant 0.016% increase in inframarginal take-up (SE: 0.2%) (Column 1). The average outreach spending is $563,000 per year. As a back-of-the-envelope calculation, to obtain a 7.5% increase in inframarginal take-up, the average state therefore needs to spend about $264 million per year in outreach ($264 million \(\approx 563,000 \times 7.5 / 0.016\)). Indeed, our null estimate is precise enough that, using the upper bound of the \(\hat{\beta}\) coefficient from the 95% confidence interval, we obtain that a state must spend...
8.5 million dollars per year to replicate the take-up effect we find from the modal eligibility expansion. The 99th percentile of outreach spending across state-years is 6.6 million dollars, so we can reject with 95% confidence that spending as much as the 99th-percentile state-year would yield a take-up effect as large as the inframarginal effects in Table 1.

When we estimate this effect using raw dollars, rather than the inverse hyperbolic sine transform, we obtain an even more pessimistic conclusion about the efficacy of outreach spending, and we reject that outreach spending has a positive effect on take-up (Column 2, $p < 0.05$). To summarize, the effects on inframarginal take-up that we find are large in magnitude relative to what the government can achieve simply by reaching out to people who are already eligible.

### 2.5 Characterizing Compliers and Uncertainty about Eligibility

Who is most affected by eligibility expansions? To the extent that inframarginal effects are driven by reductions in barriers to take-up (“ordeals”), they may affect the targeting properties of the expansions (Nichols and Zeckhauser, 1982). If inframarginal effects influence SNAP’s screening capacity, we expect the people who join the program after an eligibility expansion to look different on observables (e.g., be lower-income, more likely to have children, or so on) than the previously enrolled. Table 4 assesses whether the eligibility threshold affects the characteristics of SNAP enrollees earning 50–115% FPL. We find precise null effects on the share who are female, their average age, and the share with any children. We do not find significant effects on the fraction Black, although we cannot rule out effects of up to 2% for a 10% FPL increase in the eligibility threshold. We do not find significant effects on net income (adjusted to 2016 dollars). We do find a significant positive effect on the average poverty level of enrollees, suggesting that new enrollees from the inframarginal effects have higher gross income than those who were previously enrolled. However, the magnitude of these effects is small: increasing the eligibility threshold from 130% FPL to 140% FPL, for example, would imply a 0.07% FPL increase in the average gross income of SNAP recipients. Together, these results suggest that whatever the ordeals behind inframarginal effects, they do not have substantial screening effects.

**Uncertainty about eligibility.** One reason for inframarginal effects could be *uncertainty* about eligibility: when the eligibility threshold rises, people who were uncertain they were eligible may now join the program. One reason for uncertainty may be volatile income, if people whose income changes frequently do not find it worth it to gain information about their eligibility. We provide two pieces of evidence against this mechanism in our setting. First, noting that households with volatile income must generally recertify their eligibility more frequently than other households, we show in Column 6 of Table 4 that there is no detectable effect of eligibility expansions on certification periods among enrolled households. Specifically, there is no evidence of an increase in the share of enrollees whose SNAP certification period is less than 6 months.

---

42The inverse hyperbolic sine specification shows that the specification in raw dollars is driven by a small number of state-years with large amounts of spending.

43We caution that the QC data files have unreliable race data after 2006.
months. We have sufficient precision to rule out (with 95% confidence) increases of 2.3 pp on this outcome for every 10% FPL increase in the eligibility threshold. This suggests that the new inframarginal SNAP enrollees do not have more volatile income than those previously enrolled. Second, we purposefully restrict the sample to 50–115% of the FPL, even though the minimum threshold is 130% of the FPL, to avoid effects driven by people with income very close to the threshold who are not sure if they meet the eligibility requirements or not. We discuss our two main classes of mechanisms in more detail in Sections 3 and 4.

2.6 Robustness

Contemporaneous policy changes. One may be concerned that changes to the eligibility threshold are bundled with contemporaneous policy changes. We note that, following Ganong and Liebman (2018), all our regressions control linearly for an index of eight other SNAP policies that occur during the same period (measured by the SNAP Policy Database). As Table 1 shows, including this index makes little difference, which gives additional confidence that unobserved policies affect the results. Moreover, when the index is separated into its component parts (Column 2), the magnitude of the effect is not diminished. One might nevertheless worry that the eligibility expansions were bundled with informal policies (e.g., flyer campaigns) that the SNAP Policy Database does not measure. To allay this concern, we show that the index of observable policies does not move around the time of the BBCE expansion. In Figure A.4, we present a second placebo event study with the SNAP policy index as the dependent variable. We find no evidence that the SNAP index increases after the eligibility expansions. Overall, the test is inconsistent with economically material bundling of SNAP policies. Finally, because we control for this index, this objection requires unobserved policies to affect the outcomes even after residualizing by the index.

Measurement error. To mitigate concern about measurement error, our key empirical fact (inframarginal effects) uses the QC numerators as the dependent variable, as in Table 1. Even so, we control for the size of the eligible population, which may be measured imperfectly in the CPS. Appendix A.1.3 presents two auxiliary tests that support our claim that measurement error is unlikely to affect our results substantially. First, we estimate Equation (1) with the log of the CPS counts of the people at 50–115% of the FPL as the dependent variable. We find a slight pre-trend in the CPS populations three years before the event, but the effects are modest and we control for the CPS populations in our specifications. Second, we conduct simulations that show that only an implausible amount of measurement error, exactly coinciding with the event and only in treated states, could explain our results.

Finally, there may also be measurement error in the timing of the policy implementation. We use data at the annual level in our main specification because we measure the number of people who are eligible from the March CPS, which is only available annually. Moreover, the QC data contain relatively few people at the month-state-income group level. However, BBCE policies can be implemented mid-year. In Figure A.5A, we show our event study using monthly data to estimate Equation (2). It looks broadly similar, although

44We show the main event study with take-up rates on the left-hand side in Figures A.5C and A.5D.
the inframarginal response is slightly slower to appear. This reflects the fact that in our main specification, we index policy implementation to the beginning of the first fully treated year.

Other effects of the BBCE. A related concern is that some states grant extra eligibility through the BBCE together with explicit referrals or brochures to SNAP. As a part of the BBCE, states sometimes use the budget from the Temporary Assistance for Needy Families to fund referrals to state services, including SNAP. As Figure[5] shows, states which adopted the BBCE but did not expand eligibility did not see similar effects on SNAP enrollment. This placebo test thus constitutes strong evidence that only the eligibility threshold, and not ancillary BBCE-related policies, are responsible for the take-up effect.

The BBCE also waives some rules on the maximum assets that block families in states without BBCE from obtaining SNAP. First, the above placebo study also rejects this concern, since BBCE states that maintain eligibility at 130% of the FPL but do not change their asset limits do not exhibit a take-up increase. Second, in practice, these asset rules affect a small number of families. Eslami[2015] finds that 4 percent of inframarginal people who participate in SNAP are eligible only due to state asset eligibility rules.45 There are a host of such asset waivers, including many not linked to the BBCE. But even assuming all these households were only eligible due to the BBCE, the asset waivers could not explain even half of the inframarginal effects we find.

Two-way fixed effects and negative weights. A new econometrics literature has documented that two-way fixed effects designs can be subject to concerns about negative weights. Callaway and Sant’Anna[2020] de Chaisemartin and D’Haultfoeuille[Forthcoming]. We implement the correction in de Chaisemartin and D’Haultfoeuille[Forthcoming] and obtain very similar results (Appendix A.1.4).

Policy salience. An additional concern is that the inframarginal effects arise in our setting because the expansions are salient to people, but they are not steady-state responses. First, we show that eligibility expansions boost take-up up to five years after the expansion, so they at least have effects in the medium-term. Second, the event study plots also show that the jump in take-up does not coincide with the expansion but grows over time. While we cannot reject that the effect may decrease over time, we also find no evidence to suggest that they will disappear.

3 Model

With our empirical evidence of inframarginal effects in hand, this section presents a simple normative theoretical framework for studying optimal eligibility. Our model emphasizes the following trade-off. The planner has a fixed budget and wishes to transfer to types with higher welfare weights. The planner can give a large transfer and impose a stringent eligibility cutoff — but in that case, take-up among inframarginal types will be low. If take-up among inframarginal types responds endogenously to the eligibility cutoff, then the planner may wish to increase the cutoff, at the cost of granting a smaller transfer (conditional on

45See computation in Ratcliffe et al. [2016].
take-up). We highlight two explanations for why the cutoff might affect take-up. According to one perspective, consumers behave rationally, and adjusting the cutoff might decrease costs such as stigma costs. On the other hand, take-up decisions might be affected by optimization failures — for example, changing the eligibility cutoff might increase information about the program. In particular, we model information as awareness about the program. Our framework accommodates both of these mechanisms.

3.1 Benchmark Model

We begin by analyzing a benchmark model where take-up responds endogenously to the eligibility threshold, but all consumers optimize. We add optimization failures due to imperfect information in Section 3.2.

We start by assuming take-up costs are normatively relevant (i.e., consumers are perfectly rational optimizers with respect to the take-up decision). In the following discussion, we often refer to these costs as “stigma costs,” since we are especially interested in the case in which raising the eligibility threshold can reduce stigma and therefore boost take-up. However, the costs refer to any cost that inhibits take-up, e.g. hassle costs. Other possible mechanisms that might be cast as changing costs include (i) decreasing uncertainty about eligibility (which we argue in Section 2 does not play a large role in our setting) or (ii) transaction costs, for instance if more stores accept SNAP once more people become eligible.

There is a continuum of individual types $\theta \sim F$, which correspond to ability or higher consumption. Types are perfectly observable, but we consider an environment in which the government cannot give a type-specific transfer (e.g., due to political economy or implementation constraints). The government offers a social program with a lump-sum consumption benefit $B$. The government provides $B$ only to types $\theta < m$, where $m$ is the eligibility threshold (or means test/income cutoff) also chosen by the government. We normalize the distribution of types to be quantiles of the distribution used to determine program eligibility (for example, the income distribution), i.e. $F := U[0, 1]$.47

Denote the welfare weight on type $\theta$ by $\lambda_\theta$, which refers to the welfare weight of quantile $\theta$ in the type distribution. For example, $\lambda_0$ refers to the weight that the planner places on the lowest-quantile person. We assume that the welfare weights are weakly decreasing in $\theta$.

Assume all people have the same twice continuously differentiable and concave utility function from taking up the benefit, denoted by $u(B)$. Normalize individuals’ outside income to be 0 and outside utility to be $u(0) = 0$. We already permit differences in realized consumption utility for each type to enter the planner’s problem through $\lambda_\theta$. We can simply redefine a type’s welfare weights to capture the different consumption utility that the type experiences.

46Alternative models of information frictions have different welfare implications, which we discuss in Section 3.4.

47Note that this normalization is innocuous: it amounts to letting type $m$ simply refer to the $m$-th quantile of the type distribution. The planner chooses what fraction of people are eligible, rather than the threshold type who is eligible.
Individuals choose whether to take up the benefit. We incorporate inframarginal effects by allowing the take-up probability to depend on the eligibility threshold $m$. In particular, every individual faces a take-up utility cost $c$, drawn from a continuously differentiable distribution $H$ (which we additionally assume has a finite first moment). We suppose $H$ depends on $m$, so $H(\cdot | m)$ and $h(\cdot | m)$ are the CDF and PDF of $c$.

We assume separability between the consumption benefits and take-up cost. Write realized utility as $U(B, c) = u(B) - c$. Then individuals participate in the program if $u(B) - u(0) > c$, i.e. $u(B) > c$. Because $H(u(B) | m)$ is the take-up probability, define $p(B, m) := H(u(B) | m)$. We sometimes suppress arguments and write $p(B, m)$ as $p$, so that the probability an individual of type $\theta$ takes up the program is $p_{\theta}$. We also assume that each type takes a cost draw from the same distribution, so that $p_{q} = p$.

**Labor supply.** We assume households’ labor supply is fixed: there are no labor supply responses to the threshold. We relax this assumption in Appendix C and show how a general problem with endogenous labor supply nests the key insights in this framework. Assuming fixed labor supply simplifies the framework considerably and permits us to focus on our novel mechanism (inframarginal effects). In this case, the fixed labor supply assumption permits us to think of the government budget $T$ as fixed.

**Planner’s problem.** The planner faces a budget constraint $T$. She maximizes the welfare-weighted consumption utility, net of participation costs, among people who are eligible for the social program and who ultimately participate.\(^{49}\)

The planner solves:

$$\max_{B, m, p(B, m)} \left( \int_0^m \lambda_B u(B) d\theta - \int_0^m \int_{c \leq u(B)} \lambda_B ch(c | c < u(B)) dc d\theta \right)$$

subject to

$$p(B, m) \int_0^m B d\theta \leq T$$

$$m \in [0, 1].$$

Let $\eta_m := \frac{\partial p}{\partial m} \frac{m}{p(B, m)}$ be the take-up elasticity with respect to the eligibility threshold. The parameter $\eta_m$ is the inframarginal effect, represented as an elasticity. We assume throughout that increases in $m$ reduce costs, so $\frac{\partial p(B, m)}{\partial m} > 0$ for all $B$.

Define $\eta_B$ as the elasticity of take-up with respect to the benefit size, $B$: $\eta_B := \frac{\partial p(B, m)}{\partial B} \frac{B}{p(B, m)}$. Let $\gamma(B, m) := \frac{E[c | c < u(B), m]}{u(B)}$, noting $\gamma(B, m) < 1$. The parameter $\gamma$ is the expected cost-benefit ratio conditional on take-up. It represents the share of the welfare gain from the benefit dissipated by the cost of taking up the benefit. For instance, if $\gamma = 0.5$, then costs represent half the utility gain (at $u(B)$).

\(^{48}\)Of course, utility is also a function of other consumption. The model takes this consumption as exogenous and normalizes $u(0) = 0$. In Appendix C we show that the model can accommodate different consumption across types, at the cost of notational complexity.

\(^{49}\)We can think of the planner in this model as the welfare program administrator who faces a fixed budget but can set program parameters like the transfer or eligibility threshold.
Let $\lambda_{av}(m)$ be the average welfare weight up to type $m$:

$$
\lambda_{av}(m) := \frac{\int_0^m \lambda_{d\theta}}{\int_0^m d\theta} = \frac{\int_0^m \lambda_{d\theta}}{m}.
$$

Let $\rho$ be the coefficient of relative risk aversion: $\rho := -\frac{u''(B)}{u'(B)} B$.\(^{50}\)

Then the first-order conditions yield the following benchmark:

**Proposition 1.** At an interior optimum, $m$ and $B$ satisfy:

$$
\frac{u(B)/B}{u'(B)} = \frac{1 + \eta_m}{(1 + \eta_B) \left( \left( \eta_m + \frac{\Lambda_m}{\lambda_{av}} \right) (1 - \gamma) - m \frac{\partial \gamma}{\partial m} \right)}.
$$

All proofs are in Appendix C.5.\(^{51}\) We can apply a Taylor expansion as in Gruber (1997) to simplify the left-hand side of Equation (8). At an interior optimum where $m \in (0, 1)$, socially optimal levels of $B$ and $m$ approximately satisfy:

$$
1 + \frac{1}{2} \rho \approx \frac{1 + \eta_m}{(1 + \eta_B) \left( \left( \eta_m + \frac{\Lambda_m}{\lambda_{av}} \right) (1 - \gamma) - m \frac{\partial \gamma}{\partial m} \right)}.
$$

The approach for solving and analyzing Equation (8) is similar to that in Baily (1978), Gruber (1997), and Chetty (2006)'s work on optimal unemployment insurance. Observe that the formula holds at an optimum. Hence, assuming that current policy is in the neighborhood of the (globally optimal) solution, if the LHS $\neq$ RHS for a given social program, we can infer how to adjust $m$ and $B$ to obtain an interior solution.

**Case 1: no inframarginal effects.** Equation (8) nests the case where $\eta_m = 0$ and there are no inframarginal effects. In that case, the planner seeks to equalize welfare-weighted marginal utility across people. Due to the concavity of the utility function, she does not give the entire budget to the lowest type so the solution is interior. On the other hand, as the welfare weight schedule is decreasing, the planner values marginal utility of the lower types more than higher types’. The solution will thus depend on the utility function’s curvature as well as the schedule of welfare weights.

**Case 2: adding inframarginal effects.** If raising the eligibility threshold decreases costs, leading more people to enroll, the planner also considers the standard fiscal externality involved in transferring to people who newly take up the program but are just indifferent. She also weighs the benefit of reducing stigma costs for people who always take up the program. Equation (8) shows how the planner trades off these forces. Starting from $\eta_m = 0$, a small increase in $\eta_m$ can increase or reduce the RHS, depending on the magnitude of the other parameters. As a result, to re-equate Equation (8), the planner adjusts $m$ and $B$.\(^{52}\)

---

\(^{50}\)We use the coefficient of relative risk aversion in $B$, evaluated at the sub-utility $u(B)$, since $c$ is just an additive shifter and does not affect curvature.

\(^{51}\)This statement refers to a necessary but possibly not sufficient condition for an interior optimum. We describe the statement in more detail in Appendix C.

\(^{52}\)While not the focus of this paper, we can also now ask: are there any programs that should be universal (i.e., $m = 1$)? This discussion amounts to considering the cases in which the planner cannot choose an interior solution. The planner clearly sets $m > 0$, so the only threat is that $m = 1$. Even if $\lambda_1 = 0$ (the planner places no weight on the wealthiest people), if $\eta_m$ is large enough, the planner may still find it optimal to choose the corner solution $m = 1$. Thus, inframarginal effects can serve as an argument in favor of
Connection to Baily (1978)-Chetty (2006). As in the Baily (1978)-Chetty (2006) equation, the planner equates the fiscal costs with the utility gains. Notably, our model embeds Baily (1978)-Chetty (2006) logic in the context of a redistributive program, rather than as an analysis of social insurance against risk. Similar intuitions appear regardless because the curvature of the utility function gives the planner a motive to smooth consumption across individuals. As is standard, our optimality condition is governed by an envelope argument: people who take up the program due to a reduction in costs are just indifferent. As a result, they impose a fiscal externality, because they take up the program (thus reducing how much the planner can transfer to others), but they experience no first-order utility gain. On the other hand, in this setting, the planner has an additional way to raise the utilities of people who always take up the program. She can reduce stigma by raising the eligibility threshold. Since these people are not indifferent, they do experience first-order utility gains. A change in the eligibility threshold itself also has first-order implications for social welfare, as those who are newly eligible enjoy the benefit of the program.

Appendix C works through additional cases of Equation (8) and gives a more detailed discussion of how Baily (1978)-Chetty (2006) logic applies in this setting. It shows that the LHS of Equation (8) is monotonically increasing in $B$ as long as $\rho \geq 1$; that argument formally grounds the comparative statics described above.

Connection to the Marginal Value of Public Funds. The optimality condition can be derived from equating the Marginal Value of Public Funds (MVPF) of an increase in $m$ and an increase in $B$, where each MVPF is weighted by the social marginal utility of money for the group affected by the policy change. Appendix A.3.2 shows how, with no externalities or labor supply response, the MVPF of raising $B$ is

$$\text{MVPF}_B = \frac{1}{1 + \eta_B}$$

while the MVPF of raising $m$ is

$$\text{MVPF}_m = \frac{(1 - \gamma) \left( \frac{u'(B)}{u'(0)} + \eta_m \right) - m \frac{\partial g}{\partial m} \left( \frac{u(B)}{u'(B)} \right)}{1 + \eta_m}$$

and Equation (8) can be derived from these. In Section 5, we use the MVPF framework to gain additional insight into the welfare implications of inframarginal effects.

3.2 Incorporating Information Frictions

In this section, we present our main optimality condition. We now permit some share of consumers not to optimize. Assume share $s \in [0, 1]$ of consumers are “stigma-only” agents who behave as in the prev-making programs universal without relying on normative assumptions about the welfare weight schedule.

53 This requires recognizing that raising the eligibility threshold benefits two groups of people — those inframarginal to the change and those newly eligible — whose willingness-to-pay for the policy is weighted differently by the social planner. Details are in Appendix A.3.2.
ous section. On the other hand, share \((1 - s)\) of consumers suffer from optimization frictions: raising the eligibility threshold for these consumers raises take-up because it increases information, so we call these consumers “information(-only)” agents. We assume that the probability of being a stigma-only agent is independent of \(m\).

Let the take-up probability for stigma agents be \(p^s\) and for information agents be \(p^i\). For information agents, costs are distributed:

\[
c = \begin{cases} 
\infty, & \text{with probability } 1 - p^i(m) \\
0, & \text{with probability } p^i(m)
\end{cases},
\]

for continuously differentiable \(p^i(m)\). Put another way, information agents always participate if they know about the program. If they know about the program, the cost they face is 0 and they take it up. An agent’s awareness does not depend on her type.

Let \(\eta^i_m\) and \(\eta^s_m\) represent the take-up elasticities with respect to the eligibility threshold for information and stigma agents, respectively. In Appendix C.5, we set up the planner’s problem and obtain the following optimality condition:

**Proposition 2.** At an interior optimum, \(m\) and \(B\) satisfy:

\[
\frac{u(B)}{Bu'(B)} = \frac{(1 - s) p^i (\eta^i_m + 1) + \frac{s}{p^i} (\eta^s_m + 1)}{(1 - s) p^i (\eta^i_m + 1) + \frac{s}{p^i} (\eta^s_m + 1)} = \frac{1}{\lambda_m + \frac{(1 - s) p^i (\eta^i_m + 1) + \frac{s}{p^i} (\eta^s_m + 1)}{(1 - s) p^i (\eta^i_m + 1) + \frac{s}{p^i} (\eta^s_m + 1)}}.
\]

Applying the same Taylor expansion, we have:

\[
1 + \frac{1}{s} \approx \frac{(1 - s) p^i (\eta^i_m + 1) + \frac{s}{p^i} (\eta^s_m + 1)}{(1 - s) p^i (\eta^i_m + 1) + \frac{s}{p^i} (\eta^s_m + 1)} = \frac{1}{\lambda_m + \frac{(1 - s) p^i (\eta^i_m + 1) + \frac{s}{p^i} (\eta^s_m + 1)}{(1 - s) p^i (\eta^i_m + 1) + \frac{s}{p^i} (\eta^s_m + 1)}}.
\]

Equation (14) is the main condition that we examine empirically. Under reasonable assumptions about \(\rho\) and the other auxiliary parameters, we can take Equation (14) to the data by estimating \(\eta_m\) for a given social program.

Note that if \(s = 1\), Equation (13) nests Equation (8). Moreover, if \(s = 0\) so all people are information agents, the distribution of stigma costs captured by \(\gamma\) and \(\frac{\partial \gamma}{\partial m}\) no longer enter the expression; in this case, since \(B\) does not affect the take-up rate for information agents, \(\eta_B\) no longer enters the planner’s optimality conditions. Then, the information-only case has the especially parsimonious expression:

\[
\frac{u(B)}{Bu'(B)} = \frac{1 + \eta_m}{\lambda_m + \eta_m}.
\]

We conclude by remarking that, unlike the stigma-only case, the information-only case \((s = 0)\) is not
governed by an envelope condition. People who newly take-up the program because of information costs
are not indifferent to taking up the program. As a result, inframarginal types who join the program due
to the inframarginal effect in the information-only case are more valuable to the planner than those who
join the program in the stigma-only case. The insights from this formula are captured well by the Baize
Chetty framework; we can re-arrange the expression to capture that the optimal policy equates
the welfare-weighted willingness to pay for an eligibility expansion with the total fiscal impact on the
government budget (Appendix C.3).

3.3 Comparative Statics

We next develop a proposition that inframarginal effects often serve as a force to increase the eligibility
threshold. Fix a vector of parameters $\Xi = (p^i, p^s, \lambda_B, s, \gamma, \eta_B)$. Notice that, for any $\Xi$, any given $\eta^i_m$, $\eta^s_m$, and
function $\frac{\partial g}{\partial m}(B, m)$ induce a pair $(B^*(\eta^i_m, \eta^s_m, \frac{\partial g}{\partial m}), m^*(\eta^i_m, \eta^s_m, \frac{\partial g}{\partial m}))$ that satisfies Equation (13).

We call $(B^n, m^n) := (B^*(0, 0, 0), m^*(0, 0, 0))$ the naïve choice of $(B, m)$: this is the choice of eligibility
threshold and benefit size if (i) the planner neglects inframarginal effects arising from either agent, and (ii)
does not realize that the eligibility threshold affects stigma. $^{54}$ Similarly call $(B^w, m^w) := (B^*(\eta^i_m, \eta^s_m, \frac{\partial g}{\partial m}), m^*(\eta^i_m, \eta^s_m, \frac{\partial g}{\partial m}))$
the sophisticated choice of $(B, m)$, where superscripts are indexed with $w$ because the sophisticated planner
understands there exist woodwork/welcome-mat effects.

To rule out edge cases, we assume $\lambda_m / \lambda_{avg} < 1$ at the naïve solution: the welfare weight schedule is
strictly decreasing for some $m' \leq m^n$. We also assume $\gamma > 0. \ ^{55}$

We also employ the following assumption:

Assumption 1. $\frac{\partial g}{\partial m} = \frac{\partial (E[c|c < u(B), m]/u(B))}{\partial m} \leq 0$.

This assumption imposes that the average cost-benefit ratio, conditional on taking up the program, does
not rise with a looser eligibility threshold. As $m$ rises, stigma costs fall, which tends to reduce $\gamma$. On the
other hand, new people may take up the program. Since they are nearly indifferent, they have relatively
high draws of $c$, which raises $\gamma$. The assumption is true as long as the mass of just-indifferent people
who newly sign up for the program as a result of reduced stigma costs do not raise the cost-benefit ratio
more than the reduction in inframarginal stigma costs. For instance, in the case where costs are distributed
uniformly, $\frac{\partial g}{\partial m} = 0$.

Assumption 1 is sufficient but not necessary. We provide a thorough evaluation of the assumption in
Appendix C.4. There, we give a substantially weaker (but less concise) necessary condition. We also prove
that the assumption always holds for costs that are distributed normally or exponentially.

We then arrive at the following proposition.

---

$^{54}$ We focus on the case where the planner ignores how the eligibility threshold affects stigma because a planner who understands
that $\frac{\partial g}{\partial m} \neq 0$ but does not acknowledge inframarginal effects has an internally inconsistent worldview.

$^{55}$ If there are stigma types ($s > 0$), we assume they face some positive stigma cost. We do not require that there are stigma types,
however.
Proposition 3. If either $\eta^s_i > 0$ or $\eta^s_m > 0$, then $m^i > m^s$ for all $\Xi$, under the condition that: (i) $\eta^s_i \leq \eta^s_m$, or (ii) $\eta^s_i$ lies in a neighborhood of $\eta^s_m$.

Under condition (i), stigma agents are weakly less elastic than information agents. Condition (ii) implies that each vector $\Xi$ yields an interval $\eta^s_m \in [\eta^s_i, \eta^s_i + \epsilon]$ for $\epsilon > 0$ in which the proposition holds. This condition ensures that Proposition holds for some $\eta^s_m > \eta^s_i$ for all parameterizations. These conditions are also sufficient but not necessary. In practice there are many parameterizations such that the statement holds for many even all $\eta^s_m > \eta^s_i$, which we further discuss in the appendix.

Proposition 3 states that the planner who ignores inframarginal effects but otherwise solves Equation (13) using the true values of all the parameters may choose an eligibility threshold that is too low. Put another way, acknowledging inframarginal effects often raises the optimal eligibility threshold. Because of the fixed budget, an immediate corollary is that the sophisticated benefit size is smaller than the naive benefit size. Assumption guarantees that the fiscal cost of types who are just indifferent to taking up the program do not exceed the benefit to the planner.

The normative conclusion that inframarginal effects serve as a motive to raise the eligibility threshold holds for any degree of stigma or information (i.e., $s \in [0, 1]$), under the sufficient condition in Assumption In fact, as $s \to 0$, Appendix C.4 shows that we can substantially relax Assumption. In the limit case where $s = 0$, none of Assumption condition (i), or condition (ii) is required at all (as is intuitive, since $\frac{\partial f}{\partial m}$ vanishes from the optimality condition). Thus, as we noted in the introduction, the case with $s > 0$ is conservative for the model’s normative conclusions. If all people are information types, then Proposition 3 holds under weaker conditions.56

3.4 Discussion of Model Assumptions

Our framework yields a tractable benchmark for welfare analysis that we can take to the data. Even so, it involves several stark assumptions, which we discuss below.

Lump-sum benefits. We consider a lump-sum benefits schedule. In reality, many social programs, including SNAP, have non-linear benefits schemes that vary based on income and household size. If the planner could give non-linear benefits, she might extend a small benefit to a larger share of people, to reduce stigma costs and boost take-up without incurring as large a fiscal externality. Our model abstracts from this choice, but we view $B$ as representing the (appropriately weighted) average benefit given to inframarginal types. Relatedly, we assume that people are perfectly informed about the benefit to which they are entitled.57

56The statement that welfare analysis is conservative if $s = 1$ does not mean that the planner should increase the eligibility threshold by a greater amount as $s \to 0$. It means that Proposition holds without Assumption. Proposition deals with infinitesimal changes in the eligibility threshold. Analyzing non-marginal changes requires more structure, which we develop in Section. Moreover, if $s = 1$ and the reduction in always-takers’ stigma costs are large, then that serves as another motive to increase the eligibility threshold.

57An alternative model, as in Finkelstein and Notowidigdo (2019), casts information frictions as a noisy (mis)perception of benefits. Even if misperceptions are symmetric, correcting them can still increase take-up in our model, since benefits enter a concave utility function. The welfare implications of this model are different: the utility gain to the newly enrolled inframarginals is bounded above in accordance with the size of the misperception, while the previously enrolled do not gain.
**Information campaigns.** Our model does not feature an instrument by which the planner can spread information about SNAP directly, e.g. through information campaigns. We make this choice for two reasons. First, Section 2 shows that outreach spending not especially effective among the inframarginal population in SNAP to begin with. Second, even if the planner has other means of spreading information (i.e., the eligibility expansion is not the most effective way to do so), the model highlights that eligibility expansions could nevertheless increase information. The planner needs to contend with the costs and benefits of the resulting increase in take-up.

**Identical take-up probabilities.** If in fact $p$ varies with $\theta$, it is possible to undo some of our normative conclusions. For example, suppose most of the increase in take-up from inframarginal effects is concentrated in types for whom $\lambda_\theta$ is small. Then inframarginal effects can yield a smaller transfer to the types for whom $\lambda_\theta$ is large. A fruitful extension of the model could consider different take-up probabilities.

### 4 Online Experiment: Evidence for Stigma

The model in the preceding section makes clear that the welfare analysis depends on the extent to which inframarginal effects are driven by information frictions or changing costs such as decreasing stigma.

In the economics literature on barriers to welfare take-up, evidence of incomplete information has been more forthcoming than evidence of stigma (Currie, 2004; Bhargava and Manoli, 2015; Finkelstein and Notowidigdo, 2019). Still, there is some research in economics (Friedrichsen et al., 2018) and in sociology (e.g., Gilens, 1999; Sykes et al., 2015) suggesting welfare stigma may affect take-up, and it seems natural to think that stigma may underlie inframarginal effects.

In this section, we present evidence from an online experiment that the eligibility threshold may affect perceived stigma around SNAP take-up. Our goal in presenting this evidence is not to convince the reader that stigma is the predominant underlying channel for inframarginal effects, nor is it to measure the magnitude of the role of stigma. Rather, it is to defend our inclusion of stigma as a possible mechanism in the model. We are not aware of research that studies the mechanisms behind woodwork effects, so we view this experiment as a useful first step. That said, we emphasize that isolating a stigma mechanism is not critical to advance the normative conclusion that the planner should raise the eligibility threshold. In general, suggesting that a larger share of people are governed by stigma is conservative for the ensuing normative analysis.

#### 4.1 Experiment Overview

We conduct an experiment that provides evidence that stigma can change when the eligibility threshold increases. While we show the effects are reasonably large, we do not map these magnitudes into the welfare calculations due to concerns about external validity.
In our pre-registered online survey, we experimentally manipulated respondents’ beliefs about the income eligibility threshold for SNAP and asked them to rate their agreement to a series of statements relating to stigma surrounding SNAP. We ran this experiment in mid-March 2020 with about 2,000 U.S. participants from a nationally representative sample.⁵⁸,⁵⁹ In the welfare analysis, we use the results of the experiment to argue that at least some people are induced to join SNAP due to reductions in stigma costs, although we present robustness to this conclusion.

4.2 Experiment Design

The objective of the experiment is to induce variation in participants’ beliefs about the share of people who are income-eligible for SNAP. In particular, we study how raising people’s beliefs about the share eligible affects self-reported stigma.⁶⁰ Figure A.8 summarizes the experiment design.

**Main experiment.** Our main experiment was embedded in a question asking respondents to report what share of Americans they thought were income-eligible for SNAP in 2016. On this page of the survey, all respondents were given a truthful hint:

“In 2016, in one of the U.S. states, roughly [X] of the population had low enough income that they could qualify for SNAP.”

X was randomly either 15% or 38%, which were the highest and lowest state-level eligibility shares we see in the administrative SNAP data from 2016. We refer to those participants who saw the 38% hint as those in the “high-share” treatment.

**Belief elicitation.** After implementing the treatment, we conduct a manipulation check by eliciting people’s beliefs about the share of people eligible for SNAP. We asked participants: “In 2016, how many out of every 100 people (in all U.S. states) do you think have low enough income that they could qualify...?”

To facilitate eliciting beliefs about the share eligible, we showed respondents a grid of 100 small squares; as the respondent entered and edited their answer about the number of people eligible for SNAP, the corresponding number of squares dynamically turned maroon.

We incentivized reports about the number of people eligible for SNAP as follows. At the beginning of the experiment, each participant selected among four charities to which we could donate $50. We informed the participant that there would be at least one factual question about SNAP. If she was correct about the question, we would include her in a lottery, and one participant in the lottery’s donation choice would be implemented.

**Auxiliary experiment.** Following the belief elicitation, we included an auxiliary randomization: we informed a random subset of participants about the correct share (27%, as per our calculations combining the CPS and the SNAP Policy Database). As a result, depending on their prior beliefs, this treatment (which

---

⁵⁸ Although this was a time of great uncertainty at the onset of the coronavirus in the United States, we chose to run the experiment then because we thought it would be harder to interpret data collected after the coronavirus affected the economy. For example, it would be difficult to obtain accurate estimates of how many Americans are eligible for SNAP.

⁵⁹ We used the survey provider Lucid; other papers using Lucid include Wood and Porter 2019 and Bursztyn et al. 2020.

⁶⁰ The complete survey instruments are available from the authors’ websites.
we call the “belief-correction” treatment) is intended to cause participants to update up or down about the share of people who are eligible for treatment. For example, people who report low beliefs about the share eligible may update up. This treatment informs us whether correcting people’s beliefs up or down about the share of people who are eligible affects stigma.

The main experiment has a tighter theoretical connection to inframarginal effects than the auxiliary experiment. The main treatment manipulates people’s perceptions about the eligibility threshold, and shows how that affects stigma. The auxiliary belief correction treatment does not have a tight connection to changes in an eligibility threshold. As a result, we relegate discussion of the belief-correction treatment to the appendix.

**Stigma elicitation.** We elicited people’s self-reported beliefs about the stigma associated with SNAP. We asked respondents to rate their agreement, on a scale from 1 to 9, to a series of eight statements about SNAP that adapt language from the literature on welfare stigma (Gilens, 1999; Bartlett and Burstein, 2004); the statements are the following (and elicited in this order):

1. I would prefer not to use food stamps because I would rather be self-reliant and not accept help from the government.
2. I believe that people should do what they can to avoid being on food stamps; it is better to make it on your own.
3. Most people believe that someone who uses food stamps is just as hard-working as the average citizen.
4. If I used food stamps, I would be concerned that people would treat me disrespectfully at stores.
5. Most people believe that someone who uses food stamps does so because of circumstances outside their control.
6. Most people think less of a person who uses food stamps.
7. Most people who use food stamps would go out of their way to prevent others knowing about their food stamp receipt.
8. If I used food stamps, I would avoid telling other people about it.

We aggregate the statements into two indices: (i) “first-order stigma,” which ask respondents about their own attitudes (statements 1, 2, 4, and 8 above), and (ii) “second-order stigma,” which asks respondents about others’ attitudes (statements 3, 5, 6, and 7). We standardize these outcomes using the mean and standard deviation of the control group and then average the standardized values as in Kling et al. (2007). We also show the effects on an aggregated index.

---

61We originally included the auxiliary experiment because recent papers, e.g. Bursztyn et al. (Forthcoming), use similar belief corrections to manipulate people’s prior beliefs.

62We reverse the scale for questions 3 and 5 so that positive numbers always indicate more stigma. We asked these statements with the opposite valence so that participants would think carefully about their responses.
We chose to elicit both first- and second-order stigma for several reasons. First, eliciting both types of stigma gives a richer picture of the psychological forces that might affect take-up.\footnote{Relatedly, the literature on social image often distinguishes between first- and second-order beliefs \cite{Bursztyn2017}.} More practically, first-order stigma may be difficult to move in a light-touch online experiment. It is plausible that people’s beliefs about others are less deeply held than one’s personal views about (say) the desirability of accepting aid from the government. Moreover, if people believe that the policies implemented in a democracy provide signals about others’ beliefs, then the treatment (which moves beliefs about policy) may directly affect second-order stigma.

Either first- or second-order stigma could play a role in inframarginal effects, depending on the model. Using the terminology of \cite{Butera2020}, either “characteristics-signaling” models of social image \cite{Benabou2006, Bursztyn2017} or “action-signaling” models \cite{Lindbeck1999} can deliver that these second-order beliefs matter for take-up. In either class of models, if people care about social image and take-up is partly observable, the extent to which others condone or sanction SNAP may affect take-up costs. On the other hand, with first-order stigma, people may have a hedonic aversion to SNAP that does not depend on others’ views. Such aversion could easily influence take-up if modeled as a direct take-up cost.

**Attention checks, sample construction, and attrition.** We drop participants who fail either of two pre-registered attention checks.\footnote{The attention checks are the following. First, before treatment, we tell people: “In this survey, we will ask you about your beliefs and attitudes about the Supplemental Nutrition Assistance Program (SNAP), also known as food stamps.” After eliciting the preferred charity (the incentive), we ask: “What does SNAP stand for?” There are four multiple choice responses: “Sufficiently Noisy Animal Parties”; “Supplementary Names Artful Program”; “Supplemental Nutrition Assistance Program”; “Salty Noodles And Pasta.” We drop the 106 participants who answer the acronym question incorrectly. Second, we drop the 145 participants who report that either 0 or 100% of people in the U.S. are eligible for SNAP.} To keep a consistent sample in all regressions, we drop also participants who did not provide a prior or respond to all stigma questions. Our final sample has 2,131 participants.

A total of 567 participants (21%) consented to the survey but attrited, or were dropped for the above reasons. Table A.4 summarizes the attrition and data cleaning steps and confirms that attrition, inattention, and non-response were balanced between treatment and control. Appendix B provides more information about cleaning the demographic data.

**Balance.** The sample is balanced across the high-share treatment (joint $p$-value: 0.94) and has a relatively similar composition as the U.S. on average (Table 5).\footnote{We also show balance in the auxiliary treatment (Table A.3).} In some tests, we restrict the sample only to the 512 people below 130% of the FPL, because this subgroup — the inframarginal SNAP sample — is of particular interest for inframarginal effects. Among this subgroup only, a joint $F$-test suggests experimental imbalance ($p$-value: 0.02).\footnote{The most imbalanced covariate is that the high-share treatment is less concentrated in the Northeast region than the low-share treatment ($p$-value of difference: 0.02).} The experiment was randomized but not stratified, and any imbalance in this subgroup occurred by chance. To address the lack of balance when studying treatment effects in this subgroup, we present robustness tests that control for available demographics. We stress that the experimental treatment is balanced in the full sample, and we emphasize results from the full sample as a result.
**Econometric strategy.** In our primary specification, we simply compare the difference in means across treatments:

\[ y_i = \beta I(\text{high})_i + \gamma I(\text{truth})_i + \epsilon_i, \]  

(16)

for individual \( i \), where \( \beta \) represents the coefficient of interest. In robustness exercises, we estimate a version of Equation (16) with additional demographic controls. We conduct inference using robust standard errors.

### 4.3 Experiment Results

**Beliefs about eligibility.** The high-share treatment successfully moved beliefs about eligibility (Figure 7). The distribution of beliefs for the low-share treatment (blue bars) lies below the distribution of beliefs for the high-share treatment (white bars). Both groups report beliefs that are slightly overestimated but reasonable; the mean for the low group is that about 30% of people are eligible, and the mean for the high group is that about 39% are eligible. The raw difference in means is 9.21 pp (SE: 0.80, \( p \)-value < 0.001). The standard deviation of beliefs in the control group is 19.8 pp, so the treatment raised the beliefs by a sizable 0.47 standard deviations. Moreover, while the low- and high-share treatments anchored a large fraction of people toward the numbers we provided them (15% and 38%), it also moved beliefs for others throughout the distribution.

**Stigma.** We begin by presenting a correlation matrix of the responses to the stigma questions across all participants (Figure A.9). We label questions that pertain to first-order or second-order stigma in red and blue, respectively. While the correlations across questions are typically positive, they are not overwhelmingly so. As a result, each question may contain independent information about the participants’ views. A concern is that participants simply anchor to their responses on the first question since the question order was not randomized. In fact, while we find that responses to the second question are relatively correlated with the first question (correlation \( \approx 0.65 \)), other questions do not display a large correlation with the first question. Moreover, anchoring is not especially likely because we reversed several of the questions’ valences.

Next, we turn to investigating the treatment effects. Increasing individuals’ beliefs about the share of Americans eligible for SNAP decreases their self-reported second-order stigma (Figure 8A). Aggregating the results into indices, the high-state treatment reduced second-order stigma by \(-0.050\) standard deviations (SE: 0.027, \( p = 0.061 \)). To benchmark this result, we note that the difference in the second-order stigma

---

67 We do not control for the auxiliary treatment in this specification because that treatment occurred after eliciting prior beliefs and is therefore uncorrelated by design. Reassuringly, if we do control for the auxiliary treatment, the results are almost identical.

68 Alternatively, responses are noisy, so we gain power from aggregating them into indices.

69 We believe anchoring/order effects are not likely to affect our results for several additional reasons. First, we merely seek to test the null hypothesis that the treatment does not affects either first- or second-order stigma. Unless anchoring is more prevalent in either treatment or control, anchoring only affects the magnitude of our results but not the sign. Since we do not use the magnitude of the experiment results in the welfare analysis, anchoring is unlikely to complicate the rest of the paper. Second, at worst, if participants anchor to the response to questions 1 or 2, that would imply the treatment effects on second-order stigma would be similar to those for first-order stigma. In fact, we find no effect on first-order stigma and that the high-share treatment reduces second-order stigma. Thus, anchoring would only attenuate our treatment effects.
between Democrats and Republicans is 0.020 (with Republicans reporting higher stigma); thus the treatment effect is more than double the difference between Democrats and Republicans’ second-order stigma. Effects are larger in magnitude among the 512 participants below 130% FPL (point estimate: −0.109, SE: 0.058, \( p = 0.061 \)). Because the inframarginal effect operates through people who are already eligible (i.e., this subgroup), it is reassuring that we find a large reduction in stigma among this group.

The treatment effects for second-order beliefs are similar across questions that form the second-order index. In the full sample, the high-share treatment reduces stigma the most in the question about whether most people believe recipients “go out of their way to prevent others knowing about their food stamp receipt.” We find larger effects aggregate among people who have ever taken up SNAP, men, and Democrats, although treatment effect heterogeneity is not generally significant (Figure A.10).

On the other hand, we find positive but statistically insignificant results on first-order stigma (Panel B). Two questions for first-order stigma have positive treatment effects: “I would rather be self-reliant and not accept help from the government,” and “It’s better to make it on your own.” Respondents agree more with these statements when they think more Americans are eligible for SNAP. It is possible that the high-share treatment raises the stigma measured by these questions due to a backlash effect: people who already believe the welfare state is too generous are affirmed in their beliefs in the importance of self-sufficiency when their beliefs about SNAP generosity increase. Particularly since we find no statistical effect on the first-order beliefs index (aggregating across questions), we interpret the question-by-question results with some caution due to multiple hypothesis testing.

We summarize these results in Table A.5, and we find very similar results when we include demographic controls (Table A.6). Moreover, when we aggregate the second- and first-order stigma results into a combined index, we find no statistical evidence of an average effect on stigma, although the point estimate is negative. The null result is mechanically driven by the null or slightly positive effect on first-order stigma.

4.4 Experiment Conclusions and Caveats

This experiment provides new empirical evidence about the mechanism underlying inframarginal effects. The health literature on inframarginal effects has not provided clean evidence that either information frictions or stigma costs contribute to inframarginal effects. Methodologically, our treatment contributes to a new literature in experimental economics on how to implement a treatment that moves people’s prior beliefs while still adhering to the norm of only providing truthful information. In our case, we provide truthful information about eligibility in one state, which affects people’s beliefs about the mean across states. Finally, evidence about stigma in social welfare programs remains elusive (Currie, 2004; Bhargava and Manoli, 2015). Our experiment suggests key aspects of program design, e.g. the eligibility threshold, indeed have the potential to affect program stigma.

While the experiment suggests that stigma could, in principle, affect inframarginal effects, the evidence we provide is not dispositive. We note several caveats. First, we find no effects on first-order stigma. We
note that first-order beliefs about, say, whether one should accept help from the government may represent deep-seated aspects of one’s identity. It is therefore not surprising that people’s first-order beliefs may be harder to move in a light-touch survey experiment than second-order beliefs. For instance, if people see government policy as reflecting the will of the people, then the treatment might directly affect beliefs about others’ views (second-order stigma) without affecting one’s own views (first-order stigma).

Second, an important caveat about our design is that we presented the high- and low-share treatments before a belief correction exercise. The belief-correction exercise itself does not provide evidence in favor of the stigma interpretation, which itself suggests some caution (see results in Appendix A.2). Moreover, the belief-correction treatment had the potential to “undo” the treatment effects from the initial randomization, since for some participants, we present the high-share treatment and then, after eliciting beliefs about the share eligible, we present the true share eligible before eliciting self-reported stigma. In Table A.9, we present effects on second-order stigma when we split the sample by whether the participant is shown the truth (before eliciting stigma). The treatment effects on second-order stigma are negative for both groups (Column 2), but larger in magnitude for the group who is ultimately shown the truth (Column 4). The larger treatment effect could arise if (i) being shown the truth reinforces the fact before eliciting stigma, or (ii) the combined treatments affect beliefs about the distribution of states with a large share of people eligible for SNAP. Regardless, on average, the high-share treatment reduces second-order stigma.

Third, as with any online experiment, one may worry about external validity. We cannot experimentally manipulate the actual SNAP eligibility threshold, only people’s perceptions of it. Relatedly, we do not have a measure of whether the intervention affects SNAP take-up, and it is possible that self-reported stigma falls but does not affect take-up.70,71 Nevertheless, the experiment provides evidence that some beliefs about stigma are sensitive to the eligibility threshold. At the very least, stigma remains a plausible mechanism for inframarginal effects. We conclude by emphasizing that, in the theory section, we show that relying on the stigma mechanism yields conservative normative recommendations. To the extent that one is unpersuaded that the experiment provides evidence that stigma contributes to woodwork effects, then that only exacerbates the normative conclusion that the planner should raise the eligibility threshold.

5 Welfare Analysis

Here we use the general model of Section 3 to conduct welfare analysis in the context of SNAP.

70 We note that stigma costs for those who take up SNAP regardless do enter the social planner’s problem. Take-up is not the only relevant margin for welfare.
71 Before treatment, we ask participants to report whether they currently or ever took up SNAP. When we present associations between people who ever took up SNAP in the past or currently take up SNAP, first-order stigma is more negatively associated with current or prior take-up than second-order stigma (Table A.10, Columns 1 and 2). Controlling for first-order stigma, second-order stigma is positively correlated with take-up (Column 3). Without experimental evidence, however, these correlations could reflect omitted variables, e.g. that low-income people have different social networks than high-income people and low-income social networks tend to stigmatize SNAP more.
We begin by introducing an assumption that we maintain throughout the section:

**Assumption 2.** At the planner’s solution, \( p^i = p^s \) and \( \frac{\partial p^i}{\partial m} = \frac{\partial p^s}{\partial m} \).

Assumption\(^2\) implies \( \eta^i_m = \eta^s_m \). As we make clear in Proposition\(^3\) this assumption does rule out some potentially important economic phenomena. On the other hand, we do not observe \( \eta^i_m \) and \( \eta^s_m \) separately, so Assumption\(^2\) is a natural benchmark.

Under this assumption, Proposition\(^3\) establishes that for \( \eta_m > 0 \), the naïve planner typically sets the eligibility threshold too low. We provide strong evidence of such inframarginal effects in Section\(^2\).\(^72\)

But how wrong will the planner be if she ignores inframarginal effects? Are the normative implications of inframarginal effects economically meaningful? Armed with our parameter estimates, we conduct several exercises that document that the magnitude of the planner’s mistake can be substantial. We consider the “naïve” planner who erroneously believes \( \eta_m = 0 \), but otherwise has perfect knowledge of auxiliary parameters and behaves optimally within the context of our model. We compare the naïve choices of eligibility threshold and benefit sizes to the choices of the “sophisticated” planner who knows the true value of \( \eta_m \).

The model does not capture every relevant economic force, so we view these welfare analyses as illustrative: woodwork effects of the magnitude that we find among the sample at 50–115% of the FPL in SNAP have the potential to meaningfully affect the optimal policy.

### 5.1 Calibration

In this section, we discuss additional assumptions and calibrate the parameters that govern optimality in Proposition\(^2\). We use \( \eta_m \) estimated from Section\(^2\) and calibrate the remaining parameters.

#### 5.1.1 Estimating \( \eta_m \)

We use the full-sample estimate of the effect of eligibility expansions on the 50–115% take-up rate from Table\(^2\). The event-study estimate is smaller, and we show robustness to \( \eta_m \). As we noted, the model involves one take-up elasticity (rather than separate take-up elasticities for 0–50, 50–115% and 115–130%). We use the 50–115% elasticity since this is the group in which we can reliably detect woodwork effects. The 0–50% group has high take-up so has less scope for take-up gains, and we exclude the 115–130% since some of them may be ineligible due to asset restrictions.

#### 5.1.2 Other Parameters

**Cost-benefit ratio of taking up the program (\( \gamma \)).** To gain traction, we employ an assumption about the distribution of costs \( c \) under the stigma-only case. Suppose \( c \sim U[0, \overline{A}(m)] \) for \( \overline{A} > u(B) \). Economically, this

\(^{72}\)We have no way of calibrating the relationship between \( \eta^i_m \) and \( \eta^s_m \), so we maintain Assumption\(^2\) throughout. If \( \eta^i_m > \eta^s_m \), that will typically increase the motivation to raise the eligibility threshold. If \( \eta^i_m > \eta^s_m \), that will typically reduce the motivation to raise the eligibility threshold.
assumption posits that: (i) changing \( m \) does not change the shape of the cost distribution, and (ii) there exist people for whom the take-up cost exceeds the utility gain. In that case, \( \gamma = \frac{1}{2} \), as \( E[c|c < u(B), m] = \frac{1}{2}u(B) \).

Moreover, \( \frac{\partial g}{\partial m} = 0 \). These properties both follow as immediate implications of the uniform distribution.

Assuming uniform costs serves as a useful benchmark: it implies that stigma costs, conditional on take-up, always erode half the benefit to the planner.

**Take-up elasticity with respect to benefit size \((\eta_B)\).** The SNAP benefits schedule \( B \) is set nationally. As a result, we cannot use a two-way fixed-effects design to estimate \( \eta_B \). Instead, we collect estimates of the typical elasticity of take-up with respect to benefit size for related programs. [Krueger and Meyer (2002)] review papers estimating \( \eta_B \) for UI and worker’s compensation and conclude that, for these programs, \( \eta_B \) ranges from 0.3 to 0.6. We choose \( \eta_B = 0.5 \) as a sensible midpoint and show robustness to other values.\(^3\)

**Coefficient of relative risk aversion \((\rho)\).** As is standard in welfare analysis of social insurance programs, we need an estimate of the coefficient of relative risk aversion \( \rho \). [Chetty and Finkelstein (2013)] note that this parameter is notoriously difficult to calibrate, but review other papers that test values of \( \rho \in [1, 4] \) (e.g., [Gruber, 1997]). We use \( \rho = 3 \) as a benchmark and test robustness to other values in this range.

**Share of stigma-only types \((s)\).** The results from the experiment suggest that perceptions about the eligibility threshold affect SNAP stigma. We therefore prefer \( s > 0 \) in our benchmark for the welfare analysis; we use \( s = 0.5 \) as a benchmark but show robustness to this choice.

**Welfare weights \((\lambda_m/\lambda_{avg})\).** We use an inverse-optimum approach to calibrate \( \lambda_m/\lambda_{avg} \) ([Bourguignon and Spadaro, 2012; Hendren, 2020]). In particular, we assume that today’s planner incorrectly believes \( \eta_m = 0 \) but otherwise has perfect knowledge of all other parameters. As a result, we can solve for \( \lambda_m/\lambda_{avg} \) at today’s value from Equation (13) as a function of the other parameters. For instance, using the primary estimates of all other parameters, we obtain \( \lambda_m/\lambda_{avg} = 0.427 \). The advantage of this approach is that, by using today’s policy as evidence about social preferences, we do not need to take a stand about what the normative welfare weight schedule should be. The disadvantage is that it assumes today’s policy is being set with our model in mind.

### 5.1.3 Calibration Summary

In Table 6, we summarize the parameter values and their source. We note that, as is common in welfare analysis, some of the economic primitives have a high degree of uncertainty. As a result, we show robustness to the particular choice of parameter.

### 5.2 Local Policy Changes

Our first two exercises use local changes around current policy to benchmark the size of the naïve planner’s mistake with few structural assumptions about the shape of the utility function or how the behavioral

---

\(^3\) Kroft (2008) also cites the Krueger and Meyer (2002) review and uses \( \eta_B = 0.5 \). Auray and Fuller (2020) is an example of a recent paper that finds a similar \( \eta_B \) in later years. In their data from 2002–2015, \( \eta_B = 0.63 \) (SE: 0.23), where \( \eta_B \) is the elasticity of UI take-up with respect to the replacement rate.
responses will change out of sample.

5.2.1 Sufficient Statistics Exercise

Our first exercise, conducted in the spirit of the sufficient statistics approach (Chetty, 2009; Kleven, Forthcoming), uses the main optimality condition we derived in Section 3. We study the implied value of $\rho$ that would be required to make the current means test optimal (when inframarginal effects are present). We establish that, if the planner assumes $\eta_m = 0$ but otherwise optimizes according to the theory, she will treat people as if they are much less risk averse than they really are.\footnote{We use the standard interpretation of $\rho$ as risk-aversion. However, it also corresponds to the planner’s unweighted valuation of transferring $B$ to someone who is ineligible from someone who takes up (has a benefit of $B$). To see this, note that $\rho = -B \frac{u'(0)}{u'(B)}$.}

We impose some ground-truth value of $\rho$, say $\rho = 3$. Consider the naïve planner who chooses $m$ and $B$ to solve Equation (13) assuming $\eta_m = 0$ and that $p^i = p^s$. Given $\eta_m \neq 0$, what is the implied $\tilde{\rho}$ that keeps the optimality condition equated? We use the following algorithm:

1. Obtain inverse-optimum weights: Assuming $\eta_m = 0$, solve for $\frac{l_m}{l_{avg}}$ that satisfies Equation (13).

2. Obtain implied $\tilde{\rho}$: Given the inverse-optimum weights $\frac{l_m}{l_{avg}}$ and true value of $\eta_m$, solve for the $\tilde{\rho}$ that satisfies Equation (13).

Intuitively, because the planner ignores $\eta_m$, she treats people “as-if” they have risk aversion $\tilde{\rho}$, when they really have risk aversion $\rho$. Put another way, there is some value of $\tilde{\rho}$ that satisfies the optimality condition even under the (incorrect) assumption that $\eta_m = 0$. We are interested in the value of $100 \times \frac{\tilde{\rho} - \rho}{\rho}$, which is a measure of the bias in the coefficient of relative risk aversion.

**Results.** We show that the magnitude of the planner’s mistake can be substantial for the range of $\eta_m$ that we estimate (Figure 9A). The x-axis plots values of $\eta_m$. On the y-axis, we plot the bias in the “as-if” risk aversion parameter relative to the true risk aversion parameter, assuming $\rho = 3$. If $\eta_m = 0$, there is no bias: the naïve and sophisticated solutions coincide by construction. As $\eta_m$ grows, the bias rises; for $\eta_m = 0.05$, the bias is about 10%. For our primary estimate of $\eta_m = 0.16$, we find that the bias can be quite large: the naïve planner’s solution will treat people as if they are almost 30% less risk averse than they really are.

Intuitively, the planner who ignores inframarginal effects transfers too much to inframarginal types who already take up the program. She overvalues inframarginal types’ marginal utility and undervalues the gain in utility from those who would take up the program if she raised the eligibility threshold. As a result, she optimizes as if the coefficient of relative risk aversion were smaller than it really is.

**Robustness.** While the mechanisms underlying inframarginal effects have different welfare implications, as discussed in Section 3 the conclusions of our sufficient statistics exercise are not substantially altered under different assumptions of $s$, the share of individuals responding to stigma (versus information). Holding inframarginal effects fixed at $\eta_m = 0.16$ and repeating the above algorithm for different
values of $s$ (Figure 9B), we find that the bias in the “as-if” $\rho$ is larger than 15% for all $s \in [0,1]$. The bias shrinks in magnitude as $s$ grows because if more people are stigma-only types, then they are governed by an envelope condition that makes the inframarginal effect less valuable to the planner. The bias remains large if we use $\rho = 2$ or $\eta_B = 0.3$ (Figure A.11).

### 5.2.2 Marginal Value of Public Funds Approach

As a related alternative, we consider the policy of expanding eligibility within the context of its Marginal Value of Public Funds (MVPF) ([Hendren], 2016; [Hendren and Sprung-Keyser], 2020). This approach lets us relax the assumption that today’s policy constitutes the naive solution to Equation (13). It also permits us to probe other assumptions about agents’ behavior and utility.

In this framework, the planner considers the ratio of benefits (willingness to pay for the policy) to the net cost to the government. Because our focus is on the redistributive nature of the policy, we ultimately consider welfare-weighted MVPFs, also known as welfare impacts (per dollar of government expenditure). As we noted above, equating the welfare-weighted MVPF of raising the means test to that of raising the benefit size recovers our main optimality condition.

We analyze the size of the bias in the welfare-weighted MVPF, which we define as:

$$bias := 100 \times \frac{\tilde{\lambda}^n MVPF^n - \tilde{\lambda}^w MVPF^w}{\tilde{\lambda}^w MVPF^w},$$

where $\text{MVPF}^n$ is the MVPF when the eligibility threshold does not affect inframarginal recipients and $\text{MVPF}^w$ is the MVPF when it does. $\tilde{\lambda}^n$ and $\tilde{\lambda}^w$ correspond to the average welfare weights of the beneficiaries of the policy in each case (denoted $\bar{\eta}$ in [Hendren and Sprung-Keyser], 2020). Let $\Delta$ be the size of the eligibility threshold increase; for instance, $\Delta = 0.01$ when we study the welfare effect of letting 1 pp more people become eligible. In Appendix A.3.2, we show that without externalities onto other beneficiaries (e.g., children), the bias reduces to:

$$bias = 100 \times \left( \frac{p_m + \frac{m}{\Delta} w \frac{\kappa_{avg}}{\kappa_m}}{p_m + \frac{m}{\Delta(1-s\gamma)} \frac{\lambda_{avg}}{\lambda_m} (1-w + sp_{avg}\alpha)} - 1 \right),$$

where $p_m$ ($p_{avg}$) is the take-up rate for those newly eligible (previously eligible); $w$ is the percentage point increase in the take-up rate for information-types (i.e., the inframarginal effect); $\kappa_m$ ($\kappa_{avg}$) is the total fiscal cost of an additional 1 pp of take-up, including fiscal externalities, for those newly eligible (previously eligible); $\gamma$ represents participation (stigma) costs as a share of WTP to take up the benefit; and $\alpha$ corresponds to the reduction in costs when the eligibility threshold rises, as a share of total WTP for the policy.$^{76}$

$^{75}$At $s = 0.5$, the bias coincides with Figure 9A (evaluated at $\eta_m = 0.16$), since we assume $s = 0.5$ in Panel A.

$^{76}$The advantage of focusing on the proportion bias in the welfare impact is that the expression does not require an estimate of willingness to pay or separate estimates of the costs $\kappa_{avg}$ and $\kappa_m$. The magnitudes of these costs are difficult to estimate, because
With this approach, we relax several assumptions imposed in Section 3. While our model in Section 3 defines the gains to inframarginal stigma agents in relation to the size of the woodwork effects, this exercise decouples them (since \( a \) and \( w \) enter separately). Moreover, this expression permits inframarginal program participants to have different costs from participants who are newly eligible.\(^{77}\) This is at the expense of additional assumptions (e.g., on the size of \( a \)), of course, but it is easier to see robustness to those assumptions. We also emphasize the role of the welfare weights here: the policy evaluation is as much about its incidence as it is about its utility gains and fiscal costs.

**Results.** We evaluate the welfare impact per dollar of government expenditure of expanding eligibility by 1 pp, i.e. we set \( \Delta = 0.01 \). For the other parameters, we assume values consistent with the data used in Section 2 (detailed in Appendix A.3.2). To be conservative, we assume that the willingness to pay for a reduction in stigma costs is small, so we set \( a = 0.02 \). Finally, we use \( \lambda_{\text{avg}} \lambda_{\text{m}} \) derived from inverse-optimum weights, although we note that these employ an assumption MVPFs usually relax — that current policy is optimal under a certain model. However, our results are robust to a range of values for \( \lambda_{\text{avg}} \lambda_{\text{m}} \).

We find that, if \( \frac{k_{\text{avg}}}{k_{\text{m}}} = 1 \), the naïve MVPF can be about 20% below the sophisticated MVPF for the information-only case, with even larger results in the information and stigma case (Figure A.12A). However, the planner may overvalue the welfare impact for larger values of \( \frac{k_{\text{avg}}}{k_{\text{m}}} \), say \( \frac{k_{\text{avg}}}{k_{\text{m}}} = 10 \) (Figure A.12B). This is because with \( \frac{k_{\text{avg}}}{k_{\text{m}}} \gg 1 \), the cost of new participants who are costly may exceed their value to the planner. As a result, the naïve planner sets the eligibility threshold too high.\(^{78}\)

### 5.3 Global Optimum: Numerical Simulations

What is the optimal eligibility threshold and take-up rate? To solve for these globally optimal parameters, we require assumptions about non-local behavior.

#### 5.3.1 Set-up

**Equilibrium values.** We take the population-weighted average of states’ share eligible across years (accounting for varying eligibility thresholds) to obtain \( m^* = 0.27 \). Our data suggests an equilibrium take-up probability of 0.53.\(^{79}\) From the planner’s budget and take-up probability, we obtain the average equilibrium SNAP benefit \( B^* \).

\(^{77}\)Note that inframarginal types tend to have higher benefits, the higher benefit may yield a reduced fiscal externality because people with higher SNAP benefits receive better educations or are less likely to be incarcerated; Bailey et al. (2020) show that these benefits reduce the denominator of the MVPF for a benefit increase.

\(^{78}\)Note that if \( \frac{k_{\text{avg}}}{k_{\text{m}}} \gg 1 \), the MVPF bias is negative for the stigma and information case (\( s = 0.5 \)) and positive for the information-only case (\( s = 0 \)). Here, unlike in the model, the normative conclusion that the planner may wish to raise the eligibility threshold can be stronger if there is stigma.

\(^{79}\)This number is below the number the USDA reports because our denominator includes some people who are not eligible for SNAP due to work requirements, asset thresholds, or other tests; moreover, it is not clear that the USDA number includes people with incomes above 130 if they live in states with an eligibility threshold beyond 130. We use the number for illustrative purposes in this exercise, but the results are not sensitive to adjusting the equilibrium \( p^* \).
Welfare weights. We need to extrapolate about the welfare weight schedule $\lambda_\theta$ out of equilibrium. We assume $\lambda_\theta$ is linear in $m$ and satisfies the value of $\lambda_\theta/\lambda_m$ obtained from the inverse optimum exercise. We assume that $\lambda_0 = 1$. These two assumptions pin down a unique linear welfare weight schedule of inverse optimum welfare weights.

Linear take-up probabilities. We assume the take-up probability is linear in $m$. We assume a representative part-stigma part-information agent who obeys: $p^i(m) = p^i_0 + \frac{\partial p}{\partial m} m + s \frac{\partial p}{\partial B} B$.

Using the values for the elasticities $\eta_m$ and $\eta_B$, we obtain the slope $\frac{\partial p}{\partial m} = \eta_m \frac{\partial \nu}{\partial m}$, which is equal for stigma and information types by Assumption 2. We obtain $\frac{\partial p}{\partial m} = \eta_B \frac{\partial \nu}{\partial B}$.

Obtaining optimal $m$ and $B$. We invert Equation (13), which, together with the linearity assumptions above, delivers a unique value of optimal $m_{opt}$ and $B_{opt}$. Intuitively, this approach obtains the values of $m$ and $B$ that satisfy the planner’s optimality conditions we derived in Section 3. We solve this problem numerically using Matlab.

5.3.2 Results

We first ask how much larger is $m_{opt}$ relative to today’s $m^* = 0.27$. Because we use the inverse optimum approach to calibrate the welfare weights, Proposition 3 guarantees that the optimal $m_{opt}$ will exceed today’s $m^*$: $m_{opt} - m^* > 0$.

We present the percent increase in $m_{opt}$ relative to $m^*$, i.e. percent increase := $100 \times \frac{m_{opt} - m^*}{m^*}$. This value represents the percent increase in the optimal eligibility threshold relative to today’s threshold. Because the threshold is measured in terms of the share eligible, it equivalently represents the percent increase in the share of people who should be eligible relative to today.

We present our estimates of the percent increase in $m$ as a function of $\eta_m$ (Figure 10A) for both $s = 0.5$ (black line) and $s = 1$ and $s = 0$ (gray dashed lines). By construction, if $\eta_m = 0$, we find the optimal eligibility threshold coincides with today’s threshold. As $\eta_m$ rises, the optimal $m$ rises too. At our preferred value of $\eta_m = 0.16$, about 10% more people should be eligible than are eligible today. The value of $s$ makes relatively little difference in these computations.

We also show the optimal take-up rate (Figure 10B). Here, the mechanism makes a larger difference: while the optimal take-up rate is increasing in $\eta_m$ for $s = 0$ (gray line), the optimal take-up rate is not monotone in $\eta_m$ for $s = 0.5$ or $s = 1$. What explains the non-monotonicity? Recall that $\eta_B = 0.5$ in these simulations. As a result, a small increase in $m$, which decreases $B$ due to the fixed budget, results in lower take-up. Ultimately, once take-up falls enough, $B$ rises (because fewer people take up the benefit, which the planner perfectly anticipates). This dynamic does not exist for $s = 0$ since $\eta_B$ has no effect for information agents.

---

80The figure also shows that the magnitude of the percent increase eligibility threshold does not fall monotonically as $s$ rises. As $s \to 1$, the planner may want to raise the eligibility threshold beyond the case where $s = 0$, in order to reduce stigma costs for people who always take up the program. When we say that the normative conclusions are conservative as $s \to 1$, we mean that Proposition 3 holds without Assumption 1.
We emphasize that Panel A and Proposition 3 show that the sophisticated planner will still expand eligibility beyond today’s \( m \). However, Panel B highlights that we cannot conclude that take-up today is suboptimally low.\(^{81}\) The naïve planner erroneously believes take-up will fall more than it actually would for a small increase in \( m \) because she does not account for \( \eta_m \) and only accounts for \( \eta_B \). She therefore sets \( m \) too low. We show similar conclusions for \( \rho \in \{1, 2, 4\} \) (Figure A.13A) and \( \eta_B = 0.3 \) (Figure A.13B). Notably, the magnitude of \( \rho \) does not have a large effect on the percent change in \( m \). In this setting, \( \eta_m \) is much more important than \( \rho \).

The planner has a fixed budget, so the increase in the optimal eligibility threshold and small change in optimal take-up rates imply that the optimal benefit is decreasing in the inframarginal effect, even for non-local changes (Figure A.14). For various \( s \), at our preferred estimate of \( \eta_m \), the optimal benefit is 5–10% lower than the current optimum.

A weakness of our numerical approach is that we assume that \( \rho = 3 \) both in today’s equilibrium and also at an optimum, but that third-order utility terms vanish (in order that Equation (14) holds). We conduct a second exercise where we assume a quadratic utility function that imposes that \( \rho = 3 \) at today’s \( B \) (Figure A.15).\(^{82}\) This exercise gives similar results, although the magnitudes of the increase in eligibility are attenuated because risk aversion changes rapidly for quadratic utility, so a small change in benefits is sufficient to equate the LHS and RHS of Equation (13).

6 Conclusion

This paper documents the existence of inframarginal effects in SNAP. Motivated by this fact, we develop a general model for incomplete take-up of social welfare programs when the planner can control program eligibility. We apply our model to SNAP and assesses the implications for the optimal eligibility threshold given the inframarginal effects. The inframarginal effects are large, which leads the planner to increase SNAP eligibility. We additionally provide new experimental evidence about how policy tools can affect welfare stigma.

All social programs, even universal ones, make some determination about eligibility. This threshold is often chosen by the planner and thus is not an exogenous feature of the policy environment. As a result, our normative insights have applications in many areas in public economics. When inframarginal effects are present, our theoretical framework highlights that they may serve as a motive to raise the eligibility threshold. Future work could enrich the model to include a larger set of policy instruments and more heterogeneity in individual responses.

---

\(^{81}\)Note that take-up depends on the benefit size as well as the eligibility threshold, so the higher threshold does not necessarily imply higher take-up on net.

\(^{82}\)Together with \( u(0) = 0 \), this assumption yields that utility is: \( u(B) = -(B - k)^2 + k^2 \) for \( k := \frac{\rho + 1}{\rho} B^* \).
References


Federal Register, “Revision of Categorical Eligibility in the Supplemental Nutrition Assistance Program (SNAP),” July 2019, 84 (142), 35570–35581.


Harris, Timothy F., “Do SNAP Work Requirements Work?,” Economic Inquiry, Forthcoming.


7 Figures

The figure shows income eligibility thresholds as a percent of the Federal Poverty Level (FPL) for the largest U.S. means-tested social programs against estimates of their national take-up rates, compiled from different sources. Take-up rates are estimated out of the eligible population for each program. In programs with different eligibility thresholds per state, the level plotted is the population-weighted average of those thresholds. The data sources and assumptions involved are described in Appendix B. The SNAP take-up rate displayed here is higher than that used in our paper because the USDA uses a more involved and restrictive method for assessing eligibility than we do; our empirical results are not affected by a denominator that is too large. Where the eligibility threshold is defined in dollars (e.g., EITC, TANF), the figure shows the threshold as in terms of percent of the FPL for a family of three. Some programs (e.g., WIC, TANF) are restricted to certain subgroups in addition to imposing income thresholds—for example, families only—or have additional requirements like work requirements. Given Head Start’s capacity constraints, additional assumptions were made to estimate a take-up rate. These are also documented in Appendix B.
The figure presents the results from our literature review of papers in the *Quarterly Journal of Economics* (2010–2019) and the *American Economic Review* (2010–2018). Appendix B.6 provides details about the sampling frame. The first row shows the total number of papers that we concluded were about welfare programs, after reading the abstract and introduction. The second row shows the number of papers that considered instruments with which the planner could enact optimal policy, e.g. the benefit size or duration. The third row shows the number of papers that considered the eligibility threshold as an instrument with which the planner could enact optimal policy. The fourth row shows the number of papers that use the eligibility threshold as a source of variation with which the authors estimated a treatment effect for the program.
This figure presents the relationship between the eligibility threshold and SNAP take-up and enrollment. Panel A shows SNAP enrollment per 1,000 people in states and years where the eligibility threshold is 130% of the Federal Poverty Level (FPL) versus above 130% (and up to 200%). Each bar takes the number of people in the USDA Quality Control data whose household income is in each income bin, divided by the total population (i.e., all people, with any household income) in all state-years with the indicated eligibility regime. The data are limited to the sample we use in the main event study, and household income is top-coded at 200% FPL. Panel B shows average take-up among those earning 50–115% of the FPL in states with each eligibility threshold observed in the data. The USDA Quality Control data provide estimates of the numerator for the outcome (take-up counts, by state-year), and the Current Population Survey data provide estimates of the denominator (total counts of individuals within this sample).
This figure presents the event-study version of our estimate of $\eta$ (Equation (1)), the effect of the eligibility rate on inframarginal take-up. Panel A presents results for the sample of individuals from 50–115% of the Federal Poverty Level (FPL), while Panel B presents results for the sample of individuals from 0–130% of the FPL. The minimum eligibility in all states is 130% of the FPL. The USDA Quality Control data provide estimates of the outcome (enrollment counts, by state-year). Standard errors are robust to heteroskedasticity and clustered by state.
This figure presents a placebo event study, modifying Equation (1) as described in Section 2. The event in this sample represents the nine states that adopt the Broad Based Categorical Eligibility policy, but do not expand eligibility. The USDA Quality Control data provide estimates of the outcome (enrollment counts, by state-year). The figure shows there is no discrete change to inframarginal enrollment at the event time (those earning 50–115% FPL). Standard errors are robust to heteroskedasticity and clustered by state.
This figure presents estimates of Equation (2) using take-up rates as the outcome variable. The bars show the effect of the eligibility threshold on SNAP take-up by income group, and the whiskers show the 95% confidence intervals. While the regression specification is the same for all bars (with only the reference group changing), they are colored blue and gray to distinguish the effects on the inframarginal population versus the effects on the newly eligible population. Take-up rates are calculated using the enrollment counts from the USDA Quality Control (QC) data in the numerator and total counts of individuals within the income group from the Current Population Survey (CPS) in the denominator. Standard errors are robust to heteroskedasticity and clustered by state.
This figure presents the distribution of beliefs from the online experiment, split by treatment group, about the fraction of people who are eligible for SNAP. The y-axis shows the share of people within each treatment group who report a given fraction are eligible for SNAP. The blue bars show the values for the low-share treatment. The white bars show the values for the high-share treatment. The light blue shaded area shows the overlap.
Figure 8: Effect of High-Share Treatment on Stigma

(A) Second-order beliefs

Second-order beliefs index

- Most people believe [SNAP recipients] are just as hard-working as the average citizen. (Reversed)
- Most people believe that someone who uses food stamps does so because of circumstances outside their control. (Reversed)
- Most people think less of a person who uses food stamps.
- Most people believe [SNAP recipients] would go out of their way to prevent others knowing about their food stamp receipt.

Coefficient: -0.050
SE: 0.027
p-value: 0.061

(B) First-order beliefs

First-order beliefs index

- I would rather be self-reliant and not accept help from the government.
- It’s better to make it on your own.
- I would be concerned that people would treat me disrespectfully at stores.
- If I used food stamps, I would avoid telling other people about it.

Coefficient: 0.025
SE: 0.031
p-value: 0.421

This figure presents results from the online experiment; it shows the effect of the “high-share” treatment (where respondents were randomly given a hint that increased their reported beliefs about the share of Americans who are eligible for SNAP) on agreement with each statement in the stigma instrument (Equation 15). Outcomes marked with “(Reversed)” were reverse-coded so that for all items, a higher score indicates more stigma. The coefficients correspond to a reduced-form (intent-to-treat) estimate and do not account for the amount by which the treatment moved people’s beliefs about the share of Americans who are eligible for SNAP. Each outcome is in units of standard deviations, and the indices average the set of outcomes displayed in each panel. Bars plot 95% confidence intervals.
This figure shows the percent bias between the planner’s “as-if” risk aversion ($\tilde{\rho}$) and the ground-truth risk aversion ($\rho$) (black line). Negative numbers indicate that the planner is behaving as if people are less risk averse than they really are. Panel A plots the bias as a function of the inframarginal effect; the vertical gray line plots the empirical inframarginal effect presented in Table 9. Panel B fixes $\eta_m$ at the empirical inframarginal effect from Table 9 and varies $s$, the share of stigma agents.
This figure shows the results from our numerical simulation exercise: it presents the change in the percent of people who are eligible relative to current policy if the planner were to acknowledge inframarginal effects (Panel A) and the optimal take-up rate (Panel B), as a function of the inframarginal effect $\eta_m$, using our preferred optimality condition (Equation (13)). Auxiliary parameters are set according to the values in Table 6.
8 Tables

<table>
<thead>
<tr>
<th>Table 1: Estimates of the Infrahmarginal Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Main estimate</td>
</tr>
<tr>
<td>Income limit (% FPL) / 100</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>N states</td>
</tr>
</tbody>
</table>

This table shows the effect of the eligibility threshold on log enrollment among the inframarginal population (50–115% FPL). Column 1 estimates Equation (2), and the following columns present various extensions to show robustness. Column 2 separates the “Ganong-Liebman” policy index into separate indicators. Column 3 includes a control for the previous year’s unemployment rate in each state and a control for the population-weighted average number of months a state had ABAWD work requirement waivers in effect. Column 4 excludes years 2008–2011, during the Great Recession. Column 5 weights observations by the state-year population. Column 6 presents the difference between the average pre- and post-period event study coefficients. Finally, Column 7 estimates Equation (2) using all the data available instead of only the event study sample. Standard errors are robust to heteroskedasticity and clustered by state. *, **, and *** indicate $p < 0.1$, 0.05, and 0.01, respectively.
<table>
<thead>
<tr>
<th>Panel A. All data</th>
<th>OLS</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.161**</td>
<td></td>
</tr>
<tr>
<td>ln(Share eligible)</td>
<td>-0.100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td></td>
</tr>
<tr>
<td>Income limit (% FPL) / 100</td>
<td>0.728***</td>
<td>0.117**</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Observations</td>
<td>1071</td>
<td>1071</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B. Event study sample</th>
<th>OLS</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.097</td>
<td></td>
</tr>
<tr>
<td>ln(Share eligible)</td>
<td>-0.194***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td></td>
</tr>
<tr>
<td>Income limit (% FPL) / 100</td>
<td>0.756***</td>
<td>0.073</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>Observations</td>
<td>705</td>
<td>705</td>
</tr>
</tbody>
</table>

This table presents estimation results for $\eta_m$, the elasticity of take-up with respect to the share of the population who are eligible, controlling for the covariates included in Equation (3). We estimate this elasticity using the eligibility threshold as an instrument for the share of residents in a state who are eligible for SNAP. The first column shows results from a naïve OLS regression of ln(take-up) on ln(share eligible). The second column presents the first stage — the coefficient from a regression of ln(share eligible) on the eligibility threshold as a % of the Federal Poverty Level (FPL). The third column, the reduced form, gives the relationship between the eligibility threshold and ln(take-up). The final column gives the 2SLS estimate, our final estimate for $\eta_m$. Standard errors are robust to heteroskedasticity and clustered by state. ** and *** indicate $p < 0.05$ and 0.01, respectively.
Table 3: Effect of Outreach on Take-Up (50–115% FPL)

<table>
<thead>
<tr>
<th></th>
<th>(1) ln(count)</th>
<th>(2) ln(count)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sinh^{-1}(Outreach)</td>
<td>0.000160</td>
<td>-2.86e-08**</td>
</tr>
<tr>
<td></td>
<td>(0.00241)</td>
<td>(1.08e-08)</td>
</tr>
<tr>
<td>Outreach ($)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,071</td>
<td>1,071</td>
</tr>
<tr>
<td>Mean $ to obtain 7.5% take-up (millions)</td>
<td>264</td>
<td></td>
</tr>
<tr>
<td>Lower bound on $ to obtain 7.5% take-up (millions)</td>
<td>8.46</td>
<td></td>
</tr>
</tbody>
</table>

This table presents the results of Equation (4). Column 1 shows the effect for sinh^{-1}(Outreach). Column 2 shows the effect for raw outreach. The outreach variable is measured in units of dollars of outreach spending per state-year. We control for the baseline controls in Equation (2) (the state unemployment rate and an index of other SNAP policies), the SNAP eligibility threshold, and state and year fixed effects. Standard errors are robust to heteroskedasticity and clustered by state. *, **, and *** indicate p < 0.1, 0.05, and 0.01, respectively.
Table 4: Effects on Demographic Composition (50–115% FPL)

<table>
<thead>
<tr>
<th></th>
<th>(1) Female</th>
<th>(2) Black</th>
<th>(3) Age</th>
<th>(4) Has child</th>
<th>(5) Avg net income</th>
<th>(6) % FPL Certification</th>
<th>(7) Income limit (% FPL) / 100</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.001</td>
<td>0.059</td>
<td>0.391</td>
<td>-0.002</td>
<td>-28.557</td>
<td>0.732**</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.064)</td>
<td>(0.420)</td>
<td>(0.010)</td>
<td>(20.033)</td>
<td>(0.299)</td>
<td>(0.105)</td>
</tr>
<tr>
<td>Baseline mean</td>
<td>0.59</td>
<td>0.22</td>
<td>28.94</td>
<td>0.71</td>
<td>817.41</td>
<td>79.62</td>
<td>0.40</td>
</tr>
<tr>
<td>Observations</td>
<td>705</td>
<td>705</td>
<td>705</td>
<td>705</td>
<td>705</td>
<td>705</td>
<td>705</td>
</tr>
<tr>
<td>R²</td>
<td>0.70</td>
<td>0.81</td>
<td>0.85</td>
<td>0.84</td>
<td>0.89</td>
<td>0.70</td>
<td>0.67</td>
</tr>
</tbody>
</table>

This table presents results from estimating the effect of the SNAP eligibility threshold on the composition of enrollees earning 50–115% FPL. The columns present estimates of Equation (2) with the indicated outcome variable. The independent variable is the eligibility threshold as a ratio of the Federal Poverty Level, so that increasing by 1 corresponds to increasing the eligibility threshold from, e.g., 130% FPL to 230% FPL. In each column, we use the specification described in Equation (2), where the outcome is indicated by the column header: Column 1 shows the effect of the eligibility threshold on the fraction of the 50–115% FPL enrollee sample who are female, and so on. “Baseline mean” refers to the average of the outcome indicated by the column in state-years where the eligibility threshold is 130% FPL. Outcomes are calculated using the USDA's Quality Control (QC) data, limiting the data to households earning 50-115% FPL. Standard errors are robust to heteroskedasticity and clustered by state. *, **, and *** indicate p < 0.1, 0.05, and 0.01, respectively.
Table 5: Experiment Sample Composition and Balance for High vs. Low Treatment

<table>
<thead>
<tr>
<th></th>
<th>CPS Sample</th>
<th>Full Sample</th>
<th>Below 130% FPL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low-share</td>
<td>High-share</td>
<td>p-value</td>
</tr>
<tr>
<td>Female</td>
<td>0.517</td>
<td>0.531</td>
<td>0.658</td>
</tr>
<tr>
<td>White</td>
<td>0.776</td>
<td>0.727</td>
<td>0.623</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.165</td>
<td>0.109</td>
<td>0.824</td>
</tr>
<tr>
<td>At least some college</td>
<td>0.611</td>
<td>0.778</td>
<td>0.737</td>
</tr>
<tr>
<td>Age</td>
<td>47.714</td>
<td>45.679</td>
<td>0.526</td>
</tr>
<tr>
<td>Any Children</td>
<td>0.254</td>
<td>0.537</td>
<td>0.790</td>
</tr>
<tr>
<td>Single</td>
<td>0.291</td>
<td>0.366</td>
<td>0.927</td>
</tr>
<tr>
<td>Household Size</td>
<td>2.296</td>
<td>2.519</td>
<td>0.973</td>
</tr>
<tr>
<td>Democrat</td>
<td>-</td>
<td>0.541</td>
<td>0.275</td>
</tr>
<tr>
<td>On Food Stamps (Currently or Ever)</td>
<td>-</td>
<td>0.383</td>
<td>0.648</td>
</tr>
<tr>
<td>Household Income (000's)</td>
<td>-</td>
<td>59.007</td>
<td>0.680</td>
</tr>
</tbody>
</table>

*Census regions*
- Northeast: 0.175, 0.208, 0.191, 0.308, 0.169, 0.095, 0.014
- Midwest: 0.207, 0.190, 0.198, 0.617, 0.189, 0.209, 0.565
- South: 0.379, 0.344, 0.372, 0.176, 0.369, 0.441, 0.100
- West: 0.238, 0.259, 0.240, 0.311, 0.273, 0.255, 0.639

Joint F-test p-value: 0.941, 0.018
Observations: 2131, 512

Income uses the midpoint of a set of bins and is top-coded at $250,000. Household size is top-coded at 6. The CPS sample uses the 2019 NBER MORGs.
Table 6: Summary of Parameters for Welfare Analysis

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Primary Value</th>
<th>Range of Reasonable Values</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta_m$</td>
<td>Take-up elasticity with respect to eligibility threshold (woodwork effect)</td>
<td>0.16</td>
<td>[0.04, 0.28]</td>
<td>Table 2 (and 95% CI)</td>
</tr>
<tr>
<td>$\eta_B$</td>
<td>Take-up elasticity with respect to benefit size</td>
<td>0.5</td>
<td>[0.3, 0.6]</td>
<td>Krueger and Meyer (2002)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Coefficient of relative risk aversion</td>
<td>3</td>
<td>[1, 4]</td>
<td>Chetty and Finkelstein (2013)</td>
</tr>
<tr>
<td>$s$</td>
<td>Share of stigma-only types</td>
<td>0.5</td>
<td>[0, 1]</td>
<td>Experiment results suggest $s &gt; 0$ (Figure 8)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Cost-benefit ratio, conditional on take-up</td>
<td>0.5</td>
<td>0.5</td>
<td>Uniform costs assumption</td>
</tr>
<tr>
<td>$\frac{\partial \gamma}{\partial m}$</td>
<td>Change in cost-benefit ratio, conditional on take-up</td>
<td>0</td>
<td>0</td>
<td>Uniform costs assumption</td>
</tr>
<tr>
<td>$\lambda_m / \lambda_{avg}$</td>
<td>Ratio of marginal to inframarginal welfare weights</td>
<td>0.427</td>
<td></td>
<td>Inverse-optimum approach</td>
</tr>
</tbody>
</table>

This table summarizes the parameters used in the welfare analysis. We note the preferred value and source, but also show robustness to the range of values. The uniform costs assumption implies that $\gamma$ and $\frac{\partial \gamma}{\partial m}$ are precisely 0.5 and 0, respectively.
For Online Publication: Appendix Materials

A Empirical Appendix

A.1 SNAP Analysis Appendix
   A.1.1 Additional Figures
   A.1.2 Additional Tables
   A.1.3 Measurement Error Robustness
   A.1.4 de Chaisemartin and D’Haultfoeuille (Forthcoming) Correction

A.2 Online Experiment Appendix
   A.2.1 Auxiliary Experiment
   A.2.2 Additional Figures
   A.2.3 Additional Tables

A.3 Welfare Analysis Appendix
   A.3.1 Additional Figures
   A.3.2 Marginal Value of Public Funds (MVPF) Details

B Data and Institutional Context

B.1 SNAP Sample Construction
B.2 Broad Based Categorical Eligibility
B.3 Components of SNAP Policy Index
B.4 Experiment Sample Construction
B.5 Figure 1 Details
B.6 Figure 2 Details

C Theory Extensions and Proofs

C.1 Endogenous labor supply
   C.1.1 Simplifications
C.2 Additional Discussion of Equation (8)
C.4 Discussion of Assumption 1 and Conditions (i) and (ii)
   C.4.1 Necessary Condition for Proposition 3
   C.4.2 Discussion of Assumption 1
C.5 Proofs
   C.5.1 Proofs of Propositions 1 and 2
   C.5.2 Proof of Taylor Expansion (Equation 9)
   C.5.3 Proof of Lemma 1
   C.5.4 Proof of Proposition 3
   C.5.5 Proof of Lemma 2
   C.5.6 Proof of Proposition 4
   C.5.7 Proof of Proposition 5
A Empirical Appendix

A.1 SNAP Analysis Appendix

A.1.1 Additional Figures

Figure A.1: Rollout of Eligibility Changes Per Year

This figure presents the number of states in each year that increased (blue bars) or decreased (gray bars) eligibility to the Supplemental Nutrition Assistance Program. Four states are counted twice, because they exhibit multiple changes. Source: SNAP Policy Database.
This figure presents the maximum gross income eligibility threshold in a state from 1996–2016. The color coding refers to the maximum gross income eligibility threshold as a percent of the FPL; e.g., states colored in dark blue have maximum eligibility threshold of 200%. In two states that increase and then reduce the eligibility threshold, we present the largest eligibility threshold in the data. Source: SNAP Policy Database.
Figure A.3: Event Study of Changes to Eligibility Threshold: Without Controls

(A) Sample: 50 to 115% of FPL, No Controls

(B) Sample: 50 to 115% of FPL, Only Denominator Control

This figure is similar to Figure 4, but Panel A presents the specification with no controls beyond state and year fixed effects. Panel B presents the specification with state and year fixed effects, only controlling for the log of the total number of people between 50 and 115% of the FPL (from the CPS).
This figure presents a placebo event study with the main specification from Equation (1) but using the “Ganong-Liebman” index of SNAP policies (which are found in the USDA’s SNAP Policy Database) as the outcome. The event time is indexed around changes to state eligibility thresholds. Standard errors are robust to heteroskedasticity and clustered by state.
Panel A presents the results of estimating Equation (2) with monthly data instead of annual data. Panel B includes only SNAP recipients with any dependents—households that will not be affected by ABAWD work waivers. Panels C and D use the take-up share instead of the log of enrollment as the regressand, where the numerator in the take-up share comes from the USDA Quality Control data and the denominator uses the CPS. Panel C has no controls for state-year CPS population, while Panel D controls for the log of count of individuals in the CPS with household income in 50-115% FPL. Panel E uses the main specification and weights by population size in each state-year.
## A.1.2 Additional Tables

### Table A.1: Estimates of the Infra-marginal Effect in Alternate Samples

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Main estimate</td>
<td>Extra controls</td>
<td>Waivers, lag unemp.</td>
<td>Excludes recession</td>
<td>Weighted</td>
<td>Avg of coefficients</td>
<td>All data</td>
</tr>
<tr>
<td><strong>Panel A. 0–130% FPL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income limit (% FPL) / 100</td>
<td>0.085</td>
<td>0.087</td>
<td>0.074</td>
<td>0.086</td>
<td>0.082</td>
<td>0.076</td>
<td>0.091*</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.055)</td>
<td>(0.055)</td>
<td>(0.059)</td>
<td>(0.072)</td>
<td>(0.064)</td>
<td>(0.048)</td>
</tr>
<tr>
<td><strong>Panel B. Any dependents</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income limit (% FPL) / 100</td>
<td>0.105*</td>
<td>0.112*</td>
<td>0.096*</td>
<td>0.114*</td>
<td>0.123*</td>
<td>0.104*</td>
<td>0.133***</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.060)</td>
<td>(0.056)</td>
<td>(0.059)</td>
<td>(0.071)</td>
<td>(0.062)</td>
<td>(0.048)</td>
</tr>
</tbody>
</table>

Observations: 705 705 680 628 705 705 1071
N states: 45 45 45 45 45 45 51

This table presents Table 1 with different samples, using the specification in Equation (2). See notes to Table 1 for details. Panel A uses the sample of people at 0–130% of the Federal Poverty Line (FPL). Panel B presents estimates for the sample of households with dependents, who are not subject to ABAWD rules. The outcome is SNAP enrollment as estimated from the USDA Quality Control data. Standard errors are robust to heteroskedasticity and clustered by state. *, **, and *** indicate $p < 0.1, 0.05$, and 0.01, respectively.
A.1.3 Measurement Error Robustness

We conduct several exercises to study whether measurement error could explain our main results.

First, we present an event-study Equation 1 using the CPS values (log number eligible) as the dependent variable for the 50–115% sample (Figure A.6A) and 0–130% sample (Panel B). In both cases, we find limited reason for the concern that changes to the denominator are introduced around the time of the event. The 50–115% FPL denominator is relatively smooth around the event time; we can reject even modest changes to the denominator in all years after three years before the event. We note that the negative coefficient many years before treatment arises due to a higher value for the denominator in the reference period. Such noise in these estimates is what motivates our main specification using log enrollment counts as the outcome of interest. Moreover, we control for the CPS values in our main specifications.

Especially given that we control for the denominator, it is implausible that state populations grow fast enough only in treated state-years, beginning exactly at the time of the eligibility increase, that this measurement error could explain our event study results. Any threat to identification requires that the mismeasured portion of the denominator grows in a way that is correlated with treatment, beginning precisely at the date of treatment.

To formalize this point, we obtain the following bound on the magnitude of measurement error in the denominator required to explain our results. In state-years with an eligibility threshold above 130% of the FPL, we simulate systematic measurement error in the denominator using an “inflated” denominator that we define as:

\[
\text{simulated denominator} := \text{observed denominator} \times \text{inflation factor},
\]

where the inflation factor represents the magnitude of simulated measurement error. For instance, an inflation factor of 1.05 represents the case where we replace the treated state-years’ denominators as being 5% larger than what we observe in the CPS.

We then estimate Equation 2 with the simulated denominator in treated state-years. We find that the inframarginal effect vanishes only if the denominator in treated state-years is inflated by more than 30% (Figure A.7). Put another way, only when we add an additional 30% of the population to the denominator (and impose that this measurement error only exists in treated state-years) can we eliminate the inframarginal effect. As a benchmark, we note that the average state population between 50 to 115% of the FPL (i.e., the denominator) grew by 26% between 2001 and 2016. Thus the measurement error required to reverse our result would need to be larger than the entire observed population growth in the sample period. It is implausible that only treated states are subject to measurement error that is this extreme.

Altogether, while our denominator obtained from the CPS may be subject to some measurement error, it would have to be systematically correlated with treatment to an implausible degree in order to explain our results.
This figure presents a placebo event study, as described in Section 2, modifying Equation (1). Panel A uses the sample of people from 50–115% of the Federal Poverty Level (FPL); Panel B uses the sample of people from 0–130% of the FPL. The federal minimum eligibility threshold is 130% of the FPL. We present the take-up denominator, i.e. the CPS counts of individuals in each sample, as the outcome variable. The figure shows there is no discrete change to the denominator at the event time. The USDA Quality Control (QC) data provide estimates of the numerator for the outcome (take-up counts, by state-year), and the Current Population Survey (CPS) data provide estimates of the denominator (total counts of individuals within this sample). Standard errors are robust to heteroskedasticity and clustered by state.
This figure presents a bound on the amount of measurement error in the denominator that would be required to reverse our results. In states where the eligibility threshold exceeds 130% of the FPL, we inflate the observed population between 50–115% of the FPL by the factor on the $x$-axis. We then present the estimate of the inframarginal effect from Equation (2), estimated using the simulated denominator. Only if the population is inflated by 30% can we reverse the inframarginal effect.
We implement the estimator from de Chaisemartin and D’Haultfoeuille (Forthcoming) (Table A.2). We use the event-study sample and specification (Equation (1)), but recode treatment as a binary indicator for whether an eligibility expansion occurred. We employ this adjustment because we have a small number of treated units who receive each eligibility threshold and therefore assumption S8 in de Chaisemartin and D’Haultfoeuille (Forthcoming)’s web appendix may not hold. If this assumption does not hold, de Chaisemartin and D’Haultfoeuille (Forthcoming) recommend re-binning a continuous treatment variable.

We also implement de Chaisemartin and D’Haultfoeuille (Forthcoming)’s placebo tests, which examine whether take-up trends in treated units from year $t-1$ to $t$ were parallel. We show results for the main event-study sample (Column 1) and a sample that additionally controls for lag unemployment (Column 2). Because we implement a large number of placebo tests, we focus on the $p$-value from a joint test of $H_0 : \text{all placebos} = 0$, and a test of $H_0 : \text{the average effect across placebos} = 0$. The test of average effects tells us whether there was a differential trend that persisted across years, rather than one year that violated a differential trend due to noise. Across both placebo tests we fail to find evidence of differential take-up prior to treatment, although one placebo (a test for differential trends between $-4$ and $-3$ years) is individually significant. Given the large number of tests, failure to reject once we aggregate the placebos, and small economic magnitude of the violation of the placebo (relative to the treatment effects in the post period), we find limited reason for concern.
Table A.2: Negative Weight Correction to Event Study Design

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>-4</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>-3</td>
<td>0.039**</td>
<td>0.039**</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>-2</td>
<td>0.012</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>-1</td>
<td>-0.011</td>
<td>-0.020</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.019)</td>
</tr>
<tr>
<td>0</td>
<td>0.021</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>1</td>
<td>0.060**</td>
<td>0.060**</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>2</td>
<td>0.078**</td>
<td>0.070**</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>3</td>
<td>0.082**</td>
<td>0.072**</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>4</td>
<td>0.108***</td>
<td>0.103***</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.036)</td>
</tr>
<tr>
<td>5</td>
<td>0.133***</td>
<td>0.134***</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.039)</td>
</tr>
</tbody>
</table>

Placebo p-value: average coefficient $\neq 0$ 0.174 0.366
Placebo p-value: any coefficient $\neq 0$ 0.178 0.121

Lag unemployment controls No ✓

Notes: This table implements the estimator from [de Chaisemartin and D’Haultfoeuille (Forthcoming)] for a version of Equation (1), using a binary variable for whether a state expands eligibility. The negative periods represent the placebo tests recommended by [de Chaisemartin and D’Haultfoeuille (Forthcoming)]; for example, the row labeled -1 represents a test for a difference in trends 2 and 1 year before treatment. We aggregate placebo tests using: (i) a joint test of significance across placebos, (ii) and a test for whether the average difference across placebos is different than 0. Column 1 shows the effect on the main event-study sample. Column 2 additionally controls for lag unemployment. * *, **, and *** indicate $p < 0.1, 0.05,$ and 0.01, respectively.

A.2 Online Experiment Appendix

A.2.1 Auxiliary Experiment

The results of the second experiment are mixed (Table A.7). We find no evidence for an effect of a belief correction exercise on first-order beliefs. We find a positive effect of the belief correction on second-order beliefs: for people whose priors were below the truth, correcting beliefs raises the stigma they report (point estimate: 0.069, SE: 0.041, $p = 0.091)$.

We note that the treatment effect is positive for people whose beliefs are corrected down (point estimate: 0.018, SE: 0.035). This point estimate is consistent with the results from the high-state treatment. Alternatively, it may suggest that any belief correction may simply cause participants to report more stigma, e.g. because they do not like being corrected after receiving an initial hint. We also present effects with demographic controls (Table A.8), which are similar. In this case, the positive effect on second-order stigma for correcting beliefs upward is very slightly attenuated.

We are more cautious about interpreting the results from auxiliary experiment for the following reasons. First, people who are shown multiple pieces of information might simply end up confused, which could attenuate or undo its effects. Because we did not elicit beliefs after being shown the belief correction, we do
not have a way of checking how the correction actually shifted posteriors. The inconclusive results suggest that providing the second piece of information might have had an unintended consequence of causing participants to tune out the second piece of information, perhaps because it was perceived as contradicting the first piece of information.

Second, the auxiliary belief correction only operates on people after they have been shown a hint. As a result, because it is cross-randomized, it affects the group of people that do or do not comply with the high or low treatment. The staggered nature of the design complicates this interpretation: people who have low prior beliefs after treatment are a selected group, since they have been exposed to a hint that causes them to update.  

Third, the belief-correction treatment, when paired with the high-share treatment, affects people’s beliefs about the distribution of eligibility thresholds across states. If stigma is linked to people’s beliefs about the distribution of eligibility thresholds, it is not clear how the combination of experiments affects stigma.

Altogether, the main experiment provides a somewhat cleaner test of the null hypothesis that stigma plays no role in inframarginal effects. Nevertheless, the inconclusive results from the auxiliary experiment lead us to interpret the experiment with some caution.

\footnote{Consistent with this point, the positive treatment effect on second-order stigma from correcting beliefs upward attenuates once we add demographic controls (Table A.3).}
This figure presents the experiment design. The donation choice was to one of four charities (used to incentivize belief elicitation). We elicited several demographics (in addition to those provided by Lucid).
This figure presents correlations between the stigma questions in the order they were elicited. Section 4 provides the question texts. We classify questions 1, 2, 4, and 8 (labeled in red) as first-order stigma. We classify questions 3, 5, 6, and 7 (labeled in blue) as second-order stigma.
This figure presents treatment effects and 95% confidence intervals of the high-share treatment on the second-order stigma index (Equation (16)), split by demographic group.
### A.2.3 Additional Tables

#### Table A.3: Online Experiment: Randomization Balance for Belief Correction

<table>
<thead>
<tr>
<th></th>
<th>CPS Sample</th>
<th>Full Sample</th>
<th>Below 130% FPL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Correction</td>
<td>Belief Correction</td>
<td>p-value</td>
</tr>
<tr>
<td>Female</td>
<td>0.517</td>
<td>0.515</td>
<td>0.537</td>
</tr>
<tr>
<td>White</td>
<td>0.776</td>
<td>0.742</td>
<td>0.722</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.165</td>
<td>0.109</td>
<td>0.113</td>
</tr>
<tr>
<td>At least some college</td>
<td>0.611</td>
<td>0.783</td>
<td>0.767</td>
</tr>
<tr>
<td>Age</td>
<td>47.714</td>
<td>45.770</td>
<td>46.048</td>
</tr>
<tr>
<td>Any Children</td>
<td>0.254</td>
<td>0.528</td>
<td>0.540</td>
</tr>
<tr>
<td>Single</td>
<td>0.291</td>
<td>0.366</td>
<td>0.368</td>
</tr>
<tr>
<td>Household Size</td>
<td>2.296</td>
<td>2.530</td>
<td>2.507</td>
</tr>
<tr>
<td>Democrat</td>
<td>-</td>
<td>0.525</td>
<td>0.533</td>
</tr>
<tr>
<td>On Food Stamps (Currently or Ever)</td>
<td>-</td>
<td>0.377</td>
<td>0.398</td>
</tr>
<tr>
<td>Household Income (000's)</td>
<td>-</td>
<td>61.526</td>
<td>57.476</td>
</tr>
</tbody>
</table>

*Income uses the midpoint of a set of bins and is top-coded at $250,000. Household size is top-coded at 6. The CPS sample uses the 2019 NBER MORGs.*
### Table A.4: Online Experiment: Attrition Balance

<table>
<thead>
<tr>
<th></th>
<th>Total N</th>
<th>High-share treatment</th>
<th>Beliefs correction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All &lt;= 130% FPL</td>
<td>All &lt;= 130% FPL</td>
</tr>
<tr>
<td><strong>1. Any attrition or drops</strong></td>
<td>567</td>
<td>0.009 (0.016)</td>
<td>0.018 (0.033)</td>
</tr>
<tr>
<td>2. Bad priors</td>
<td>237</td>
<td>0.002 (0.011)</td>
<td>0.008 (0.024)</td>
</tr>
<tr>
<td>3. Attrited before share treatment</td>
<td>49</td>
<td>0.004 (0.005)</td>
<td>-0.004 (0.010)</td>
</tr>
<tr>
<td>4. Attrited at or after treatment</td>
<td>126</td>
<td>0.006 (0.008)</td>
<td>0.007 (0.019)</td>
</tr>
<tr>
<td>5. Omitted any stigma answers</td>
<td>107</td>
<td>0.002 (0.008)</td>
<td>0.005 (0.016)</td>
</tr>
<tr>
<td>6. Inattentive</td>
<td>106</td>
<td>0.000 (0.008)</td>
<td>0.021 (0.013)</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>2,698</td>
<td>689</td>
<td>2,698</td>
</tr>
</tbody>
</table>

This table shows that attrition and drops were balanced across treatment and control. Each row tests for balance between treatment and control on a different dummy outcome. The first column gives the total number of respondents who were dropped for the reason indicated by the row. Note that respondents could be dropped for multiple reasons. The next two columns show balance for the main experiment, where respondents were provided a random hint about the share of Americans eligible for SNAP. The last two columns show balance for the secondary experiment, where respondents beliefs were corrected with the true share. Row 1’s outcome is a dummy for attriting or being dropped from the sample. Row 2’s outcome is a dummy for providing prior beliefs about the share of Americans eligible for SNAP that were below 1% or above 99%, or skipping this question entirely. Row 3’s outcome is a dummy for dropping out of the survey before the treatment screen. The second two columns of this row are empty because individuals who attrited before the treatment screen were not randomized into treatment or control for the beliefs correction. Row 4’s outcome is a dummy for attriting at or after the share treatment screen. Row 5’s outcome is a dummy for not answering any of the stigma questions. Row 6’s outcome is a dummy for failing an attention check. *, **, and *** indicate $p < 0.1$, 0.05, and 0.01, respectively.
Table A.5: Online Experiment: High-Share Effect on Reported Stigma, without Demographic Controls

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Subindices</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First-Order</td>
<td>Second-Order</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under 130% FPL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-share treatment</td>
<td>-0.032</td>
<td>0.046</td>
<td>-0.109*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.065)</td>
<td>(0.058)</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.530</td>
<td>0.485</td>
<td>0.061</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>512</td>
<td>512</td>
<td>512</td>
<td></td>
</tr>
<tr>
<td>Full Sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-share treatment</td>
<td>-0.013</td>
<td>0.025</td>
<td>-0.050*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.031)</td>
<td>(0.027)</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.598</td>
<td>0.421</td>
<td>0.061</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2,131</td>
<td>2,131</td>
<td>2,131</td>
<td></td>
</tr>
</tbody>
</table>

The table shows the effect of the “high-share” hint on individuals’ level of agreement to statements measuring stigma around food stamps and welfare for individuals under 130% FPL (top panel) and the full sample (bottom panel) (Equation (16)). The estimates are identical to Figure 8. *, **, and *** indicate $p < 0.1, 0.05,$ and 0.01, respectively.
Table A.6: Online Experiment: High-Share Effect on Reported Stigma, with Demographic Controls

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Subindices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First-Order</td>
<td>Second-Order</td>
</tr>
<tr>
<td>Under 130% FPL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-share treatment</td>
<td>-0.023</td>
<td>0.049</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>p-value</td>
<td>0.640</td>
<td>0.448</td>
</tr>
<tr>
<td>Observations</td>
<td>512</td>
<td>512</td>
</tr>
<tr>
<td>Full Sample</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-share treatment</td>
<td>-0.016</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>p-value</td>
<td>0.489</td>
<td>0.580</td>
</tr>
<tr>
<td>Observations</td>
<td>2,131</td>
<td>2,131</td>
</tr>
</tbody>
</table>

The table shows the effect of the “high-share” hint on individuals’ level of agreement to statements measuring stigma around food stamps and welfare for individuals under 130% FPL (top panel) and the full sample (bottom panel) (Equation (16)). It is identical to Table A.5 and Figure 8 except we include demographic controls for: an age quadratic, income, political party, gender, region, household size, marital status, having children, being on or ever having been on food stamps, and education and race/ethnicity fixed effects. *, **, and *** indicate $p < 0.1$, 0.05, and 0.01, respectively.
This table shows results from the second experiment embedded in our online survey, where respondents were informed of the true share of Americans eligible for SNAP after previously being asked to report their beliefs (and given a hint, which is the primary experiment discussed in the text). It presents treatment effect estimates from Equation (16). Panel A restricts the sample to those who initially underestimated the eligibility share, so that the treatment should have led them to revise upwards. Panel B restricts the sample to those who initially overestimated the eligibility share, so that the treatment should have decreased their beliefs. *, **, and *** indicate $p < 0.1, 0.05,$ and $0.01,$ respectively.
### Table A.8: Online Experiment: Belief Correction, With Demographic Controls

<table>
<thead>
<tr>
<th>Overall</th>
<th>Subindices</th>
<th>First-Order</th>
<th>Second-Order</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Priors &lt; Truth</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beliefs Correction Treatment</td>
<td>0.036</td>
<td>0.006</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.045)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>Observations</td>
<td>868</td>
<td>868</td>
<td>868</td>
</tr>
<tr>
<td>p-value</td>
<td>0.301</td>
<td>0.900</td>
<td>0.103</td>
</tr>
<tr>
<td><strong>Panel B. Priors ≥ Truth</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beliefs Correction Treatment</td>
<td>0.032</td>
<td>0.034</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.038)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,263</td>
<td>1,263</td>
<td>1,263</td>
</tr>
<tr>
<td>p-value</td>
<td>0.290</td>
<td>0.375</td>
<td>0.389</td>
</tr>
</tbody>
</table>

This table shows results from the second experiment embedded in our online survey, where respondents were informed of the true share of Americans eligible for SNAP after previously being asked to report their beliefs (and given a hint, which is the primary experiment discussed in the text). It presents treatment effect estimates from Equation (16). Panel A restricts the sample to those who initially underestimated the eligibility share, so that the treatment should have led them to revise upwards. Panel B restricts the sample to those who initially overestimated the eligibility share, so that the treatment should have decreased their beliefs. This table is identical to Table A.7 except we additionally include demographic controls for: an age quadratic, income, political party, gender, region, household size, marital status, having children, being on or ever having been on food stamps, and education and race/ethnicity fixed effects. *, **, and *** indicate \( p < 0.1, 0.05, \) and \( 0.01 \), respectively.
Table A.9: Online Experiment: Treatment Effect by Belief-Correction Randomization

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First-order index</td>
<td>Second-order index</td>
<td>First-order index</td>
<td>Second-order index</td>
</tr>
<tr>
<td>High-share treatment</td>
<td>0.048</td>
<td>-0.019</td>
<td>0.003</td>
<td>-0.081**</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.037)</td>
<td>(0.044)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>Observations</td>
<td>1050</td>
<td>1050</td>
<td>1081</td>
<td>1081</td>
</tr>
<tr>
<td>Sample</td>
<td>Not shown truth</td>
<td>Not shown truth</td>
<td>Shown truth</td>
<td>Shown truth</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

This table presents treatment effects on first- and second-order stigma from Equation (16) the sample by whether the sample’s beliefs were not truthfully corrected (Columns 1 and 2) or were truthfully corrected (Columns 3 and 4). *, **, and *** indicate p < 0.1, 0.05, and 0.01, respectively.
Table A.10: Online Experiment: Association Between Take-Up and Stigma

<table>
<thead>
<tr>
<th></th>
<th>(1) On SNAP (currently or ever)</th>
<th>(2) On SNAP (currently or ever)</th>
<th>(3) On SNAP (currently or ever)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First-order index</strong></td>
<td>-0.133***</td>
<td>-0.145***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td><strong>Second-order index</strong></td>
<td>-0.012</td>
<td>0.044**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.018)</td>
<td></td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>0.391***</td>
<td>0.388***</td>
<td>0.391***</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.010)</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>2131</td>
<td>2131</td>
<td>2131</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

This table presents associations between first- and second-order stigma and participants’ reports about taking up SNAP (now or in the past). We elicit the take-up questions before treatment. *, **, and *** indicate $p < 0.1$, 0.05, and 0.01, respectively.
A.3 Welfare Analysis Appendix

A.3.1 Additional Figures

Figure A.11: Robustness: Naïve Planner’s Biased Risk Aversion

(A) True $\rho = 2$

This figure shows the percent bias between the planner’s “as-if” risk aversion ($\tilde{\rho}$) and the ground-truth risk aversion ($\rho$) (black line). It is identical to Figure 9A except it sets $\rho = 2$ (Panel A) or $\eta_B = 0.3$ (Panel B). Negative numbers indicate that the planner is behaving as if people are less risk averse than they really are. Panel A plots the bias as a function of the inframarginal effect; the vertical gray line plots the empirical inframarginal effect presented in Table 6.
This figure shows the percent bias in the welfare impact per dollar of government expenditure (Equation \( A.2 \)) for an inframarginal to marginal cost ratio of 1 (Panel A) and 10 (Panel B). The vertical red line plots our preferred estimate of the inframarginal effect \( \bar{w} \) in terms of take-up. See Section A.3.2 for details.
Figure A.13: Numerical Simulations: Robustness

(A) Varying the Coefficient of Relative Risk Aversion ($\rho$)

(B) Varying Take-Up Elasticity with Respect to Benefit Size ($\eta_B$)

This figure shows the results from our numerical simulation exercise, which uses the optimality condition in Equation (13). It presents changes in the percent of people who are eligible if the planner acknowledges inframarginal effects. It shows robustness to different $\rho$ (Panel A) and take-up elasticities with respect to the benefit size ($\eta_B$). Auxiliary parameters are set according to the values in Table 6.
This figure shows the results from our numerical simulation exercise, which uses the optimality condition in Equation (13). It presents the percent change in the optimal benefit size if the planner acknowledges inframarginal effects. Auxiliary parameters are set according to the values in Table 6.
This figure shows the results from our numerical simulation exercise, which uses the optimality condition in Equation (13). It presents the change in the percent of people who are eligible if the planner acknowledges inframarginal effects. It is identical to Figure 10A except the simulations impose quadratic utility with $\rho = 3$ at equilibrium, using Equation (1) with $\eta_m = 0$ to infer the welfare weights.
A.3.2 Marginal Value of Public Funds (MVPF) Details

Relation to Optimality Condition in Section 3.1 The MVPF of policy $P$ is the total willingness to pay (WTP) for the policy (where WTP = $\frac{dU}{dp}$) divided by its net cost to the government. In Section 3, we simply make clear how this relates to the optimality condition in our benchmark model, when we remain committed to the specific modeling assumptions used in that section. Later, in Section 5, we give a more general approach to understanding the welfare implications of woodwork effects through an MVPF lens. In this appendix, we show how we derive the MVPFs in Section 3 and how they could be used to derive Equation 8.

Equation (10): Consider a policy that increases the benefit size by $dB$. An individual is willing to pay for this policy if they would take up the program regardless (otherwise, if it leads them to join the program, they are indifferent by the Envelope Theorem). Their WTP is $u'(B)dB / u'(B) = dB$. Integrating over the mass of people who are eligible and taking up, total WTP is $p * m * dB$. The cost to the government is the mechanical effect $(p * m * dB)$ plus our main effect, the inframarginal effect $(dp * m * B)$. Thus we have

$$\text{MVPF}_B = \frac{\text{WTP}}{\text{Cost}} = \frac{p * m * dB}{p * m * dB + dp * m * B} = \frac{1}{1 + \eta_B}.$$

Equation (11): Consider a policy that increases the income threshold of eligibility by $dm$. An individual is willing to pay for this policy if they are newly eligible or if they would take up regardless (and benefit from decreasing costs). Those who are newly eligible have WTP $= \frac{u(B)(1-\gamma)}{u'(0)}$. Integrating over the mass of newly eligible people who take up gives $\frac{m}{u'(B)} * \int_0^u \frac{H(c|m)}{dm} dc * dm$. The cost to the government is the mechanical effect $dm * p * \frac{u(B)}{u'(0)}$. Those who would take up regardless enjoy a decrease in costs, and integrating over the mass of these people gives $\frac{m}{u'(B)} * \int_0^u \frac{H(c|m)}{dm} dc$. Thus we have

$$\text{MVPF}_M = \frac{\text{WTP}}{\text{Cost}} = \frac{\left(p * \frac{u(B)(1-\gamma)}{u'(0)} + \frac{m}{u'(B)} * \int_0^u \frac{H(c|m)}{dm} dc \right) dm}{dm * p * B + dp * m * B} = \frac{\left(\frac{u(B)}{u'(0)} (1-\gamma) + \frac{m}{u'(B)} \int_0^u \frac{H(c|m)}{dm} dc \right)}{1 + \eta_m}.$$

From Lemma C.3 in Appendix C.3, we have that $\eta_m (1-\gamma) - m \frac{\gamma}{dm} = \frac{\int_0^u \frac{H(c|m)}{dm} dc}{u(B)}$. Now we can write

$$\text{MVPF}_M = \frac{1}{u'(B)} \left(\frac{u'(B)}{u'(0)} (1-\gamma) - m \frac{\gamma}{dm} \right).$$

Finally, to derive Equation 8, weight the WTPs by the appropriate social marginal utilities of money to get welfare impacts $WI$. Individuals who were previously eligible and taking up have weights $\lambda_{avg} u'(B)$. Those who are newly eligible have weights $\lambda_m u'(0)$.

$$WI^B = WI^M \lambda_{avg} u'(B) * 1 \frac{1}{1 + \eta_B} = \frac{\lambda_{avg} u'(B) (u(B) / B) \left(\frac{\lambda_{avg} u'(0)}{u'(B)} \frac{u'(B)}{u'(0)} (1-\gamma) - m \frac{\gamma}{dm} \right)}{1 + \eta_m}.$$

$$\lambda_{avg} u'(B) = \lambda_{avg} \left(\frac{u(B)}{B} \left(\frac{\lambda_{avg} u'(0)}{u'(B)} \frac{u'(B)}{u'(0)} (1-\gamma) - m \frac{\gamma}{dm} \right)\right).$$

$$\lambda_{avg} u'(B) = \lambda_{avg} \left(\frac{u'(B)}{B} (1-\gamma) - m \frac{\gamma}{dm} \right).$$

$$u'(B) = \frac{(u'(B) / B) \left(\frac{\lambda_{avg} u'(0)}{u'(B)} \frac{u'(B)}{u'(0)} (1-\gamma) - m \frac{\gamma}{dm} \right)}{1 + \eta_m}.$$
SNAP. Let WTP be the willingness to pay for the benefit B.

The naïve welfare impact per dollar of government expenditure of an eligibility increase is: $W^\text{naive} := \frac{\lambda m \Delta \mu_m (1 - \gamma) WTP}{\Delta \mu_m k_m}$, with $\lambda m := \mu_m \lambda_m$, where $\mu_m$ is the marginal utility of income for newly eligible population. The denominator $k_m$ is the entire fiscal cost per person of providing food stamps to the next share $\Delta$ of the population (including the fiscal externality).\footnote{In a standard MVPF case we might think of this as 1 + the Fiscal Externality, but here, the cost of providing SNAP to some share of the population is more than $1. Assume the WTP to take up the benefit is the same for all $\theta$.}

Let $WTP^s$ be the willingness to pay for the reduction in stigma costs from increasing the eligibility threshold to the next share $\Delta$ of the population. Let $\tilde{\alpha}$ satisfy $WTP^s := \tilde{\alpha} WTP$, where $\tilde{\alpha} < 1$, and $\mu_B$ and $\mu_0$ are the marginal utilities of income of previously and newly enrolled, respectively.

Stigma agents who newly take up due to the inframarginal effect are just indifferent due to the Envelope Theorem.\footnote{We assume that there are not utility gains to individuals who are not decision-makers (e.g., the children of SNAP recipients). Otherwise, while newly-enrolled, inframarginal stigma agents have no first-order welfare gains, there would be utility gains from their children.} Stigma agents who previously took up the benefit have a positive willingness to pay for the reduction in stigma costs from increasing the eligibility (which we assume is constant across all types), and $\text{Utility of income}$

<table>
<thead>
<tr>
<th>Mechanical effect for marginal types</th>
<th>Inframarginal effect for info agents</th>
<th>Reduction in stigma costs for stigma agents</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W^\text{naive} := \lambda m \Delta \mu_m (1 - s\gamma) WTP$</td>
<td>$\int_0^m \mu_\theta WTP(1 - s)wd\theta$</td>
<td>$\Delta p_m K_m$</td>
</tr>
<tr>
<td>$\int_0^m \mu_\theta WTP(1 - s)wd\theta$</td>
<td>$\Delta p_m K_m$</td>
<td></td>
</tr>
</tbody>
</table>

Cost of inframarginal effect

Cost of expansion to marginal types

(A.1)

with

$$\tilde{\lambda} := \frac{\Delta \lambda m \mu_m p_m (1 - s\gamma) WTP + \int_0^m \mu_\theta \lambda_\theta WTP(1 - s)wd\theta + \int_0^m \mu_\theta \lambda_\theta \tilde{\alpha} p_{avg}s WTPd\theta}{\Delta \mu_m (1 - s\gamma) WTP + \int_0^m \mu_\theta WTP(1 - s)wd\theta + \int_0^m \tilde{\alpha} p_{avg}s WTPd\theta}$$

where $\mu_\theta$ is the marginal utility of income for person $\theta$ who is not previously enrolled on the program, $\mu_\theta^E$ is the marginal utility of income for person $\theta$ who is an ‘always-taker,’ $w$ is the proportion increase in inframarginal take-up (inframarginal effect), $s_{avg}$ is the take-up rate for inframarginal types prior to the eligibility expansion (which we assume is constant across all types), and $k_\theta$ is the total fiscal cost of an additional 1 pp inframarginal take-up of type $\theta$ (including the fiscal externality).

Assume that all individuals who would newly enroll in the program have marginal utility of income $\mu_\theta = \mu_0$ before the policy change, and individuals who would previously enrolled have marginal utility of income $\mu_\theta^E = \mu_B$. Define $\alpha$ as $\tilde{\alpha} \approx \frac{\mu_0}{\mu_B}$.

Defining the bias as $\frac{W^\text{naive} - W^{\text{naive}}}{W^\text{naive}}$, algebra gives:

$$\text{bias} = \frac{p_m + \frac{m \Delta \mu \mu_{avg}}{\Delta \mu m} k_{avg}}{p_m + \frac{m \Delta \mu \mu_{avg}}{\Delta \mu m} \frac{\lambda_{avg}}{\lambda_m} \frac{1}{(1 - s)w} + s p_{avg} \tilde{\alpha}} - 1,$$

(A.2)

where $k_{avg} := \frac{\int_0^m k_\theta wd\theta}{m}$, by analogy to $\lambda_{avg}$.

\footnote{This is analogous to the $\gamma$ in the model in Section 3; stigma agents face costs which erode some fraction of their WTP.}
Note also that, if \( s = 0, \ w > 0, \) and \( \frac{\lambda_{avg}}{\lambda_m} > \frac{k_{avg}}{k_m}, \) then bias < 0. Intuitively, as long as the planner’s valuation of inframarginal types exceeds their fiscal cost, inframarginal effects raise the welfare impact of an eligibility increase. In the stigma case, the planner also values the reduction in costs to inframarginal types.

**Parameters.** We estimate \( w \) using instrumental variables, as in Table 2, where we instrument for the share of a state’s population that is eligible using the eligibility threshold; here, however, we regress the take-up rate on the share eligible, instead of estimating an elasticity. The result is \( w = 0.0038 \): take-up increases by 0.38 percentage points for every 1 percentage point increase in the share eligible. We continue to assume today’s take-up rate, \( p_{avg} = p_m = 0.53, \) and eligibility threshold \( m = 0.27. \) We assume \( \gamma = 0.5, \) analogous to the calibration used elsewhere in the paper. We assume \( \frac{\lambda_{avg}}{\lambda_m} = 0.43 \) (Table 6).
B Data and Institutional Context

B.1 SNAP Sample Construction

We build off the sample in Ganong and Liebman (2018), and adapt their public-use code and data associated with the published paper. We extend the sample to 2016. Our main outcome (the number of people enrolled in SNAP, for different income groups) uses the USDA’s Quality Control (QC) data from 1996–2016. The QC data provides information on the household’s income as a fraction of the FPL. We use the QC data (together with its household weights) to obtain counts of the number of people in a given state-year that enroll in SNAP who are within some income band (as a fraction of the FPL).

In our welfare exercise and in some supplemental analyses, we are interested in SNAP take-up rates. For these, we treat the QC data as the numerator in the take-up rate, and form the denominator from the CPS, which contains the count of people within a household income band in each state and year.

Our data on state-level SNAP policies, including the income eligibility threshold and other policies (e.g., outreach spending), come from the USDA’s SNAP Policy Database (2019).

The QC data include individuals in the household who are not in the SNAP unit. As in Ganong and Liebman (2018), we include these individuals as taking up SNAP. Many of these individuals are relatives of the individuals in the SNAP unit and may, in practice, have their consumption subsidized by SNAP. Results are very similar if we limit only to individuals in the SNAP unit.

B.2 Broad Based Categorical Eligibility

We provide more information about the BBCE provision that permits states to expand SNAP eligibility.

Broad Based Categorical Eligibility permits states to expand eligibility using Temporary Assistance for Needy Families (TANF) or State Maintenance of Effort (MOE) budgets. States cannot expand eligibility beyond 200% of the FPL.

There are two concerns about other effects of the BBCE that could affect our analysis of inframarginal effects. In practice, states are legally required to fund small auxiliary services (e.g., telephone hotlines) using TANF/MOE funds in order to grant eligibility to more people in SNAP. Congressional Research Service (2019) writes:

“As of July 2019, 42 jurisdictions have implemented what the U.S. Department of Agriculture (USDA) has called “broad-based” categorical eligibility. These jurisdictions generally make all households with incomes below a state-determined income threshold eligible for SNAP. States do this by providing households with a low-cost TANF-funded benefit or service such as a brochure or referral to a telephone hotline. There are varying income eligibility thresholds within states that convey “broad-based” categorical eligibility, though no state may have a gross income limit above 200% of the federal poverty guidelines.”

The first concern, which we address in Section 2.6, is that this policy requires that SNAP administrators must notify households that they are eligible. In practice, the policy discussion around BBCE centers around the eligibility expansion, and the notification of receipt may not be much different than typical state efforts to notify recipients, especially for households below 115% of the FPL. The core of our robustness tests uses states that are treated with BBCE but do not expand eligibility. We find no evidence take-up increases in these states.

A secondary concern is that BBCE expansions sometimes waive asset rules. We also address this concern in Section 2.6.

B.3 Components of SNAP Policy Index

We use the SNAP policy index defined in Ganong and Liebman (2018), but without the BBCE. It is the average of dummies for each of seven policies. Six policies are directly from the SNAP Policy Database (2019). These are defined to be 1 if at least some parts of the state use the policy:

- At least one household vehicle is exempted from the asset test.
• Households with at least one recipient of Supplemental Security Income can use a simplified application for SNAP.

• Households can recertify with a telephone interview instead of a face-to-face interview.

• Households can apply to SNAP online.

• The state has fewer requirements for reporting changes in household earnings.

• There are call centers in the state for households to ask questions about SNAP, and in some places, recertify.

The final policy is a dummy if fewer than 20% of households have a certification period of 3 months or less, indicating that only a low share of SNAP households in the state must recertify at frequent intervals.
The index averages all seven policies except for when information about vehicle exemptions is unavailable; in this case, we average the remaining six.
In cases in which the index varies throughout the year, we use the minimum of the index in that year.

B.4 Experiment Sample Construction
We document several data cleaning decisions.

• A small number of participants had missing information about their household size or composition. We assume people with missing information were single, non-married, with no children (so had a household size of 1).

• A small number of participants had missing income. We assume they were in the bottom income bin and therefore had an income of $7,500.

• We top-coded household size at 6 because the most number of children that participants could report was 4.

• Incomes were top-coded at $250,000. We assume these participants had incomes of $300,000.

• Fewer than five participants took the experiment multiple times, and we drop them.

• **Below 130% FPL Sample.** To form the “inframarginal” sample of experiment respondents, we predicted the relevant 2020 poverty threshold for each respondent using (1) the midpoint of their household income bin and (2) their household size, constructed via their marital status and number of kids. Anyone who reported a household income bin with a midpoint below $1.3 \times \text{result is included in the sample of respondents under 130% FPL. This may have excluded some respondents from the inframarginal sample if they were also living with or supporting parents or elders.}

B.5 Figure Details
We collected income eligibility rules and take-up rates from various sources for a subset of U.S. social programs. To the extent possible, all values are from 2016. The set of programs was determined by the following process: We began by limiting to programs with FY 2016 budgets over $5 billion. We eliminated tax credits. Then we eliminated the following programs for specific reasons. We eliminated Section 8 Housing because the notion of participation is difficult to define where there are long wait lists and barriers to take-up are very high (often requiring moving). We eliminated Old Age Assistance and Social Security because income-based means tests are not meaningful for a population that often does not work and lives in households with other earners. Finally, we eliminated Pell Grants because eligibility is not based on a specific income threshold.

• CHIP
– Eligibility data are from Brooks et al. (2016), Table 1, which gives income thresholds for children’s eligibility to receive Medicaid or CHIP benefits, assuming a family of 3. In some states, the income threshold varies for different subgroups. The figure uses a population-weighted average of all the states’ highest income thresholds.

– The take-up rate is from Appendix Exhibit 1 of Haley et al. (2018), also as referenced by The Kaiser Family Foundation (KFF).

• EITC

– Eligibility is calculated using the IRS.gov EITC maximum allowable AGI for a family of three.

– The take-up rate is from the IRS.gov “About EITC” webpage (Internal Revenue Service 2020), estimated by the Census Bureau using the CPS.

• Head Start

– Eligibility is generally 100% of the FPL (HHS).

– The take-up rate was calculated as follows:

  1. Participation rates are 35% (Child Trends 2018), calculated using the total number of children enrolled in Head Start divided by the total number of children in poverty (ages 3-5).

  2. However, Head Start is oversubscribed. We use details from the Head Start Impact Study (U.S. Department of Health and Human Services 2010): this study found that 85% of Head Start centers were oversubscribed. Within oversubscribed Head Start centers, the study randomized 60% of applicants into acceptance, while the remaining 40% were wait listed. In some centers, not all applicants were included in the randomization; in others, there were not enough applicants to attain this ratio in the randomization. We assume that take-up is 35*(1)*(15%) + 35*(10/6)*(85%). That is, the take-up rate is 35% among the 15% of centers which were not oversubscribed and 35*(10/6) in the oversubscribed centers, on average.

• Medicaid (parents only)

– Eligibility data are from Brooks et al. (2016), Table 5, which gives income thresholds for parents’ eligibility to receive Medicaid or CHIP benefits, assuming a family of 3. In some states, the income threshold varies for different subgroups. The figure uses a population-weighted average of all the states’ highest income thresholds for parents.

– The take-up rate is from Appendix Exhibit 2 of Haley et al. (2018), as referenced by KFF.

• NSLP (National School Lunch Program)

– Eligibility for free lunch is 130% FPL in most districts; eligibility for reduced-price lunch is 185% FPL in most districts.

– The take-up rate is calculated as follows:

  1. First, we take the total number of students eligible for free or reduced-price lunch in the 2015-2016 school year, according to Table 204.10 in National Center for Education Statistics (2017). This is around 26 million.

  2. We take the average number of free and reduced-price meals served daily in 2016, provided by the USDA Food and Nutrition Service: around 22 million (United States Department of Agriculture Food and Nutrition Service 2020a).

  3. The take-up rate is 22 / 26

• SNAP

– Eligibility data use a population-weighted average of states’ eligibility thresholds.

– The take-up rate is from Cunningham (2019), which gives estimates of 2016 take-up rates.

• TANF (Temporary Assistance for Needy Families)
Eligibility data are from Giannarelli et al. (2017), which provides, for all states, the income cutoff in dollars for TANF initial eligibility for a family of three. These cutoffs were converted to percent of the 2016 Federal Poverty Level for a family of three. The final eligibility level is the population-weighted average of these.

The take-up rate estimate comes from Giannarelli (2019).

- WIC (The Special Supplemental Nutrition Program for Women, Infants, and Children)
  - Eligibility is capped at 185% of the FPL.
  - The take-up rate is an estimate from the USDA FNS (United States Department of Agriculture Food and Nutrition Service 2020b).

### B.6 Details

Using JSTOR and EBSCO, a research assistant collected all AER and QJE papers that met one of 33 search terms according to the search engine. The search terms were: “welfare program,” “social insurance,” “social program,” “social assistance,” “social welfare,” “social benefit,” “income threshold,” “participation threshold,” “means-testing threshold,” “means-tested program,” “means-tested welfare,” “means-tested benefit,” “means-tested subsidy,” “income means testing,” “eligibility rule,” “eligibility threshold,” “eligibility criteria,” “eligibility criterion,” “eligibility requirement,” “woodwork effect,” “program eligibility,” “program benefit,” “program subsidy,” “program duration,” “optimal program,” “optimal provision,” “benefit schedule,” “program schedule,” “public insurance,” “program take-up,” “incomplete take-up,” “welfare take-up,” “benefit take-up.”

We limit the sampling frame to the 2010–2018 AER and 2010–2019 QJE. Appendix B.6 provides the search terms. On the authors’ websites, we also provide a spreadsheet of all the papers, their inclusion criteria, and how we classified them. We also provide a list of judgment calls involved in this exercise and our rationale for our decision. We exclude the papers and proceedings but include comments. We exclude the 2019 AER because it was not available on JSTOR or EBSCO. We then read the abstract and/or introduction of each of the 278 papers that met at least one of the 33 search terms. We determine whether a paper was about a social welfare program.

We impose the following additional criteria when categorizing papers.

- We exclude papers that are principally about optimal income or capital taxation.
- We exclude transfers that are not intended to alleviate poverty (e.g., the effects of giving people computers).
- We exclude papers about credit market restrictions only, such as papers about mortgage deductions. We do include papers about consumer bankruptcy.
- We exclude papers about search and matching in labor markets if they do not have a substantial social insurance angle (e.g., UI).
- Because of the important theoretical connection between optimal social insurance and welfare design, we include papers that are about private insurance markets (including health insurance), as long as they have a significant angle about optimal policy.
- We define “program eligibility” as rules that determine whether a person has access to a social program. We do not consider eligibility to include access to different plan choices within a health program; our decision to exclude these papers is conservative, since they would only estimate a treatment effect using eligibility but not use optimal eligibility as an instrument.

A spreadsheet is available from the authors’ websites that lists all the papers and how they were categorized. In the spreadsheet, we also document a list of judgment calls involved in classifying the papers.

---

87 The research assistant also searched the downloaded PDFs to see which search terms were most often met. Two of the papers that the search engines specified met the search terms did not actually include the search terms in the downloaded PDF, perhaps due to a bug in the search engine. Neither paper was deemed to be about social welfare programs so this issue does not substantively affect the conclusions.
C Theory Extensions and Proofs

C.1 Endogenous labor supply

Section 3 develops a proposition that gives that $B$ and $m$ satisfy

\[ \frac{u(B)}{u'(B)B} = \frac{1 + \eta_m}{\lambda_m/\lambda_{avg} + \eta_m}, \]  

(C.1)

if $\eta_B = 0$ and $s = 0$.

We show how this expression can be microfounded in a more elaborate environment with endogenous labor supply. We focus on this parsimonious expression, nested by the more general case, for simplicity; this analysis captures many of the relevant insights.

**Model environment.** There is a continuum of types $\theta \sim F$, where $F$ has support $\Theta$. People earn labor income $y$ from hours worked $h$, depending on their type $\theta$. Let labor income $y = \theta h(\theta)$; we use this parametric form for simplicity, but the model can easily be generalized. People with labor income below $r$ (the “eligibility threshold”) earn a benefit $B$. People have utility $\tilde{v}(h, B, \theta)$ over labor supply and the benefit amount. This utility function induces an indirect utility function $v$ over labor supply, the benefit amount, and the eligibility threshold:

\[ v(h^*(B, r, \theta), B, r, \theta) = \max_h \left\{ \tilde{v}(h, B, \theta) \right\} \begin{aligned} \theta h(\theta) \leq r & \quad \theta h(\theta) > r. \end{aligned} \]  

(C.2)

The Envelope Theorem gives the following intermediate results, which we will invoke later:

\[ \frac{dv}{dB}(h^*(B, r, \theta), B, r, \theta) = \frac{\partial v}{\partial B} \]  

(C.3)

\[ \frac{dv}{dr}(h^*(B, r, \theta), B, r, \theta) = 0 \text{ if } h^*(B, r, \theta) \neq r/\theta \]  

(C.4)

Equation (C.4) states that if the benefit constraint does not bind, there is no value to the agent to relaxing the constraint. Intuitively, for people who are very poor or very rich, adjustments to the eligibility threshold have no effect on behavior. However, the existence of a lump-sum benefit and discrete eligibility threshold can induce bunching at the threshold. A small change in eligibility will have first-order effects on utility for bunchers.

**Take-up probabilities.** Agents are aware of the program with probability $p(r)$ and get $v(h^*(B, r, \theta), B, r)$ if they take up. Otherwise they optimize as if the program does not exist, do not take up the program, and get $v(h^*(0, r, \theta); 0, r, \theta)$ (the “outside option”). Moreover, this outside option does not depend on $r$: $v(h^*(0, r, \theta); 0, r, \theta) = v(h^*(0, 0, \theta); 0, \theta)$ for all $r$.

**Planner’s problem.** We begin with a technical assumption. Assume that income $\theta h^*(B, r, \theta)$ is weakly increasing in $\theta$: higher types always earn weakly more labor income even though the existence of the benefit distorts labor supply. This assumption amounts to a standard single-crossing condition: even if the tax system affects labor supply or causes bunching, it will not cause high types to earn strictly less income than low types (or vice-versa).

This assumption yields a threshold type $\tilde{\theta}(B, r)$ such that all $\theta \leq \tilde{\theta}$ will choose a labor supply that is low enough that they will be eligible for the benefit. All types $\theta > \tilde{\theta}$ are not eligible.

Next, we assume that the planner has a budget $T$ which depends on the amount of money raised through taxes on labor income. Assume the income tax schedule is exogenous, but make no other restrictions on this schedule. In that case, we can parameterize $T$ as depending on $B$ and $r$ alone:

\[ T(B, r). \]

---

88 We can think of utility over the benefit as the indirect utility of the agent’s inner problem of allocating the benefit to consumption of various goods.

89 Formally, let

\[ I(B, r) := \{ (\theta h^*(B, r, \theta), \theta h^*(0, r, \theta, \theta) : \theta \in \Theta \}. \]

Here, $I$ is the set of triples of: (i) labor incomes chosen if a given type $\theta$ receives the benefit, (ii) labor incomes chosen if the type $\theta$ does not receive the benefit, and (iii) the type $\theta$, which then yields a density $f(\theta)$ and a labor supply $h^*(B, r, \theta)$. These values uniquely determine the taxes raised for a generic tax schedule that only depends on labor income, even if there is incomplete take-up of the
Altogether, the planner’s problem is:

$$\max_{r,B} \int_0^{\bar{\theta}(B,r)} \lambda \alpha p(r) v(h^*(B,r,\theta); B,r,\theta) f(\theta) d\theta + \int_0^{\bar{\theta}(B,r)} \lambda \alpha (1-p(r)) v(h^*(0,r,\theta); 0,\theta) f(\theta) d\theta + \int_{\bar{\theta}(B,r)}^\infty \lambda \alpha v(h^*(0,r,\theta); 0,\theta) f(\theta) d\theta$$

subject to

$$\int_0^{\bar{\theta}(B,r)} p(r) B f(\theta) d\theta \leq T(B,r).$$

(C.5)

Noting that $$\int_0^\infty \lambda \alpha v(h^*(0); 0,\theta) f(\theta) d\theta$$ is a constant, we can re-write the planner’s problem as:

$$\max_{r,B} \int_0^{\bar{\theta}(B,r)} \lambda \alpha p(r) (v(h^*(B,r,\theta); B,r,\theta) - v(h^*(0,r,\theta); 0,\theta)) f(\theta) d\theta$$

subject to

$$\int_0^{\bar{\theta}(B,r)} p(r) B f(\theta) d\theta \leq T(B,r).$$

(C.7)

(C.8)

Then, let $$V(h^*(B,r,\theta); B,r,\theta) := v(h^*(B,r,\theta); B,r,\theta) - v(h^*(0,\theta); 0,\theta)$$ be the net utility gain from taking up the program. Note that for types $$\theta > \tilde{\theta}$$, $$V = 0$$; these types choose labor supply that renders them ineligible for the benefit. For other types, $$\theta \leq \tilde{\theta}$$, $$V > 0$$ assuming they earn positive utility from the benefit.

**Solving for the optimum.** Letting $$\sigma$$ represent the Lagrange multiplier, take the first-order condition with respect to $$r$$:

$$\frac{\partial \bar{\theta}}{\partial r} \left( \lambda \alpha p(r) V(h^*(B,r,\tilde{\theta}); B,r,\tilde{\theta}) f(\tilde{\theta}) \right) + \int_0^{\bar{\theta}} \lambda \alpha \left( \frac{dV}{dr} \left( V(h^*(B,r,\theta); B,r,\theta) + p \left( \frac{dV}{dr} \cdot 1(\theta = \tilde{\theta}) \right) \right) f(\theta) d\theta \right) -$$

$$\lambda \alpha p(r) B f(\tilde{\theta}) + \frac{d\sigma}{d\theta} \lambda \alpha B f(\tilde{\theta}) - \frac{dT}{dr} = 0. \tag{C.9}$$

Take the first-order condition with respect to $$B$$:

$$\frac{\partial \bar{\theta}}{\partial B} \left( \lambda \alpha p(V(h^*(B,r,\tilde{\theta}); B,r,\tilde{\theta}) f(\tilde{\theta}) \right) + \int_0^{\bar{\theta}} \lambda \alpha p(r) \frac{\partial V}{\partial B} f(\theta) d\theta$$

$$= \frac{d\sigma}{d\theta}, \text{ by Equation } \text{C.3}$$

$$\sigma \left( \frac{\partial \bar{\theta}}{\partial B} p(r) B f(\tilde{\theta}) + p(r) F(\tilde{\theta}) - \frac{dT}{d\theta} \right) = 0. \tag{C.10}$$

Solving for $$\sigma$$, we obtain:

$$\sigma = -\frac{\frac{\partial \bar{\theta}}{\partial B} \lambda \alpha p(r) V(h^*(B,r,\tilde{\theta}); B,r,\tilde{\theta}) f(\tilde{\theta}) + \int_0^{\bar{\theta}} \lambda \alpha p(r) \frac{\partial V}{\partial B} f(\theta) d\theta}{\frac{\partial \bar{\theta}}{\partial B} p(r) B f(\tilde{\theta}) + p(r) F(\tilde{\theta}) - \frac{dT}{d\theta}} \tag{C.11}$$

Plugging into Equation (C.9) yields:

benefit, assuming the planner knows $$p(r)$$. This notation shows that we can write $$T(B,r) = T(I(B,r))$$. Intuitively, holding $$F$$ fixed, any $$(B, r)$$ pair induces a distribution of labor incomes chosen across types.
represent the share who is eligible for the benefit: of labor income, we raise the eligibility threshold by one quantile of the population that is eligible. Let

\[
\frac{\partial \tilde{\theta}}{\partial r} \left( \lambda \tilde{\theta} p(r) V(h^*(B, r, \tilde{\theta}); B, r, \tilde{\theta}) f(\tilde{\theta}) \right) + \int_{0}^{\tilde{\theta}} \lambda \tilde{\theta} \left( \frac{dp}{dr} V(h^*(B, r, \theta); B, r, \theta) + p(r) \left( \frac{dV}{dr} \cdot 1(\theta = \tilde{\theta}) \right) \right) \, f(\theta) \, d\theta
\]

Value of \( r \uparrow \) to otherwise ineligible people

\[
- \left( \frac{\partial \tilde{\theta}}{\partial B} \lambda \tilde{\theta} p(r) V(h^*(B, r, \theta); B, r, \theta) f(\tilde{\theta}) \right) + \int_{0}^{\tilde{\theta}} \lambda \tilde{\theta} p(r) \frac{\partial V}{\partial B} f(\tilde{\theta}) \, d\theta
\]

Value of \( B \uparrow \) to bunchers

\[
\frac{\partial \tilde{\theta}}{\partial r} (pBf(\tilde{\theta})) + \frac{dp}{dr} BF(\tilde{\theta}) - \frac{dT}{dr}
\]

Mechanical cost of \( r \uparrow \)

\[
\frac{\partial \tilde{\theta}}{\partial B} p(r) Bf(\tilde{\theta}) + p(r) F(\tilde{\theta}) - \frac{dT}{dB}
\]

Mechanical cost of \( B \uparrow \)

Indirect cost of \( r \uparrow \) from changes in take-up

\[
\text{Indirect cost of } r \uparrow \text{ from changes in taxes via change in labor supply} = 0.
\]

Discussion. Equation (C.12), while involved, captures the following intuitions. At an optimum, the planner equates the following trade-offs.

- Raising \( r \) has benefits. First, it brings in more people to the program who were previously ineligible. Second, it has the value of raising take-up among inframarginal types. Third, it has a direct effect on welfare for people who bunch at the eligibility threshold, who can then adjust their labor supply (which was not necessarily at an optimum).

- Raising \( B \) has benefits. First, it brings value by affecting bunching. Second, it also has value to inframarginal types who take up the program, because it is a transfer.

- Raising \( r \) has costs. First, there is a mechanical cost of bringing more people into the program because more people are eligible. Second, there is an indirect cost of raising take-up. Third, there is an indirect cost of changing people’s labor supply, which then affects the income taxes collected.

- Raising \( B \) has costs. First, there is a mechanical cost of raising the transfer to people who take up the program. Second, there is an indirect cost of bringing more people into the program via changes in labor supply. Third, there is an indirect cost of changing people’s labor supply, which then affects the income taxes collected.

C.1.1 Simplifications

In this subsection, we show how this more general solution nests the solution in the paper.

First, we apply a change of units. Instead of considering raising the eligibility threshold by one dollar of labor income, we raise the eligibility threshold by one quantile of the population that is eligible. Let \( m \) represent the share who is eligible for the benefit: \( m := F(\tilde{\theta}) \).

Use the chain rule to observe that:

\[
\frac{\partial p}{\partial r} = \frac{\partial p}{\partial F(\tilde{\theta})} \frac{\partial F(\tilde{\theta})}{\partial \tilde{\theta}} \frac{\partial \tilde{\theta}}{\partial r} = \frac{\partial p}{\partial m} f(\theta) \frac{\partial \tilde{\theta}}{\partial r}.
\]

(C.13)

Next, we invoke the following assumption:
Assumption 1: No Bunching. Assume that \( h^*(\cdot) = \tilde{h}(\theta) \) for all \( B, r \), i.e. that the amount of labor supply chosen depends only on one’s type.

This assumption has three implications. First, \( \int_0^\delta 1(\theta = \tilde{\theta})f(\theta) = 0 \), assuming there are no atoms in the type distribution. Second, \( \frac{\partial \tilde{\theta}}{\partial B} = 0 \). Third, since labor supply is constant for all \( \theta \) and the budget \( T \) only depends on \( r \) and \( B \) via \( h, \frac{dT}{\partial R} = \frac{d}{\partial B} = 0 \).

As a result, employing the No Bunching assumption and dividing by \( f(\tilde{\theta}) \frac{\partial \tilde{\theta}}{\partial B} \) gives:

\[
\lambda_\theta p(r) V(h^*(B, r, \tilde{\theta}); B, r, \tilde{\theta}) + \int_0^\delta \lambda_\theta \frac{\partial p}{\partial m} V(h^*(B, r, \theta), B, r, \theta)f(\theta)d\theta = \left( \int_0^\delta \lambda_\theta \frac{\partial V}{\partial B} f(\theta)d\theta \right) \left( pB + \frac{\partial p}{\partial m} BF(\tilde{\theta}) \right) \frac{1}{F(\tilde{\theta})}.
\]

Finally, under the No Bunching assumption, observe that for any fixed \((B, r)\) pair, there exists \( \kappa(\theta; B, r) \) such that \( V(h(B, r, \theta); B, r, \theta) = \kappa(\theta; B)u(B) \) for all \( B \) in a neighborhood of the solution, for all types that are eligible for the benefit:

Assumption 2: Multiplicative Separability. Suppose \( V(B, r, \theta) = \kappa(\theta)u(B) \) for all \( B \) in a neighborhood of \( B^* \) and for all \( \theta \leq \tilde{\theta} \).

The Multiplicative Separability assumption states that utility gains from the benefit can be multiplicatively rescaled by the schedule \( \kappa(\theta) \). Note that this assumption always holds if utility is homogeneous across types and all types have the same outside option; in that case, \( \kappa(\theta) = 1 \) for all \( \theta \). In the body of the paper, we start directly from that more demanding homogeneity assumption.

For other utility functions, the assumption holds as long as slight changes to the benefit around the optimum do not change the relative differences in the net utility that the different types experience from receiving the benefit. These relative differences are parameterized by the \( \kappa(\theta) \) schedule, which must be invariant around the optimum. This assumption fails if, e.g., high types’ marginal utility from receiving \( B \) diminishes at a faster rate than low types’ marginal utility even in a neighborhood around the optimum.

The Multiplicative Separability assumption permits us to rescale differences in net utility with a (re-written) \( \lambda_\theta \) welfare weight schedule.

Define \( \bar{\lambda}_\theta := \lambda_\theta \kappa(\theta) \). Moreover, let

\[
\bar{\lambda}_{\text{avg}} = \frac{\int_0^\delta \lambda_\theta \kappa(\theta) f(\theta)d\theta}{F(\tilde{\theta})}.
\]

Intuitively, these \( \bar{\lambda}_\theta \) weights capture both: (i) the differences in the planner’s value for one unit given to each type (parameterized via the \( \lambda_\theta \) weights), and (ii) the differences in utility each type experiences when given \( B \) in benefits (parameterized via the \( \kappa(\theta) \) schedule).

Then, working from Equation \((C.14)\), applying the Multiplicative Separability assumption, dividing by \( p \) and using that \( F(\tilde{\theta}) = m \), we obtain:

\[
\frac{u(B)}{u'(B)B} = \frac{1 + \eta_m}{\bar{\lambda}_m + \eta_m} \tag{C.15}
\]

for \( \eta_m := \frac{\partial p}{\partial m} \frac{m}{p} \), which is the equation we target.

C.2 Additional Discussion of Equation (8)

We begin the additional discussion by stating the following Proposition, proven in Appendix C.5.

One way to gain additional insight into the sufficient statistics derived in Section 3 is to rearrange terms so that they equate the willingness to pay for a higher means test with the fiscal externality associated with it, à la Baily (1978) and Chetty (2006). To simplify the analysis here, we focus on the case where \( p^i = p^s \) and \( \eta_m^i = \eta_m^s \). Then, from Equation (3), we have:

\[
\frac{u(B)}{u'(B)} = \frac{1 + \eta_m}{(1 + s \eta_B) \left( \eta_m + \frac{\lambda_m}{\lambda_{avg}} (1 - s \gamma) - m \frac{\partial q}{\partial m} s \right)}.
\]

(C.16)

Rearranging terms gives:

\[
\frac{\lambda_m}{\lambda_{avg}} (1 - s \gamma) u(B) + \left( \eta_m (1 - s \gamma) - m \frac{\partial q}{\partial m} s \right) u(B) = B \left( 1 + \eta_m \right) \frac{u(B)}{u'(B)} (1 + s \eta_B).
\]

(C.17)

Equation C.17 shows the willingness to pay for a higher means test on the left-hand size and the fiscal externality on the right-hand side. However, it is instructive to separate the willingness-to-pay of the inframarginal population by whether they are stigma agents or information agents:

\[
\frac{\lambda_m}{\lambda_{avg}} (1 - s \gamma) u(B) + (1 - s) \eta_m u(B) + s \left( \eta_m (1 - \gamma) - m \frac{\partial q}{\partial m} s \right) u(B) = B \left( 1 + \eta_m \right) \frac{u(B)}{1 + s \eta_B}.
\]

(C.18)

We invoke the following lemma:
Lemma 2. \( \eta_m(1 - \gamma) - m \frac{\partial \gamma}{\partial m} = \frac{m}{\rho(B,m)} \int_0^{u(B)} \frac{\partial H(c|m)}{\partial m} dc \).

The proof is below, in Appendix C.5. Now we can rewrite (and label) the expression as:

\[
\frac{\lambda_m}{\lambda_{avg}} (1 - s) u(B) + \frac{(1 - s) \eta_m u(B)}{\lambda_{avg}} + \frac{sm}{p(B,m)} \int_0^{u(B)} \frac{\partial H(c|m)}{\partial m} dc = B(1 + \eta_m) \frac{1}{(1 + s \eta_B)}. \tag{C.19}
\]

Discussion. The left-hand side gives the willingness to pay for an eligibility expansion. All newly eligible types receive the utility gains from the program, which is \((1 - \gamma)u(B)\) for those who were previously optimizing (stigma types) and \(u(B)\) for those who were not (information types).

However, the inframarginal population gain differently from an eligibility expansion depending on their type. Stigma types gain the amount that the higher means test shifts the distribution of stigma costs, decreasing the costs they pay when they take up. These are gains accrue to stigma types who would always take up the program; those who newly enroll because of the lower costs do not gain, by the Envelope Theorem. Meanwhile, the inframarginal information types only gain if they newly enroll: those driving the woodwork effects capture the full utility gains from the enrollment. The information types who would always enroll regardless of the means test do not gain from the eligibility expansion.

The right-hand side gives the fiscal cost of an eligibility expansion. The government must pay benefits to the newly enrolled, as well as to those in the inframarginal population who are prompted to enroll by the expanded eligibility. Meanwhile, it saves some amount of money by the loss of enrollment from optimizing agents who find the new, lower benefits level not worthwhile to enroll.

C.4.1 Necessary Condition for Proposition 3

C.4 Discussion of Assumption 3 and Conditions (i) and (ii)

Assumption 3 states that the change in the eligibility threshold reduces the average stigma costs among the fraction of people who take up the program. Assumption 3 is difficult to validate empirically without granular information on the treatment effect of changing the eligibility threshold on people’s perceived stigma cost at every part of the stigma cost distribution.

First we show that this assumption is sufficient but not necessary. Equation (C.45) from the proof of Proposition 3 gives that the necessary and sufficient condition is:

\[
\left( \frac{1 - s}{p^s} \left( \eta_m + 1 \right) \right) \left( \frac{(1 - s) \eta_m u(B)}{\lambda_{avg}} + \frac{(1 - s) \eta_m}{\lambda_{avg}} \right) < \left( \frac{(1 - s)}{p^s} \right) \left( \frac{(1 - s) \eta_m u(B)}{\lambda_{avg}} \right) \left( \frac{(1 - s) \eta_m}{\lambda_{avg}} \right).
\tag{C.20}
\]

As long as Equation (C.20) holds, it is true that for all \( \Xi, m^\Xi > m^\Xi \). Put another way, Equation (C.20) is a necessary condition that encodes the combination of Assumption 1 and either condition (i) or condition (ii) in Proposition 3. Thus, Equation (C.20) is weaker than Assumption 3 and condition (i) or condition (ii).

Equation (C.20) encodes the observation that as \( s \to 0 \), Proposition 3 always holds, because the necessary condition then reduces to:

\[
\left( \eta_m + 1 \right) \left( \frac{\lambda_m}{\lambda_{avg}} \right) < \left( \eta_m + \frac{\lambda_m}{\lambda_{avg}} \right),
\tag{C.21}
\]

which always holds since \( \frac{\lambda_m}{\lambda_{avg}} \leq 1 \). Intuitively, because information-only types capture the full benefit of the program, the fully naive planner undervalues the social value of raising \( m \) more with more information-only types. As a result, she can tolerate a larger violation of \( \frac{\partial \gamma}{\partial m} > 0 \).

We also note that for various configurations of \( \lambda_m / \lambda_{avg}, \gamma, \) and \( \frac{\partial \gamma}{\partial m}, \) as well as the other parameters, the necessary condition may hold. For instance, as \( \gamma \to 0 \), the necessary condition always holds. Intuitively, as
stigma agents become more like information agents, we no longer need a separate condition governing the behavior of \( \eta^l_m \) and \( \eta^r_m \).

C.4.2 Discussion of Assumption 1

How could Assumption 1 fail? Suppose there are no information-only types \( \sigma = 1 \). Suppose moving the eligibility threshold reduces costs for people who are just indifferent to taking up the program (i.e., for whom \( c \approx u(B) \)). Suppose it has no effect on people for whom \( c < u(B) \). Then, changing the eligibility threshold will first-order stochastically reduce the cost distribution. However, \( \frac{\partial g}{\partial m} \) will perhaps counterintuitively rise. Intuitively, the average cost conditional on taking up the program will feature a larger density at \( c \approx u(B) \).

For a concrete example of the assumption failing, suppose \( c \sim H_{\text{pre}} = N(1, \sigma) \) for known \( \sigma \). Suppose \( u(B) = 2 \). Suppose raising \( m \) changes moves all costs larger than 2 to be at 2:

\[
H_{\text{post}} = \begin{cases} 
N(1, \sigma) & c \leq 2 \\
2 & \text{otherwise} 
\end{cases}
\]  

(C.22)

where we denote this truncated distribution by \( H_{\text{post}} \). Changing the eligibility threshold induces a first-order stochastic reduction in the cost distribution. It raises the share of people in the population who take up the program. However, it also raises the average cost conditional on taking up the program. The average cost before raising \( m \) is \( E[c|c \leq 2, H_{\text{pre}}] \approx 0.71 \), whereas the average cost after raising \( m \) is \( E[c|c \leq 2, H_{\text{post}}] \approx 0.91 \).

However, violations of Assumption 1 are unlikely in practice. To see why, note that the counter-example above requires a large change in the cost distribution only for draws of the cost distribution that are about as large as \( u(B) \). If raising \( m \) also affects the draws of the cost distribution for \( c < u(B) \), that serves as a force pushing \( \frac{\partial g}{\partial m} \) downward.

Second, Equation (C.20) shows that the necessary and sufficient condition for Proposition 3 to fail is much weaker than \( \frac{\partial g}{\partial m} \leq 0 \).

Third, the specific counter-example changed the shape of the cost distribution. \( H_{\text{pre}} \) is normal; \( H_{\text{post}} \) is a truncated normal. We develop propositions showing that for the normal and exponential cost distributions, any change in the (unconditional) mean costs that maintains the distributional family from which the costs are drawn will feature \( \frac{\partial g}{\partial m} < 0 \).

Proposition 4. Let \( c \sim N(\mu(m), \sigma) \) with \( \mu'(m) < 0 \). Then \( \frac{\partial g}{\partial m} < 0 \).

We prove Proposition 4 in Appendix C.5. A change in \( m \) reduces the mean (unconditional) cost but the cost distribution remains normal. Then the change in ratio of costs to benefits, conditional on taking up the program, will shrink in \( m \); i.e., Assumption 1 holds. We develop a similar proposition if costs are exponentially distributed:

Proposition 5. Let \( c \sim \text{Exp}(\theta(m)) \), where \( 1/\theta \) is the mean of the exponential distribution \( \text{Exp} \) and \( \theta'(m) > 0 \). Then \( \frac{\partial g}{\partial m} < 0 \).

Note that \( \theta'(m) > 0 \) implies the average unconditional cost \( 1/\theta \) falls in \( m \), so Proposition 5 is qualitatively similar to Proposition 4.

\(^{90}\text{Note that we suppose all types receive draws from the same cost distribution. Thus, the violation of Assumption 1 is not that changing } m \text{ only affects costs for } \theta = m \text{ at the marginal of eligibility. Rather, Assumption 1 is only likely to fail if changing } m \text{ affects people for whom } c \approx u(B), \text{ i.e. they are indifferent to signing up (regardless of their income).}\)
C.5 Proofs

C.5.1 Proofs of Propositions 1 and 2

Proof. Note that Proposition 1 is a special case of Proposition 2. We therefore prove Proposition 2 only. The planner’s problem is:

$$\max_{B, m} \left( sp^s(B, m) \left( \int_0^m \lambda_d u(B) d\theta - \int_0^m \int_{c \leq u(B)} c \lambda_d h(c|m) dcd\theta \right) \right)$$

subject to

$$(1 - s) p^i(m) \int_0^m B d\theta + sp^s(B, m) \int_0^m B d\theta \leq T$$

$$m \in [0, 1]$$

We inspect interior solutions using the Karush-Kuhn-Tucker conditions where the constraint $m \in [0, 1]$ is slack. We consider cases in which such an interior solution exists; there are possible corner solutions where $m = 1$ (i.e., the program is universal). Proposition 1 and 2 give necessary conditions for local optimality. To obtain that the statement in the proposition is sufficient for a global maximum, it is sufficient to additionally impose that the maximand is concave and the constraint is convex.

The first-order condition for $B$ is:

$$s \left( \frac{\partial p^s}{\partial B} \lambda_{avg} u(B) + p^s(B, m) u'(B) \lambda_{avg} m - \left( \frac{\partial}{\partial B} \int_0^m \int_{c \leq u(B)} c \lambda_d h(c|m) dcd\theta \right) \right) + (1 - s) p^i(m) u'(B) \lambda_{avg} m = \sigma \left( (1 - s) p^i(m) m + s \frac{\partial p^s}{\partial B} B m + sp^s m \right), \quad (C.26)$$

where $\sigma$ denotes the Lagrange multiplier, and we note that

$$\int_0^m \int_{c \leq u(B)} \lambda_d h(c < u(B), m) dcd\theta = \frac{1}{H(u(B)|m)} \int_0^m \int_{c \leq u(B)} \lambda_d h(c|m) dcd\theta. \quad (C.27)$$

Leibniz’s rule gives that:

$$\frac{\partial}{\partial B} \int_0^m \int_{c \leq u(B)} c \lambda_d h(c|m) dcd\theta = u(B) \lambda_{avg} m h(u(B)|m) u'(B)$$

$$= \lambda_{avg} m u'(B) \frac{\partial p^s}{\partial B}. \quad (C.28)$$

We collect terms to obtain:

$$\lambda_{avg} m (sp^s + (1 - s) p^i) u'(B) = \sigma \left( s \frac{\partial p^s}{\partial B} B m + sp^s m + (1 - s) p^i m \right). \quad (C.30)$$

We divide by $p^s m$ and rearrange, recalling that $\eta_B = \frac{\partial p^s}{\partial B} p^i$:

$$\frac{sp^s + (1 - s) p^i}{p^s} u'(B) \lambda_{avg} = \sigma \left( s \eta_B + sp^s + (1 - s) p^i \right). \quad (C.31)$$

Next we take the first-order condition with respect to $m$ and use the shorthand $E := E[c|c < u(B)]$ to be
succinct:
\[ s \frac{\partial p^s}{\partial m} \left( \lambda_{\text{avg}}mu(B) - \lambda_{\text{avg}}mE \right) + sp^s \left( \lambda_mu(B) - \lambda_mE - \lambda_{\text{avg}}m \frac{\partial E}{\partial m} \right) \]
\[ + (1-s) \left( \frac{\partial p^i}{\partial m} \lambda_{\text{avg}}mu(B) + p^i\lambda_mu(B) \right) = \sigma \left( 1-s \left( \frac{\partial p^i}{\partial m} Bm + p^iB \right) \right) + s \left( \frac{\partial p^s}{\partial m} Bm + p^sB \right). \tag{C.32} \]

Noting that \( \frac{\partial E}{\partial m} = \frac{\partial p^s}{\partial m} u(B), \) we collect terms to obtain:
\[ u(B)\lambda_{\text{avg}} \left( (1-\gamma)s \frac{\partial p^s}{\partial m} m + (1-s) \frac{\partial p^i}{\partial m} m - sp^s \frac{\partial \gamma}{\partial m} m \right) + \lambda_mu(B)(1-\gamma)sp^s \]
\[ + \lambda_mu(B)p^i(1-s) = \sigma \left( 1-s \left( \frac{\partial p^i}{\partial m} Bm + p^iB \right) \right) + s \left( \frac{\partial p^s}{\partial m} Bm + p^sB \right). \tag{C.33} \]

We divide by \( p^s \) and \( p^i \) to get:
\[ u(B)\lambda_{\text{avg}} \left( \frac{(1-\gamma)s\eta_m^s + (1-s)\eta_i^i}{p^i} - \frac{s \frac{\partial \gamma}{\partial m} m}{p^i} \right) \]
\[ + s \lambda_mu(B)(1-\gamma) + \lambda_mu(B)(1-s) = B\sigma \left( \frac{(1-s)\eta_i^i + \eta_m^s + 1}{p^i} + s \frac{\partial \gamma}{\partial m} \right). \tag{C.34} \]

Then, substituting for \( \sigma \) from Equation (C.31) and rearranging gives:
\[ \frac{u(B)}{Bu''(B)} = \frac{(1-s)\eta_i^i + \eta_m^s + 1}{(sp^s + (1-s)p^s)\eta_B + 1} \left( \frac{(1-\gamma)s\eta_m^s + 1}{p^i} + \frac{s \frac{\partial \gamma}{\partial m} m}{p^i} \right). \tag{C.35} \]

An important special case is where \( p^i = p^s \) and \( \frac{\partial p^i}{\partial m} = \frac{\partial p^s}{\partial m}. \) Then, multiplying both the numerator and denominator by \( p := p^i = p^s, \) and noting that \( \eta_m^i = \eta_m^s, \) we get:
\[ \frac{u(B)}{Bu''(B)} = \frac{\eta_m^s + 1}{(s\eta_B + 1) \left( \eta_m^s + \lambda_m \lambda_{\text{avg}} (1-s) - s \frac{\partial \gamma}{\partial m} \right)}. \tag{C.36} \]

\[ \Box \]

C.5.2 Proof of Taylor Expansion (Equation (9)).

Proof. Throughout the paper, we use the second-order Taylor approximation:
\[ u(0) = 0 \approx u(B) - u'(B)B + \frac{u''(B)B^2}{2}, \tag{C.37} \]
which gives
\[ u(B) \approx u'(B)B - \frac{u''(B)B^2}{2}. \tag{C.38} \]
We substitute this approximation into the LHS of Equation (9) to obtain
\[ \frac{u(B)}{Bu'(B)} \approx 1 + \frac{B}{2}. \tag{C.39} \]

Note that \( u'(B) = \frac{\partial}{\partial B} (u(B) - c), \) so \( \rho \) represents the coefficient of relative risk aversion for people who would take up the program if informed. \[ \Box \]
C.5.3 Proof of Lemma

Proof. The quotient rule gives
\[ \frac{\partial}{\partial B} \left( \frac{u(B)}{u'(B)} \right) > 0 \] (C.40)
iff
\[ (u'(B))^2B - u(B)u'(B) - Bu''(B)u(B) > 0. \] (C.41)
Dividing by \( u'(B) \) (which is always greater than 0), we conclude that the left-hand side is always positive as long as
\[ u'(B)B + u(B)(\rho - 1) > 0, \] (C.42)
which completes the proof.

Proof. The quotient rule gives
\[ \frac{\partial}{\partial B} \left( \frac{u(B)}{u'(B)} \right) > 0 \] (C.40)
iff
\[ (u'(B))^2B - u(B)u'(B) - Bu''(B)u(B) > 0. \] (C.41)
Dividing by \( u'(B) \) (which is always greater than 0), we conclude that the left-hand side is always positive as long as
\[ u'(B)B + u(B)(\rho - 1) > 0, \] (C.42)
which completes the proof.

C.5.4 Proof of Proposition

To make it clear what we seek to prove, we restate the second condition in Proposition 3:

\[ \text{Proposition.} \quad \text{If} \quad \eta_i^m > 0 \quad \text{or} \quad \eta_i^s > 0, \quad \text{then} \quad m^w > m^n \quad \text{for all} \quad \Xi \quad \text{as long as} \quad \eta_i^m \leq \eta_i^m \quad \text{(condition (i))}. \quad \text{Moreover, there exists} \quad \varepsilon > 0 \quad \text{such that} \quad m^w > m^n \quad \text{for all} \quad \Xi \quad \text{as long as} \quad \eta_i^m \in [\eta_i^m, \eta_i^m + \varepsilon] \quad \text{(condition (ii)).} \]

Proof. We want to show that the naïve planner would raise the eligibility threshold \( m \) and lower the benefit size \( B \). First, we note that for a given \( (B, m) \) pair, the budget constraint ensures that raising \( B \) requires lowering \( m \), and raising \( m \) requires lowering \( B \). Thus, it is sufficient to argue that the naïve planner sets \( B \) too high. Using Lemma 1, we note that the LHS of Equation (C.35) is increasing in \( B \). Thus, noting that the naïve planner solves Equation (C.35) for \( B^n \) and the sophisticated planner solves the equation for \( B^s \), we want to show:
\[ \frac{u(B^s)}{B^s u'(B^s)} - \frac{u(B^n)}{B^n u'(B^n)} < 0, \] (C.43)
and substituting Equation (C.35), we have that we want to show:
\[ \left( \frac{1-s}{p^s} \right) (\eta_i^m + 1) + \frac{s}{p^i} (\eta_i^m + 1) - \left( \frac{1-s}{p^i} \right) (0 + 1) + \frac{s}{p^i} (0 + 1) \] (C.44)
Observe that
\[ \left( \frac{p^s}{sp^s + (1-s)p^i} sB_i + 1 \right) - \left( 1-s \right) p^s \eta_i^m + 1 \] (C.46)
Cross-multiplying, it is then sufficient to show:
\[ \left( \frac{1-s}{p^s} \right) (\eta_i^m + 1) + \frac{s}{p^i} (\eta_i^m + 1) \left( \frac{\lambda_m}{\lambda_{avg}} \left( \frac{1-s}{p^i} (\eta_i^m + 1) \right) \right) - \left( \frac{1-s}{p^i} \right) (0 + 1) + \frac{s}{p^i} (0 + 1) \] (C.45)
Under Assumption 1, \( -\frac{sm \partial \gamma_i}{p^i \partial m} \geq 0 \), so it is sufficient to show that:
\[ \left( \frac{1-s}{p^s} \right) (\eta_i^m + 1) + \frac{s}{p^i} (\eta_i^m + 1) \left( \frac{\lambda_m}{\lambda_{avg}} \left( \frac{1-s}{p^i} (\eta_i^m + 1) \right) \right) - \left( \frac{1-s}{p^i} \right) (0 + 1) + \frac{s}{p^i} (0 + 1) \] (C.46)
Rearranging gives that this condition is equivalent to:

\[
\left(1 - s \frac{\eta^i_m}{\nu} + s \frac{\eta^i_m}{\nu} \right)\left(\frac{\lambda_m}{\lambda_{\text{avg}}} \left(\frac{1 - s\gamma}{\nu} + 1 - s \frac{\nu}{\nu}\right)\right) < \left(1 - s \frac{\eta^i_m}{\nu} + s \frac{\eta^i_m}{\nu} \right)\left(\frac{1 - s\gamma}{\nu} \eta^i_{\text{avg}} + 1 - s \eta^i_m\right)
\]

\[
\iff \left(1 - s \frac{\eta^i_m}{\nu} + s \frac{\eta^i_m}{\nu} \right)\left(\frac{\lambda_m}{\lambda_{\text{avg}}} \left(\frac{1 - s\gamma}{\nu} + 1 - s \frac{\nu}{\nu}\right)\right) < \left(\frac{1 - s\gamma}{\nu} \eta^i_{\text{avg}} + 1 - s \eta^i_m\right).
\]

The statement holds strictly if \(\eta^i_m = \eta^i_{\text{avg}}\) as long as \(\lambda_m < \lambda_{\text{avg}}\), by factoring the LHS and canceling. Moreover, holding \(\eta^i_m\) fixed, the LHS is strictly decreasing in \(\eta^i_m\). Thus, if the statement holds for \(\eta^i_m = \eta^i_{\text{avg}}\), it also holds for \(\eta^i_m < \eta^i_{\text{avg}}\). This shows that the desired statement holds under condition (i). To argue that the desired statement holds under condition (ii), notice that the LHS is strictly increasing in \(\eta^i_m\), holding \(\eta^i_{\text{avg}}\) fixed. As a result, there exists \(\eta^i_m > \eta^i_{\text{avg}}\) such that the statement holds with equality. Since the LHS is increasing in \(\eta^i_m\), the statement holds strictly for \(\eta^i_m < \eta^i_{\text{avg}}\). Thus, there exists an interval \(\eta^i_m \in [\eta^i_{\text{avg}}, \eta^i_{\text{avg}} + \epsilon]\) for \(\epsilon > 0\) such that Equation (C.48) holds, which completes the proof.

C.5.5 Proof of Lemma 2

Proof. Multiplying both sides by \(u(B)\) and recalling that

\[
\frac{\partial E_{\text{mc}}[c < u(B), m]}{\partial m} = \frac{\partial}{\partial m} E_{\{c < u(B), m\}},
\]

we have

\[
\frac{\partial}{\partial m} E_{\{c < u(B), m\}} = \frac{\partial}{\partial m} \int_0^{u(B)} \partial H(c|m) dc.
\]

Below, we show that

\[
\frac{\partial}{\partial m} E_{\{c < u(B), m\}} = \frac{\partial}{\partial m} \int_0^{u(B)} \partial H(c|m) dc,
\]

which completes the proof.

We apply integration by parts to the first integral. We apply that

\[
\frac{\partial}{\partial m} \frac{\partial}{\partial m} dc = E_{\{c < u(B), m\}}
\]

to the second integral. Suppressing arguments of \(h\) and \(H\) to be concise, this yields:

\[
m \left(\frac{1}{H} \int_0^{u(B)} \frac{\partial h(c|m)}{\partial m} dc - \frac{1}{H^2} \int_0^{u(B)} \frac{\partial h(c|m)}{\partial m} dc \right)
\]

\[
= m \left(\frac{1}{H} \left(\int_0^{u(B)} \frac{\partial H(c|m)}{\partial m} dc - \int_0^{u(B)} \frac{\partial H(c|m)}{\partial m} dc \right) - \frac{\partial}{\partial m} \frac{\partial}{\partial m} E_{\{c < u(B), m\}} \right)
\]

\[
= \eta_m (u(B) - E_{\{c < u(B), m\}}) - \frac{m}{p(B, m)} \int_0^{u(B)} \frac{\partial H(c|m)}{\partial m} dc,
\]

recalling that \(\frac{m}{p(B, m)} = \eta_m\).
Proof. It suffices to prove that \( \frac{\partial}{\partial m} (E[c|c \leq u(B), \mu(m)]) < 0 \). First, let \( \chi(Z) := \phi(Z)/\Phi(Z) \) for normal PDF \( \phi \) and normal CDF \( \Phi \). Equation (3) in Sampford (1953) gives that
\[
0 < \frac{\partial}{\partial Z} \left( \frac{\phi(Z)}{1 - \Phi(Z)} \right) < 1 \tag{C.56}
\]
for all \( Z \). Thus
\[
-1 < \frac{\partial \chi}{\partial Z} < 0 \tag{C.57}
\]
since the normal PDF is even and \( 1 - \Phi(Z) = \Phi(-Z) \).
The usual properties of the normal distribution give:
\[
E[c|c \leq u(B), \mu(m)] = \mu(m) - \sigma \frac{\Phi(Z(m))}{\Phi(Z(m))} \tag{C.58}
\]
for \( Z(m) := (u(B) - \mu(m))/\sigma \).
The chain rule gives
\[
\frac{\partial}{\partial m} \left( \frac{\phi(Z(m))}{\Phi(Z(m))} \right) = -\frac{\partial \chi}{\partial Z} \frac{\mu'(m)}{\sigma}. \tag{C.59}
\]
Then evaluating Equation (C.58) at the bounds in Equation (C.57) gives
\[
\mu'(m) < \frac{\partial}{\partial m} (E[c|c < u(B), \mu(m)]) < 0. \tag{C.60}
\]
\[ \square \]

Proof. It suffices to prove that \( \frac{\partial}{\partial m} (E[c|c < u(B), \mu(m)]) < 0 \). The mean of the truncated exponential distribution is:
\[
\mu(\theta(m)) = \frac{1}{\theta} - u(B) (\exp(\theta u(B)) - 1)^{-1} \tag{C.61}
\]
for \( u(B) > 0 \). This function is monotonically decreasing for all \( u(B) \) (Al-Athari 2008). \[ \square \]