Online Appendix to

"Take-up and Targeting: Experimental Evidence from SNAP"

Amy Finkelstein and Matthew J. Notowidigdo

February 2019

A: Interventions

Description of "assistance" component

An individual who responds to an outreach letter by calling into BDT is connected to a BDT employee –a "Benefits Outreach Specialist" (BOS) - who provides assistance over the phone. BOS's are highly knowledgeable of available benefits. They receive 4 weeks of classroom and experiential learning to become well-versed in the public benefits application process and policies. The up-front training includes coaching and training on phone-based assistance skills so that the caller receives a person-centered and results-driven experience. After this initial training, the BOS continues to receive continuous monitoring and coaching.

The BOS has real-time access to a searchable history of information on the caller from previous interactions with BDT and administrative data sources; in PA these administrative data sources include identified information BDT regularly receives on individuals enrolled in Medicaid, LIHEAP and PACE, and individuals who have exhausted unemployment compensation benefits. BDT has built an internal software platform that stores all this data in a household "portfolio" and allows for the collection of additional self-reported information for each individual linked to the portfolio. The software provides a clickable interface through which BOS can access notes on previous calls, question prompts to determine likely eligibility, an estimated benefits calculator, and a platform for scheduling follow-up actions. BDT customizes question prompts and the benefit calculator to each state's benefit regulations, to ensure that all of the necessary information is collected to estimate eligibility and benefits amounts. This software also allows for direct submission of the application and related verification documents.

Upon being connected to a caller, the BOS asks a series of intake questions designed to collect information relevant for eligibility and benefit screening. Information collected include demographic characteristics (e.g., number of people in the household, current enrollment in other public benefit programs, sex, ethnicity, disability, etc.), legal information (citizenship, marital status, etc.), self-reported monthly income (including pension), other financial resources when necessary (e.g.,

checking and savings account balances), and expenses by category (rent, utility bills, medical expenses, etc.). Collection of detailed information on expenses may increase the amount of benefits the individual is likely eligible for by increasing their allowable deductions. BDT's custom screening tool allows the BOS to use this self-reported information to inform the caller of whether they are likely eligible for SNAP and their estimated benefit amount.

If the caller decides based on this that they are interested in potentially applying, the BOS then provides information and assistance with the application process. The full set of assistance (which about half of applicants in this intervention arm avail themselves of), includes several stages. BDT completes the application for the caller based on the information received over the telephone and, in that same phone call, informs the applicant of required verification documents. Leveraging state policy options and technology, BDT also minimizes paper verification requirements by proactively informing individuals that they can self-declare shelter expenses (unless questionable) and that DHS can electronically verify Social Security income, identity, residency, and certain medical expenses. BDT then mails an envelope to recipients to collect verification documents, reviews the verification documents it receives, and re-contacts the individual if documentation is inadequate. BDT can then submit the application on behalf of the individual. The individual themselves however must participate in a phone interview with DHS.

BDT may also provide assistance after the application is submitted by reviewing and submitting any follow-up verification documentation requested by DHS, or working with DHS to troubleshoot issues with individual cases. The BDT custom software stores digital records of all received documents in an individual's record, including those submitted to DHS, which allows BDT to keep a detailed history of all application information and to advise applicants on how to advocate for themselves if there are issues with their application. For example, DHS may request a document that has already been provided or that is not necessary. In addition, some applicants miss their interview, or fail to receive an interview call, but still wish to apply. These incidences delay the application process, or even worse, can result in DHS rejecting an application. If contacted by a client about such an issue, BDT advises on how to navigate DHS customer services, and as a last resort, may elevate these issues to their point of contact at DHS to find a solution.

Comparison of standard outreach materials: "Information Plus Assistance" and "Information Only" interventions

The outreach materials in the two main treatment arms were designed to be as similar as possible. The outreach materials in the baseline Information Plus Assistance treatment were the standard materials BDT uses. Appendix Figure A1 shows the letter, envelope and postcard that were sent in this treatment arm. Appendix Figure A2 shows the analogous letter, envelope and postcard. They are designed to be as similar as possible - including the sender (the Secretary of the Pennsylvania Department of Human Services), the layout, and the content. There were, however, some unavoidable differences in the letters which we detail here.

First, the Information Plus Assistance letters reference the PA Benefits Center (the local name of

BDT), while the Information Only letter, naturally, does not. Specifically in the former the outreach materials say "We are working closely with the PA Benefits Center to help you get SNAP" and "Please call the PA Benefits Center today" while the information-only outreach materials say ""We want to help you get SNAP" and "Please call the Department of Human Services today." Second, and relatedly, the PA benefits center logo was included - in addition to the PA Department of Human Services logo - on the outreach materials in the Information Plus Assistance interventions, while only the PA Department of Human Services logo was included in the Information Only materials. Third, the hours of operation provided for the call in numbers were slightly different, reflecting the practical reality that BDT hours are 9:00am to 5:00pm while the DHS HELPLINE hours are 8:45am - 4:45pm. Finally, the phone numbers to call naturally differed (although all phone numbers were "1-800" numbers) and the PO box for the return address on the envelope also differed.

Sub-treatments

Appendix Figure A3 shows the study design with all of the treatments and sub-treatments.[33]

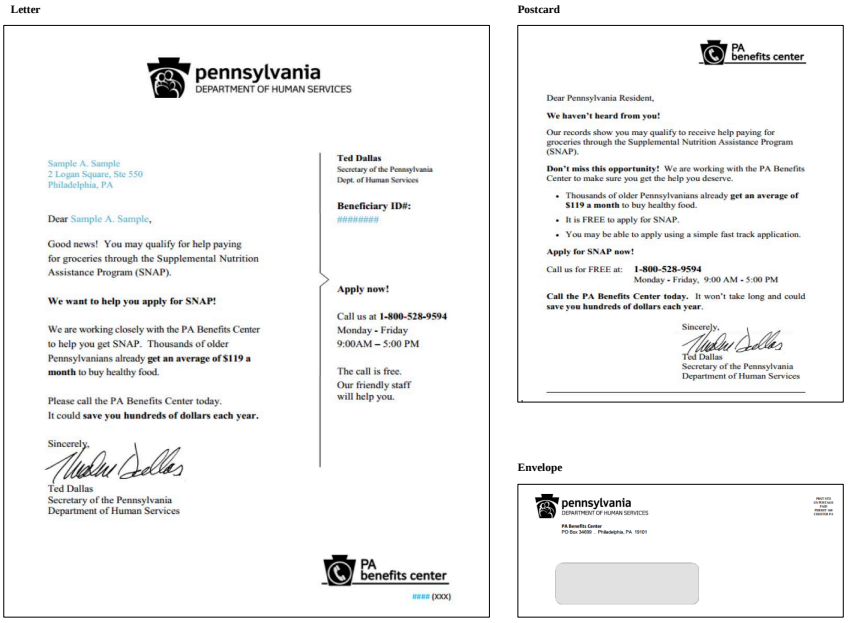
One sub-treatment was a "marketing" intervention. One-quarter of each treatment was randomized into an arm with a variant of the outreach letters and postcards designed to attract clients by using a "marketing" approach that borrowed language and graphics from credit card solicitations in an attempt to grab potential applicants' attention and potentially reduce stigma surrounding applying for SNAP. To grab attention, it included a catchy banner that read "Need help buying groceries? Apply today!", bolded text to highlight key information followed by an informative explanation, were printed in color rather than black and white, and included a PA benefits "ACCESS" card image. To try to reduce stigma, it included language such as "Join thousands of older Pennsylvanians already claiming their SNAP benefits" and did not explicitly define SNAP as food stamps. This design was motivated in part by the finding in Schanzenbach (2009) that individuals randomized to outreach materials describing the Food Stamp program in arguably more positive terms (including emphasizing a "Gold State Advantage" card) expressed somewhat higher rates of interest in receiving information about food stamp benefits than those whose outreach materials reflected standard USDA materials.

In the Information Plus Assistance treatment the remaining three-quarters received the standard outreach ("standard"). In the Information Only treatment, one-quarter received the standard outreach, while another one-quarter received the standard letter but no follow-up postcard ("no postcard") and another one-quarter received a letter that varied the description of the expected benefit amounts ("framing") to try to make them appear larger by focusing on the maximum benefit amount the individual could receive, rather than the average benefit amount received.

Our main analysis compares across three groups: the (pooled) standard (with postcard) and marketing treatments in the Information Only arm, the (pooled) standard and marketing treat-

[33] Appendix Table A3 showed balance across each the three main groups we study on the study characteristics we examined in Table 1; for completeness, Appendix Table A4 shows balance of characteristics across the sub-treatments.

Figure A1: Standard Outreach Materials: Information Plus Assistance



ments in the Information Plus Assistance arm, and the control. We down-weight the individuals in the standard treatment in the Information Plus Assistance arm so that the (weighted) share in standard vs. marketing is the same (50 percent) in the Information Plus Assistance and Information Only arms.

B: DHS Data

Data sharing protocols

To construct our study population, DHS supplied BDT with a Medicaid outreach file of approximately 230,000 individuals aged 60 and older who were enrolled in Medicaid as of October 31, 2015. BDT removed the Medicaid recipient ID and created a unique, non-identifying scrambled study ID that uniquely identifies each individual. We received de-identified data files from DHS for all individuals on the initial outreach list (see Table 1, column 1). The data consist of: Medicaid enrollment and claims data, SNAP applications and enrollment data, and SNAP benefits data.

BDT provided DHS with the crosswalk between these de-identified study IDs and their unique Medicaid recipient ID. DHS then attached information on SNAP applications, SNAP enrollment, SNAP benefits, and Medicaid enrollment and claims. For the SNAP data, DHS sent the data to BDT who removed all personally-identifying information (i.e., full name, social security number, full address, and Medicaid recipient ID) and transmitted the de-identified data to us via a secure FTP process. For the Medicaid enrollment and claims files, DHS removed the same identifying information and directly transmitted the data to us.

Figure A2: Standard Outreach Materials: Information Only

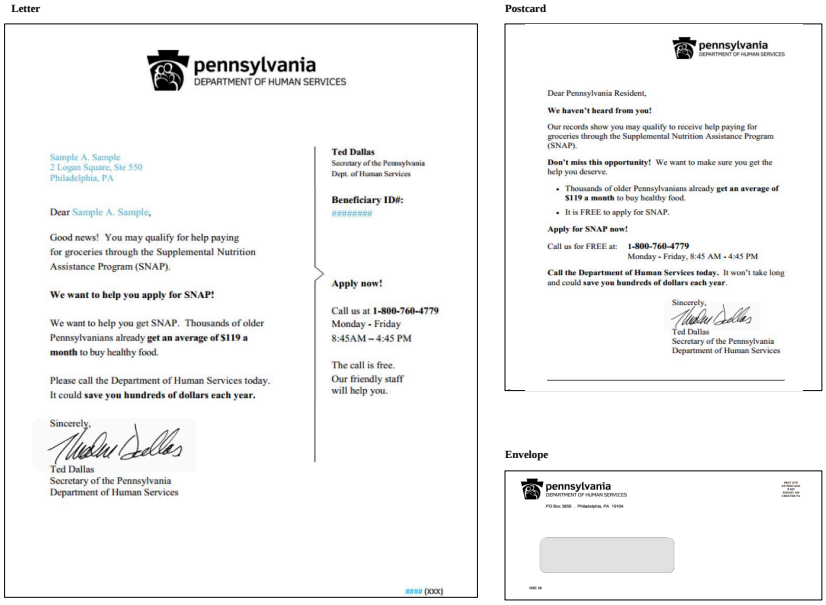
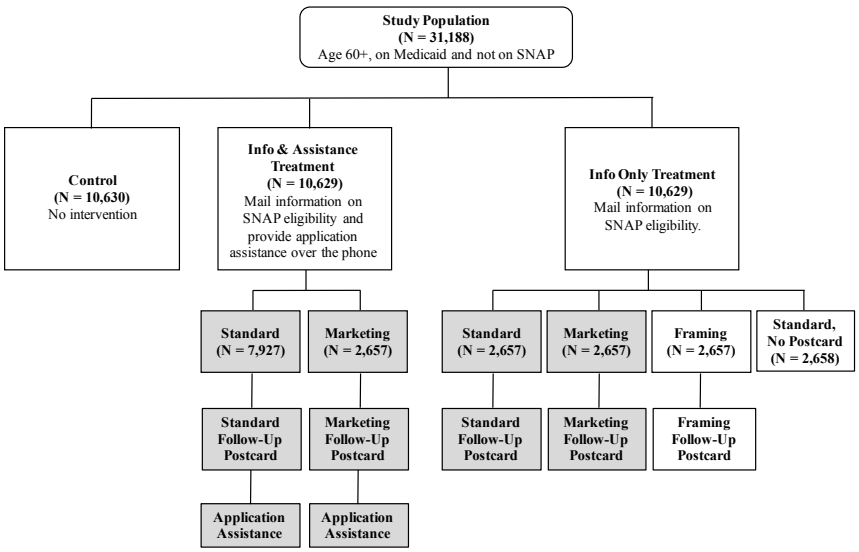


Figure A3: Experimental Design



Notes: Figure shows experimental design. Grey arms are the ones included in the main analyses.

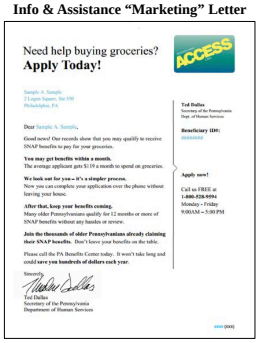
Figure A4: Timing and Sample Sizes for Mail Batches

Treatment	Subtreatment	1	2	3	4	5	6	7	8	9	10	11	Total
Date of Initial Mailing		1/6/2016	1/13/2016	1/20/2016	1/27/2016	2/3/2016	2/10/2016	2/17/2016	2/24/2016	3/2/2016	3/9/2016	3/16/2016	
Date of Follow Up Postcard mailing		3/2/2016	3/9/2016	3/16/2016	5/26/2016	5/27/2016	4/6/2016	4/13/2016	4/20/2016	4/27/2016	5/4/2016	5/11/2016	
Info & Assistance	Standard	750	750	750	750	750	750	750	750	750	750	472	7972
Info & Assistance	"Marketing"	250	250	250	250	250	250	250	250	250	250	157	2657
Info Only	Standard	250	250	250	250	250	250	250	250	250	250	157	2657
Info Only	Marketing	250	250	250	250	250	250	250	250	250	250	157	2657
Info Only	No Postcard	250	250	250	250	250	250	250	250	250	250	158	2658
Info Only	Framing	250	250	250	250	250	250	250	250	250	250	157	2657
Info & Assistance (Pooled)		1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	629	10629
Info Only (Standard + Marketing Pooled)		500	500	500	500	500	500	500	500	500	500	314	5314
Info Only (All Subtreatments Pooled)		1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	629	10629
		3000	3000	3000	3000	3000	3000	3000	3000	3000	3000	1888	31888

NOTE: Due to an implementation error, postcards for mail batches 4 and 5 were not sent as planned (eight weeks after the mail date) and therefore were sent following the last planned mailings in May 2016.

Figure A5: "Marketing" and "Framing" Outreach Materials

Info & Assistance "Marketing" Letter


Need help buying groceries?
Apply Today!

ACCESS

Sample A. Sample
2 Logan Square, Ste 150
Philadelphia, PA

Dear Sample A. Sample,

Good news! Our records show that you may qualify to receive SNAP benefits to pay for your groceries.

You may get benefits within a month.
The average applicant gets $119 a month to spend on groceries.

We look out for you – it's a simpler process.
Now you can complete your application over the phone without leaving your house.

After that, keep your benefits coming.
Many older Pennsylvanians qualify for 12 months or more of SNAP benefits without any hassles or review.

Join the thousands of older Pennsylvanians already claiming their SNAP benefits. Don't leave your benefits on the table.

Please call the PA Benefits Center today. It won't take long and could **save you hundreds of dollars each year**.

Sincerely,

Ted Dallas
Secretary of the Pennsylvania
Department of Human Services

Ted Dallas
Secretary of the Pennsylvania
Dept. of Human Services

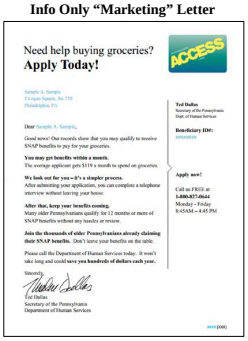
Beneficiary ID#:
#######

Apply now!

Call us FREE at
1-800-528-9994
Monday - Friday
9:00AM – 5:00 PM

[000]

Info Only "Marketing" Letter


Need help buying groceries?
Apply Today!

ACCESS

Sample A. Sample
2 Logan Square, Ste 150
Philadelphia, PA

Dear Sample A. Sample,

Good news! Our records show that you may qualify to receive SNAP benefits to pay for your groceries.

You may get benefits within a month.
The average applicant gets $119 a month to spend on groceries.

We look out for you – it's a simpler process.
After submitting your application, you can complete a telephone interview without leaving your house.

After that, keep your benefits coming.
Many older Pennsylvanians qualify for 12 months or more of SNAP benefits without any hassles or review.

Join the thousands of older Pennsylvanians already claiming their SNAP benefits. Don't leave your benefits on the table.

Please call the Department of Human Services today. It won't take long and could **save you hundreds of dollars each year**.

Sincerely,

Ted Dallas
Secretary of the Pennsylvania
Department of Human Services

Ted Dallas
Secretary of the Pennsylvania
Dept. of Human Services

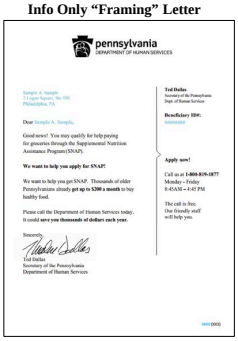
Beneficiary ID#:
#######

Apply now!

Call us FREE at
1-800-827-0664
Monday - Friday
8:45AM – 4:45 PM

[000]

Info Only "Framing" Letter


pennsylvania
DEPARTMENT OF HUMAN SERVICES

Sample A. Sample
2 Logan Square, Ste 150
Philadelphia, PA

Dear Sample A. Sample,

Good news! You may qualify for help paying for groceries through the Supplemental Nutrition Assistance Program (SNAP).

We want to help you apply for SNAP!

We want to help you get SNAP. Thousands of older Pennsylvanians already **get up to $200 a month** to buy healthy food.

Please call the Department of Human Services today. It could **save you thousands of dollars each year**.

Sincerely,

Ted Dallas
Secretary of the Pennsylvania
Department of Human Services

Ted Dallas
Secretary of the Pennsylvania
Dept. of Human Services

Beneficiary ID#:
#######

Apply now!

Call us at **1-800-819-1877**
Monday - Friday
8:45AM – 4:45 PM

The call is free.
Our friendly staff
will help you.

[000]

Info & Assistance "Marketing" Postcard


Need help buying groceries?
Apply Today!

ACCESS

Dear Pennsylvania Resident,

Good news! Our records show that you may qualify to receive SNAP benefits to help pay for your groceries. **It is FREE to apply!**

You may get benefits within a month. The average applicant gets $119 a month to spend on groceries.

We look out for you – it's a simpler process. Now you can complete your application over the phone without leaving your house.

After that, keep your benefits coming. Get benefits for 12 months or more with one application.

Call us for FREE at: **1-800-528-9994**
Monday - Friday, 9:00 AM - 5:00 PM

Call the PA Benefits Center today. It won't take long and could **save you hundreds of dollars each year**.

Sincerely,

Ted Dallas
Secretary of the Pennsylvania
Department of Human Services

pennsylvania
DEPARTMENT OF HUMAN SERVICES

Info Only "Marketing" Postcard


Need help buying groceries?
Apply Today!

ACCESS

Dear Pennsylvania Resident,

Good news! Our records show that you may qualify to receive SNAP benefits to help pay for your groceries. **It is FREE to apply!**

You may get benefits within a month. The average applicant gets $119 a month to spend on groceries.

We look out for you – it's a simpler process. After submitting your applications, you can complete a phone interview without leaving your house.

After that, keep your benefits coming. Get benefits for 12 months or more with one application.

Call us for FREE at: **1-800-827-0664**
Monday - Friday, 8:45 AM - 4:45 PM

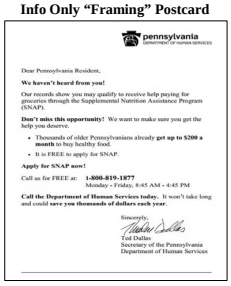
Call the Department of Human Services today. It won't take long and could **save you hundreds of dollars each year**.

Sincerely,

Ted Dallas
Secretary of the Pennsylvania
Department of Human Services

pennsylvania
DEPARTMENT OF HUMAN SERVICES

Info Only "Framing" Postcard


pennsylvania
DEPARTMENT OF HUMAN SERVICES

Dear Pennsylvania Resident,

We haven't heard from you!

Our records show you may qualify to receive help paying for groceries through the Supplemental Nutrition Assistance Program (SNAP).

Don't miss this opportunity! We want to make sure you get the help you deserve.

- Thousands of older Pennsylvanians already **get up to $200 a month** to buy healthy food.
- It is FREE to apply for SNAP.

Apply for SNAP now!

Call us for FREE at: **1-800-819-1877**
Monday - Friday, 8:45 AM - 4:45 PM

Call the Department of Human Services today. It won't take long and could **save you thousands of dollars each year**.

Sincerely,

Ted Dallas
Secretary of the Pennsylvania
Department of Human Services

NOTE: Envelopes (not shown) were identical to the standard envelopes for the respective arm (Information Plus Assistance, or Information Only) shown in Appendix Figures A1 and A2 respectively.

Medicaid outreach list

The Medicaid outreach file we analyze contains the individual's birth year, gender, city, primary language, an indicator of SNAP enrollment, and information on which Medicaid program the individual is enrolled in. All that information was provided by DHS; in addition, BDT supplemented it with a pseudo "household" ID that BDT created to denote people in the outreach file with the same last name and full address.

Medicaid enrollment and claims data

We received Medicaid enrollment and claims data from DHS for everyone on the outreach list. The Medicaid data contains seven files. There is an enrollment file that contains Medicaid enrollment spells from 1981 - 2016; we use the enrollment file to define the start date of the individual's last enrollment spell in Medicaid, and the days enrolled in 2015. We also use the enrollment file to construct a measure of race (since we do not have that in the outreach file).

In addition, there are six claims files that contain claims in 2015 for everyone on the outreach list. The claims include not only payments by Medicaid but also payments by Medicare if the individual is eligible for both. Only insurer payments are included; out-of-pocket spending is not observed but is unlikely to be large in this population.

Three of the claims files contain Fee-For-Service claims for outpatient, inpatient, and pharmaceutical services, respectively. The other three files contain analogous Managed Care encounters. Our claims files are therefore a mix of encounter data from Medicaid Managed Care and Fee for Service claims. In the data we received, we can only distinguish between managed care and fee-for-service based on claims filed. Although there are well-known measurement issues with encounter data - and comparability issues with fee for service claims data (e.g., Lewin Group 2012) - such measurement issues should not bias our comparisons of these measures across randomly assigned arms. For our study population (see Table 1, column 5), we estimate that about 60 percent of claims and about 80 percent of spending was in fee for service in 2015 .

We use the 2015 claims data to construct healthcare utilization and health measures. The healthcare measures are all measures of annual spending or healthcare use. However about one-quarter of our study population was not enrolled in Medicaid for the entirety of 2015. We therefore annualize the healthcare utilization and healthcare spending measures by multiplying our raw measures by the ratio of 365 to the number of days enrolled in Medicaid in 2015.

Below we describe the construction of specific variables.

Start of Last Medicaid Enrollment: We use the enrollment file to define the start date of the last consecutive enrollment period in Medicaid.

Indicator of Full Year Enrollment in 2015: Following the construction of enrollment spells as above, an individual is indicated as full-year enrollment in 2015 if she has the entirety of 2015 enrolled in Medicaid.

Total Health Care Spending: Total healthcare spending is defined as the sum of inpatient, outpatient, and pharmaceutical spending paid by Medicare or Medicaid. We winsorize spending at twice of the 99.5th percentile of the study population, which is $371,620.

Hospital Days: We measure the number of hospital days based on the total length of inpatient stays in the inpatient file. Stays with a discharge date earlier than the admission date are dropped, and overlapping periods are removed. By construction, the maximum value of this measure is 365.

Emergency Room (ER) Visits: We measure the number of emergency room visits in the outpatient file. ER visits are identified by HCPCS codes 99281-99285, 99291-99292, G0380-G0384, and G0390. We count the total number of ER visits for each individual, allowing a maximum of one ER visit per individual per day.

Doctor Visits: We measure the number of doctor visits as the sum of the number of primary care visits and the number of specialist visits, allowing a maximum of one primary care visit and one specialist visit per individual per day. To identify primary and specialist visits in the outpatient files, we match provider type and provider specialty to taxonomy codes using a crosswalk from DHS[34]. The taxonomy codes are then matched to HCFA specialty codes using a crosswalk from CMS[35]. Finally, HFCA codes are matched to a primary care or specialist classification using a crosswalk from Finkelstein et al. (2016)[36].

Skilled Nursing Facility (SNF) Days: We identify SNF stays in the inpatient files by provider type code 03 and provider specialty codes 030 or 031. Stays with a discharge date earlier than the admission date are dropped, and overlapping periods are removed. The total number of SNF days for each individual is calculated as the sum of length of stays. By construction the maximum value of this measure is 365.

Total Number of Visits and Days: This is the sum of number of hospital days, number of ER visits, number of doctor visits, and number of SNF days.

Weighted Total Number of Visits and Days: This is the (weighted) sum of hospital days, ER visits, doctor visits, and SNF days, where the weights are based on the average cost per encounter in our study population. The average cost per hospital day or SNF day is calculated by dividing total spending over the period of hospital stays or SNF stays experienced in the data by our study population by the total number of hospital days or SNF days experienced by our study population. The average cost per ER visit or doctor visit is calculated by averaging spending in our study population across ER visits or doctor visits. The resulting

[34]See http://www.dhs.pa.gov/cs/groups/webcontent/documents/document/p_002941.pdf.

[35]See https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/MedicarePr oviderSupEnroll/Downloads/TaxonomyCrosswalk.pdf.

[36]Finkelstein et al. 2016. Sources of Geographic Variation in Health Care: Evidence From Patient Migration. Quarterly Journal of Economics. 131 (4): 1681-1726.

estimates of average costs are $1,607 for a hospital day, $197 for an ED visit, $147 for a SNF day, and $79 for a doctor visit.

Chronic Conditions: We measure the number of chronic conditions recorded for each individual using the claims files. Each claim has between one and nine ICD-9 or ICD-10 codes for diagnoses. We unify diagnoses codes to ICD-9[37] and identify which ICD-9 codes correspond to chronic conditions following the method developed by Hwang et al. (2001).[38] We measure the number of chronic conditions for each individual by counting the number of unique ICD-9 codes matched to chronic conditions for each individual across their 2015 claims.

SNAP application and enrollee data

We received SNAP application data from DHS for all individuals on the original Medicaid outreach list. We obtained data on SNAP applications from March 2008 through February 2018 . For each SNAP application, the data contain the date the application was received, a disposition code, a disposition date, and a reason for rejection (if rejected). We use the date the application was received to define the date the individual applied for SNAP - our primary application measure is whether the individual applied for SNAP within 9 months of her initial mail date.

We use the disposition code to determine whether the application was approved, rejected or pending. The disposition date tells us when this disposition occurs. We define an application as rejected if they applied after the initial mail date and were rejected with a disposition date within 9 months after the initial mail date.

We define an individual as enrolled in SNAP if her application is approved with a disposition date within 9 months after the initial mail date. Note that because the definition of enrollee does not depend on the date the application was received, it is possible for us to record someone as a SNAP enrollee but not a SNAP applicant, if they applied before their initial mail date but enroll after it. In practice, this applies to about 2 percent of our SNAP enrollees, with the earliest application date 40 days before the initial mail date.

SNAP enrollee benefits

We received monthly benefit information for anyone on the original Medicaid outreach list enrolled in SNAP from March 2008 through February 2018 . We use these data to measure monthly benefits for our SNAP enrollees (defined in the previous section as individuals with an application approved during our nine-month observation window) in any months they are enrolled in the 9 months after the initial mail date. In principle the monthly benefit amount should be constant. However, many enrollees have a different benefit amount in the first month they are enrolled, presumably reflecting some pro-rating of benefits that (partial) month. We therefore measure benefits based on subsequent months if they disagree; for about two-thirds of enrollees, benefits are the same for

[37] Retrived from http://www.nber.org/data/icd9-icd-10-cm-and-pcs-crosswalk-general-equivalence-mapping.html.

[38] Hwang et al. 2001. Out-Of-Pocket Medical Spending for Care of Chronic Conditions. Health Affairs. 20 (6): 267-278.

all remaining months; when they are not, we use the modal benefit amount (or in rare cases of two modes, the most recent modal amount).

As noted in the text, while in principle we should be able to observe benefits for all individuals whom we measure as enrolled, in practice we are missing benefit information for about 4 percent of enrollees. If we instead define enrollment based on receiving positive benefits in any month since enrollment, this slightly increases our enrollment estimates (because we do not require the application to be approved during our nine-month observations window) but does not otherwise affect our results. Specifically, compared to our baseline findings in Table 2 of enrollment rates of 5.8% for control, 10.5% for Information Only, and 17.6% in Information Plus Assistance, we now estimate 8.8% for control, 14.2% for Information Only, and 21.0% for Information Plus Assistance.

Because SNAP is a benefit at household level, we also receive the number of individuals in the household linked to a given case; we use this to define the household size for enrollees.

C: Call-in data and call forwarding service

This section provides more detail on the call-in data, the call forwarding service, and the script for the call receptionists, which was provided in English and in Spanish.

We report the "raw" call-in rates in each study arm. Because the call forwarding service is not as good at determining the identity of callers as our BDT partner, the information-only treatment has a non-trivial number of callers without a valid study ID. We therefore also report an "adjusted" call-in rate for the Information Only treatment, which adjusts the measured call-in rate to account for our estimate of the rate of unrecorded callers. We also provide details on this adjustment procedure here.

Call-in Data for Information plus Assistance Treatment

BDT tracks all calls that it receives, and this forms the basis for the measure of call-ins to the BDT number in response to the outreach letters. We define a caller as someone who calls in to the appropriate phone number during business hours (9am-5pm) in the 9 months after the mail date. We exclude very small amount of cross-arm call contamination (e.g., individuals in control group calling in to the BDT number), and we also exclude calls beyond the 9 month window. BDT uses internal software that attempts to automatically determine the identity of caller. If the software is not able to determine identity automatically in real time, the BOS handling the phone call will ask the individual for additional identifying information (e.g., name, address) to try to determine identity. The BOS will also ask for the identification number on the printed letter or postcard to determine identity, if necessary. BDT provided us with both "caller-level" data as well as "call-level" data, and we take union of two files to determine individual-level call-in rates.

Call-in Data for Information Only Treatment

In order to capture comparable information on which individuals call in to DHS in response to the Information Only treatment, we contracted with a call forwarding service, HostedNumbers (HN) (www.hostednumbers.com). We arranged for a different 1-800 number for each treatment arm. If an individual called into one of these numbers, it would be directed to a call receptionist employed by HostedNumbers. The call receptionists were asked to read from a standard script and were asked to record the individual's identification number (printed on the outreach materials) before forwarding the call to DHS.

We worked with HostedNumbers (HN) to design a protocol that would try to capture comparable information to what BDT captures on which individuals call in to DHS in response to the Information Only treatment. We provided a "call script" in English and Spanish which the call receptionists were instructed to follow. The receptionists were instructed to ask for the nine digit beneficiary ID number on the letter (or postcard) that was received. The receptionists were instructed not to ask for any other information, and were told to interrupt callers if the caller was providing any other information. The goal was simply to collect the ID number and then forward the call to the Pennsylvania Department of Human Services.

The call receptionists used HN software to record the date and time of call, the length of call, ID number, and whether or not the call was successfully transferred to DHS. This data is transferred to us as "call-level" data, which we use to define valid calls during business hours (8:45am-4:45pm). Note that the valid call time is very slightly different from BDT hours (9am-5pm) because BDT hours are slightly different from DHS hours.

Calculating an adjusted call-in rate

Because HN does not have access to the same software as BDT to determine the identity of caller, we expected that the HN data would be less successful at measuring call-in rates when using only calls with a valid ID number. This is one explanation for the 3 percentage point lower raw caller rate in the Information Only treatment arm shown in Table 2.

As a result, we also developed an adjusted measure of the call-in rate that adjusts the "raw" call-in rate. The adjustment uses information on the number of calls from callers who call in and are successfully transferred to DHS, but do not provide a valid ID number in the HN database. This might occur, for example, if the call receptionist was not able to record it properly or the individual did not find the number of the letter or postcard.

To construct an adjusted call-in rate for the Information Only treatment arm, we pool the data across each of the Information Only sub-treatments, and we make the following assumptions and calculations:

1. Let N be the total number of individuals in Information Only treatment.

2. Let A be the total number of calls with a valid ID during the time period from the first day after the first mail batch to 9 months after the last letter batch (i.e., between 1/6/2016

and 12/15/2016). We use this period rather than 9 months after the mail date because for unknown calls, we do not know the mail date. Since distribution of calls is heavily skewed towards the beginning of time period, we still expect this to be a good estimate of the actual number of calls with a valid ID during the "study period" (defined as 9 months after mail date).

3. Let B be the total number of callers in Information Only treatment without any adjustment (the "raw" call-in rate).

4. Let C be the total number of calls *without* a valid ID. We only include calls between 1/6/2016 and 12/15/2016, during business hours, and calls that were successfully transferred to DHS.

5. Let D be the estimated number of calls that are not part of study, either because they are "test calls" that we made ourselves to HN during study period or because the calls are unsolicited calls from individuals outside of study population. To construct estimate of unsolicited calls, we calculate number of unsolicited calls in 4 months before study period and we multiply by 2.875 (11.5/4) to scale up this estimate to match period 1/6/2016-12/15/2016. This assumes that rate of unsolicited calls from outside study population is same during study period as in the 4 months before study period.

6. Let E = C - D, which is number of calls without valid ID that we belief are callers from study population.

7. Let F = B/A represent estimated probability of an unknown call coming from a caller, adjusting from repeated calls. This assumes that rate of repeat calls from population that provides valid ID is same as for callers who do not provide a valid ID.

8. Let G = B + E*F be estimate of number of "adjusted" callers for Information Only treatment.

As can be seen in Table 2, the adjusted caller rate in the information-only treatment is about 2 percentage points higher than the raw caller rate. To construct adjustment for each Information Only sub-treatment, we assume that the adjustment ratio is the same across arms and use the overall adjustment ratio for each arm.

Cross-contamination across arms

In processing the call-in data from both BDT and HN, we find a very small amount cross-contamination across all arms, meaning that we find calls from individuals to phone numbers different from the phone number that they are assigned. In some cases, this could reflect the fact that individuals find out about BDT through other channels. In other cases, this could reflect mistakes in the mail room in assigning letter batches. In either case, we proceed by only analyzing calls to the appropriate phone number, and we ignore cross-contamination calls. The extent of cross contamination is extremely small; see Appendix table A8.

D: Additional Analyses

Lee (2009) Bounds for Missing Benefits

In principle, we should observe benefits for all individuals whose applications have been approved during our nine-month observation window (our measure of "enrollee"). In practice, we are missing such information for about 4 percent of enrollees, and this missing rate is not balanced across arms. As shown in Table 4, 7.3 percent of control enrollees are missing benefit information, compared to 4.3 percent of the Information Only enrollees and 2.8 percent of the Information Plus Assistance enrollees; differences in missing benefit rates are statistically significantly different between either intervention arm and the control group. Such non-random attrition could bias our comparison of enrollee benefits across arms.

Therefore we implement the fairly conservative procedure of Lee (2009) to bound the potential bias arising from differential rates of missing benefits across study arms. We apply this approach to the the average monthly benefit of enrollees in each arm, as well as the share of enrollees with certain benefit amounts. Since we found that enrollees in both treatment arms had lower benefit amounts than the control arm, and enrollees in both treatment arms have lower missing benefit rates than the control arms, we remove the lowest observed benefits from the enrollees in each treatment arm until the share of enrollees with observed benefits in each treatment arm is equal to the share in the control arm. Lee (2009) shows that under a monotonicity assumption - any enrollee for whom we observe benefits in the treatment arm we would have observed benefits for had he/she been in the control arm - a lower bound for the effect of the treatment arm can be calculated for dropping treatment individuals with the lowest outcome values until the effective response rates are equalized between the two groups.

The computation follows Tauchmann (2014)[39]. We use STATA command: leebounds. The results - shown in Table A1 below, are similar to the unadjusted results in Table 4, which for comparison purposes we reproduce here. The adjusted benefit amounts for each intervention arm are still lower than for the control arm, and the differences are statistically significantly if they were in the unadjusted results in Table 4..

Generating Predicted Benefits and Predicted Enrollment

We estimate predicted benefits using everyone in the study population who enrolled in SNAP in the 9 months following the initial mail date and for whom we observe a benefit amount. The covariates are all pre-randomization variables taken from Table 1. Specifically, our predictors are dummies for age 80+, white, black, other race (unknown is the omitted race category), male, primary language non-English, residence in Pittsburgh, last Medicaid enrollment spell started before 2011, and enrolled in Medicaid for the full year of 2015; we also include as predictors continuous measures for the number of recorded chronic conditions in the Medicaid claims in 2015 and 2015 annualized

[39]The Stata Journal (2014) 14, Number 4, pp. 884–894, retrieved from https://www.stata-journal.com/sjpdf.html?articlenum=st0364.

Table A1: Bounding bias from missing benefit amount

	Control	Information Only		Information Plus Assistance	
		Unadjusted	Lee Bound	Unadjusted	Lee Bound
	(1)	(2)	(3)	(4)	(5)
Benefit Amount	145.94	115.38	118.67	101.32	105.51
		[0.000]	[0.000]	[0.000]	[0.000]
Share $16 Benefit	0.192	0.312	0.322	0.367	0.385
		[0.000]	[0.000]	[0.000]	[0.000]
Share $194 Benefit	0.206	0.164	0.170	0.147	0.154
		[0.076]	[0.102]	[0.003]	[0.003]
Share $357 Benefit	0.060	0.052	0.054	0.040	0.042
		[0.587]	[0.691]	[0.077]	[0.135]
N	613	559		1,861	

Notes: Table shows benefit amounts among enrollees in different arms. Enrollees with missing benefits are excluded. Column (1) shows the control means. Columns (2) and (3) show the means (with the p-value (relative tothe control) in [square brackets] show the means for the Information Only intervention. Column 2 reports the unadjusted mean (as shown in Table 4); Column 3 reports an upper bound on the mean, based on the Lee (2009) bounding procedure described in the text. Columns (4) and (5) present parallel results for the Information Plus Assistance Intervention. All p-values are calculated via the bootstrap.

health measures for health care spending, number of hospital days, SNF days, ED visits, and doctor visits.

There are clear modes in the distribution of benefits received, corresponding to minimum and maximum benefit amounts. To address this, we classify benefits into one of seven categories as shown in Appendix Table A2 . We use the "One-Vs-All" method for multi-class classification (Rifkin and Klautau 2004). Specifically, we estimate seven separate Logit models, where each model has dependent variable that takes on value of 1 for a given category and 0 otherwise. We then compute fitted value from each of these Logit models and we assign predicted category based on which fitted value is highest (e.g., if the fitted value is highest from the Logit model with the category 3 indicator as the dependent variable, then we assign category 3 as the predicted category). In order to avoid systematically underpredicting extreme categories, we adjust fitted values by adding and subtracting constant terms in each of the 7 Logit models and we iteratively adjust these constant terms until we have the overall predicted category shares that match the actual data. This does not adjust any of the Logit coefficients themselves, but ensures that the predicted category assignments are "unbiased" (i.e., for each category we predict the same number of observations as actually appear in that category in the data). We then convert each category to a predicted benefit amount in dollars by using the average actual benefit level in each category in the actual data.

Appendix Figures A6 shows our fit, which shows very close match across categories by design. This confirms that the algorithm appears to be unbiased in its predictions. To assess accuracy of the predictions, we calculate that roughly 38% of the observations are categorized correctly, and 78% of the predicted categories are only "off by one" category. Thus, we conclude that the machine

Table A2: Enrollee Monthly Benefits Categorization

Category	Criteria	Observations	Share of Observations
1	Monthly Benefit < $16	31	0.90
2	Monthly Benefit = $16	1,084	31.62
3	Monthly Benefit > $16 and Monthly Benefit < $194	1,346	39.26
4	Monthly Benefit = $194	559	16.31
5	Monthly Benefit > $194 and Monthly Benefit < $357	178	5.19
6	Monthly Benefit = $357	166	4.84
7	Monthly Benefit > $357	64	1.87
Total		3,428	

learning algorithm appears to have limited bias and a high degree of accuracy.

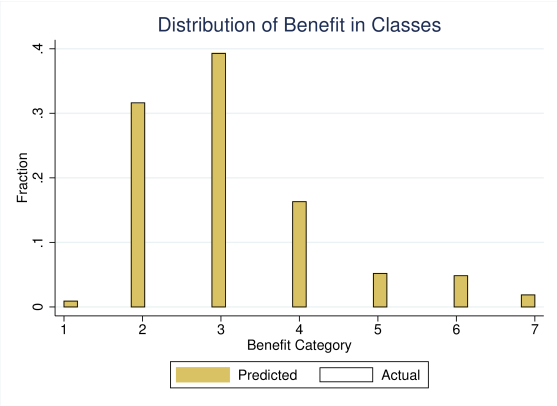
We estimate predicted enrollment in SNAP in the 9 months after the initial mail date using everyone in the control group. The predictors are the same as the estimation of predicted benefits. Since enrollment is a binary outcome, we estimate a Logit model, with the dependent variable taking on value of 1 for enrollee and 0 otherwise. We then compute the fitted value from the Logit model. To achieve an unbiased prediction, we adjust a "threshold" value interactively by adding or subtracting constant terms so that by assigning the observations whose fitted values are greater than the "threshold" to predicted enrollees, we have the overall predicted enrollment rate that matches the actual enrollment rate in the control group. We then apply this "threshold" to the whole study population.

We assess the accuracy of our predictions. We calculate that roughly 90% of the observations have the correct predicted enrollment. In addition, we calculate sensitivity to be 10% and specificity to be 94%.[40] Thus, we conclude that the algorithm appears to have a high degree of accuracy when prevalence is low, which is the case we expect for the study population.

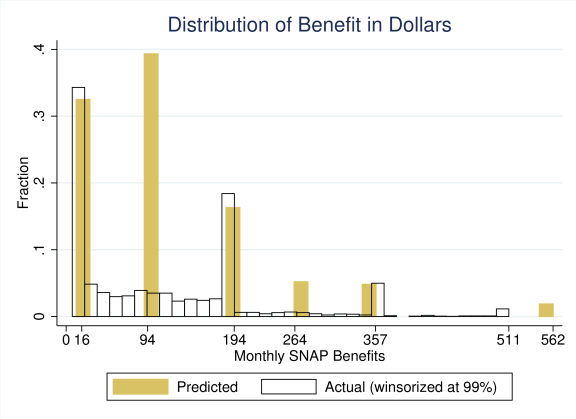
[40] We calculate sensitivity as the ratio of true positive and the sum of true positive and false negative; and we calculate specificity as the ratio of true negative and the sum of true negative and false positive. Note that the overall accuracy is the prevalence weighted sum of these two values.

Figure A6: Predicted and Actual Enrollee Monthly Benefits

Panel A: In Categories



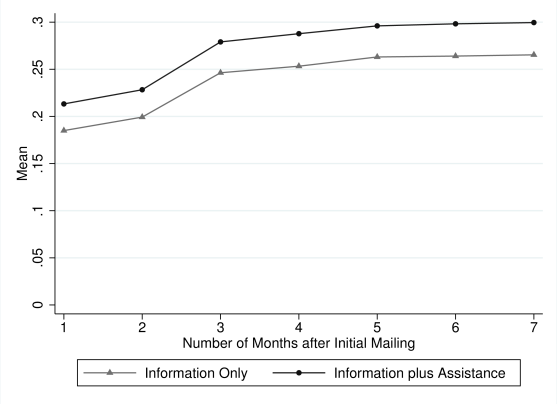
Panel B: In Dollars



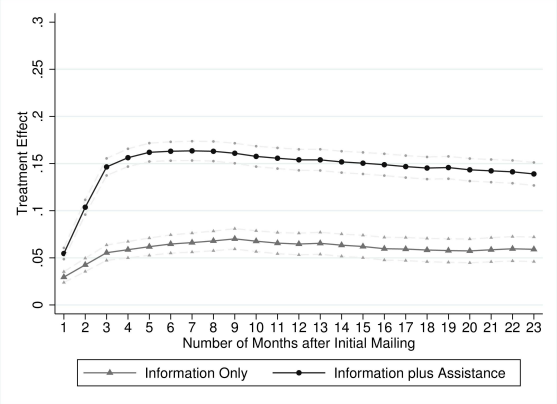
Additional Results

Figure A7: Time pattern of callers and applications

Panel A: Callers

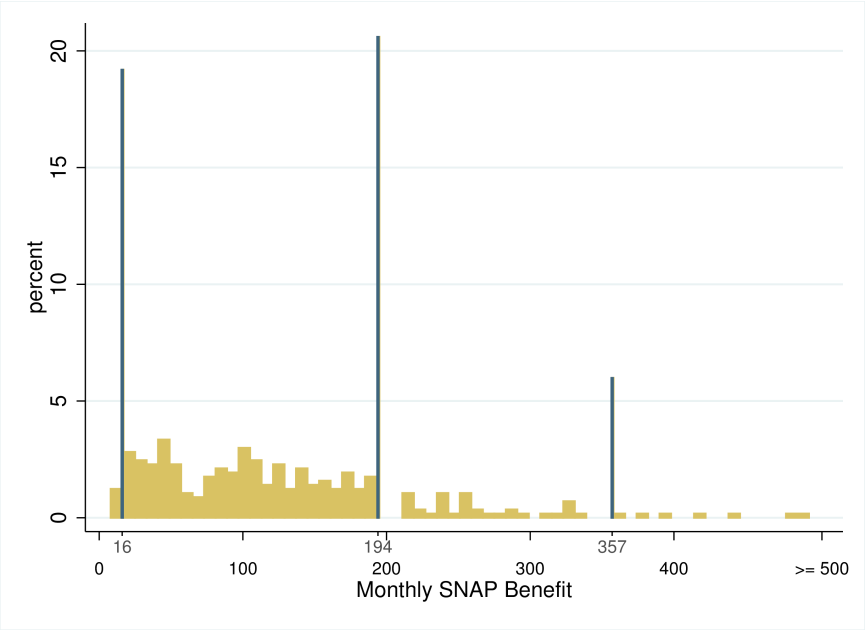


Panel B: Applicants



Notes: Figure shows (cumulative) outcomes by month relative to the initial mailing. Panel A shows the mean (cumulative) caller rates by month; the Information Only caller rate shown is unadjusted for under-measurement (see text for more details). Panel B shows the estimated treatment effects (relative to the control) for the Information Only arm and the Information Plus Assistance arm; 95 percent confidence intervals on these estimates are shown in the dashed light gray lines.

Figure A8: Distribution of enrollee benefits: control group enrollees



Notes: Figure plots the monthly enrollee benefit amount among enrollees in the control group (N=568). The three modes - which are the minimum benefit for a categorically eligible household of size 1 or 2 ($16), the maximum monthly benefit for a household of size 1 ($194), and the maximum monthly benefit for a household of size 2 ($357) - are binned separately from the other values; values greater than $500 are set to $500.

Table A3: Balance of Characteristics of Study Population Across Arms

	Control	Information Only	Information Plus Assistance	P Value of Difference (Column 2 vs 3)
	(1)	(2)	(3)	(4)
Panel A - Demographics				
Age (as of October 31, 2015)	68.80	68.93	68.80	
		[0.425]	[0.975]	[0.434]
Share Age 80+	0.16	0.17	0.16	
		[0.349]	[0.861]	[0.459]
Male	0.38	0.38	0.38	
		[0.965]	[0.702]	[0.718]
Share White [a]	0.76	0.76	0.75	
		[0.634]	[0.089]	[0.330]
Share Black [a]	0.08	0.07	0.08	
		[0.371]	[0.281]	[0.079]
Share Primary Language not English	0.04	0.04	0.04	
		[0.377]	[0.574]	[0.191]
Share Living in Pittsburgh	0.06	0.06	0.06	
		[0.737]	[0.871]	[0.854]
Share Last Medicaid Spell Starting before 2011	0.33	0.33	0.34	
		[0.629]	[0.287]	[0.665]
Share Enrolled in Medicaid for 2015 Full Year	0.73	0.73	0.72	
		[0.738]	[0.515]	[0.820]
Panel B - (Annual) Health Care Measures, 2015				
Total Health Care Spending ($) [b]	11,755	11,517	12,197	
		[0.632]	[0.325]	[0.201]
Number of Hospital Days	2.09	1.93	2.29	
		[0.470]	[0.378]	[0.151]
Number of ER Visits	0.47	0.59	0.50	
		[0.160]	[0.532]	[0.297]
Number of Doctor Visits	7.20	7.01	7.23	
		[0.515]	[0.920]	[0.514]
Number of SNF Days	2.65	2.39	2.81	
		[0.459]	[0.623]	[0.269]
Number of Chronic Conditions	5.46	5.34	5.44	
		[0.337]	[0.809]	[0.477]
F Statistic		0.560	0.660	0.746
P Value		[0.906]	[0.825]	[0.752]
Observations (N)	10,630	5,314	10,629	

NOTE: Table shows means of pre-randomization characteristics for the study population. Columns 1 through 3 show means by intervention arm with the p-value (relative to the control arm) in [square brackets] Column 4 reports the p-value for the difference between the two intervention arms. The bottom rows of the table report the F-statistic (and the p-value on that F-statistic) from the joint test of equality across all of the pre-randomization characteristics shown. P-values of the differences in characteristics are based on heteroskedasticity-robust standard errors. The F-statistic (and associated p-value) is calculated based a regression in which we "stack" all of the variable values into a single left-hand side outcome variable and interact the treatment indicator with variable fixed effects; the F-distribution is simulated using permutation with 1,000 iterations.

Table A4: Balance of Characteristics of Study Population: By Sub treatments

	Control	Information Only				Information Plus Assistance		P Value of Difference between			
		Standard	No-Postcard	Marketing	Framing	Standard	Marketing	Control vs Treatment (col 1 vs 2+4+6+7)	Standard vs Marketing (col 2+6 vs 4+7)	Information Only Standard vs Framing (col 2 vs 5)	Information Only Standard vs No Follow-up Postcard (col 2 vs 3)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A - Demographics											
Age (as of October 31, 2015)	68.80	68.95	69.10	68.90	68.89	68.68	68.91				
		[0.472]	[0.136]	[0.608]	[0.662]	[0.388]	[0.589]	[0.622]	[0.583]	[0.817]	[0.559]
Share Age 80+	0.16	0.17	0.17	0.16	0.17	0.16	0.17				
		[0.153]	[0.518]	[0.997]	[0.402]	[0.630]	[0.574]	[0.484]	[0.734]	[0.637]	[0.532]
Male	0.38	0.38	0.37	0.38	0.38	0.38	0.37				
		[0.987]	[0.391]	[0.959]	[0.814]	[0.602]	[0.378]	[0.851]	[0.463]	[0.843]	[0.490]
Share White [a]	0.76	0.76	0.75	0.75	0.75	0.75	0.75				
		[0.728]	[0.590]	[0.698]	[0.407]	[0.154]	[0.179]	[0.205]	[0.798]	[0.703]	[0.880]
Share Black [a]	0.08	0.07	0.07	0.07	0.08	0.08	0.08				
		[0.599]	[0.463]	[0.387]	[0.651]	[0.848]	[0.186]	[0.954]	[0.576]	[0.439]	[0.871]
Share Primary Language not English	0.04	0.04	0.05	0.05	0.04	0.04	0.04				
		[0.897]	[0.296]	[0.149]	[0.254]	[0.815]	[0.342]	[0.790]	[0.723]	[0.433]	[0.347]
Share Living in Pittsburgh	0.06	0.05	0.05	0.06	0.06	0.06	0.05				
		[0.338]	[0.468]	[0.685]	[0.852]	[0.552]	[0.520]	[0.760]	[0.853]	[0.370]	[0.858]
Share Last Medicaid Spell Starting before 2011	0.33	0.33	0.33	0.33	0.36	0.33	0.34				
		[0.849]	[0.888]	[0.577]	[0.006]	[0.974]	[0.142]	[0.362]	[0.253]	[0.043]	[0.969]
Share Enrolled in Medicaid for 2015 Full Year	0.73	0.73	0.74	0.73	0.76	0.73	0.72				
		[0.857]	[0.276]	[0.737]	[0.005]	[0.599]	[0.596]	[0.561]	[0.841]	[0.020]	[0.317]
Panel B - (Annual) Health Care Measures, 2015											
Total Health Care Spending ($) [b]	11,755	11,514	12,630	11,520	11,860	11,561	12,833				
		[0.711]	[0.205]	[0.710]	[0.871]	[0.654]	[0.109]	[0.796]	[0.230]	[0.674]	[0.193]
Number of Hospital Days	2.09	1.88	2.33	1.97	2.27	2.29	2.29				
		[0.474]	[0.474]	[0.680]	[0.565]	[0.359]	[0.563]	[0.914]	[0.848]	[0.308]	[0.262]
Number of ER Visits	0.47	0.59	0.44	0.59	0.44	0.54	0.46				
		[0.180]	[0.420]	[0.392]	[0.459]	[0.271]	[0.819]	[0.144]	[0.699]	[0.108]	[0.103]
Number of Doctor Visits	7.20	6.81	6.85	7.22	6.67	7.23	7.23				
		[0.219]	[0.275]	[0.973]	[0.103]	[0.900]	[0.955]	[0.736]	[0.549]	[0.728]	[0.903]
Number of SNF Days	2.65	2.02	2.97	2.76	2.96	2.55	3.07				
		[0.138]	[0.511]	[0.820]	[0.525]	[0.753]	[0.391]	[0.862]	[0.098]	[0.106]	[0.101]
Number of Chronic Conditions	5.46	5.28	5.63	5.40	5.31	5.52	5.35				
		[0.255]	[0.345]	[0.719]	[0.337]	[0.634]	[0.503]	[0.456]	[0.886]	[0.890]	[0.102]
F Statistic		0.856	1.038	0.399	1.367	0.555	0.972	0.617	0.676	1.044	1.143
P Value		[0.644]	[0.450]	[0.980]	[0.197]	[0.918]	[0.481]	[0.867]	[0.816]	[0.377]	[0.336]
Observations (N)	10,630	2,657	2,658	2,657	2,657	7,972	2,657				

Notes: Table shows means of pre-randomization characteristics for the study population. Columns 1 through 7 show means by intervention sub-arm with the p-value (relative to the control arm) in [square brackets] Columns 8 through 11 report the p-value for differences between various groups of sub-arms, as indicated. The bottom rows of the table report the F-statistic (and the p-value on that F-statistic) from the joint test of equality across all of the pre-randomization characteristics shown. P-values of the differences in characteristics are based on heteroskedasticity-robust standard errors. The F-statistic (and associated p-value) is calculated based a regression in which we "stack" all of the variable values into a single left-hand side outcome variable and interact the treatment indicator with variable fixed effects; the F-distribution is simulated using permutation with 1,000 iterations.

Table A5: Behavioral Responses to Interventions: All sub-treatments

	Control	Information Only				Information Plus Assistance		P Value of Difference between			
		Standard	No-Postcard	Marketing	Framing	Standard	Marketing	Control vs Treatment (col 1 vs 2+4+6+7)	Standard vs Marketing (col 2+6 vs 4+7)	Information Only Standard vs Framing (col 2 vs 5)	Information Only Standard vs No Follow-up Postcard (col 2 vs 3)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
SNAP Enrollees	0.058	0.112	0.092	0.098	0.111	0.174	0.179				
		[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.481]	[0.896]	[0.016]
SNAP Applicants	0.077	0.151	0.120	0.143	0.157	0.236	0.239				
		[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.730]	[0.543]	[0.001]
SNAP Rejections among Applicants	0.233	0.224	0.216	0.311	0.281	0.261	0.250				
		[0.751]	[0.536]	[0.005]	[0.071]	[0.116]	[0.442]	[0.115]	[0.133]	[0.065]	[0.777]
Callers	0.000	0.278	0.212	0.256	0.300	0.298	0.303				
		[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.288]	[0.079]	[0.000]
Adjusted Callers	0.000	0.300	0.234	0.278	0.322	0.298	0.303				
		[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.295]	[0.086]	[0.000]
SNAP Applicants among Non-Callers	0.077	0.089	0.074	0.084	0.093	0.085	0.077				
		[0.079]	[0.593]	[0.295]	[0.025]	[0.066]	[0.953]	[0.069]	[0.262]	[0.681]	[0.071]
SNAP Applicants among Callers	0.000	0.311	0.295	0.315	0.306	0.592	0.612				
		[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.238]	[0.830]	[0.524]
SNAP Enrollees among Non-Callers	0.058	0.064	0.054	0.058	0.062	0.060	0.058				
		[0.284]	[0.492]	[0.934]	[0.437]	[0.578]	[0.908]	[0.467]	[0.449]	[0.824]	[0.172]
SNAP Enrollees among Callers	0.000	0.237	0.234	0.215	0.225	0.442	0.457				
		[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.847]	[0.571]	[0.921]
Observations (N)	10,630	2,657	2,658	2,657	2,657	7,972	2,657				

Notes: Columns 1 through 6 show means of outcomes by intervention arm, with the p-value (relative to the control arm) in [square brackets]. Columns 8 through 11 report p-values for comparisons shown in column heading. In column 8, sub-treatments are weighted so that within the Information Plus Assistance arm the standard and marketing sub-treatments receive equal weight, and the Information Plus Assistance treatments receive equal weight as the Information Only treatments. In column 9, sub-treatments are weighted so that Information Plus Assistance and Information Only are equally weighted in Standard and Marketing arms. All outcomes are binary rates measured during the nine months from the initial mail date. All p-values are based on heteroskedasticity-robust standard errors. Callers are measured for the relevant call number and are therefore mechanically zero for the control; see text for a description of the adjusted caller rate.

Table A6: Enrollee Benefits and Predicted Benefits: All sub-treatments

	Control	Information Only				Information Plus Assistance		P Value of Difference between			
		Standard	No-Postcard	Marketing	Framing	Standard	Marketing	Control vs Treatment (col 1 vs 2+4+6+7)	Standard vs Marketing (col 2+6 vs 4+7)	Information Only Standard vs Framing (col 2 vs 5)	Information Only Standard vs No Follow-up Postcard (col 2 vs 3)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Benefit Amount	145.94	112.60	119.72	118.55	132.33	103.96	98.72				
		[0.000]	[0.004]	[0.003]	[0.174]	[0.000]	[0.000]	[0.000]	[0.754]	[0.065]	[0.471]
Share $16 Benefit	0.192	0.309	0.295	0.316	0.264	0.367	0.368				
		[0.000]	[0.003]	[0.000]	[0.021]	[0.000]	[0.000]	[0.000]	[0.802]	[0.247]	[0.732]
Share $194 Benefit	0.206	0.161	0.184	0.168	0.156	0.148	0.146				
		[0.107]	[0.467]	[0.193]	[0.070]	[0.003]	[0.011]	[0.006]	[0.981]	[0.856]	[0.504]
Share $357 Benefit	0.060	0.035	0.060	0.072	0.069	0.038	0.041				
		[0.094]	[0.999]	[0.527]	[0.622]	[0.056]	[0.176]	[0.166]	[0.127]	[0.073]	[0.193]
Share Missing Benefit	0.073	0.044	0.045	0.042	0.064	0.021	0.036				
		[0.061]	[0.093]	[0.056]	[0.613]	[0.000]	[0.005]	[0.001]	[0.352]	[0.264]	[0.943]
Predicted Benefit for Enrollees w/ Nonmissing Benefit	140.20	111.06	126.11	114.13	131.06	106.19	99.72				
		[0.000]	[0.130]	[0.003]	[0.293]	[0.000]	[0.000]	[0.000]	[0.510]	[0.038]	[0.140]
Predicted Benefit for All Enrollees	138.65	112.99	126.04	115.17	130.07	106.57	101.56				
		[0.001]	[0.166]	[0.007]	[0.302]	[0.000]	[0.000]	[0.000]	[0.581]	[0.067]	[0.193]
Share of Enrollees in Household Size of 1	0.657	0.742	0.673	0.682	0.695	0.752	0.769				
		[0.008]	[0.652]	[0.479]	[0.256]	[0.000]	[0.000]	[0.000]	[0.637]	[0.207]	[0.084]
Benefit Amount for Enrollees in Household Size of 1	116.97	95.32	98.01	90.88	96.45	85.61	86.03				
		[0.004]	[0.010]	[0.001]	[0.005]	[0.000]	[0.000]	[0.000]	[0.686]	[0.895]	[0.754]
Observations (N)	613	298	245	261	295	1,385	476				

Notes: Sample is individuals who enrolled in the 9 months after their initial mailing. Columns 1 through 7 shows means by intervention arm with the p-value (relative to the control arm) in [square brackets] for SNAP enrollees. Column 8 - 11 report p-values for comparisons shown in the column headings. In column 8, sub-treatments are weighted so that within the Information Plus Assistance arm the standard and marketing sub-treatments receive equal weight, and the Information Plus Assistance treatments receive equal weight as the Information Only treatments. In column 9, sub-treatments are weighted so that Information Plus Assistance and Information Only are equally weighted in Standard and Marketing arms. First 5 measures are based on actual benefit amounts received by SNAP enrollees; see text for a description of the predicted benefits. All p-values are based on heteroskedasticity-robust standard errors.

Table A7: Always Taker and Complier Enrollee Benefits and Predicted Benefits

	Always Takers	Compliers		P Value of Difference (Column 2 vs 3)
		Information Only	Information Plus Assistance Arms	
	(1)	(2)	(3)	(4)
Benefit Amount	145.94	78.31	79.66	
		[0.000]	[0.000]	[0.910]
Share $16 Benefit	0.192	0.458	0.453	
		[0.000]	[0.000]	[0.911]
Share $194 Benefit	0.206	0.114	0.119	
		[0.079]	[0.002]	[0.908]
Share $357 Benefit	0.060	0.043	0.030	
		[0.595]	[0.074]	[0.599]
Share Missing Benefit	0.073	0.006	0.007	
		[0.023]	[0.000]	[0.955]
Predicted Benefit for Enrollees w/ Actual Benefit	140.20	78.87	84.83	
		[0.000]	[0.000]	[0.629]
Predicted Benefit for All Enrollees	138.65	84.10	87.21	
		[0.001]	[0.000]	[0.788]
Share of Enrollees in Household Size of 1	0.657	0.782	0.810	
		[0.035]	[0.000]	[0.581]
Benefit Amount for Enrollees in Household Size of 1	116.97	64.69	70.70	
		[0.000]	[0.000]	[0.587]
Share of Sub-Population	0.058	0.048	0.119	

Notes: Sample is individuals who enrolled in the 9 months after their initial mailing. Variables reported are the same as in Table 4. Column 1 shows the mean of the always takers (individuals who enroll regardless of intervention), while columns 2 and 3 show the means for compliers (individuals who enroll if and only if they receive the intervention) for each intervention; p-value (relative to the always takers) is in [square brackets] for SNAP enrollees. Column 4 reports the p-value of the difference between the compliers in the two intervention arms. In column 2 the two equally-sized sub-treatments are pooled; in column 3 the two pooled sub-treatments are weighted so that they receive equal weight. Standard errors and p-values are computed with 10,000 replications of the bootstrap. Appendix F provides more detail on how the objects in this table were calculated.

Table A8: Cross-Group Caller Rates

Call from: \ Call to:	Info Plus Assistance (1)	Info Only (Standard) (2)	Info Only (No Postcard) (3)	Info Only (Marketing) (4)	Info Only (Framing) (5)	Observations (N) (6)
Control	0.395	0.019	0.000	0.000	0.009	10630
Info Plus Assistance (Standard)	29.767	0.013	0.013	0.000	0.050	7972
Info Plus Assistance (Marketing)	30.335	0.000	0.000	0.000	0.000	2657
Info Only (Standard)	0.414	27.813	0.000	0.038	0.000	2657
Info Only (No Postcard)	0.376	0.000	21.181	0.150	0.075	2658
Info Only (Marketing)	0.414	0.000	0.075	25.555	0.000	2657
Info Only (Framing)	0.489	0.000	0.038	0.188	29.996	2657

Notes: Table reports the percent of the study population in each arm who calls into the phone line for each arm. An individual will be counted multiple times if she calls into more than one phone line; however in practice less than 1 percent of callers who call the number they are supposed to call also call another group's number.

Table A9: Rejection reasons

	Control	Information Only	Information Plus Assistance	P Value of Difference (Column 2 vs 3)
	(1)	(2)	(3)	(4)
Insufficient Interest	0.511	0.433	0.680	
		[0.121]	[0.000]	[0.000]
Ineligible After Review	0.389	0.486	0.232	
		[0.054]	[0.000]	[0.000]
Other Reasons	0.100	0.082	0.088	
		[0.529]	[0.630]	[0.791]
Observations (N)	190	208	650	

Notes: Table reports the percent of rejected applicants rejected for different reasons. Columns 1 through 3 shows share of rejected applicants that were rejected for a given group of reasons, by intervention arm with the p-value (relative to the control arm) in [square brackets]. Column 1 shows the control. Column 2 shows the Information Only arm (for the same two equally-sized pooled sub-treatments). Column 3 shows the Information Plus Assistance arms (weighted so that the two sub-treatments received equal weight). Column 4 reports the p-value of the difference between the Information Plus Assistance and Information Only treatment arms. We grouped rejections by reason given. “Insufficient interest” includes "failure to furnish required information", "failure to sign required forms", "failure to supply identification proof", "voluntary withdrawal", and "failure to keep appointments". “Ineligibility after review” includes "failing income, resources, or public assistance tests", "failure to meet citizenship or residence requirements", "categorical ineligibility", "failure to meet employment tests", "failure to meet household composition requirements", and "institutionalization or imprisonment". There other reasons such as "duplicate application", "application entered in error" that we cannot categorize into the preceding groups are reported as “other reasons”.

Table A10: Age and Health Characteristics of Applicants and Enrollees: Additional Detail

	Applicants				Enrollees			
	Means			P Value	Means			P Value
	Control	Info Only	Info Plus Assistance	Info Plus Assistance vs Info Only	Control	Info Only	Info Plus Assistance	Info Plus Assistance vs Info Only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A - Individual (Annual) Health Care Measures, 2015								
Number of Hospital Days	2.24	1.39	1.21		2.64	1.48	1.19	
		[0.099]	[0.028]	[0.576]		[0.075]	[0.015]	[0.448]
Number of ER Visits	0.75	0.57	0.41		0.84	0.64	0.41	
		[0.308]	[0.034]	[0.037]		[0.388]	[0.042]	[0.025]
Number of Doctor Visits	8.88	7.40	6.39		9.75	7.42	6.38	
		[0.102]	[0.002]	[0.067]		[0.045]	[0.001]	[0.125]
Number of SNF Days	1.47	2.30	1.91		1.56	1.36	1.94	
		[0.357]	[0.517]	[0.647]		[0.824]	[0.646]	[0.472]
Panel B - Demographics								
Age (as of October 31, 2015)	66.07	67.32	67.91		65.94	67.06	68.03	
		[0.001]	[0.000]	[0.095]		[0.011]	[0.000]	[0.022]
Observations (N)	817	781	2,519		613	559	1,861	

Notes: Table reports additional characteristics that are shown in different form in Table 5. Columns 1 - 3, and 5 - 7 show means by intervention arm with the p-value (relative to the control arm) in [square brackets] for the study population, SNAP applicants who applied within 9 months of their initial mailing, and SNAP enrollees who enrolled within 9 months of their initial mailing, respectively. Column 1 and 5 show the control. Columns 2 and 6 show the Information Only arm (for the same two equally-sized pooled sub-treatments). Columns 3 and 7 show the Information Plus Assistance arms (weighted so that the two pooled sub-treatments received equal weight). Columns 4 and 8 report the p-value of the difference between the Information Plus Assistance and Information Only treatment arms. All p-values are based on heteroskedasticity-robust standard errors.

Table A11: Demographic and Health Characteristics of Enrollees

	Enrollees						
	Always Takers	Compliers		P Value of Difference (Column 2 vs 3)	Never Takers		P Value of Difference between Compliers and Never Takers (pooling 2 & 3 vs pooling 5 & 6)
		Info Only	Info Plus Assistance		Info Only	Info Plus Assistance	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A - Predicted Benefits							
Predicted Benefits	138.65	84.10	87.21	[0.788]	111.68	115.00	[0.000]
		[0.001]	[0.000]				
Panel B - (Annual) Health Care Measures, 2015							
Total Health Care Spending ($) [a]	10,238	8,676	7,809	[0.767]	11,750	12,968	[0.001]
		[0.675]	[0.213]				
Total Number of Visits and Days	14.79	6.19	7.56	[0.632]	12.04	13.45	[0.001]
		[0.049]	[0.004]				
Weighted Total Number of Visits and Days	5,407	716	1,504	[0.604]	4,214	5,181	[0.000]
		[0.054]	[0.004]				
Number of Chronic Conditions	6.54	4.07	4.80	[0.381]	5.33	5.45	[0.044]
		[0.020]	[0.006]				
Panel C - Demographics							
Share Age above Median (=65)	0.39	0.46	0.49	[0.569]	0.52	0.51	[0.214]
		[0.299]	[0.007]				
Share Age 80+	0.07	0.17	0.18	[0.815]	0.17	0.17	[0.615]
		[0.005]	[0.000]				
Male	0.39	0.44	0.37	[0.155]	0.38	0.38	[0.669]
		[0.443]	[0.436]				
Share White [b]	0.71	0.87	0.82	[0.213]	0.75	0.74	[0.000]
		[0.003]	[0.000]				
Share Black [b]	0.11	0.02	0.10	[0.003]	0.07	0.08	[0.823]
		[0.011]	[0.845]				
Share Primary Language not English	0.06	0.03	0.01	[0.482]	0.04	0.04	[0.008]
		[0.232]	[0.002]				
Share Living in Pittsburgh	0.05	0.08	0.09	[0.696]	0.06	0.05	[0.016]
		[0.366]	[0.029]				
Share Last Medicaid Spell Starting before 2011	0.26	0.42	0.34	[0.087]	0.33	0.34	[0.308]
		[0.006]	[0.024]				
Share of Individuals in Household Size of 1	0.66	0.78	0.81	[0.581]	0.08	0.07	[0.000]
		[0.035]	[0.000]				
Share of Sub-Population	0.058	0.048	0.119		0.89	0.82	

Notes: Sample is individuals who enrolled in SNAP within 9 months of their initial mailing. Column 1 shows the mean of the always takers (individuals who enroll regardless of intervention), while columns 2 and 3 show the means for compliers (individuals who enroll if and only if they receive the intervention) for each intervention; p-value (relative to the always takers) is in [square brackets]. Column 4 reports the p-value of the difference between the compliers in the two intervention arms. Columns 5 and 6 show the means for never takers (individuals who never enroll regardless of intervention) for each intervention. Column 7 reports the p-value of the difference between the compliers and never takers (pooling two intervention arms together). In columns 2 and 5 the two equally-sized sub-treatments are pooled; in columns 3 and 6 the two pooled sub-treatments are weighted so that they receive equal weight. Standard errors and p-values are computed with 10,000 replications of the bootstrap. Appendix F provides more detail on how the objects in this table were calculated. Variables reported are the same as in Table 5.

[a] Total spending is truncated at twice 99.5th percentile of study population, which is 371,620 (99.5th percentile in study population is 185,810). Amounts greater than the threshold are set to missing.

[b] Omitted category is other or missing race.

Table A12: Demographic and Health Characteristics of Applicants

	Applicants						
	Always Takers	Compliers		P Value of Difference (Column 2 vs 3)	Never Takers		P Value of Difference between Compliers and Never Takers (pooling 2 & 3 vs pooling 5 & 6)
		Info Only	Info Plus Assistance		Info Only	Info Plus Assistance	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A - Predicted Benefits							
Predicted Benefits	148.26	100.85	99.65	[0.910]	109.56	112.35	[0.028]
		[0.000]	[0.000]				
Panel B - (Annual) Health Care Measures, 2015							
Total Health Care Spending ($) [a]	9,424	7,707	7,813	[0.937]	12,019	13,403	[0.000]
		[0.533]	[0.314]				
Total Number of Visits and Days	13.33	9.84	8.29	[0.576]	11.96	13.73	[0.004]
		[0.331]	[0.011]				
Weighted Total Number of Visits and Days	4,661	1,752	1,938	[0.857]	4,261	5,362	[0.000]
		[0.117]	[0.011]				
Number of Chronic Conditions	6.21	4.83	4.83	[0.999]	5.31	5.49	[0.079]
		[0.094]	[0.006]				
Panel C - Demographics							
Share Age above Median (=65)	0.41	0.51	0.49	[0.680]	0.52	0.51	[0.277]
		[0.074]	[0.016]				
Share Age 80+	0.06	0.16	0.18	[0.646]	0.18	0.17	[0.811]
		[0.000]	[0.000]				
Male	0.41	0.40	0.37	[0.394]	0.38	0.38	[0.891]
		[0.994]	[0.228]				
Share White [b]	0.67	0.80	0.78	[0.540]	0.76	0.75	[0.120]
		[0.004]	[0.000]				
Share Black [b]	0.10	0.05	0.11	[0.011]	0.07	0.07	[0.058]
		[0.101]	[0.587]				
Share Primary Language not English	0.08	0.04	0.02	[0.232]	0.04	0.04	[0.145]
		[0.133]	[0.000]				
Share Living in Pittsburgh	0.05	0.07	0.08	[0.796]	0.06	0.05	[0.025]
		[0.389]	[0.068]				
Share Last Medicaid Spell Starting before 2011	0.25	0.35	0.31	[0.269]	0.34	0.35	[0.291]
		[0.018]	[0.014]				
Share of Individuals in Household Size of 1	0.53	0.62	0.63	[0.748]	0.08	0.07	[0.000]
		[0.067]	[0.000]				
Share of Sub-Population	0.077	0.070	0.161		0.853	0.762	

Notes: Sample is individuals who applied for SNAP within 9 months of their initial mailing. Column 1 shows the mean of the always takers (individuals who apply regardless of intervention), while columns 2 and 3 show the means for compliers (individuals who apply if and only if they receive the intervention) for each intervention; p-value (relative to the always takers) is in [square brackets]. Column 4 reports the p-value of the difference between the compliers in the two intervention arms. Columns 5 and 6 show the means for never takers (individuals who never apply regardless of intervention) for each intervention. Column 7 reports the p-value of the difference between the compliers and never takers (pooling two intervention arms together). In columns 2 and 5 the two equally-sized sub-treatments are pooled; in columns 3 and 6 the two pooled sub-treatments are weighted so that they receive equal weight. Standard errors and p-values are computed with 10,000 replications of the bootstrap. Appendix F provides more detail on how the objects in this table were calculated. Variables reported are the same as in Table 5.

[a] Total spending is truncated at twice 99.5th percentile of study population, which is 371,620 (99.5th percentile in study population is 185,810). Amounts greater than the threshold are set to missing.

[b] Omitted category is other or missing race.

Table A13: Caller Demographic and Health Characteristics: By Treatment Arm

	Information Only	Information Plus Assistance	P Value of Difference
	(1)	(2)	(3)
Panel A - Predicted Benefits			
Predicted Benefits	104.99	108.76	[0.286]
Predicted Enrollment	0.05	0.06	[0.330]
Panel B - (Annual) Health Care Measures, 2015			
Total Health Care Spending ($) [a]	6779	7792	[0.074]
Total Number of Visits and Days	10.16	8.96	[0.194]
Weighted Total Number of Visits and Days	3167	2575	[0.265]
Number of Chronic Conditions	5.16	5.15	[0.982]
Panel C - Demographics			
Share Age 80+	0.16	0.16	[0.895]
Male	0.38	0.37	[0.561]
Share White [b]	0.79	0.76	[0.044]
Share Black [b]	0.08	0.09	[0.189]
Share Primary Language not English	0.03	0.03	[0.389]
Share Living in Pittsburgh	0.06	0.06	[0.654]
Share Last Medicaid Spell Starting before 2011	0.34	0.31	[0.076]
Observations (N)	1,418	3,179	

Notes: Table shows the demographic and health characteristics of caller in each intervention arm (based on the unadjusted caller measure shown in Table 2). The Information Only arm pools the two equally-sized sub-treatments; the Information Plus Assistance pools the two sub-treatments and weights them so that they receive equal weight. All p-values are based on heteroskedasticity-robust standard errors. All demographic and health characteristics are the same as shown in Table 5.

[a] Total spending is truncated at twice 99.5th percentile of study population, which is 371,620 (99.5th percentile in study population is 185,810). Amounts greater than the threshold are set to missing.

[b] Omitted category is other or missing race.

Table A14: Demographic and Health Characteristics by Sub-Treatment: Applicants

	Control	Information Only				Information Plus Assistance		P Value of Difference between			
		Standard	No-Postcard	Marketing	Framing	Standard	Marketing	Control vs Treatment (col 1 vs 2+4+6+7)	Standard vs Marketing (col 2+6 vs 4+7)	Information Only Standard vs Framing (col 2 vs 5)	Information Only Standard vs No Follow-up Postcard (col 2 vs 3)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A - Predicted Benefits											
Predicted Benefits	148.26	128.66	133.83	122.46	136.87	116.61	114.13				
		[0.009]	[0.078]	[0.000]	[0.118]	[0.000]	[0.000]	[0.000]	[0.369]	[0.328]	[0.573]
Panel B - (Annual) Health Care Measures, 2015											
Total Health Care Spending ($)[b]	9,424	10,848	7,155	6,238	5,797	8,333	8,335				
		[0.405]	[0.124]	[0.012]	[0.004]	[0.312]	[0.375]	[0.341]	[0.044]	[0.002]	[0.044]
Total Number of Visits and Days	13.33	12.78	9.12	10.49	10.01	9.18	10.65				
		[0.793]	[0.023]	[0.149]	[0.073]	[0.003]	[0.112]	[0.054]	[0.994]	[0.189]	[0.081]
Weighted Total Number of Visits and Days	4,661	3,735	2,513	2,785	3,634	2,607	3,026				
		[0.381]	[0.037]	[0.070]	[0.367]	[0.012]	[0.068]	[0.036]	[0.836]	[0.929]	[0.233]
Number of Chronic Conditions	6.21	5.93	5.45	5.15	4.68	5.15	5.40				
		[0.562]	[0.130]	[0.024]	[0.000]	[0.002]	[0.054]	[0.011]	[0.612]	[0.012]	[0.388]
Panel C - Demographics											
Share Age 80+	0.06	0.10	0.12	0.13	0.11	0.13	0.14				
		[0.050]	[0.006]	[0.001]	[0.008]	[0.000]	[0.000]	[0.000]	[0.174]	[0.541]	[0.359]
Male	0.41	0.40	0.37	0.41	0.39	0.37	0.39				
		[0.903]	[0.299]	[0.928]	[0.628]	[0.070]	[0.642]	[0.436]	[0.364]	[0.757]	[0.417]
Share White[a]	0.67	0.73	0.70	0.73	0.66	0.74	0.74				
		[0.025]	[0.263]	[0.020]	[0.758]	[0.000]	[0.001]	[0.000]	[0.866]	[0.027]	[0.407]
Share Black[a]	0.10	0.09	0.07	0.07	0.12	0.11	0.11				
		[0.464]	[0.055]	[0.040]	[0.312]	[0.855]	[0.469]	[0.710]	[0.883]	[0.131]	[0.296]
Share Primary Language not English	0.08	0.06	0.06	0.07	0.06	0.04	0.03				
		[0.087]	[0.213]	[0.457]	[0.058]	[0.000]	[0.000]	[0.001]	[0.905]	[0.892]	[0.773]
Share Living in Pittsburgh	0.05	0.08	0.06	0.04	0.05	0.07	0.06				
		[0.093]	[0.603]	[0.612]	[0.920]	[0.022]	[0.294]	[0.105]	[0.062]	[0.156]	[0.340]
Share Last Medicaid Spell Starting before 2011	0.25	0.31	0.30	0.28	0.29	0.30	0.28				
		[0.020]	[0.095]	[0.181]	[0.093]	[0.005]	[0.127]	[0.008]	[0.254]	[0.552]	[0.667]
Observations (N)	817	401	320	380	417	1,883	636				

Notes: Table shows demographic and health characteristics of applicants (as shown in Table 5) separately for each sub-treatment. In column 8, sub-treatments are weighted so that within the Information Plus Assistance arm the standard and marketing sub-treatments receive equal weight, and the Information Plus Assistance treatments receive equal weight as the Information Only treatment. In column 9, sub-treatments are weighted so that Information Plus Assistance and Information Only are equally weighted in Standard and Marketing arms. All p-values are based on heteroskedasticity-robust standard errors. All demographic and health characteristics are the same as shown in Table 5.

[a]Omitted category is other or missing race.

[b]Total spending is truncated at twice 99.5th percentile of study population, which is 371,620 (99.5th percentile in study population is 185,810). Amounts greater than the threshold are set to missing.

Table A15: Demographic and Health Characteristics by Sub-Treatment: Enrollees

	Control	Information Only				Information Plus Assistance		P Value of Difference between			
		Standard	No-Postcard	Marketing	Framing	Standard	Marketing	Control vs Treatment (col 1 vs 2+4+6+7)	Standard vs Marketing (col 2+6 vs 4+7)	Information Only Standard vs Framing (col 2 vs 5)	Information Only Standard vs No Follow-up Postcard (col 2 vs 3)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A - Predicted Benefits											
Predicted Benefits	138.65	112.99	126.04	115.17	130.07	106.57	101.56				
		[0.001]	[0.166]	[0.007]	[0.302]	[0.000]	[0.000]	[0.000]	[0.581]	[0.067]	[0.193]
Panel B - (Annual) Health Care Measures, 2015											
Total Health Care Spending ($) [b]	10,238	11,938	7,391	6,785	5,394	8,058	9,131				
		[0.429]	[0.113]	[0.037]	[0.002]	[0.095]	[0.468]	[0.317]	[0.246]	[0.002]	[0.045]
Total Number of Visits and Days	14.79	12.48	10.08	9.10	9.52	8.84	10.96				
		[0.352]	[0.047]	[0.009]	[0.016]	[0.001]	[0.070]	[0.012]	[0.977]	[0.193]	[0.329]
Weighted Total Number of Visits and Days	5,407	4,122	2,936	2,335	3,821	2,687	2,868				
		[0.345]	[0.064]	[0.010]	[0.275]	[0.010]	[0.023]	[0.017]	[0.356]	[0.834]	[0.367]
Number of Chronic Conditions	6.54	5.60	5.78	5.23	4.55	5.13	5.60				
		[0.089]	[0.205]	[0.025]	[0.000]	[0.001]	[0.063]	[0.004]	[0.676]	[0.061]	[0.791]
Panel C - Demographics											
Share Age 80+	0.07	0.10	0.13	0.13	0.11	0.15	0.14				
		[0.083]	[0.006]	[0.008]	[0.076]	[0.000]	[0.000]	[0.000]	[0.585]	[0.966]	[0.276]
Male	0.39	0.41	0.35	0.42	0.43	0.37	0.38				
		[0.640]	[0.247]	[0.437]	[0.331]	[0.239]	[0.771]	[0.889]	[0.510]	[0.662]	[0.163]
Share White [a]	0.71	0.79	0.73	0.77	0.69	0.77	0.79				
		[0.007]	[0.504]	[0.038]	[0.688]	[0.005]	[0.001]	[0.000]	[0.638]	[0.009]	[0.118]
Share Black [a]	0.11	0.07	0.07	0.06	0.12	0.10	0.11				
		[0.086]	[0.063]	[0.009]	[0.735]	[0.695]	[0.978]	[0.216]	[0.922]	[0.085]	[0.842]
Share Primary Language not English	0.06	0.04	0.05	0.05	0.05	0.03	0.02				
		[0.232]	[0.442]	[0.464]	[0.357]	[0.010]	[0.001]	[0.012]	[0.545]	[0.823]	[0.769]
Share Living in Pittsburgh	0.05	0.08	0.05	0.03	0.06	0.08	0.07				
		[0.056]	[0.806]	[0.312]	[0.356]	[0.007]	[0.205]	[0.059]	[0.025]	[0.366]	[0.153]
Share Last Medicaid Spell Starting before 2011	0.26	0.34	0.31	0.33	0.32	0.32	0.30				
		[0.029]	[0.225]	[0.044]	[0.116]	[0.007]	[0.166]	[0.007]	[0.506]	[0.598]	[0.465]
Observations (N)	613	298	245	261	295	1,385	476				

Notes: Table shows demographic and health characteristics of enrollees (as shown in Table 5) separately for each sub-treatment. In column 8, sub-treatments are weighted so that within the Information Plus Assistance arm the standard and marketing sub-treatments receive equal weight, and the Information Plus Assistance treatments receive equal weight as the Information Only treatment. In column 9, sub-treatments are weighted so that Information Plus Assistance and Information Only are equally weighted in Standard and Marketing arms. All p-values are based on heteroskedasticity-robust standard errors. All demographic and health characteristics are the same as shown in Table 5.

[a] Omitted category is other or missing race.

[b] Total spending is truncated at twice 99.5th percentile of study population, which is 371,620 (99.5th percentile in study population is 185,810). Amounts greater than the threshold are set to missing.

Table A16: Behavioral Responses to Interventions: Robustness to covariates

	Control	Information Only	Information Plus Assistance	P Value of Difference (Column 2 vs 3)
	(1)	(2)	(3)	(4)
SNAP Enrollees	0.058	0.105	0.176	
		[0.000]	[0.000]	[0.000]
SNAP Applicants	0.077	0.147	0.238	
		[0.000]	[0.000]	[0.000]
SNAP Rejections among Applicants	0.233	0.266	0.255	
		[0.041]	[0.016]	[0.796]
Callers	0.000	0.267	0.301	
		[0.000]	[0.000]	[0.000]
Adjusted Callers	0.000	0.289	0.301	
		[0.000]	[0.000]	[0.133]
SNAP Applicants among Non-Callers	0.077	0.086	0.081	
		[0.058]	[0.256]	[0.397]
SNAP Applicants among Callers	0.000	0.313	0.602	
		[0.000]	[0.000]	[0.000]
SNAP Enrollees among Non-Callers	0.058	0.061	0.059	
		[0.380]	[0.592]	[0.679]
SNAP Enrollees among Callers	0.000	0.226	0.450	
		[0.000]	[0.000]	[0.000]
Observations (N)	10,630	5,314	10,629	

Notes: Table shows robustness of our main estimates of behavioral responses (see Table 2) to controlling for indicator variables for which of the 11 mail batches the individual was assigned to and for the baseline covariates shown in Table 5. As in Table 2, columns 1 through 3 shows means by intervention arm with the p-value (relative to the control arm) in [square brackets]. Column 1 shows the control. Column 2 shows the Information Only arm (for the same two equally-sized pooled sub-treatments). Column 3 shows the Information Plus Assistance arm (weighted so that the two pooled sub-treatments received equal weight). Column 4 reports the p-value of the difference between the Information Plus Assistance and Information Only treatment arms. All outcomes are binary rates measured during the nine months from the initial mail date. All p-values are based on heteroskedasticity-robust standard errors. Callers are measured for the relevant call number and are therefore mechanically zero for the control; see text for a description of the adjusted caller rate.

Table A17: Health Characteristics of Enrollees and Applicants: Robustness to restriction to full year of Medicaid

	Applicants				Enrollees			
	Means			P Value	Means			P Value
	Control	Info Only	Info Plus Assistance	Info Plus Assistance vs Info Only	Control	Info Only	Info Plus Assistance	Info Plus Assistance vs Info Only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(Annual) Health Care Measures, 2015								
Total Health Care Spending ($) [a]	10,304	9,313	9,366		10,932	10,585	9,684	
		[0.532]	[0.469]	[0.966]		[0.862]	[0.427]	[0.575]
Total Number of Visits and Days	11.06	10.41	10.05		11.74	10.24	10.14	
		[0.682]	[0.451]	[0.791]		[0.421]	[0.324]	[0.951]
Weighted Total Number of Visits and Days	3,673	2,506	2,966		4,182	2,771	2,909	
		[0.215]	[0.413]	[0.480]		[0.250]	[0.243]	[0.867]
Number of Chronic Conditions	6.54	5.90	5.78		6.88	5.61	5.80	
		[0.182]	[0.074]	[0.758]		[0.029]	[0.039]	[0.669]
Observations (N)	565	562	1,836		425	410	1,396	

Notes: Table shows robustness of our main estimates of the health characteristics of applicants and enrollees (see Table 5) to restricting to the approximately three-quarters of the sample who is enrolled in Medicaid for all of 2015. As in Table 5, columns 1 - 3 and 5 - 7 show means by intervention arm with the p-value (relative to the control arm) in [square brackets] for SNAP applicants who applied within 9 months of their initial mailing, and SNAP enrollees who enrolled within 9 months of their initial mailing, respectively. Column 1 and 5 show the control. Column 2 and 6 show the Information Only arm (for the same two equally-sized pooled sub-treatments). Columns 3 and 7 show the Information Plus Assistance arm (weighted so that the two pooled sub-treatments received equal weight). Columns 4 and 8 report the p-value of the difference between the Information Plus Assistance and Information Only treatment arms. All p-values are based on heteroskedasticity-robust standard errors.

Table A18: Behavioral Responses to Interventions: By Age

	Outcomes			No. of Observations
	SNAP Enrollees	SNAP Applicants	Callers	
	(1)	(2)	(3)	(4)
Control Mean	0.058	0.077	NA	10,630
Info Only (at Mean Age)	0.105	0.147	0.267	
	[0.000]	[0.000]	[0.000]	
				15,944
Info Only × Age	-0.00035	-0.00028	-0.00093	
	[0.449]	[0.589]	[0.149]	
Info Plus Assistance (at Mean Age)	0.177	0.238	0.300	
	[0.000]	[0.000]	[0.000]	
				21,259
Info Plus Assistance × Age	0.00034	-0.00001	-0.00087	
	[0.491]	[0.984]	[0.120]	

Notes: "Control Mean" row shows mean of control group. The next two rows show results from a regression on the Info Only and Control sample of the outcome in the column on a treatment indicator, age, and the interaction; "Info Only" (at Mean Age) shows the implied mean (and p-values in brackets relative to the control mean, or relative to 0 for callers) at the mean age of the study population (69); and Info Only x Age shows the coefficient on the interaction of treatment with age. The final two rows show results from an analogous analysis with the Info Plus Assistance treatment.

Table A19: Benefit Outreach Specialist (BOS) Characteristics

	Mean (SD)	Min	25 Percentile	Median	75 Perentile	95 Percentile	Max
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age	28.82 (6.919)	21	24	27	31	42	57
Experience (in Month)	8.57 (9.893)	0	2	4	15	22	64
Gender (Share of Male)	0.39 (0.492)	0	0	0	1	1	1
Number of Calls Received as the First Call of Some Caller	48.17 (38.031)	1	12	49	76	107	137
Observations (N)				66			

Notes: Table summarizes BOS characteristics. It is limited to BOS who received calls as the first call of some caller in our study population.

Table A20: Relationship between Benefit Outreach Specialist (BOS) Characteristics and Applications or Enrollment

	Applications	Enrollment
	(1)	(2)
Mean in Callers	0.5974	0.4461
Effects of BOS characteristics		
BOS Age	0.0004	-0.0016
	[0.756]	[0.174]
BOS Age above Median (=27)	0.0022	-0.0281
	[0.902]	[0.116]
BOS Experience (in Month)	0.0012	0.0001
	[0.386]	[0.955]
BOS Experience above Median (=4)	0.0313	0.0158
	[0.075]	[0.374]
BOS Gender (Male)	0.0026	0.0062
	[0.885]	[0.731]
BOS Male × Caller Male	0.0369	0.0521
	[0.316]	[0.162]
Observations (N)	3,179	3,179

Notes: Table shows coefficients and p-values in [square-brackets] from regressing the outcome indicator shown in the column heading on the characteristic of the BOS receiving the first call within 9 months of the initial mailing for the each caller in the Information Plus Assistance arm. Each row shows results from a different regression. In the last row, the regression includes separate indicator variables for whether the BOS is male, whether the caller is male, and the interaction of the two; it is the interaction effect that is shown.

E: Conceptual Framework - Details and Extensions

This section provides detailed proofs of the main propositions and additional extended results summarized in the main text.

E.1 More detail on Conceptual Framework in Section 2

This subsection goes through the model setup from the main text in more detail, and includes proofs of the Propositions in the main text.

The model features two types of individuals $j \in \{L, H\}$. Each type has an unobserved wage of θ_j, with $\theta_H > \theta_L$. This is the key source of heterogeneity in the model; it can be interpreted as heterogeneity in ability or labor market productivity, and it creates a potential motive for redistribution. We assume throughout that there is a unit mass of each type of individual.

Individuals choose hours of work h_j (which produces labor income $\theta_j h_j$) and whether or not to apply to the supplemental income program, which provides benefits B if income is below an earnings cutoff we denote by r^*. We allow each type to misperceive the benefit amount by ϵ_j, so that the *perceived* benefit from applying is $(1 + \epsilon_j)B$. We refer to the special case of no misperceptions – i.e., $\epsilon_j = 0$ for $j \in \{L, H\}$ – as the "neoclassical" benchmark case. With $\epsilon_j < 0$, misperceptions reduce the perceived benefit from applying.

There is a (potentially non-linear) income tax system $\tau(\theta_j h_j)$, which maps pre-tax labor earnings to taxes owed to the government. We denote net of tax earnings by $y_j \equiv \theta_j h_j - \tau(\theta_j h_j)$.

Individuals share a common utility function: $u(x_j) - v(h_j)$ if they don't apply and $u(x_j) - v(h_j) - (\bar{\Lambda}\kappa_j + c)$ if they apply. Individuals get utilty from consumption (x), disutility from hours worked (h_j), and disutility from applying ($\bar{\Lambda}\kappa_j + c$).

Disutility from applying can include the time and effort spent compiling documents, filling out forms, and participating in an interview, as well as any associated stigma. This disutility depends on three terms: c is an individual-specific utility cost of applying and is distributed according to a type-specific distribution $f_j(c)$, $\bar{\Lambda}$ is a parameter that affects the utility cost to applying that is common across individuals (and is under control of the social planner or researcher), and κ_j is how the utility cost varies with $\bar{\Lambda}$ for individuals of type j. This formulation nests ordeals that impose a greater utility cost on H types ($\kappa_H > \kappa_L$), or on L types ($\kappa_L > \kappa_H$). The former case includes utility costs $\kappa_j = \theta_j$, which might correspond to a common time cost that has higher utility costs for H types due to higher wages (see, e.g., Nichols and Zeckhauser 1982). The latter case includes the possibility that L types having greater difficulty filling out forms (see, e.g., Bertrand et al. 2004).

Individual choices and private welfare

Individuals make application and labor supply choices to maximize private utility, given their (possibly incorrect) perceptions. We denote the optimal hours choice for type j individuals who apply as h_j^A and the optimal hours choice for individuals who do not apply as $h_j^{\neg A}$. For low-ability

individuals, we assume that either hours choice would leave them with labor earnings below the income eligibility threshold r^* needed to qualify for the supplemental income program, with optimal solutions given by the following maximization problems:

$$\begin{aligned} h_L^A &= \arg\max_{h_L} u(h_L\theta_L - \tau(h_L\theta_L) + (1+\epsilon_L)B) - v(h_L) - (\bar{\Lambda}\kappa_L + c) \\ h_L^{\neg A} &= \arg\max_{h_L} u(h_L\theta_L - \tau(h_L\theta_L)) - v(h_L) \end{aligned} \tag{7}$$

For high-ability individuals, we assume that the hours choice if they do not apply puts their income above the eligibility threshold r^*. Therefore, if they do apply their hours choice is given by $h_H^A = r^*/\theta_H$, so that they are at the income threshold. Their hours choice when not applying is given by the same maximization problem as above; i.e,

$$h_H^{\neg A} = \arg\max_{h_H} u(h_H\theta_H - \tau(h_H\theta_H)) - v(h_H) \tag{8}$$

Given these optimal hours choices, expected utility from applying (A) is:

$$u(h_j^A\theta_j - \tau(h_j^A\theta_j) + (1+\epsilon_j)B) - v(h_j^A) - (\bar{\Lambda}\kappa_j + c)$$

and the utility of not applying ($\neg A$) is given by:

$$u(h_j^{\neg A}\theta_j - \tau(h_j^{\neg A}\theta_j)) - v(h_j^{\neg A})$$

Individuals of type j apply if the expected utility of applying is greater than the utility of not applying. We define c_j^* to be the threshold level of c such that a type j individual is indifferent between applying and not applying. For $c < c_j^*$, the individual chooses to apply. This threshold is defined as follows:

$$\begin{aligned} c_j^* &= u(h_j^A\theta_j - \tau(h_j^A\theta_j) + (1+\epsilon_j)B) - v(h_j^A) - \bar{\Lambda}\kappa_j \\ &\quad -u(h_j^{\neg A}\theta_j - \tau(h_j^{\neg A}\theta_j)) + v(h_j^{\neg A}) \end{aligned} \tag{9}$$

Note that an intervention that reduces transaction costs of applying (i.e., decreases $\bar{\Lambda}$) will increase c_j^* more for types with higher utility costs of applying (i.e., larger κ_j). This will lead to greater increases in applications if $f_j(c)$ is the same across types.

We use this application decision rule in equation (9) and integrate across the distribution of

private costs to get the total private welfare of type j individuals:

$$\begin{aligned} V_j &= Pr(apply) * E[u()|apply] + Pr(\neg apply) * E[u()|\neg apply] \\ &= \int_0^{c_j^*} [u(h_j^A \theta_j - \tau(h_j^A \theta_j) + B) - v(h_j^A) - (\bar{\Lambda}\kappa_j + c)] dF_j(c) \\ &+ \int_{c_j^*}^{\infty} [u(h_j^{\neg A} \theta_j - \tau(h_j^{\neg A} \theta_j)) - v(h_j^{\neg A})] dF_j(c) \end{aligned}$$

Note that ϵ_j affects the individual application decision but not realized utility, since the ϵ_j term only changes the perceived benefits, not the actual benefits.

Social welfare

Because of the fiscal externality on the government budget, privately optimal application decisions may not be socially optimal. We consider a redistributive social welfare function, which is natural given the redistributive purpose of the transfer program. Specifically, we consider a utilitarian social welfare function, although we could easily accommodate alternative individualistic social welfare functions at the cost of some additional notation.

The social welfare function W is therefore the sum of total private welfare of both types of individuals, minus public expenditures on benefits, plus net government revenue:

$$\begin{aligned} W \quad = \quad & \underbrace{V_L + V_H}_{\text{Private Welfare}} \\ & - \underbrace{[B(A_L + A_H)]}_{\text{Mechanical Program Costs}} \\ & + \underbrace{[A_L \tau(h_L^A \theta_L) + (1 - A_L)\tau(h_L^{\neg A} \theta_L) + A_H \tau(h_H^A \theta_H) + (1 - A_H)\tau(h_H^{\neg A} \theta_H)]}_{\text{Fiscal Externality from Program}} \end{aligned} \tag{10}$$

where A_j is the expected number of applications from type j individuals; this is equal to $A_j = F_j(c_j^*)$ from the take-up decision in equation (9). Note that in the main text we denoted net government revenue from applicants (i.e., tax revenue $\tau(h_j^A \theta_j)$) by G_j^A and equivalently for non applicants. Note also that rather than add net government revenue to the social welfare function, we could instead "close" the government budget by having net government revenue and program expenditures "paid for" out of individual consumption. Our approach assumes that the costs of the government budget are born by someone with the average marginal utility of consumption in society; implicitly, our W expression in equation (10) is thus a "money metric" social welfare expression, normalized by the average marginal utility of consumption in the population

The social planner chooses the income tax system and the income transfer program (including the "ordeal" parameter $\bar{\Lambda}$) to maximize social welfare. As has been shown (see e.g. Currie and

Gahvari 2008), if $\kappa_H > \kappa_L$, the social optimum will involve a non-zero ordeal utility cost ($\bar{\Lambda} > 0$) even in the presence of an arbitrary optimal nonlinear income tax. Intuitively, with unobserved ability θ_j and endogenous hours choices, the government is not able to achieve the first best desired amount of redistribution (equal consumption across types); redistribution to low ability types is limited by the binding incentive compatibility constraint that high ability types not want to "mimic" the hours choice of low ability types. Adding ordeals that are more costly for the high ability types (i.e. $\kappa_H > \kappa_L$) can relax the incentive compatibility constraint on the H type and thus allow for more redistribution. Our goal, however, is not to characterize the globally optimal system of taxes, transfers, and ordeals, but rather to characterize the *marginal* social welfare gain (or loss) from interventions that may affect information about eligibility and the private cost of application.

Welfare Effects of Interventions

Proposition 1. *Let $\mu_j \equiv u(h_j^A\theta_j - \tau(h_j^A\theta_j) + B) - u(h_j^A\theta_j - \tau(h_j^A\theta_j) + (1+\epsilon_j)B)$, which equals the difference between the actual and perceived utility when applying. The effect of the Information Only treatment on welfare is given by the following expression:*

$$\begin{aligned}\frac{dW}{dT}^{Info\,Only} = & \underbrace{\mu_L\frac{dA_L}{dT} + \mu_H\frac{dA_H}{dT}}_{\textit{Change in Private Welfare}} \\ & - \underbrace{\left[B(\frac{dA_L}{dT} + \frac{dA_H}{dT})\right]}_{\textit{Change in Public Expenditure on Benefits}} \\ & + \underbrace{\left[(\tau(\theta_L h_L^A) - \tau(\theta_L h_L^{\neg A}))\frac{dA_L}{dT} + (\tau(\theta_H h_H^A) - \tau(\theta_H h_H^{\neg A}))\frac{dA_H}{dT}\right]}_{\textit{Change in Net Government Revenue}}\end{aligned}$$

and the effect of the Information Plus Assistance treatment on welfare is given by the following expression:

$$\begin{aligned}\frac{dW}{dT}^{Info+Assistance} = & \underbrace{\mu_l\frac{dA_l}{dT} + \mu_h\frac{dA_h}{dT} + \kappa_L A_L + \kappa_H A_H}_{\text{Change in Private Welfare}} \\ & - \underbrace{\left[B(\frac{dA_L}{dT} + \frac{dA_H}{dT})\right]}_{\text{Change in Program Costs}} \\ & + \underbrace{\left[(\tau(\theta_L h_L^A) - \tau(\theta_L h_L^{\neg A}))\frac{dA_L}{dT} + (\tau(\theta_H h_H^A) - \tau(\theta_H h_H^{\neg A}))\frac{dA_H}{dT}\right]}_{\text{Change in Net Government Revenue}}\end{aligned}$$

Proof: Welfare is given by

$$
\begin{aligned}
W =& V_L + V_H \\
& - [B(A_L + A_H)] \\
& + \Big[A_L \tau(h_L^A \theta_L) + (1 - A_L)\tau(h_L^{\neg A}\theta_L) + A_H \tau(h_H^A \theta_H) + (1 - A_H)\tau(h_H^{\neg A}\theta_H)\Big] \\
=& u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) - v(h_L^{\neg A}) \\
& + \int_0^{c_L^*} \Big[(u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - v(h_L^A) - (\bar{\Lambda}\kappa_L + c)) - (u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) - v(h_L^{\neg A}))\Big] dF_L(c) \\
& + u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A}) \\
& + \int_0^{c_H^*} \Big[(u(h_H^A\theta_H - \tau(h_H^A\theta_H) + B) - v(h_H^A) - (\bar{\Lambda}\kappa_H + c)) - (u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A}))\Big] dF_H(c) \\
& - [B(A_L + A_H)] \\
& + \Big[A_L \tau(h_L^A \theta_L) + (1 - A_L)\tau(h_L^{\neg A}\theta_L) + A_H \tau(h_H^A \theta_H) + (1 - A_H)\tau(h_H^{\neg A}\theta_H)\Big],
\end{aligned}
$$

where $c_j^* = u(h_j^A\theta_j - \tau(h_j^A\theta_j) + (1+\epsilon_j)B) - v(h_j^A) - \bar{\Lambda}\kappa_j - (u(h_j^{\neg A}\theta_j - \tau(h_j^{\neg A}\theta_j)) - v(h_j^{\neg A}))$.

Information Only ($dT = d\epsilon$): Taking the derivative with respect to to ϵ yields

$$
\begin{aligned}
\frac{dW}{d\epsilon} =& \frac{d}{d\epsilon}\int_0^{c_L^*} \Big[u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - v(h_L^A) - (\bar{\Lambda}\kappa_L + c) - u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) + v(h_L^{\neg A})\Big] f_L(c)dc \\
& + \frac{d}{d\epsilon}\int_0^{c_H^*} \Big[u(h_H^A\theta_H - \tau(h_H^A\theta_H) + B) - v(h_H^A) - (\bar{\Lambda}\kappa_H + c) - u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) + v(h_H^{\neg A})\Big] f_H(c)dc \\
& - \Big[B(\frac{dA_L}{d\epsilon} + \frac{dA_H}{d\epsilon})\Big] \\
& + \Big[(\tau(h_L^A\theta_L) - \tau(h_L^{\neg A}\theta_L))\frac{dA_L}{d\epsilon} + (\tau(h_H^A\theta_H) - \tau(h_H^{\neg A}\theta_H))\frac{dA_H}{d\epsilon}\Big]
\end{aligned}
$$

Applying Leibniz's Rule, we get

$$
\begin{aligned}
& \frac{d}{d\epsilon}\int_0^{c_L^*} \Big[(u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - v(h_L^A) - (\bar{\Lambda}\kappa_L + c)) - u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) + v(h_L^{\neg A})\Big] f_L(c)dc \\
=& \Big[(u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - v(h_L^A) - (\bar{\Lambda}\kappa_L + c_L^*)) - (u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) - v(h_L^{\neg A}))\Big] f_L(c_L^*)\frac{dc_L^*}{d\epsilon} \\
=& (u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - u(h_L^A\theta_L - \tau(h_L^A\theta_L) + (1+\epsilon_L)B))f_L(c_L^*)\frac{dc_L^*}{d\epsilon}
\end{aligned}
$$

Similarly,

$$\frac{d}{d\epsilon}\int_0^{c_H^*}\left[(u(h_H^A\theta_H-\tau(h_H^A\theta_H)+B)-v(h_H^A)-(\bar{\Lambda}\kappa_H+c))-u(h_H^{\neg A}\theta_H-\tau(h_H^{\neg A}\theta_H))+v(h_H^{\neg A})\right]f_H(c)dc$$
$$=(u(h_H^A\theta_H-\tau(h_H^A\theta_H)+B)-u(h_H^A\theta_H-\tau(h_H^A\theta_H)+(1+\epsilon_H)B))f_H(c_H^*)\frac{dc_H^*}{d\epsilon}$$

Since the number of applicants is given by $A_H=F_H(c_H^*)$ and $A_L=F_L(c_L^*)$,

$$\frac{dA_H}{d\epsilon}=f_H(c_H^*)\frac{dc_H^*}{d\epsilon}$$

and

$$\frac{dA_L}{d\epsilon}=f_L(c_L^*)\frac{dc_L^*}{d\epsilon}$$

Therefore, we can re-write

$$\frac{d}{d\epsilon}\int_0^{c_L^*}\left[(u(h_L^A\theta_L-\tau(h_L^A\theta_L)+B)-v(h_L^A)-(\bar{\Lambda}\kappa_L+c))-(u(h_L^{\neg A}\theta_L-\tau(h_L^{\neg A}\theta_L))-v(h_L^{\neg A}))\right]f_L(c)dc$$
$$=(u(h_L^A\theta_L-\tau(h_L^A\theta_L)+B)-u(h_L^A\theta_L-\tau(h_L^A\theta_L)+(1+\epsilon_L)B))\frac{dA_L}{d\epsilon}$$

and

$$\frac{d}{d\epsilon}\int_0^{c_H^*}\left[(u(h_H^A\theta_H-\tau(h_H^A\theta_H)+B)-v(h_H^A)-(\bar{\Lambda}\kappa_H+c))-(u(h_H^{\neg A}\theta_H-\tau(h_H^{\neg A}\theta_H))-v(h_H^{\neg A})))\right]f_H(c)dc$$
$$=(u(h_H^A\theta_H-\tau(h_H^A\theta_H)+B)-u(h_H^A\theta_H-\tau(h_H^A\theta_H)+(1+\epsilon_H)B))\frac{dA_H}{d\epsilon}$$

Putting all this together (and noting that $dT=d\epsilon$), we have

$$\begin{aligned}\frac{dW}{dT}=&\mu_L\frac{dA_L}{dT}+\mu_H\frac{dA_H}{dT}\\&-\left[B(\frac{dA_L}{dT}+\frac{dA_H}{dT})\right]\\&+\left[(\tau(h_L^A\theta_L)-\tau(h_L^{\neg A}\theta_L))\frac{dA_L}{dT}+(\tau(h_H^A\theta_H)-\tau(h_H^{\neg A}\theta_H))\frac{dA_H}{dT}\right].\end{aligned}$$

This completes the first part of the proposition.

Assistance Only ($dT=-d\bar{\Lambda}$): Recall that the total private cost of applying for type j individual is $\bar{\Lambda}\kappa_j+c$. Thus, a change $-d\bar{\Lambda}$ is a downward shift in every applicant's total private cost of applying.

As in previous derivation, we can differentiate social welfare, W, with respect to $\bar{\Lambda}$:

$$\begin{aligned}\frac{dW}{d\bar{\Lambda}} =& \frac{d}{d\bar{\Lambda}} \int_0^{c_L^*} \left[u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - v(h_L^A) - (\bar{\Lambda}\kappa_L + c) - u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) + v(h_L^{\neg A})\right] f_L(c)dc \\ &+ \frac{d}{d\bar{\Lambda}} \int_0^{c_H^*} \left[u(h_H^A\theta_H - \tau(h_H^A\theta_H) + B) - v(h_H^A) - (\bar{\Lambda}\kappa_H + c) - u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) + v(h_H^{\neg A})\right] f_H(c)dc \\ &- \left[B(\frac{dA_L}{d\bar{\Lambda}} + \frac{dA_H}{d\bar{\Lambda}})\right] \\ &+ \left[(\tau(h_L^A\theta_L) - \tau(h_L^{\neg A}\theta_L))\frac{dA_L}{d\bar{\Lambda}} + (\tau(h_H^A\theta_H) - \tau(h_H^{\neg A}\theta_H))\frac{dA_H}{d\bar{\Lambda}}\right]\end{aligned}$$

Applying Leibniz's Rule, we get

$$\begin{aligned}&\frac{d}{d\bar{\Lambda}} \int_0^{c_L^*} \left[(u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - v(h_L^A) - (\bar{\Lambda}\kappa_L + c)) - u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) + v(h_L^{\neg A})\right] f_L(c)dc \\ =& \left[(u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - v(h_L^A) - (\bar{\Lambda}\kappa_L + c_L^*)) - (u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) - v(h_L^{\neg A}))\right] f_L(c_L^*)\frac{dc_L^*}{d\bar{\Lambda}} \\ &+ \int_0^{c_L^*} [-(\kappa_L)] f_L(c)dc \\ =&(u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - u(h_L^A\theta_L - \tau(h_L^A\theta_L) + (1+\epsilon_L)B))f_L(c_L^*)\frac{dc_L^*}{d\epsilon} - \kappa_L A_L\end{aligned}$$

Similarly,

$$\begin{aligned}&\frac{d}{d\bar{\Lambda}} \int_0^{c_H^*} \left[(u(h_H^A\theta_H - \tau(h_H^A\theta_H) + B) - v(h_H^A) - (\bar{\Lambda}\kappa_H + c)) - u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) + v(h_H^{\neg A})\right] f_H(c)dc \\ =&(u(h_H^A\theta_H - \tau(h_H^A\theta_H) + B) - u(h_H^A\theta_H - \tau(h_H^A\theta_H) + (1+\epsilon_H)B))f_H(c_H^*)\frac{dc_H^*}{d\bar{\Lambda}} - \kappa_H A_H\end{aligned}$$

Note that the two above expressions are similar to the Information Only change in private welfare for each type with additional term $(\kappa_j A_j)$ that reflects the change in welfare for infra-

marginal applicants.Noting that $\frac{dA_j}{d\bar{\Lambda}} = f_j(c_j^*)\frac{dc_j^*}{d\bar{\Lambda}}$, and putting this together implies

$$\begin{aligned}
\frac{dW}{dT} &= -\frac{dW}{d\bar{\Lambda}} \\
&= \kappa_L F_L(c_L^*) + \mu_L f_L(c_L^*) + (\tau(h_L^A\theta_L) - \tau(h_L^{\neg A}\theta_L) - B)f_L(c_L^*) \\
&+ \kappa_H F_H(c_H^*) + \mu_H f_H(c_H^*) + (\tau(h_H^A\theta_H) - \tau(h_H^{\neg A}\theta_H) - B)f_H(c_H^*) \\
&= \mu_L\frac{dA_L}{dT} + \kappa_L A_L + \mu_H\frac{dA_H}{dT} + \kappa_H A_H - B(\frac{dA_L}{dT} + \frac{dA_H}{dT}) \\
&+ (\tau(h_L^A\theta_L) - \tau(h_L^{\neg A}\theta_L))\frac{dA_L}{dT} + (\tau(h_H^A\theta_H) - \tau(h_H^{\neg A}\theta_H))\frac{dA_H}{dT}.
\end{aligned}$$

Information + Assistance ($dT = d\epsilon, -d\bar{\Lambda}$):

Combining information and assistance yields

$$\begin{aligned}
\frac{dW}{dT} &= \mu_L\frac{dA_L}{dT} + A_L + \mu_H\frac{dA_H}{dT} + A_H - B(\frac{dA_L}{dT} + \frac{dA_H}{dT}) \\
&+ (\tau(h_L^A\theta_L) - \tau(h_L^{\neg A}\theta_L))\frac{dA_L}{dT} + (\tau(h_H^A\theta_H) - \tau(h_H^{\neg A}\theta_H))\frac{dA_H}{dT}
\end{aligned}$$

where $\frac{dA_j}{dT}$ is the change in the number of applications from both de and $-dc$. This is the second part of the proposition, and completes the proof.

Relationship Between Targeting Impacts and Changes in Welfare

Proposition 2. *Holding constant the change in applications due to an intervention, the change in social welfare in response to an improvement in targeting ($de/dT > 0$) from an Information Only (or Information Plus Assistance) treatment is given by the following expression:*

$$\frac{\partial}{\partial(de/dT)}\left(\frac{dW}{dT}\right)\Big|_{\frac{dA}{dT}} = \left[(\mu_L - \mu_H) + (\tau(\theta_L h_L^A) - \tau(\theta_L h_L^{\neg A})) - (\tau(\theta_H h_H^A) - \tau(\theta_H h_H^{\neg A}))\right](E_L + E_H).$$

Proof: Note that $\mu_j \equiv u(h_j^A\theta_j - \tau(h_j^A\theta_j) + B) - u(h_j^A\theta_j - \tau(h_j^A\theta_j) + (1+\epsilon_j)B)$, as defined above. We assume applications are all accepted, so $E_j = A_j$. We extend model to allow for uncertainty in application process and prove analogous propositions below (see E.3).

Since $e = \frac{E_L}{E_H+E_L} = \frac{E_L}{A_L+A_H} = \frac{E_L}{A}$, can solve for de/dT as follows:

$$\begin{aligned}
\frac{de}{dT} &= \frac{d\frac{E_L}{A}}{dT} \\
&= \frac{1}{A}\cdot\frac{dE_L}{dT} - \frac{dA}{dT}\cdot\frac{E_L}{A^2}
\end{aligned}$$

From Proposition 1, we know that the change in welfare from Information Only is the following:

$$
\begin{aligned}
\frac{dW}{dT} &= \mu_L \frac{dA_L}{dT} + \mu_L \frac{dA_L}{dT} - \left[B(\frac{dA_L}{dT} + \frac{dA_H}{dT}) \right] \\
&+ \left[(\tau(\theta_L h_L^A) - \tau(\theta_L h_L^{\neg A})) \frac{dA_L}{dT} + (\tau(\theta_H h_H^A) - \tau(\theta_H h_H^{\neg A})) \frac{dA_H}{dT} \right] \\
&= \left[\mu_L - B + \tau(\theta_L h_L^A) - \tau(\theta_L h_L^{\neg A}) \right] (\frac{dA_L}{dT} + \frac{dA_H}{dT}) \\
&+ \left[(\mu_H - B + \tau(\theta_H h_H^A) - \tau(\theta_H h_H^{\neg A})) - (\mu_L - B + \tau(\theta_L h_L^A) - \tau(\theta_L h_L^{\neg A})) \right] \frac{dA_H}{dT} \\
&= \left[\mu_L - B + \tau(\theta_L h_L^A) - \tau(\theta_L h_L^{\neg A}) \right] \frac{dA}{dT} \\
&+ \left[(\mu_H - B + \tau(\theta_H h_H^A) - \tau(\theta_H h_H^{\neg A})) - (\mu_L - B + \tau(\theta_L h_L^A) - \tau(\theta_L h_L^{\neg A})) \right] \frac{dA_H}{dT}.
\end{aligned}
$$

Now, since dA/dT is held constant, and $\frac{dh_L^A}{dT}$ only depends on the private utility function, we just need to determine how the right-hand term varies with de/dT. We begin with the following derivative given model assumptions :

$$
-\frac{dA_H}{dT} = -\frac{dE_H}{dT} = \frac{dE_L}{dT} - \frac{dA}{dT}
$$

Thus, we can solve for $-\frac{dA_H}{dT}$ in terms of de/dT:

$$
\begin{aligned}
-\frac{dA_H}{dT} + \frac{dA}{dT} &= A \cdot \frac{de}{dT} + \frac{dA}{dT} \cdot \frac{E_L}{A} \\
&= A \cdot \frac{de}{dT} + e \frac{dA}{dT} \\
-\frac{dA_H}{dT} &= A \cdot \frac{de}{dT} + (e-1) \frac{dA}{dT}
\end{aligned}
$$

which can then be substituted back into the dW/dT expression above. Then, taking the partial derivative with respect to de/dT gives the expression in Proposition 2:

$$
\begin{aligned}
&\frac{\partial}{\partial(de/dT)} \left(\frac{dW}{dT} \right) \Big|_{\frac{dA}{dT}} \\
=& \frac{\partial}{\partial(de/dT)} \left[(\mu_L - \mu_H) + (\tau(\theta_L h_L^A) - \tau(\theta_L h_L^{\neg A})) - (\tau(\theta_H h_H^A) - \tau(\theta_H h_H^{\neg A}) \right] (A \cdot \frac{de}{dT} + (e-1) \frac{dA}{dT}) \\
=& \left[(\mu_L - \mu_H) + (\tau(\theta_L h_L^A) - \tau(\theta_L h_L^{\neg A})) - (\tau(\theta_H h_H^A) - \tau(\theta_H h_H^{\neg A}) \right] \frac{\partial}{\partial(de/dT)} (A \cdot \frac{de}{dT} + (e-1) \frac{dA}{dT}) \\
=& \left[(\mu_L - \mu_H) + (\tau(\theta_L h_L^A) - \tau(\theta_L h_L^{\neg A})) - (\tau(\theta_H h_H^A) - \tau(\theta_H h_H^{\neg A}) \right] \cdot A \\
=& \left[(\mu_L - \mu_H) + (\tau(\theta_L h_L^A) - \tau(\theta_L h_L^{\neg A})) - (\tau(\theta_H h_H^A) - \tau(\theta_H h_H^{\neg A}) \right] (E_L + E_H)
\end{aligned}
$$

E.1.1 Welfare effects from infra-marginal applicants

As described in the main text, Proposition 1 abstracts from income effects on infra-marginal applicants. These income effects can affect social welfare through an additional fiscal externality and –

if there are pre-existing misperceptions – then there will also be effects of the interventions on the private welfare of infra-marginal applicants. Intuitively, for individuals already applying, changes in their beliefs can cause them to change their labor supply, and this can affect both private welfare and also generate an additional fiscal externality from the labor supply response. Proposition 1 abstracts from these income effects on infra-marginal applicants. We ignore these additional terms in the main text because they scale with the magnitude of the change in beliefs from intervention, the magnitude of the misperceptions, and the magnitude of the income effect in labor supply. Unless these terms are large, this term is likely to be small.

However, for completeness we go through these additional terms in this sub-section, focusing on the Information Only intervention.

Information Only ($dT = d\epsilon$): In our baseline setting, the private welfare of type j applicants is given by

$$V_{j,\,apply} = \int_0^{c_j^*} \Big[u(h_j^A\theta_j - \tau(h_j^A\theta_j) + B) - v(h_j^A) - (\bar{\Lambda}\kappa_j + c)\Big]\, dF_j(c)$$

Since the welfare gain for infra-marginal applicants is through the adjustment of their labor supply, type H infra-marginal applicants will not experience this gain since their labor supply is fixed by $h_H^A = \frac{r*}{\theta_H}$. For type L infra-marginal applicants, however, the change in private welfare of infra-marginal applicants can be derived as follows:

$$\begin{aligned}
\frac{dV_{L,\,apply}}{d\epsilon} &= \int_0^{c_L^*} \frac{d}{d\epsilon}\Big[u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - v(h_L^A) - (\bar{\Lambda}\kappa_L + c)\Big]\, dF_L(c) \\
&= \frac{d}{d\epsilon}\Big[u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - v(h_L^A)\Big] \int_0^{c_L^*} f_L(c)dc \\
&= \frac{d}{d\epsilon}\Big[u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - v(h_L^A)\Big]\, A_L
\end{aligned}$$

From line 2 to line 3, the marginal effect is taken out of the integral, because it does not depend on cost as discussed above. Notice that the first-order condition applicants use to make choices remains the same, so $d\epsilon$ does not have first-order effect on $u(h_L^A\theta_L - \tau(h_L^A\theta_L) + (1+\epsilon_L)B) - v(h_L^A) - (\bar{\Lambda}\kappa_L + c)$. But this is the *perceived* utility of infra-marginal applicants, not the actual utility, and we can take a first-order approximation of the difference between this level of utility and the (actual) utility level in the private welfare expression. The change of this difference is approximately equal to the change of the private welfare.

$$\begin{aligned}
\triangle u^A &= u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - v(h_L^A) - (\bar{\Lambda}\kappa_L + c) - \Big[u(h_L^A\theta_L - \tau(h_L^A\theta_L) + (1+\epsilon_L)B) - v(h_L^A) - (\bar{\Lambda}\kappa_L + c)\Big] \\
&\approx -u'(h_L^A\theta_L - \tau(h_L^A\theta_L) + B)\epsilon_L B
\end{aligned}$$

This expression shows that when there is no misperception ($\epsilon_j = 0$), this term is zero, and there is no first-order effect on private welfare of infra-marginal applicants. If individuals under-estimate benefits from applying, however, then individuals gain from Information Only treatment, and this benefit scales with $d\epsilon$, ϵ, and the marginal utility of consumption.

The second term coming from infra-marginal applicants is the additional fiscal externality. This is a straightforward term to characterize and it does not depend on whether or not there are pre-existing misperceptions. Any change in labor earnings from infra-marginal applicants' changing labor supply from Information Only intervention will (in turn) affect government revenue. If the change in hours for infra-marginal type L applicants from Information Only treatment is $\frac{dh_L^A}{dT}$, then the additional fiscal externality is given by $\tau'(\theta_L h_L^A)\theta_L A_L \frac{dh_L^A}{dT}$. This term will be small if either the change in beliefs is small, or the income effect in labor supply is small.

In the remainder of the Appendix we will exclude this additional fiscal externality term as well as the change in private welfare for infra-marginal applicants from all extensions of the baseline model. These terms will be similar in the model extensions to the terms in this sub-section.

E.2 Extensions

This section goes through several extensions of the main model presented in the main text. These extensions cover non-marginal changes in beliefs and alternative ways of modeling "mistakes" – specifically, misperceptions of the private costs of applying and modeling inattention instead of misperceptions in the benefits from applying. We also refer readers to the NBER Working Paper version of this paper (#24652, www.nber.org/papers/w24652) for a number of additional model extensions, including allowing for heterogeneity in beliefs and modelling the information interventions as reducing the variance in beliefs.

In each of these extensions, we focus on deriving results that are analogous to Proposition 1 on the effect of the interventions on social welfare, and we focus on the Information Only treatment, since the derivations for Information Plus Assistance are often very similar and follow many of the same steps as in the proof to Proposition 1 above.

E.2.1 Non-Marginal Changes in Beliefs (Away From Envelope Theorem)

Our main analysis focuses on the marginal welfare gain (or loss) from "small" interventions, allowing us to make heavy use of the envelope theorem. We explore here an extension where we consider non-marginal changes, similar in spirit to Kleven (2018) who extends sufficient statistics analysis to discrete policy changes. We describe how this addition to the "standard" setting (no mistakes and the only fiscal externality occurs via labor supply) can also undo the "standard" relationship between changes in targeting and changes in welfare.

With non-marginal changes, the envelope theorem will no longer apply, and so changes in private welfare can be meaningful, even in the neoclassical benchmark case with accurate beliefs. As a result, we will still have the same lack of a general relationship between targeting properties

of the intervention and changes in welfare. This is illustrated in the following result, which focuses on the case of accurate beliefs for simplicity:

Proposition 3. *Let c'_j be the threshold application cost after the non-marginal change. Define $h_L^{\triangle\epsilon} = \arg\max_{h_L} u(h_L\theta_L - \tau(h_L\theta_L) + (1+\epsilon_L+\triangle\epsilon)B) - v(h_L) - (\bar{\Lambda}\kappa_L + c))$, and define $c'_L = u(h_L^{\triangle\epsilon}\theta_L - \tau(h_L^{\triangle\epsilon}\theta_L) + (1+\triangle\epsilon)B) - v(h_L^{\triangle\epsilon}) - \bar{\Lambda}\kappa_L - u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) + v(h_L^{\neg A})$ and $c'_H = u(h_H^A\theta_H - \tau(h_H^A\theta_H) + (1+\triangle\epsilon)B) - v(h_H^A) - \bar{\Lambda}\kappa_H - u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) + v(h_H^{\neg A})$.*

If $\epsilon_L = \epsilon_H = 0$, then the effect of a non-marginal Information Only treatment ($\Delta T = \Delta\epsilon$) on welfare is given by:

$$\begin{aligned}
\Delta W = &\int_{c_L^*}^{c'_L} \left[(u(h_L^{\triangle\epsilon}\theta_L - \tau(h_L^{\triangle\epsilon}\theta_L) + B) - v(h_L^{\triangle\epsilon}) - (\bar{\Lambda}\kappa_L + c)) - u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) + v(h_L^{\neg A})\right] f_L(c)dc \\
&+ \int_{c_H^*}^{c'_H} \left[(u(h_H^A\theta_H - \tau(h_H^A\theta_H) + B) - v(h_H^A) - (\bar{\Lambda}\kappa_H + c)) - u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) + v(h_H^{\neg A})\right] f_H(c)dc \\
&- \left[(B - \tau(h_L^{\triangle\epsilon}\theta_L) + \tau(h_L^{\neg A}\theta_L))\Delta A_L + (B - \tau(h_H^A\theta_H) + \tau(h_H^{\neg A}\theta_H))\Delta A_H\right]
\end{aligned}$$

In the expression above, $c_j^* = u(h_j^A\theta_j - \tau(h_j^A\theta_j) + B) - v(h_j^A) - \bar{\Lambda}\kappa_j - u(h_j^{\neg A}\theta_j - \tau(h_j^{\neg A}\theta_j)) + v(h_j^{\neg A})$ is the threshold cost prior to the non-marginal change in beliefs. Type H applicants do not adjust their labor supply, which is given by $h_H^A = \frac{r^*}{\theta_H}$; thus the expression above already substitutes h_H^A for $h_H^{\triangle\epsilon}$.

This expression shows that the change in social welfare can be decomposed into three parts. The first line is the change of private welfare from new type L applicants; the second line is the change in private welfare from new type H applicants. The third line represents the change of program cost and government expenditure brought by new applicants of each type.

To see why these terms are not obviously signed in the case of an intervention that increases targeting, note that in the non-marginal case the increase in private welfare from a given change in applicants depends on the shape of the type-specific cost distribution $f_j(c)$. If most of the individuals induced to apply were close to indifferent before the non-marginal change in costs, then a non-marginal change in costs can have a non-marginal change in private welfare; however, if most of the individuals induced to apply were close to indifferent to applying *after* the non-marginal change, then the non-marginal change in costs will have a negligible effect on their utility, since they are close to indifferent to applying after the intervention. Thus, the cost distribution functions – much as the misperception terms did away from the neoclassic benchmark – provide another factor that potentially breaks the relationship between improvements in targeting and changes in social welfare.

Proof:

Information Only ($\Delta T = \Delta\epsilon$): Let c'_j be the threshold in cost after the non-marginal change. Let $h_L^{\triangle\epsilon} = \arg\max_{h_L} u(h_L\theta_L - \tau(h_L\theta_L) + (1+\triangle\epsilon)B) - v(h_L) - (\bar{\Lambda}\kappa_L + c))$, and define $c'_L = u(h_L^{\triangle\epsilon}\theta_L - \tau(h_L^{\triangle\epsilon}\theta_L) + (1+\triangle\epsilon)B) - v(h_L^{\triangle\epsilon}) - \bar{\Lambda}\kappa_L - u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) + v(h_L^{\neg A})$ and $c'_H = u(h_H^{A}\theta_H - \tau(h_H^{A}\theta_H) + (1+\triangle\epsilon)B) - v(h_H^{A}) - \bar{\Lambda}\kappa_H - u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) + v(h_H^{\neg A})$.

If $\epsilon_L = \epsilon_H = 0$, then the effect of the non-marginal Information Only treatment ($\Delta T = \Delta\epsilon$) can be derived as follows:

$$\begin{aligned}
\Delta W = &\int_{c_L^*}^{c'_L} \left[(u(h_L^{\triangle\epsilon}\theta_L - \tau(h_L^{\triangle\epsilon}\theta_L) + B) - v(h_L^{\triangle\epsilon}) - (\bar{\Lambda}\kappa_L + c)) - u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) + v(h_L^{\neg A})\right] f_L(c)dc \\
&+ \int_{c_H^*}^{c'_H} \left[(u(h_H^{A}\theta_H - \tau(h_H^{A}\theta_H) + B) - v(h_H^{A}) - (\bar{\Lambda}\kappa_H + c)) - u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) + v(h_H^{\neg A})\right] f_H(c)dc \\
&- \left[\int_{c_L^*}^{c'_L} (B - \tau(h_L^{\triangle\epsilon}\theta_L) + \tau(h_L^{\neg A}\theta_L)) f_L(c)dc + \int_{c_H^*}^{c'_H} (B - \tau(h_H^{\triangle\epsilon}\theta_H) + \tau(h_H^{\neg A}\theta_H)) f_H(c)dc\right]
\end{aligned}$$

$$\begin{aligned}
\Delta W = &\int_{c_L^*}^{c'_L} \left[(u(h_L^{\triangle\epsilon}\theta_L - \tau(h_L^{\triangle\epsilon}\theta_L) + B) - v(h_L^{\triangle\epsilon}) - (\bar{\Lambda}\kappa_L + c)) - u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) + v(h_L^{\neg A})\right] f_L(c)dc \\
&+ \int_{c_H^*}^{c'_H} \left[(u(h_H^{A}\theta_H - \tau(h_H^{A}\theta_H) + B) - v(h_H^{A}) - (\bar{\Lambda}\kappa_H + c)) - u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) + v(h_H^{\neg A})\right] f_H(c)dc \\
&- \left[(B - \tau(h_L^{\triangle\epsilon}\theta_L) + \tau(h_L^{\neg A}\theta_L)) \int_{c_L^*}^{c'_L} f_L(c)dc + (B - \tau(h_H^{A}\theta_H) + \tau(h_H^{\neg A}\theta_H)) \int_{c_H^*}^{c'_H} f_H(c)dc\right] \\
&= \int_{c_L^*}^{c'_L} \left[(u(h_L^{\triangle\epsilon}\theta_L - \tau(h_L^{\triangle\epsilon}\theta_L) + B) - v(h_L^{\triangle\epsilon}) - (\bar{\Lambda}\kappa_L + c)) - u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) + v(h_L^{\neg A})\right] f_L(c)dc \\
&+ \int_{c_H^*}^{c'_H} \left[(u(h_H^{A}\theta_H - \tau(h_H^{A}\theta_H) + B) - v(h_H^{A}) - (\bar{\Lambda}\kappa_H + c)) - u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) + v(h_H^{\neg A})\right] f_H(c)dc \\
&- \left[(B - \tau(h_L^{\triangle\epsilon}\theta_L) + \tau(h_L^{\neg A}\theta_L))\Delta A_L + (B - \tau(h_H^{A}\theta_H) + \tau(h_H^{\neg A}\theta_H))\Delta A_H\right]
\end{aligned}$$

E.2.2 Alternative Models of "Mistakes": Misperceptions in Costs of Applying and Inattention

Misperceptions in Private Utility Cost of Applying (Instead of Misperceptions in Benefits from Applying) Next, instead of misperceiving the benefits from applying by misperceiving

B, now assume that individuals misperceive private utility cost of applying. Specifically, each type may also misperceive true cost of applying by δ_j, which raises the perceived cost of applying for type j.

So it is still the case that individual of type j applies if perceived expected utility of applying is greater than (certain) utility of not applying. For individual with private cost c_i this can be defined as follows (assuming there is no misperception of benefits from applying; i.e., $\epsilon_j = 0$)

$$
\begin{aligned}
& u(h_j^A\theta_j - \tau(h_j^A\theta_j) + (1+\epsilon_j)B) - v(h_j^A) - (\bar{\Lambda}\kappa_j + c_j + \delta_j) > u(h_j^{\neg A}\theta_j - \tau(h_j^{\neg A}\theta_j)) + v(h_j^{\neg A}) \\
\Rightarrow\ & c_j + \delta_j < u(h_j^A\theta_j - \tau(h_j^A\theta_j) + (1+\epsilon_j)B) - v(h_j^A) - \bar{\Lambda}\kappa_j - u(h_j^{\neg A}\theta_j - \tau(h_j^{\neg A}\theta_j)) + v(h_j^{\neg A}) \\
\Rightarrow\ & c_j < u(h_j^A\theta_j - \tau(h_j^A\theta_j) + B) - v(h_j^A) - \bar{\Lambda}\kappa_j - u(h_j^{\neg A}\theta_j - \tau(h_j^{\neg A}\theta_j)) + v(h_j^{\neg A}) - \delta_j = c_j^* - \delta_j
\end{aligned}
$$

As a result, the share of individuals of type H applying is $F_H(c_H^* - \delta_H)$. We can also define the (private + public) welfare coming from type H individuals.

$$
\begin{aligned}
W_H =& u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A}) \\
&+ \int_0^{c_H^* - \delta_H} \Big[(u(h_H^A\theta_H - \tau(h_H^A\theta_H) + B) - v(h_H^A) - (\bar{\Lambda}\kappa_H + c)) - (u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A}))\Big]\, dF_H(c) \\
&- BF_H(c_H^* - \delta_H) \\
&+ \Big[F_H(c_H^* - \delta_H)\tau(h_H^A\theta_H) + (1 - F_H(c_H^* - \delta_H))\tau(h_H^{\neg A}\theta_H)\Big].
\end{aligned}
$$

Note in above expression that the δ affects application decision but not realized utility (it's perceived cost, not an actual cost). All of these results are identical for the type L individuals, simply replace the H subscripts with L subscripts.

Proposition 4. *The effect of the Information Only treatment on welfare is given by the following expression:*

$$
\frac{dW}{dT} = \delta_L \frac{dA_L}{dT} + \delta_H \frac{dA_H}{dT} - B\left(\frac{dA_L}{dT} + \frac{dA_H}{dT}\right) + \Big[\tau(h_L^A\theta_L) - \tau(h_L^{\neg A}\theta_L)\Big]\frac{dA_L}{dT} + \Big[\tau(h_H^A\theta_L) - \tau(h_H^{\neg A}\theta_H)\Big]\frac{dA_H}{dT}
$$

Proof: Welfare is given by:

$$
\begin{aligned}
W =& u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A}) \\
&+ \int_0^{c_H^*-\delta_H} \left[(u(h_H^A\theta_H - \tau(h_H^A\theta_H) + B) - v(h_H^A) - (\bar{\Lambda}\kappa_H + c)) - (u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A}))\right] dF_H(c) \\
&+ u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) - v(h_L^{\neg A}) \\
&+ \int_0^{c_L^*-\delta_L} \left[(u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - v(h_L^A) - (\bar{\Lambda}\kappa_L + c)) - (u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) - v(h_L^{\neg A}))\right] dF_L(c) \\
&- B\left[F_H(c_H^* - \delta_H) + F_L(c_L^* - \delta_L)\right] \\
&+ \left[F_H(c_H^* - \delta_H)\tau(h_H^A\theta_H) + (1 - F_H(c_H^* - \delta_H))\tau(h_H^{\neg A}\theta_H)\right] \\
&+ \left[F_L(c_L^* - \delta_L)\tau(h_L^A\theta_L) + (1 - F_L(c_L^* - \delta_L))\tau(h_L^{\neg A}\theta_L)\right].
\end{aligned}
$$

Information Only ($dT = -d\delta$): Taking the derivative with respect to to δ (i.e., both δ_H and δ_L change by $d\delta$) yields the following:

$$
\begin{aligned}
\frac{dW}{d\delta} =& \frac{d}{d\delta} \int_0^{c_H^*-\delta_H} \left[u(h_H^A\theta_H - \tau(h_H^A\theta_H) + B) - v(h_H^A) - (\bar{\Lambda}\kappa_H + c) - u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) + v(h_H^{\neg A})\right] dF_H(c) \\
&+ \frac{d}{d\delta} \int_0^{c_L^*-\delta_L} \left[u(h_L^A\theta_L - \tau(h_L^A\theta_L) + B) - v(h_L^A) - (\bar{\Lambda}\kappa_L + c) - u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) + v(h_L^{\neg A})\right] dF_L(c) \\
&+ B\left[f_H(c_H^* - \delta_H) + f_L(c_L^* - \delta_L)\right] \\
&- \left[f_H(c_H^* - \delta_H)\tau(h_H^A\theta_H) - f_H(c_H^* - \delta_H)\tau(h_H^{\neg A}\theta_H) + f_L(c_L^* - \delta_L)\tau(h_\vdots^A\theta_L) - f_L(c_L^* - \delta_L)\tau(h_L^{\neg A}\theta_L)\right]
\end{aligned}
$$

Applying Leibniz's Rule, we get

$$
\begin{aligned}
&\frac{d}{d\delta} \int_0^{c_H^*-\delta_H} \left[u(h_H^A\theta_H - \tau(h_H^A\theta_H) + B) - v(h_H^A) - (\bar{\Lambda}\kappa_H + c) - u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) + v(h_H^{\neg A})\right] dF_H(c) \\
&= -\delta f_H(c_H^* - \delta_H)
\end{aligned}
$$

Since the number of applicants is given by $A_L = F_L(c_L^* - \delta)L)$ and $A_H = F_H(c_H^* - \delta_H)$,

$$\frac{dA_L}{d\delta} = -f_L(c_L^* - \delta_L)$$

and

$$\frac{dA_H}{d\delta} = -f_H(c_H^* - \delta_H).$$

Therefore, we can re-write

$$
\begin{aligned}
&B\left[f_H(c_H^* - \delta_H) + f_L(c_L^* - \delta_L)\right] \\
&- \Big[f_H(c_H^* - \delta_H)\tau(h_H^A\theta_H) - f_H(c_H^* - \delta_H)\tau(h_H^{\neg A}\theta_H) + f_L(c_L^* - \delta_L)\tau(h_L^A\theta_L) - f_L(c_L^* - \delta_L)\tau(h_L^{\neg A}\theta_L)\Big] \\
=& - B(\frac{dA_L}{d\delta} + \frac{dA_H}{d\delta}) + \left[\frac{dA_L}{d\delta}\tau(h_L^A\theta_L) - \frac{dA_L}{d\delta}\tau(h_L^{\neg A}\theta_L) + \frac{dA_H}{d\delta}\tau(h_H^A\theta_H) - \frac{dA_H}{d\delta}\tau(h_H^{\neg A}\theta_L)\right]
\end{aligned}
$$

Putting all this together (since $dT = -d\delta$), we have

$$
\frac{dW}{dT} = \delta_L\frac{dA_L}{dT} + \delta_H\frac{dA_H}{dT} - B(\frac{dA_L}{dT} + \frac{dA_H}{dT}) + \Big[\tau(h_L^A\theta_L) - \tau(h_L^{\neg A}\theta_L)\Big]\frac{dA_L}{dT} + \Big[\tau(h_H^A\theta_L) - \tau(h_H^{\neg A}\theta_H)\Big]\frac{dA_H}{dT}
$$

Alternative Non-Neoclassical Model: Inattention Now suppose that agents have correct beliefs (so we have $\epsilon_L = \epsilon_H = 0$), but a fraction $(1-\alpha)$ of agents are inattentive where $\alpha \in [0,1]$. This fraction is assumed to be independent of private utility cost of applying. Attentive agents make a choice about whether or not to apply for benefits, but inattentive agents simply do not apply because they forget, don't read the paperwork, etc. The Information Only treatment then reduces the fraction of inattentive agents (or increases the fraction of attentive agents). In other words, $dT = d\alpha$.

Let $U_j(\cdot)$ be individual utility regardless of whether the individual chooses to apply. As before,

$$
\begin{aligned}
W =& \int U_L(\cdot)dF_L(c) + \int U_H(\cdot)dF_H(c) \\
&- [B(A_L + A_H)] \\
&+ \Big[A_L\tau(h_L^A\theta_L) + (1-A_L)\tau(h_L^{\neg A}\theta_L) + A_H\tau(h_H^A\theta_H) + (1-A_H)\tau(h_H^{\neg A}\theta_H)\Big],
\end{aligned}
$$

Attentive agents apply if

$$
\begin{aligned}
&u(h_j^A\theta_j - \tau(h_j^A\theta_j) + (1+\epsilon_j)B) - v(h_j^A) - (\bar{\Lambda}\kappa_j + c_j) > u((1-\tau)h_j^{\neg A}\theta_j) + v(h_j^{\neg A}) \\
\Rightarrow\ & c_j < u(h_j^A\theta_j - \tau(h_j^A\theta_j) + B) - v(h_j^A) - \bar{\Lambda}\kappa_j - u((1-\tau)h_j^{\neg A}\theta_j) + v(h_j^{\neg A}) \equiv c_j^*
\end{aligned}
$$

By contrast, inattentive agents don't apply, regardless of their value of c_j. Let $\nu_j = \int_0^{c_j^*}\Big[c_j^* - c\Big]\,dF_j(c)$, which is the applicants' aggregate amount of the difference between threshold in application cost and individual application cost. In the case where there is no misperception of the probability of application being accepted, this is the aggregate amount of utility gain by turning to an applicant of those whose optimal choice it is to apply. Therefore, we can rewrite (focusing on type H):

$$
\begin{aligned}
\int U_H(\cdot)dF_H(c) =& \alpha\left[\int_A U_H(\cdot|A)f_H(c)dc + \int_{\neg A} U_H(\cdot|\neg A)f_H(c)dc\right] + (1-\alpha)\left[\int U_H(\cdot|\neg A)f_H(c)dc\right] + \\
=& \alpha\left[u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A})\right] \\
&+ \alpha\left[\int_0^{c_H^*}\left[(u(h_H^A\theta_H - \tau(h_H^A\theta_H) + B) - v(h_H^A) - (\bar{\Lambda}\kappa_H + c)) - (u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A}))\right]dF_H(c)\right] \\
&+ (1-\alpha)\left[u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A})\right] \\
=& u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A}) \\
&+ \alpha\int_0^{c_H^*}\left[(u(h_H^A\theta_H - \tau(h_H^A\theta_H) + B) - v(h_H^A) - (\bar{\Lambda}\kappa_H + c)) - (u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A}))\right]dF_H(c) \\
=& u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A}) + \alpha\int_0^{c_H^*}[c_H^* - c]\,dF_H(c) \\
=& u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A}) + \alpha\nu_H,
\end{aligned}
$$

and similarly,

$$
\int U_L(\cdot)dF_L(c) = u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) - v(h_L^{\neg A}) + \alpha\nu_L
$$

Applicants are given by

$$
A_L = \alpha F_L(c_L^*)
$$

and

$$
A_H = \alpha F_H(c_H^*)
$$

Therefore, welfare is

$$
\begin{aligned}
W =& u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) - v(h_L^{\neg A}) + \alpha\nu_H \\
&+ u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A}) + \alpha\nu_L \\
&- [\alpha B(F_L(c_L^*) + F_H(c_H^*))] \\
&+ \left[\alpha F_L(c_L^*)\tau(h_L^A\theta_L) + (1-\alpha F_L(c_L^*))\tau(h_L^{\neg A}\theta_L) + \alpha F_H(c_H^*)\tau(h_H^A\theta_H) + (1-\alpha F_H(c_H^*))\tau(h_H^{\neg A}\theta_H)\right]
\end{aligned}
$$

Just to compare, in the fully rational benchmark case, welfare is given by

$$
\begin{aligned}
W =& u(h_L^{\neg A}\theta_L - \tau(h_L^{\neg A}\theta_L)) - v(h_L^{\neg A}) + \nu_L \\
&+ u(h_H^{\neg A}\theta_H - \tau(h_H^{\neg A}\theta_H)) - v(h_H^{\neg A}) + \nu_H \\
&- [B(F_L(c_L^*) + F_H(c_H^*))] \\
&+ \left[F_L(c_L^*)\tau(h_L^A\theta_L) + (1-F_L(c_L^*))\tau(h_L^{\neg A}\theta_L) + F_H(c_H^*)\tau(h_H^A\theta_H) + (1-F_H(c_H^*))\tau(h_H^{\neg A}\theta_H)\right]
\end{aligned}
$$

We consider the effect of an Information Only intervention that increases attention:

Proposition 5. *If $\epsilon_L = \epsilon_H = 0$, and there is at least some inattention, then the effect of Information Only treatment ($dT = d\alpha$) on welfare is given by:*

$$\begin{aligned}\frac{dW}{dT} &= \nu_H + \nu_L \\ &- B(F_L(c_L^*) + F_H(c_H^*)) \\ &+ F_L(c_L^*)\left[\tau(h_L^A\theta_L) - \tau(h_L^{\neg A}\theta_L) - g_L\right] + F_H(c_H^*)\left[\tau(h_H^A\theta_H) - \tau(h_H^{\neg A}\theta_H)\right]\end{aligned}$$

Proof: Simply differentiating W with respect to α yields

$$\begin{aligned}\frac{dW}{dT} &= \nu_H + \nu_L \\ &- B(F_L(c_L^*) + F_H(c_H^*)) \\ &+ F_L(c_L^*)\left[\tau(h_L^A\theta_L) - \tau(h_L^{\neg A}\theta_L)\right] + F_H(c_H^*)\left[\tau(h_H^A\theta_H) - \tau(h_H^{\neg A}\theta_H)\right]\end{aligned}$$

Note that this result makes clear that there are effects on private welfare from a treatment that reduces inattention, but the targeting properties of this intervention will depend on relative magnitude of ν_H and ν_L, which depend on the conditional distribution of costs of applying (conditional on cost being between 0 and the threshold cost) as well as the κ_H and κ_L.

E.3 Welfare effects of Intervention in the Empirical Setting

As described in the main text, we use the conceptual framework developed in Section 2 to assess the normative implications of the empirical findings from our RCT. We tailor the framework in two specific, but minor, ways in order to apply it to our empirical setting. First, to facilitate our calibration of the model based on our empirical evidence, we allow for an exogenous probability the application is accepted, given by π_j. Second, we allow for two different benefit levels: individuals may receive either $\overline{B}$ or B_{min}, with $\overline{B} > B_{min}$. These are benefits individuals receive conditional on application being accepted.

In addition, given the partial equilibrium nature of the intervention and the elderly study population, we assume that earnings do not respond endogenously to our intervention. This does not constrain the fiscal externalities from the intervention since, as discussed, in Section 2, the framework and propositions developed apply generally to any fiscal externality. So, we will still use the G_j^A and $G_j^{\neg A}$ to denote fiscal externalities of type j individuals, as in the main text, but in our empirical setting we view these fiscal externalities as coming from application processing costs rather than endogenous earnings.

Without endogenous earnings, $h_j^A = h_j^{\neg A}$ for both types of individuals, and so we will just use

$h_j\theta_j$ to denote labor earnings for type j individuals. Importantly, since earnings do not adjust, the level of benefits that individuals receive is fully determined by their type, with low-ability types receiving higher benefits $\overline{B}$ and high-ability types receiving the minimum level of benefits B_{min}. This suggests a natural empirical definition of targeting can be based on the level of benefits received: $e = E_L/(E_L + E_H)$, where E_L is enrollees who receive minimum level of benefits and E_H is enrollees who receive higher benefits. We thus interpret benefit level as a direct proxy for type in our setting. Given this proxy, our empirical results therefore indicate that both interventions decrease targeting ($de/dT < 0$).

Lastly, since expected benefits are now a function of both probability application is accepted and the benefits (conditional on application acceptance), the misperceptions can either take the form of misperceptions in acceptance rate or misperceptions in benefits (conditional on application acceptance). In this section we focus on misperceptions in $\overline{B}$ or B_{min},, and we refer readers to the NBER Working Paper version (#24652, www.nber.org/papers/w24652) for analogous results (which hold as first-order approximations) in the case where misperception are based on probability application is accepted (i.e., the π_j). The results are first-order approximations because we need to assume that the welfare effects of receiving benefits B with probability π can be approximated by the welfare effects of receiving benefits $\pi * B$ with certainty.

With these modifications, we re-state Propositions 1 and 2 (as 1a and 2a) as follows:

Proposition 1a. *With benefit level as a direct proxy for type in our setting, now let* $\mu_j \equiv \pi_j u(h_j^A\theta_j - \tau(h_j^A\theta_j) + B_j) - \pi_j u(h_j^A\theta_j - \tau(h_j^A\theta_j) + (1+\epsilon_j)B_j)$. *The effect of the Information Only treatment on welfare is given by the following expression:*

$$\frac{dW}{dT}^{Information\ Only} = \underbrace{\mu_L\frac{dA_L}{dT} + \mu_H\frac{dA_H}{dT}}_{\text{Change in Private Welfare}} - \underbrace{\left[(\pi_H B_{min})\frac{dA_H}{dT} + (\pi_L\overline{B})\frac{dA_L}{dT}\right]}_{\text{Change in Mechanical Program Costs}} + \underbrace{\left[[G_L^A - G_L^{\neg A}]\frac{dA_L}{dT} + [G_H^A - G_H^{\neg A}]\frac{dA_H}{dT}\right]}_{\text{Change in Fiscal Externality}}$$

And the effect of the Information Plus Assistance treatment on welfare is given by the following expression:

$$\frac{dW}{dT}^{Information\ Plus\ Assistance} = \underbrace{\mu_L\frac{dA_L}{dT} + \mu_H\frac{dA_H}{dT} + \kappa_L A_L + \kappa_H A_H}_{\text{Change in Private Welfare}} - \underbrace{\left[(\pi_H B_{min})\frac{dA_H}{dT} + (\pi_L\overline{B})\frac{dA_L}{dT}\right]}_{\text{Change in Mechanical Program Costs}} + \underbrace{\left[[G_L^A - G_L^{\neg A}]\frac{dA_L}{dT} + [G_H^A - G_H^{\neg A}]\frac{dA_H}{dT}\right]}_{\text{Change in Fiscal Externality}}$$

Proof: Welfare is given by

$$
\begin{aligned}
W =& V_L + V_H \\
& - (\pi_L \overline{B} A_L + \pi_H B_{min} A_H) \\
& + G_L^A A_L + G_L^{\neg A}(1 - A_L) + G_H^A A_H + G_H^{\neg A}(1 - A_H) \\
=& u(h_L\theta_L - \tau(h_L\theta_L)) - v(h_L) \\
& + \int_0^{c_L^*} \Big[\pi_L u(h_L\theta_L - \tau(h_L\theta_L) + \overline{B}) + (1-\pi_L)u(h_L\theta_L - \tau(h_L\theta_L)) - v(h_L) - (\bar{\Lambda}\kappa_L + c)\Big]\, dF_L(c) \\
& - \int_0^{c_L^*} [u(h_L\theta_L - \tau(h_L\theta_L)) - v(h_L)]\, dF_L(c) \\
& + u(h_H\theta_H - \tau(h_H\theta_H)) - v(h_H) \\
& + \int_0^{c_H^*} \Big[\pi_H u(h_H\theta_H - \tau(h_H\theta_H) + B_{min}) + (1-\pi_H)u(h_H\theta_H - \tau(h_H\theta_H)) - v(h_H) - (\bar{\Lambda}\kappa_H + c)\Big]\, dF_H(c) \\
& - \int_0^{c_H^*} [u(h_H\theta_H - \tau(h_H\theta_H)) - v(h_H)]\, dF_H(c) \\
& - \pi_L \overline{B} A_L - \pi_H B_{min} A_H \\
& + G_L^A A_L + G_L^{\neg A}(1 - A_L) + G_H^A A_H + G_H^{\neg A}(1 - A_H),
\end{aligned}
$$

where $c_j^* = \pi_j u(h_j\theta_j - \tau(h_j\theta_j) + (1+\epsilon_j)B_j) - \bar{\Lambda}\kappa_j - \pi_j u(h_j\theta_j - \tau(h_j\theta_j))$, defining $B_j = B_{min}$ for type L individuals, and $B_j = \overline{B}$ for type Hindividuals. This welfare expression can be simplified to the following:

$$
\begin{aligned}
W =& u(h_L\theta_L - \tau(h_L\theta_L)) - v(h_L) \\
& + \int_0^{c_L^*} \Big[\pi_L u(h_L\theta_L - \tau(h_L\theta_L) + \overline{B}) - \pi_L u(h_L\theta_L - \tau(h_L\theta_L)) - (\bar{\Lambda}\kappa_L + c)\Big]\, dF_L(c) \\
& + u(h_H\theta_H - \tau(h_H\theta_H)) - v(h_H) \\
& + \int_0^{c_H^*} \Big[\pi_H u(h_H\theta_H - \tau(h_H\theta_H) + B_{min}) - \pi_H u(h_H\theta_H - \tau(h_H\theta_H)) - (\bar{\Lambda}\kappa_H + c)\Big]\, dF_H(c) \\
& - \pi_L \overline{B} A_L - \pi_H B_{min} A_H + G_L^A A_L + G_L^{\neg A}(1 - A_L) + G_H^A A_H + G_H^{\neg A}(1 - A_H)
\end{aligned}
$$

Applying Leibniz's Rule, we get:

$$\frac{d}{d\epsilon}\int_0^{c_L^*}\Big[\pi_L u(h_L\theta_L-\tau(h_L\theta_L)+\overline{B})-\pi_L u(h_L\theta_L-\tau(h_L\theta_L))-(\bar{\Lambda}\kappa_L+c)\Big]f_L(c)dc$$
$$=\Big[\pi_L u(h_L\theta_L-\tau(h_L\theta_L)+\overline{B})-\pi_L u(h_L\theta_L-\tau(h_L\theta_L))-(\bar{\Lambda}\kappa_L+c_L^*)\Big]f_L(c_L^*)\frac{dc_L^*}{d\epsilon}$$
$$=(\pi_L u(h_L\theta_L-\tau(h_L\theta_L)+\overline{B})-\pi_L u(h_L\theta_L-\tau(h_L\theta_L)+(1+\epsilon_L)\overline{B}))f_L(c_L^*)\frac{dc_L^*}{d\epsilon}$$

Since the number of applicants is given by $A_j = F_j(c_j^*)$,

$$\frac{dA_j}{d\epsilon} = f_j(c_j^*)\frac{dc_j^*}{d\epsilon}.$$

Putting all this together (and using $dT = d\epsilon$), we have

$$\begin{aligned}\frac{dW}{d\epsilon} =&\mu_L\frac{dA_L}{d\epsilon}+\mu_H\frac{dA_H}{d\epsilon}\\ &-\left[(\pi_H B_{min})\frac{dA_H}{dT}+(\pi_L\overline{B})\frac{dA_L}{dT}\right]\\ &+[G_L^A-G_L^{\neg A}]\frac{dA_L}{dT}+[G_H^A-G_H^{\neg A}]\frac{dA_H}{dT}\end{aligned}$$

This completes the first part of the proposition.

Assistance Only ($dT = -dc$): This proof combines the arguments above and the steps in the proof of Proposition 1.

Relationship Between Targeting Impacts and Changes in Welfare

Proposition 2a. *Holding constant the change in applications due to an intervention, the change in social welfare in response to an improvement in targeting ($de/dT > 0$) from an Information Only (or Information Plus Assistant) treatment is given by the following expression:*

$$\frac{\partial}{(de/dT)}\left(\frac{dW}{dT}\right)\Big|_{\frac{dA}{dT}} = \Big[(\mu_L-\mu_H)-(\pi_L\overline{B}-\pi_H B_{min})+(G_L^A-G_L^{\neg A})-(G_H^A-G_H^{\neg A})\Big]*\Gamma$$

where $\Gamma = \frac{(E_L+E_H)^2}{\pi_L E_L+\pi_H E_H} > 0$.

Application acceptance rates are exogenous and given by π_j, so $E_j = \pi_j A_j$. Since $e = \frac{E_L}{E_H+E_L} =$

$\frac{\pi_L A_L}{\pi_L A_L + \pi_H A_H}$

$$
\begin{aligned}
\frac{de}{dT} &= \frac{d}{dT}\frac{\pi_L A_L}{\pi_L A_L + \pi_H A_H} \\
&= \frac{\pi_L \frac{dA_L}{dT}(\pi_L A_L + \pi_H A_H) - (\pi_L \frac{dA_L}{dT} + \pi_H \frac{dA_H}{dT})\pi_L A_L}{(\pi_L A_L + \pi_H A_H)^2} \\
&= \frac{1}{E_L + E_H}(\pi_L \frac{dA_L}{dT} - \frac{(\pi_L \frac{dA_L}{dT} + \pi_H \frac{dA_H}{dT})E_L}{E_L + E_H}) \\
&= \frac{1}{E_L + E_H}(\pi_L \frac{dA_L}{dT} - \frac{(\pi_L \frac{dA}{dT} + (\pi_L - \pi_H)\frac{dA_L}{dT})E_L}{E_L + E_H}) \\
&= \frac{1}{E_L + E_H}(\frac{\pi_L E_L + \pi_H E_H}{E_L + E_H} \cdot \frac{dA_L}{dT} - \frac{\pi_L E_L}{E_L + E_H} \cdot \frac{dA}{dT})
\end{aligned}
$$

From Proposition 1, we know that change in welfare from Information Only is the following

$$
\begin{aligned}
\frac{dW}{dT} &= \mu_L \frac{dA_L}{dT} + \mu_H \frac{dA_H}{dT} - (\pi_H B_{min})\frac{dA_H}{dT} - (\pi_L \overline{B})\frac{dA_L}{dT} + [G_L^A - G_L^{\neg A}]\frac{dA_L}{dT} + [G_H^A - G_H^{\neg A}]\frac{dA_H}{dT} \\
&= \left[\mu_L - \pi_L \overline{B} + (G_L^A - G_L^{\neg A})\right](\frac{dA_L}{dT} + \frac{dA_H}{dT}) \\
&\quad + \left[(\mu_H - \pi_H B_{min} + (G_H^A - G_H^{\neg A})) - (\mu_L - \pi_L \overline{B} + (G_L^A - G_L^{\neg A}))\right]\frac{dA_H}{dT} \\
&= \left[\mu_L - \pi_L \overline{B} + (G_L^A - G_L^{\neg A})\right]\frac{dA}{dT} + \left[(\mu_H - \pi_H B_{min} + (G_H^A - G_H^{\neg A})) - (\mu_L - \pi_L \overline{B} + (G_L^A - G_L^{\neg A}))\right]\frac{dA_H}{dT}.
\end{aligned}
$$

Now, since dA/dT is held constant in the Proposition, we just need to derive how the second term varies with de/dT. Thus, we can solve for $-\frac{dA_H}{dT}$ in terms of de/dT:

$$
\begin{aligned}
-\frac{dA_H}{dT} + \frac{dA}{dT} &= ((E_L + E_H)\frac{de}{dT} + \frac{\pi_L E_L}{E_L + E_H} \cdot \frac{dA}{dT})\frac{E_L + E_H}{\pi_L E_L + \pi_H E_H} \\
&= \frac{(E_L + E_H)^2}{\pi_L E_L + \pi_H E_H} \cdot \frac{de}{dT} + \frac{\pi_L E_L}{\pi_L E_L + \pi_H E_H} \cdot \frac{dA}{dT} \\
-\frac{dA_H}{dT} &= \frac{(E_L + E_H)^2}{\pi_L E_L + \pi_H E_H} \cdot \frac{de}{dT} + (\frac{\pi_L E_L}{\pi_L E_L + \pi_H E_H} - 1)\frac{dA}{dT}
\end{aligned}
$$

which can then be substituted back into the dW/dT expression above. Then, taking the partial derivative with respect to de/dT gives the expression in Proposition 2a:

$$
\begin{aligned}
\frac{\partial}{\partial(de/dT)}\left(\frac{dW}{dT}\right)\Bigg|_{\frac{dA}{dT}} &= \frac{\partial}{\partial(de/dT)}\left[(\mu_L-\mu_H)-(\pi_L\overline{B}-\pi_H B_{min})+(G_L^A-G_L^{\neg A})-(G_H^A-G_H^{\neg A})\right] \\
&\quad\times\left(\frac{(E_L+E_H)^2}{\pi_L E_L+\pi_H E_H}\cdot\frac{de}{dT}+\left(\frac{\pi_L E_L}{\pi_L E_L+\pi_H E_H}-1\right)\cdot\frac{dA}{dT}\right) \\
&= \left[(\mu_L-\mu_H)-(\pi_L\overline{B}-\pi_H B_{min})+(G_L^A-G_L^{\neg A})-(G_H^A-G_H^{\neg A})\right] \\
&\quad\times\frac{\partial}{\partial(de/dT)}\left(\frac{(E_L+E_H)^2}{\pi_L E_L+\pi_H E_H}\cdot\frac{de}{dT}+\left(\frac{\pi_L E_L}{\pi_L E_L+\pi_H E_H}-1\right)\cdot\frac{dA}{dT}\right) \\
&= \left[(\mu_L-\mu_H)-(\pi_L\overline{B}-\pi_H B_{min})+(G_L^A-G_L^{\neg A})-(G_H^A-G_H^{\neg A})\right]\frac{(E_L+E_H)^2}{\pi_L E_L+\pi_H E_H}
\end{aligned}
$$

E.3.1 Derivation of MVPF formula from marginal welfare gain expression

Proposition 2a indicates that with $\epsilon_L < \epsilon_H < 0$, a benefit formula that pays higher benefits to L types, and constant fiscal externalities g across types, our finding that the interventions decrease targeting bodes poorly for their welfare impacts. However, this is merely a qualitative comparative static result. Even with $\epsilon_L < \epsilon_H < 0$, the targeting effects of the intervention are neither necessary nor sufficient to sign the overall social welfare impact of the intervention. The overall social welfare effect may be positive, if private welfare gains to individuals with misperceptions outweigh the negative externality from the public application processing costs and expenditures on benefits.

We begin with the dW/dT expression for the Information Only intervention in Proposition 1a and impose our baseline assumption $G_L^A = G_H^A \equiv -g$, and $G_L^{\neg A} = G_H^{\neg A} = 0$ (which we will relax later). We use a first-order Taylor approximation around actual utility (which we also used to calibrate ϵ) to approximate μ_H as $\xi_H \pi_H \epsilon_H B_{min}$, and μ_L as $\xi_L \pi_L \epsilon_L \overline{B}$. The first-order Taylor approximation is a useful simplification to allow us to implement this formula empirically, but we assess the sensitivity to this assumption in the next sub-section.This transforms $\frac{dW}{dT}^{Information\ Only}$ in Proposition 1a into:

$$
\begin{aligned}
\frac{dW}{dT}^{Information\ Only} &= \underbrace{\xi_L \pi_L \epsilon_L \overline{B}\frac{dA_L}{dT}+\xi_H \pi_H \epsilon_H B_{min}\frac{dA_H}{dT}}_{\textit{Change in Private Welfare}} \\
&\quad - \underbrace{\left[(\pi_L\overline{B}+g)\frac{dA_L}{dT}+(\pi_H B_{min}+g)\frac{dA_H}{dT})\right]}_{\textit{Change in Government Revenue and Public Expenditure on Benefits}}
\end{aligned}
$$

To translate the change in private welfare into a change in dollars of surplus to each type of individual, we divide the change in private welfare for each type by the marginal utility of consumption for each type (ξ_j). We also express the formula as a ratio of change in private welfare to change in costs - rather than the difference between them - so that we can interpret it (in the spirit of Hendren 2016), as the marginal value of public funds (MVPF) of the intervention. This yields:

$$MVPF^{Information\,Only} = \frac{\mu_L \frac{dA_L}{dT} * (1/\xi_L) + \mu_H \frac{dA_H}{dT} * (1/\xi_H)}{(\pi_L \overline{B} + g)\frac{dA_L}{dT} + (\pi_H B_{min} + g)\frac{dA_H}{dT}}$$

Intuitively, this expression represents the dollars of surplus transferred to each type (measured in that type's own money metric), divided by the total fiscal cost (in dollars) of the intervention. Smaller values of MVPF mean that transfering surplus to these two types requires more resources than larger MVFP values.

E.3.2 Sensitivity of Model Calibrations to Allowing for Risk Aversion

In Section 6, the model calibrations back out values for ϵ_H and ϵ_L by calculating values such that individuals (with these biased beliefs) are indifferent to applying, given time cost of applying of \$75, and no non-time costs of applying. By taking a first-order Taylor approximation around the utility of not applying, the difference for type H individuals between the expected utility of applying and not applying is given by $(1 + \epsilon_H) * \pi_H * B_{min} * u'(x_H^{\neg A})$. Dividing by $u'(x_H^{\neg A})$ turns this into a money metric that will be equal to \$75 for the marginal applicant. Solving for ϵ_H gives $\epsilon_H = 75/(432) - 1 = -0.833$. For ϵ_L the analogous calculation is $\epsilon_L = 75/4806 - 1 = -0.984$.

To see sensitivity to different assumptions on risk aversion, we assume CRRA utility function ($u(x) = x^{1-\gamma}/(1-\gamma)$ and we assume consumption is equal to income. We also assume that for type H non-applicants $y_H = \$36,000$, and for type L non-applicants $y_L = \$18,000$. These values correspond to approximately 36 months of income for a single person at 100% and 50% of the federal poverty line, and since we assume labor supply is fixed, this value is the same whether the individual applies or not. With these assumptions it is possible to back out ϵ_j for a given value of γ using the following expression: $(1+\epsilon_j)\pi_j u(y_j + B_j) + (1-(1+\epsilon_j)\pi_j)u(y_j) - c_j^* = u(y_j)$, assuming that c_j^* can be approximated as $\$75 * u'(y_j)$.

For a value of $\gamma = 2$ we calculate $\epsilon_H = -0.831$ and for $\gamma = 5$ we calculate $\epsilon_H = -0.826$. For type L individuals, for a value of $\gamma = 2$ we calculate $\epsilon_L = -0.979$ and for $\gamma = 5$ we calculate $\epsilon_L = -0.968$. These calculations show very little sensitivity of the misperception parameters to allowing for risk aversion, which means that we calculate very similar implied change in private welfare for marginal applicants as well as similar MVPFs.

F: Complier Characteristics

We report the characteristics of enrollees and applicants separately for always takers, compliers, and never takers. The compliers for a given intervention are the individuals who enroll in SNAP if and only if they receive that intervention. Always takers are individuals who will enroll in SNAP absent any intervention. Never takers are individuals who will not enroll in SNAP regardless of whether the treatment is received or not. Estimation of these objects is standard (see, e.g., Abadie 2002, Abadie 2003, and Angrist and Pischke 2009).

Suppose we observe some SNAP enrollee (or analogously, applicant) characteristic X, such as age. We denote by μ_T the characteristic's mean among treatment enrollees. Individuals in the treatment arm who enroll in SNAP include both compliers and always takers. The mean of the characteristic for treatment enrollees is therefore a weighted average of the means for always takers and compliers:

$$\mu_T = \frac{\pi_{AT}}{\pi_{AT} + \pi_C}\mu_{AT} + \frac{\pi_C}{\pi_{AT} + \pi_C}\mu_C$$

where , π_{AT} is the share of always takers, π_C is the share of compliers, and μ_{AT} and μ_C are means, respectively.

Re-arranging the equation, we get

$$\mu_C = \frac{(\pi_{AT} + \pi_C)\mu_T - \pi_{AT}\mu_{AT}}{\pi_C}. \tag{11}$$

All the parameters on the right hand side can be derived from the data. Specifically, we define D as the treatment status, $D = 1$ if treated, and 0 otherwise. We define Z as enrollee status, $Z = 1$ if enrolled and 0 otherwise. The four parameters can thus be expressed:

$$\pi_{AT} = \frac{p(D = 0,\, Z = 1)}{p(D = 0)} = p(Z = 1|\, D = 0)$$

(note that we can only observe always takers in the control group but by definition of the randomization we assume the share of always takers is the same in the treatment and control group),

$$\pi_C = p(Z = 1|\, D = 1) - p(Z = 1|\, D = 0)$$

$$\mu_{AT} = E[\, X \,|\, Z = 1,\, D = 0] \tag{12}$$

(where again we note that we can only observe always takers in the control group but by assumption the characteristics of always takers are the same in the treatment and control group), and

$$\mu_T = E[\, X \,|\, Z = 1,\, D = 1].$$

We denote by π_{NT} the share of never takers, and μ_{NT} the the mean. They can be derived from

the data:

$$\pi_{NT} = \frac{p(D = 1,\, Z = 0)}{p(D = 1)} = p(Z = 0|\, D = 1)$$

$$\mu_{NT} = E[\, X \,|\, Z = 0,\, D = 1] \tag{13}$$

(where we note that we can only observe never takers in the treatment group but by assumption the characteristics of never takers are the same in the treatment and control group).

To calculate standard errors of these estimated means and p-values of tests of their difference, we perform 10,000 replications of the bootstrap.

G: Background on SNAP Eligibility and Benefits

While SNAP program rules are mostly determined at the federal level, there is some variation across states. In PA at the time of our intervention (2016), there were three ways an elderly individual can be eligible for SNAP.[41] First, the household would be categorically eligible if all household members received a qualifying benefit - SSI, TANF, General Assistance, State Blind Pensions, or Family Works benefits. Second, the household would be eligible if its gross income were below 200 percent of the Federal Poverty Income Guidelines (FPIG) and has resources below the $3,250 resource limit.[42] Third, the household would be eligible if its gross income were above 200 percent of FPIG but its net income (gross income minus certain exempt income and deductions for certain expenses)[43] were less than 100% FPIG and it had resources below the $3,250 resource limit. Once deemed eligible, an elderly household is certified to receive SNAP benefits for 36 months, although there are exceptions that require earlier re-certification.[44]

If eligible, the benefit amount is set, based on a federally determined formula, as a decreasing function of net income, subject to a minimum and maximum. Benefits are designed so that households spend approximately 30% of their net income (i.e., gross income minus certain deductions and exemptions) on food. Specifically, the maximum benefit is set equal to the cost of food under the USDA's Thrifty Food Plan, which is the minimum amount deemed necessary to buy enough food for a household of a particular size. A family with no income receives the maximum benefit, with benefits taxed away by 30 percent of net income, up to a floor. Thus – subject to a minimum and maximum – monthly benefits are the Thrifty Food Plan Amount (which varies by household size) minus 30 percent of Net Monthly Income. During our study period, the minimum monthly benefit for a categorically eligible household of size 1 or 2 was $16; the minimum monthly benefit was $0 for other enrollees. The maximum monthly benefit was $194 for a household size of 1, $357 for a household size of 2, and $511 for a household size of 3.

The state has 30 calendar days to process an application.[45]

Given the SNAP program rules, both the individual and state's determination of eligibility and benefit amounts require the individual to actively apply with the required information. From the individual's perspective, there is uncertainty about the benefit function, the inputs into it (e.g., various shelter and medical expenses that serve as deductions to income and affect benefits), and

[41] Unless noted otherwise, all information in this section comes from Pennsylvania Department of Human Services (online).

[42] Resources counted toward the limit include bank accounts, cash on hand, cars and motorcycles; many resources are excluded from the resource limit.

[43] Net income is gross income minus a standard deduction and certain exempt income (e.g., TANF benefits, loans, and interest on savings and checking accounts) and minus certain deductions (e.g., for earned income, dependent care, utilities excess medical expenses and excess shelter expenses).

[44] At the time of the intervention, households were required to submit an annual reporting form. Additionally, these households were required to report certain changes, such as when gross monthly income exceeds 130% of FPIG. In June 2016, the state announced a policy change for eldery (age 60 and older) and disabled households, which included a change in the re-certification process, which extended the time period to 36 months for our study sample. https://www.media.pa.gov/Pages/DHS_details.aspx?newsid=209

[45] Households who – by virtue of extreme need – qualify for expedited review must have their application reviewed within 5 calendar days of application.

the potential for mistakes in the process (e.g., not showing up for the interview, not filing the appropriate documentation of expenses, etc.) which cause an otherwise eligible application to be rejected or assigned a lower benefit amount. From the government's perspective, the needed information cannot be passively obtained, even if it had access to data on the individual from tax returns and other public benefit programs. In particular, three specific types of information are not available from other sources. The first is the definition of a household, which is a SNAP-specific definition: people who "live together and customarily prepare food together" (Gray et al., 2016). The household unit is required both to assess eligibility and to determine benefit amounts. Second, the resource limit that is applied to all non-categorically eligible households requires information on resources like bank accounts and second properties that are not readily available in other administrative data. Third, the calculation of net income – which is required in some cases to determine eligibility and in all cases to determine benefits – likewise can be affected by information not otherwise available (like excess out-of-pocket medical expenses and shelter expenses), although of course one could provide less information here and receive commensurately lower benefits. Underlining the difficulty of circumventing the active application process is the experience of the tax preparer Intuit (TurboTax), which in 2015 tried - through a program called Benefits Assist - to submit applications for SNAP on behalf of their low-income clients, using the information that had been provided on their tax returns. States encountered substantially increased administrative burden in response to the noticeable increase in applications, and it appeared that many of these applications were incomplete and could not be approved as filed.[46]

H: Related Literature

Our paper relates to two strands of literature: analysis of barriers to take-up and analysis of how barriers to take-up affect the *characteristics* of applicants and enrollees. Studies of barriers to take-up have, with the exception of Bettinger et al. (2012), focused on either informational barriers or transaction costs, rather than analyzing them together as we do here.[47] This literature has also focused primarily on the descriptive, with little of the normative analysis we add here.

Barriers to Take-up

Reductions in informational barriers have been found to be quantitatively important in generating take-up in some contexts but not others. In a recent series of randomized interventions aimed at

[46] See, e.g., `https://fns-prod.azureedge.net/sites/default/files/snap/State-Guidance-on-Intuit-SNAP-Applications.pdf`; `http://www.macssa.org/memberlogin/15minutes/selfsufficiency_dec15.pdf`; and `https://benkallos.com/press-release/memorandum-automatic-benefits-using-government-data-deliver-better-citizen-services-le`.

[47] The literature has paid comparatively less attention to the role of stigma, but the limited evidence does not point to a large role for stigma (Currie 2006). Recent efforts at "stigma" interventions have proven less successful at increasing take-up than informational interventions such as reminders or simplification (Bhargava and Manoli 2015). In the specific context of SNAP, Currie (2003) describes several pieces of survey evidence consistent with both lack of awareness and transaction costs in reducing SNAP take-up, but concludes that it does not appear from the existing survey evidence that stigma is a major deterrent to SNAP enrollment.

increasing take-up of the EITC among likely eligible individuals, Day Manoli and co-authors have found that take-up is highly sensitive to both the frequency and nature of reminder letters sent by the IRS, although the effects of the reminder do not persist into the following year when the individuals would have to sign up again (Bhargava and Manoli 2015, Manoli and Turner 2014, Guyton et al. 2016). Quasi-experimental studies have also found that information is an important barrier to take up of SSDI (Armour, forthcoming) and post-secondary enrollment among unemployment insurance recipients (Barr and Turner, forthcoming). Several of these studies conclude, as we do, that the results are consistent with misperceptions by individuals (see, e.g., Bhargava and Manoli 2015, Armour forthcoming). However, Alcott and Greenstone's (2017) randomized evaluation finds that informational interventions do not affect take-up of home energy efficiency audits, concluding that lack of awareness is not a contributor to low take-up; likewise, Bettinger et al.'s (2012) randomized evaluation finds that providing low-income families with information about financial aid eligibility and nearby colleges had no effect on applications to college, and Dynarski et al. (2018) find that an information intervention that informed high-acheving students about a tuition-free college scholarship increased enrollment at a flagship state university.

In the SNAP context, a survey of likely eligible SNAP non-participants found that about half reported that they were not aware of their eligibility (Bartlett et al. 2004). And in an early and innovative small randomized trial in 1993 in Pennsylvania, Daponte et al (1999) found suggestive evidence that informing non-participating, eligible households about their SNAP eligibility affected SNAP applications; however, small sample sizes (32 households were in the treatment arm and responded to the follow-up survey) as well as loss to follow-up made definitive conclusions difficult.

Reductions in transactional barriers have been found to be important for increasing enrollment in several different programs. Bettinger et al. (2012) found that while information alone was ineffective, combining information with assistance in completing a streamlined application process increased aid applications and ultimately college attendance and persistence by low-income individuals. Our findings in the SNAP context suggest, by contrast, that information alone can have an effect, but that pairing it with assistance doubles the impact. In addition, Deshpande and Li (2017) find that the closing of local field offices where SSDI and SSI applications can be submitted substantially reduces both applications and enrollment, and Rossin-Slater (2013) finds that openings and closings of Women, Infants and Children (WIC) local program office affect program participation. Alatas et al. (2016) present evidence from a randomized evaluation across Indonesian villages that increasing the transaction cost of applying for a conditional cash transfer program reduces enrollment. At the extreme of reducing transaction barriers, defaulting to enrollment has been found to have substantial effects on outcomes such as participation in tax-subsidized 401(k) savings plans (Madrian and Shea 2001).

In the SNAP context, Schanzenbach (2009) provides evidence from one California county of a randomized evaluation in which a low-income tax preparer provided assistance to likely SNAP-eligible individuals. It found that, among those who expressed interest in learning more about SNAP, those in offices randomized into full assistance (in which the tax preparer went through a

detailed interview with the client and then filled out and filed the application on the client's behalf), were more likely to file an application than those who received help filling out the application but had to file it themselves, or those who only received a blank application (which might be viewed as analogous to our "Information Only" intervention).

In addition to the Schanzenbach (2009) RCT, there are a number of other studies looking at SNAP take-up using quasi-experimental research designs and either survey data or administrative data. Kabbani and Wilde (2003) use a within-state difference-in-difference design to estimate the effect of short recertification periods on SNAP participation rate (as well as "error rates" as measured in quality control data). They find that shorter recertification periods (which can be interpreted as a kind of transaction barrier) reduce participation rates as well as error rates, and they also look separately at working and non-working households. Changes in recertification period length is a change in the "ordeal" associated with program participation, and the model that we develop in this paper could be used calibrate the welfare consequences of these recent policy changes.

Ratcliffe et al. (2008) and Hanratty (2006) use the SIPP 1996 and 2001 individual-level panel survey to study the effect of various SNAP policies on participation, including outreach spending that can be viewed as a kind of information intervention . Consistent with Kabbani and Wilde (2003), Hanratty (2006) finds evidence that "certification requirements" affect participation rates, while Ratcliffe et al. (2008) study a large number of SNAP policies (eligibility requirements, recertification periods, outreach, biometrics). The study reports some suggestive evidence that "outreach spending" affects SNAP receipt, and our model could also be used to study the welfare effects of this outreach spending (similar to our analysis of the information treatment in our experiment). Lastly, Klerman and Danielson (2011) use a difference-in-difference research design and monthly administrative data to estimate the effect of SNAP policies, TANF policies, and local economic conditions on SNAP caseloads. They find strong evidence that economic conditions affect SNAP caseloads, and find the SNAP and TANF policies account for a large share of the overall shift to a "food-based safety net" during the 1990s and early 2000s.

Targeting

Our paper also relates to a second strand of the literature that investigates how barriers to enrollment affect the *characteristics* of applicants and enrollees. The existing "targeting" literature has been primarily descriptive, focusing on the observable characteristics of individuals affected by different barriers. Our theoretical framework, however, suggests that there is no general relationship between this targeting on observables and the impact of the intervention on either private or social welfare. We provide additional conditions that need to be examined empirically in order for an intervention's targeting properties to yield normative implications.

To our knowledge, our study is the first to examine the targeting properties of both an information intervention and an assistance intervention. From prior information interventions, there is evidence that complexity disproportionately deters EITC enrollment of lower income potential recipients (Bhargava and Manoli 2015), and that lower income employees are more likely to choose

dominated health insurance plans, due at least in part to a lack of insurance literacy (Bhargava et al., 2017). Our findings, by contrast, suggest that information about eligibility disproportionately encourages enrollment among less needy applicants.

Prior studies have tended to find that transaction costs increase targeting on some but not all dimensions. Alatas et al. (2016) find that introducing transaction costs by requiring individuals to apply for a conditional cash transfer in Indonesian villages rather than have the government automatically screen the individuals for eligibility improves targeting; specifically, it results in substantially poorer enrollees. However, marginally increasing the transaction costs does not further affect the characteristics of enrollees. Deshpande and Li (2017) find that increasing transaction costs in U.S. disability programs (SSDI and SSI) worsens targeting among applicants - by increasing the share of applicants with only moderately severe disabilities - but increases targeting among enrollees, decreasing the share of enrollees with the least severe disabilities (conditional on being severe enough to be eligible); however they also find that the increased transaction costs reduce the share of enrollees with low education levels and low pre-application earnings, suggesting a reduction in targeting. In our context, by contrast, we find that reducing transaction costs decrease targeting on all dimensions, and at all stages (application and enrollment).

I. Appendix References

Alberto Abadie, 2002. “Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models", Journal of the American Statistical Association 97(457), pp. 284-292.

Alberto Abadie, 2003. "Semiparametric instrumental variable estimation of treatment response models", Journal of Econometrics 113, pp. 231-263.

Alatas Vivi, Abhijit Banerjee, Rema Hanna, Benjamin Olken, Matthew Wai-Poi and Ririn Pernamasari. 2016. “Self-Targeting: Evidence from a Field Experiment in Indonesia.” Journal of Political Economy 124 (2): 371-427.

Alcott, Hunt and Michael Greenstone. 2017. “Measuring The Welfare Effects of Residential Energy Efficiency Programs.” NBER Working Paper 23386.

Armour, Philip. forthcoming. “The Role of Information in Disability Insurance Application: An Analysis of the Social Security Statement Phase-In.” American Economic Journal: Economic Policy.

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

Bartlett, Susan, Nancy Burstein, and William Hamilton. 2004. “Food Stamp Access Study: Final Report”. United States Department of Agriculture, Economic Research Service, November. https://www.ers.usda.gov/publications/pub-details/?pubid=43407 (last accessed July 31, 2017).

Bertrand, M., Mullainathan, S., Shafir, E., 2004. A behavioral-economics view of poverty. Am. Econ. Rev. 419–423.

Bettinger, E.P., Long, B.T., Oreopoulos, P., Sanbonmatsu, L., 2012. The role of application

assistance and information in college decisions: Results from the H&R Block FAFSA experiment. Q. J. Econ. 127, 1205–1242.

Bhargava, S., Manoli, D., 2015. Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment. Am. Econ. Rev. 105, 3489–3529. doi:10.1257/aer.20121493

Bhargava, S, Loewenstein, G and J. Sydnor. 2017. "Choose to Lose: Health Plan Choices from a Menu with Dominated Options." Quarterly Journal of Economics August. Volume 132 (3): 1319-1372.

Currie, J., 2006. The Take Up of Social Benefits. In Alan Auerbach, David Card, And John Quigley (Eds). Poverty, the Distribution of Income, and Public Policy (New York: Russell Sage), 80-148.

Currie, J., 2003. US food and nutrition programs, in: Means-Tested Transfer Programs in the United States. University of Chicago Press, pp. 199–290.

Currie, J., Gahvari, F. 2008. Transfers in Cash and In-Kind: Theory Meets the Data. Journal of Economic Literature, 46(2), 333-383.

Daponte, B.O., Sanders, S., Taylor, L., 1999. Why do low-income households not use food stamps? Evidence from an experiment. J. Hum. Resour. 612–628.

Deshpande, Manasi and Yue Li. 2017. "Who is Screened Out? Application Costs and the Targeting of Disability Programs." Unpublished Miemo. `https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxtZGVzaHBhbmRlZWNvbnxneDo3ZTU1MGYyNTdhMWFhNmY3`

Dynarski S., C. J. Libassi, K. Michelmore, S. Owen. 2018. "Closing the Gap: The Effect of a Targeted, Tuition-Free Promise on College Choices of High-Achieving, Low-Income Students," NBER Working Paper No. 25349.

Finkelstein A, Gentzkow M, Williams H. 2016. Sources of geographic variation in health care: evidence from patient migration. Q J Econ;131:1681-726.

Gray, K., Fisher, S., Lauffer, S., 2016. Characteristics of Supplemental Nutrition Assistance Program Households: Fiscal Year 2015 (No. Report No. SNAP-16-CHAR). United States Department of Agriculture.

Guyton, John, Dayanand Manoli, Brenda Schafer and Michael Sebastiani. 2016. "Reminders & Recidivism: Evidence from Tax Filing & EITC Participation Among Low-Income NonFilers." NBER Working Paper 21904.

Hanratty, M. J. 2006, Has the Food Stamp program become more accessible? Impacts of recent changes in reporting requirements and asset eligibility limits. J. Pol. Anal. Manage., 25: 603-621.

Hendren, N. (2016). The policy elasticity. Tax Policy and the Economy 30 (1), 51-89.

Hwang et al. 2001. Out-Of-Pocket Medical Spending for Care of Chronic Conditions. Health Affairs. 20 (6): 267-278.

Kabbani, N., & Wilde, P. 2003. Short Recertification Periods in the U.S. Food Stamp Program. The Journal of Human Resources, 38, 1112-1138.

Klerman, J. A. and Danielson, C. 2011. The transformation of the Supplemental Nutrition

Assistance Program. J. Pol. Anal. Manage., 30: 863-888.

Lee, D. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," Review of Economic Studies, 76(3): 1071-1102.

Lewin Group. 2012. "Evaluating Encounter Data Completeness." `https://www.ccwdata.org/documents/10280/19002254/evaluating-encounter-data-completeness.pdf`

Madrian, Brigitte and Dennis Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior" Quarterly Journal of Economics. CXVI (4): 1149-1187.

Manoli, Dayand and Nicholas Turner. 2014. "Nudges and Learning: Evidence from Informational Interventions for Low-Income Taxpayers." NBER Working Paper 20718.

Nichols, A.L., Zeckhauser, R.J., 1982. Targeting transfers through restrictions on recipients. Am. Econ. Rev. 372–377.

Ratcliffe, C., McKernan, S., & Finegold, K. 2008. Effects of Food Stamp and TANF Policies on Food Stamp Receipt. Social Service Review, 82(2), 291-334. doi:10.1086/589707

Rifkin, Ryan and Aldebaro Klautau. 2004. "In Defense of One-Vs-All Classification", Journal of Machine Learning Research, 5 (January):101-141.

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

Schanzenbach, Diane. 2009. "Experimental Estimates of the Barriers to Food Stamp Enrollment." Instituet for Research on Poverty, University of Wisconsin-Madison, Discussion Paper no. 1367-09.

Tauchmann, Harald. 2014. Lee (2009) Treatment-Effect Bounds for Nonrandom Sample Selection. The Stata Journal, 14, Number 4, pp. 884–894.