Online Appendix to

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

Amy Finkelstein and Matthew J. Notowidigdo

May 2018

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 1 shows the letter, envelope and postcard that were sent in this treatment arm. Appendix Figure 2 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 3 shows the study design with all of the treatments and sub-treatments.[42]

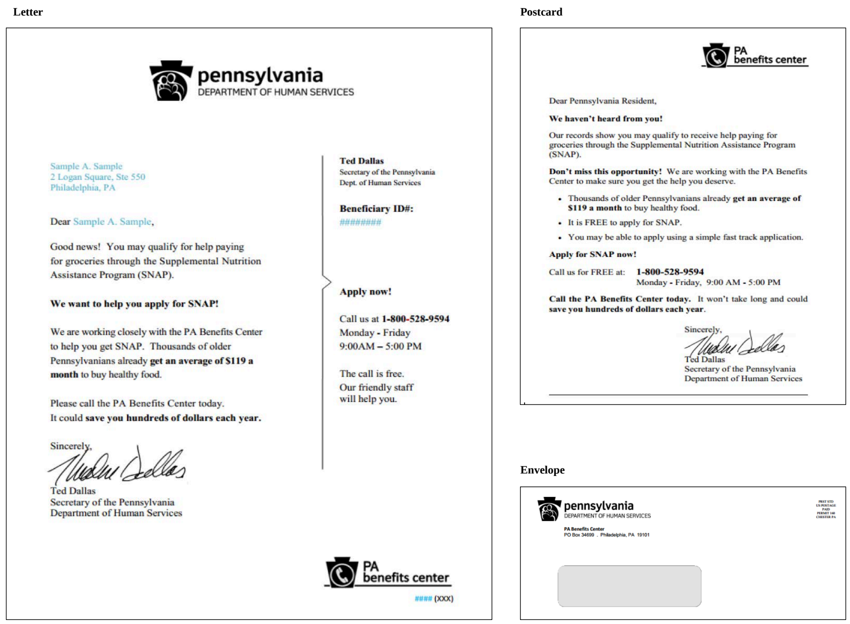
One sub-treatment was a "marketing" intervention. One-quarter of each treatment was randomized into an arm with a variant of the outreach letters and postcards designed to attract clients by using a "marketing" approach that borrowed language and graphics from credit card solicitations in an attempt 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 treatments in the Information Plus Assistance arm, and the control. We down-weight the individuals in the standard treatmentin the Information Plus Assistance arm so that the (weighted) share in stan-

[42] Appendix Table A2 showed balance across each the three main groups we study on the study characteristics we examined in Table 1; for completeness, Appendix Table A3 shows balance of characteristics across the sub-treatments.

Figure A1: Standard Outreach Materials: Information Plus Assistance



dard vs. marketing is the same (50 percent) in the Information Plus Assistance and Information Only arms.

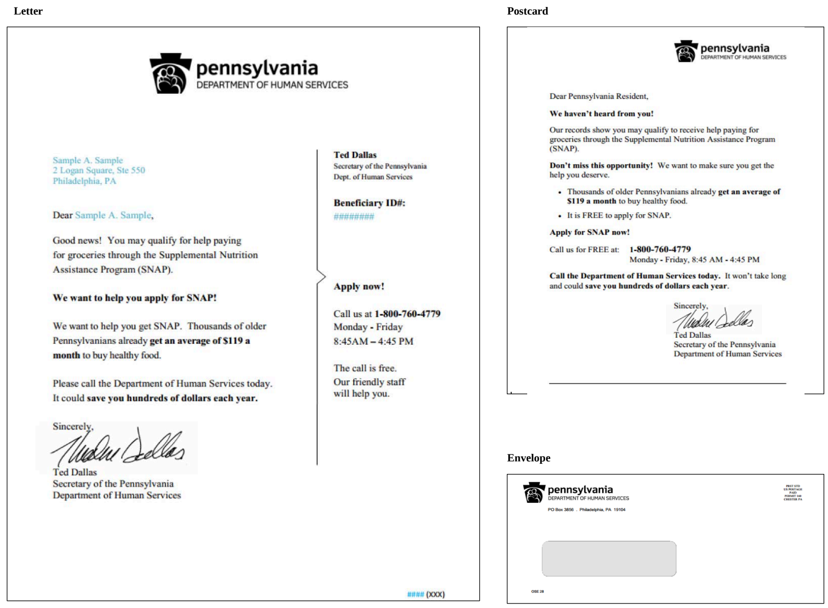
B: DHS Data

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


Letter

pennsylvania
DEPARTMENT OF HUMAN SERVICES

Sample A. Sample
2 Logan Square, Ste 550
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 an average of $119 a month** to buy healthy food.

Please call the Department of Human Services today. It 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 at **1-800-760-4779**
Monday - Friday
8:45AM – 4:45 PM

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

(XXX)

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 an average of $119 a month** to buy healthy food.
- It is FREE to apply for SNAP.

Apply for SNAP now!

Call us for FREE at: **1-800-760-4779**
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**.

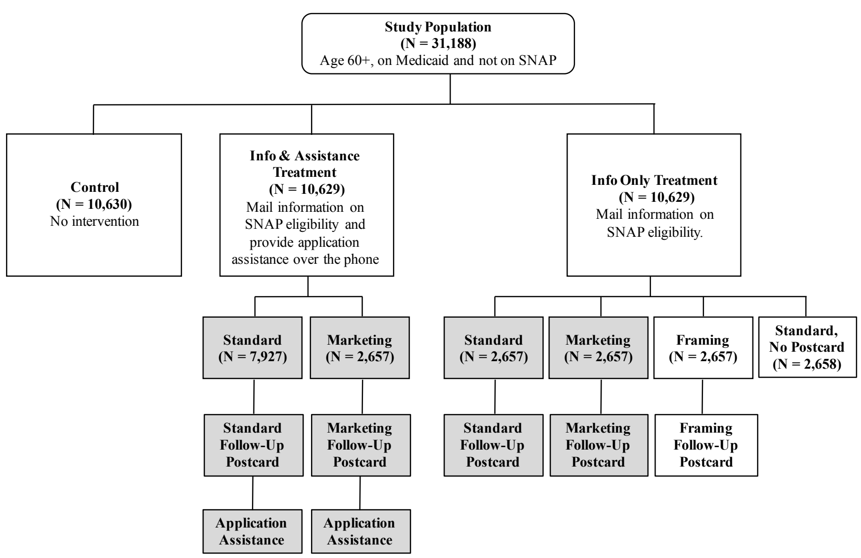
Sincerely,

Ted Dallas
Secretary of the Pennsylvania
Department of Human Services

Envelope

pennsylvania
DEPARTMENT OF HUMAN SERVICES

Figure A3: Experimental Design



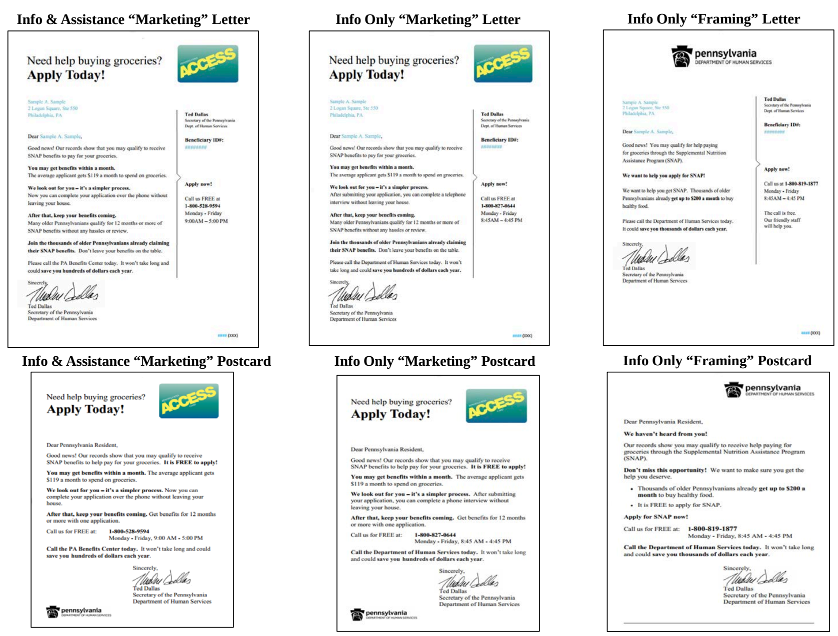
Notes: Figure shows experimental design. Grey arms are the ones included in the main analyses.

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



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[43]. The taxonomy codes are then matched to HCFA specialty codes using a crosswalk from CMS[44]. Finally, HFCA codes are matched to a primary care or specialist classification using a crosswalk from Finkelstein et al. (2016)[45].

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

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

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

[45]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 codes for diagnoses. We identify which ICD-9 codes correspond to chronic conditions following the method developed by Hwang et al. (2001)[46]. 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

[46]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 A7.

D: Additional Analyses

Generating Predicted Benefits

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 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 A1 . 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.

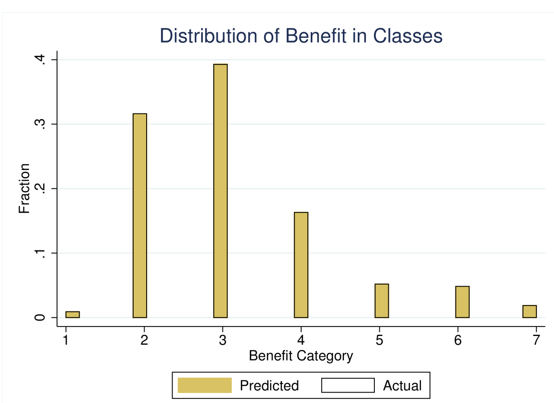
Appendix Figures A6 shows our fit, which shows very close match across categories by design. This confirms that the algorithm appears to be unbiased in its predictions. To assess accuracy of the predictions, we calculate that roughly 38% of the observations are categorized correctly, and 78% of the predicted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

Table A1: 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	

Figure A6: Predicted and Actual Enrollee Monthly Benefits

Panel A: In Categories



Panel B: In Dollars

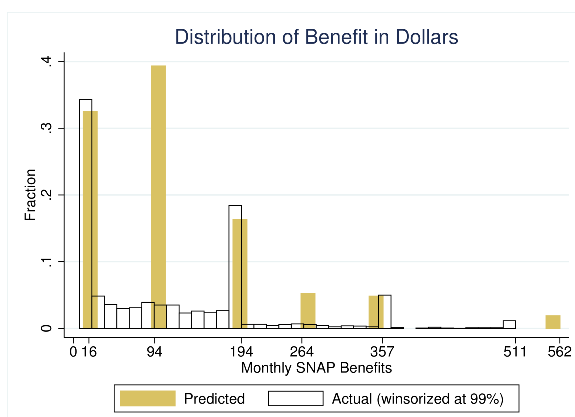
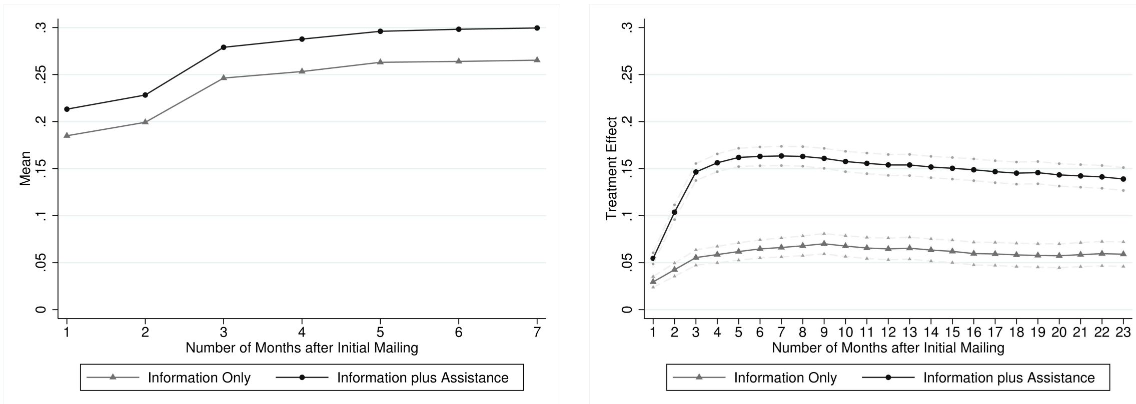


Figure A7: Time pattern of callers and applications

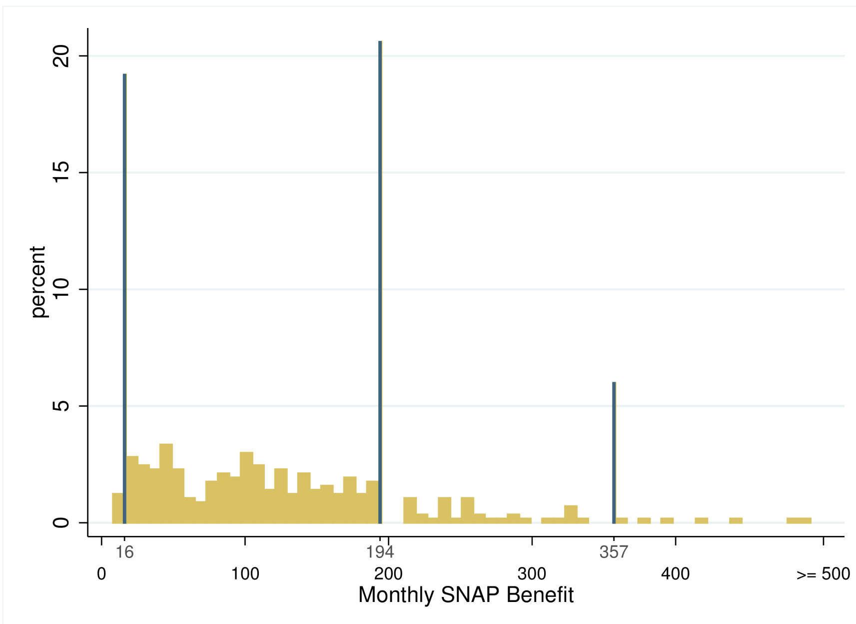
Panel A: Callers Panel B: Applicants



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.

Additional Results

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 mothly benefit for a household of size 2 ($357) - are binned separately from the other values; values greater than $500 are set to $500.

Table A2: 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	

Table A3: 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 A4: 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 A5: 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	0.178	0.295	0.282	0.303	0.247	0.359	0.355				
		[0.000]	[0.002]	[0.000]	[0.019]	[0.000]	[0.000]	[0.000]	[0.905]	[0.191]	[0.727]
Share $194 Benefit	0.191	0.154	0.176	0.161	0.146	0.145	0.141				
		[0.165]	[0.597]	[0.281]	[0.083]	[0.013]	[0.026]	[0.017]	[0.959]	[0.770]	[0.511]
Share $357 Benefit	0.055	0.034	0.057	0.069	0.064	0.038	0.040				
		[0.117]	[0.924]	[0.459]	[0.600]	[0.090]	[0.228]	[0.230]	[0.136]	[0.082]	[0.195]
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 A6: Always Taker and Complier Enrollee Benefits and Predicted Benefits

	Always Takers	Compliers: Information Only	Compliers: Information Plus Assistance Arms	P Value of Difference (Column 2 vs 3)
	(1)	(2)	(3)	(4)
Benefit Amount	145.94	78.31	79.66	
		[0.000]	[0.000]	[0.910]
Share $16 Benefit	0.178	0.445	0.444	
		[0.000]	[0.000]	[0.973]
Share $194 Benefit	0.191	0.117	0.120	
		[0.141]	[0.008]	[0.933]
Share $357 Benefit	0.055	0.044	0.031	
		[0.687]	[0.113]	[0.583]
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 A7: 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 A8: 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 A9: 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 A10: Demographic and Health Characteristics for Always Takers and Complier Applicants and Enrollees

	Applicants				Enrollees			
	Always Takers	Compliers: Info Only	Compliers: Info Plus Assistance	P Value of Difference (Column 2 vs 3)	Always Takers	Compliers: Info Only	Compliers: Info Plus Assistance	P Value of Difference (Column 6 vs 7)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A - Predicted Benefits								
Predicted Benefits	148.26	100.85	99.65		138.65	84.10	87.21	
		[0.000]	[0.000]	[0.910]		[0.001]	[0.000]	[0.788]
Panel B - (Annual) Health Care Measures, 2015								
Total Health Care Spending ($) [b]	9,424	7,707	7,813		10,238	8,676	7,809	
		[0.533]	[0.314]	[0.937]		[0.675]	[0.213]	[0.767]
Total Number of Visits and Days	13.33	9.84	8.29		14.79	6.19	7.56	
		[0.331]	[0.011]	[0.576]		[0.049]	[0.004]	[0.632]
Weighted Total Number of Visits and Days	4,661	1,752	1,938		5,407	716	1,504	
		[0.117]	[0.011]	[0.857]		[0.054]	[0.004]	[0.604]
Number of Chronic Conditions	6.21	4.83	4.83		6.54	4.07	4.80	
		[0.094]	[0.006]	[0.999]		[0.020]	[0.006]	[0.381]
Panel C - Demographics								
Share Age 80+	0.06	0.16	0.18		0.07	0.17	0.18	
		[0.000]	[0.000]	[0.646]		[0.005]	[0.000]	[0.815]
Male	0.41	0.40	0.37		0.39	0.44	0.37	
		[0.994]	[0.228]	[0.394]		[0.443]	[0.436]	[0.155]
Share White [a]	0.67	0.80	0.78		0.71	0.87	0.82	
		[0.004]	[0.000]	[0.540]		[0.003]	[0.000]	[0.213]
Share Black [a]	0.10	0.05	0.11		0.11	0.02	0.10	
		[0.101]	[0.587]	[0.011]		[0.011]	[0.845]	[0.003]
Share Primary Language not English	0.08	0.04	0.02		0.06	0.03	0.01	
		[0.133]	[0.000]	[0.232]		[0.232]	[0.002]	[0.482]
Share Living in Pittsburgh	0.05	0.07	0.08		0.05	0.08	0.09	
		[0.389]	[0.068]	[0.796]		[0.366]	[0.029]	[0.696]
Share Last Medicaid Spell Starting before 2011	0.25	0.35	0.31		0.26	0.42	0.34	
		[0.018]	[0.014]	[0.269]		[0.006]	[0.024]	[0.087]
Share of Sub-Population	0.077	0.070	0.161		0.058	0.048	0.119	

Notes: Sample in columns 1-4 is individuals who applied for SNAP within 9 months of their initial mailing, and in columns 5-8 it is individuals who enrolled in SNAP within 9 months of their initial mailing. Column 1 (respectively, 5) shows the mean of the always takers (individuals who apply (respectively, enroll) regardless of intervention), while columns 2 and 3 (respectively 6 and 7) show the means for compliers (individuals who apply (respectively, 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 (respectively, 8) reports the p-value of the difference between the compliers in the two intervention arms. In columns 2 and 6 the two equally-sized sub-treatments are pooled; in columns 3 and 7 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]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 A11: Demographic and Health Characteristics of Callers

	Information Only	Information Plus Assistance	P Value of Difference (Column 2 vs 3)
	(1)	(2)	(3)
Panel A - Predicted Benefits			
Predicted Benefits	104.99	108.76	[0.286]
Panel B - (Annual) Health Care Measures, 2015			
Total Health Care Spending ($) [b]	6,779	7,792	[0.074]
Total Number of Visits and Days	10.16	8.96	[0.194]
Weighted Total Number of Visits and Days	3,167	2,575	[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 [a]	0.79	0.76	[0.044]
Share Black [a]	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] 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 A12: Demographic and Health Characteristics by Sub-Treatment: Applicants

								P Value of Difference between			
	Control	Information Only				Information Plus Assistance		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)
		Standard	No-Postcard	Marketing	Framing	Standard	Marketing				
	(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 A13: Demographic and Health Characteristics by Sub-Treatment: Enrollees

								P Value of Difference between			
	Control	Information Only				Information Plus Assistance		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)
		Standard	No-Postcard	Marketing	Framing	Standard	Marketing				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
<u>Panel A - Predicted Benefits</u>											
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]
<u>Panel B - (Annual) Health Care Measures, 2015</u>											
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]
<u>Panel C - Demographics</u>											
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 A14: 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 A15: 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.

E: Proofs of Propositions

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

E.1 Proofs of Propositions in Main Text

Welfare Effects of Interventions

Proposition 1. *Let $\mu_j \equiv -u'(y_j)(\pi_j B_j)\epsilon_j$, which equals the marginal utility of income times the expected benefit of applying times the misperception of the application probability. Note that μ_j is positive if ϵ_j is negative, i.e., if agents under-estimate the probability of acceptance. The effect of the Information Only treatment on welfare is given by the following expression:*

$$\frac{dW}{dT}^{Info\,Only} = \underbrace{\mu_l \frac{dA_l}{dT} + \mu_h \frac{dA_h}{dT}}_{Change\ in\ Private\ Welfare} - \underbrace{\left[(\pi_l B_l + g_h)\frac{dA_l}{dT} + (\pi_h B_h + g_l)\frac{dA_h}{dT}\right]}_{Change\ in\ Public\ Cost}$$

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

$$\frac{dW}{dT}^{Info\,+\,Assistance} = \underbrace{\mu_l \frac{dA_l}{dT} + \mu_h \frac{dA_h}{dT} + u'(y_l)A_l + u'(y_h)A_h}_{Change\ in\ Private\ Welfare} - \underbrace{\left[(\pi_l B_l + g_h)\frac{dA_l}{dT} + (\pi_h B_h + g_l)\frac{dA_h}{dT}\right]}_{Change\ in\ Public\ Cost}.$$

Proof: Welfare (assuming $\epsilon_h = \epsilon_l = \epsilon$) is given by

$$\begin{aligned}
W &= V_l + V_h - [(\pi_l B_l + g_l)A_l + (\pi_h B_h + g_h)A_h] \\
&\approx u(y_l) + \int_0^{(1+\epsilon)\pi_l B_l} u'(y_l)(\pi_l B_l - c)f_l(c)dc + u(y_h) + \int_0^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c)f_h(c)dc \\
&\quad - [((\pi_l B_l + g_l)A_l + (\pi_h B_h + g_h)A_h].
\end{aligned}$$

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

$$\begin{aligned}
\frac{dW}{d\epsilon} &= \frac{d}{d\epsilon}\int_0^{(1+\epsilon)\pi_l B_l} u'(y_l)(\pi_l B_l - c)f_l(c)dc + \frac{d}{d\epsilon}\int_0^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c)f_h(c)dc \\
&\quad - \left[(\pi_l B_l + g_l)\frac{dA_l}{d\epsilon} + (\pi_h B_h + g_h)\frac{dA_h}{d\epsilon}\right].
\end{aligned}$$

Applying Leibniz's Rule, we get

$$\frac{d}{d\epsilon} \int_0^{(1+\epsilon)\pi_l B_l} u'(y_l)(\pi_l B_l - c) f_l(c) dc = u'(y_l)(-\epsilon \pi_l B_l) f_l((1+\epsilon)\pi_l B_l)\pi_l B_l.$$

Similarly,

$$\frac{d}{d\epsilon} \int_0^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc = u'(y_h)(-\epsilon \pi_h B_h) f_h((1+\epsilon)\pi_h B_h)\pi_h B_h.$$

Since the number of applicants is given by $A_h = F_h((1+\epsilon)\pi_h B_h)$ and $A_l = F_l((1+\epsilon)\pi_l B_l)$,

$$\frac{dA_h}{d\epsilon} = f_h((1+\epsilon)\pi_h B_h)\pi_h B_h$$

and

$$\frac{dA_l}{d\epsilon} = f_l((1+\epsilon)\pi_l B_l)\pi_l B_l.$$

Therefore, we can re-write

$$\frac{d}{d\epsilon} \int_0^{(1+\epsilon)\pi_l B_l} u'(y_l)(\pi_l B_l - c) f_l(c) dc = u'(y_l)(-\epsilon \pi_l B_l) \frac{dA_l}{d\epsilon}$$

and

$$\frac{d}{d\epsilon} \int_0^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc = u'(y_h)(-\epsilon \pi_h B_h) \frac{dA_h}{d\epsilon}.$$

Putting all this together, we have

$$\frac{dW}{d\epsilon} = \mu_l \frac{dA_l}{d\epsilon} + \mu_h \frac{dA_h}{d\epsilon} - \left[(\pi_l B_l + g_l)\frac{dA_l}{d\epsilon} + (\pi_h B_h + g_h)\frac{dA_h}{d\epsilon} \right].$$

Since $dT = d\epsilon$, we know that

$$\frac{dW}{dT} = \mu_l \frac{dA_l}{dT} + \mu_h \frac{dA_h}{dT} - \left[(\pi_l B_l + g_l)\frac{dA_l}{dT} + (\pi_h B_h + g_h)\frac{dA_h}{dT} \right].$$

This completes the first part of the proposition.

Assistance Only ($dT = -dc$): Define $-dc$ as a downward shift in every applicant's value of c.

First, let's focus on the l types. Suppose the entire distribution $f_l(c)$ shifts down by Δc. Define W_l as the welfare associated with the l types (the sum of private welfare V_l and public costs

$(\pi_l B_l + g_l)A_l)$. How does W_l change with $-\Delta c$?

$$\Delta W_l = \left[\underbrace{\int_0^{(1+\epsilon)\pi_l B_l+\Delta c} u'(y_l)(\pi_l B_l - (c-\Delta c))f_l^{\Delta}(c-\Delta c)dc - (\pi_l B_l + g_l)F_l((1+\epsilon)\pi_l B_l + \Delta c)}_{W_l\ after}\right]$$
$$- \left[\underbrace{\int_0^{(1+\epsilon)\pi_l B_l} u'(y_l)(\pi_l B_l - c)f_l(c)dc - (\pi_l B_l + g_l)F_l((1+\epsilon)\pi_l B_l)}_{W_l\ before}\right]$$

where $f_l^{\Delta}(\cdot)$ is the shifted cost PDF. Note that $f_l^{\Delta}(c-\Delta c) = f_l(c)$ at any value of c. Therefore, by combining the integrals on the region $[0,(1+\epsilon)\pi_l B_l]$, we can re-write

$$\begin{aligned}\Delta W_l &= u'(y_l)\Delta c \int_0^{(1+\epsilon)\pi_l B_l} f_l(c)dc + u'(y_l)\int_{(1+\epsilon)\pi_l B_l}^{(1+\epsilon)\pi_l B_l+\Delta c} (\pi_l B_l - (c-\Delta c))f_l(c)dc \\ &\quad + (\pi_l B_l + g_l)\left[F_l((1+\epsilon)\pi_l B_l) - F_l((1+\epsilon)\pi_l B_l + \Delta c)\right].\end{aligned}$$

Next, note that

$$\begin{aligned}F_l((1+\epsilon)\pi_l B_l) - F_l((1+\epsilon)\pi_l B_l + \Delta c) &= \int_0^{(1+\epsilon)\pi_l B_l} f_l(c)dc - \int_0^{(1+\epsilon)\pi_l B_l+\Delta c} f_l(c)dc \\ &= -\int_{(1+\epsilon)\pi_l B_l}^{(1+\epsilon)\pi_l B_l+\Delta c} f_l(c)dc\end{aligned}$$

and so we can re-write

$$\begin{aligned}\Delta W_l &= u'(y_l)\Delta c \int_0^{(1+\epsilon)\pi_l B_l} f_l(c)dc + u'(y_l)\int_{(1+\epsilon)\pi_l B_l}^{(1+\epsilon)\pi_l B_l+\Delta c} (\pi_l B_l - (c-\Delta c))f_l(c)dc \\ &\quad - (\pi_l B_l + g_l)\int_{(1+\epsilon)\pi_l B_l}^{(1+\epsilon)\pi_l B_l+\Delta c} f_l(c)dc.\end{aligned}$$

Taking this expression, dividing by Δc, and taking the limit as $\Delta c \to 0$ yields $-\frac{dW}{dc}$. However, both the numerator and denominator approach zero (i.e. the fraction is indeterminate), so we apply L'Hopital's Rule. The denominator is one, so we just need to differentiate each term of the numerator. This yields:

$$\frac{d}{d(\Delta c)}\left[u'(y_l)\Delta c\int_0^{(1+\epsilon)\pi_l B_l} f_l(c)dc\right] = u'(y_l)\int_0^{(1+\epsilon)\pi_l B_l} f_l(c)dc$$

$$\frac{d}{d(\Delta c)}\left[u'(y_l)\int_{(1+\epsilon)\pi_l B_l}^{(1+\epsilon)\pi_l B_l+\Delta c}(\pi_l B_l-(c-\Delta c))f_l(c)dc\right] = u'(y_l)(-\epsilon\pi_l B_l)f_l((1+\epsilon)\pi_l B_l+\Delta c)\ \text{(Leibniz)}$$

$$\frac{d}{d(\Delta c)}\left[-(\pi_l B_l+g_l)\int_{(1+\epsilon)\pi_l B_l}^{(1+\epsilon)\pi_l B_l+\Delta c} f_l(c)dc\right] = -(\pi_l B_l+g_l)f_l((1+\epsilon)\pi_l B_l+\Delta c)\ \text{(Leibniz)}$$

Taking each term as $\Delta c \to 0$ and combining yields

$$\begin{aligned}-\frac{dW_l}{dc} &= u'(y_l)\int_0^{(1+\epsilon)\pi_l B_l} f_l(c)dc + u'(y_l)(-\epsilon\pi_l B_l)f_l((1+\epsilon)\pi_l B_l) - (\pi_l B_l+g_l)f_l((1+\epsilon)\pi_l B_l)\\ &= u'(y_l)F_l((1+\epsilon)\pi_l B_l) + \mu_l f_l((1+\epsilon)\pi_l B_l) - (\pi_l B_l+g_l)f_l((1+\epsilon)\pi_l B_l).\end{aligned}$$

When you lower c, you generate three terms:

1. $u'(y_l)F_l((1+\epsilon)\pi_l B_l)$ represents the utility gained by a fraction $F_l((1+\epsilon)\pi_l B_l)$ of agents.

2. $\mu_l f_h((1+\epsilon)\pi_l B_l)$ represents the additional utility gained by the marginal agent. This is because while the marginal agent thinks he is indifferent, if $\epsilon_l \neq 0$ (he misperceives), he actually should not be indifferent. Therefore, by applying, he gains utility (the envelope theorem does not apply). $f_l((1+\epsilon)\pi_l B_l)$ is the number of marginal agents, and μ_l is the utility that each agent gains.

3. $(\pi_l B_l + g_l)f_l((1+\epsilon)\pi_l B_l)$ represents the application costs paid due to the marginal agent. $f_l((1+\epsilon)\pi_l B_l)$ is the number of marginal agents, $\pi_l B_l$ is the expected benefit per applicant, and g_l is the processing cost per applicant.

Similarly,

$$-\frac{dV_h}{dc} = u'(y_h)F_h((1+\epsilon)\pi_h B_h) + \mu_h f_h((1+\epsilon)\pi_h B_h) - (\pi_h B_h+g_h)f_h((1+\epsilon)\pi_h B_h).$$

Putting this together implies

$$\begin{aligned}\frac{dW}{dT} &= -\frac{dW}{dc}\\ &= -\frac{dW_h}{dc} - \frac{dW_l}{dc}\\ &= u'(y_l)F_l((1+\epsilon)\pi_l B_l) + \mu_l f_l((1+\epsilon)\pi_l B_l) - (\pi_l B_l+g_l)f_l((1+\epsilon)\pi_l B_l)\\ &+ u'(y_h)F_h((1+\epsilon)\pi_h B_h) + \mu_h f_h((1+\epsilon)\pi_h B_h) - (\pi_h B_h+g_h)f_h((1+\epsilon)\pi_h B_h)\\ &= \mu_l\frac{dA_l}{dT} + u'(y_l)A_l + \mu_h\frac{dA_h}{dT} + u'(y_h)A_h - \left[(\pi_l B_l+g_l)\frac{dA_l}{dT} + (\pi_h B_h+g_h)\frac{dA_h}{dT}\right].\end{aligned}$$

Information + Assistance ($dT = d\epsilon, -dc$):

Combining information and assistance yields

$$\frac{dW}{dT} = \mu_l \frac{dA_l}{dT} + u'(y_l)A_l + \mu_h \frac{dA_h}{dT} + u'(y_h)A_h - \left[(\pi_l B_l + g_l)\frac{dA_l}{dT} + (\pi_h B_h + g_h)\frac{dA_h}{dT}\right]$$

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 high-benefit targeting from an intervention (in either the Information Only or Information Plus Assistance treatments) is giving by the following:*

$$\frac{\partial}{\partial(dlog(e)/dT)} \left(\frac{dW}{dT}\right)\bigg|_{\frac{dA}{dT}} = \left[(\mu_h - \mu_l) - (\pi_h B_h + g_h) - (\pi_l B_l + g_l)\right] \frac{E_H(E_H + E_L)}{E_H(\pi_{lL} - \pi_{hL}) + E_L(\pi_{hH} - \pi_{lH})}.$$

Proof: Define $\mu_j = -u'(y_j)(\pi_j B_j)\epsilon_j$. Since $e = \frac{E_H}{E_H + E_L}$, $\log(e) = \log(E_H) - \log(E_H + E_L)$ and

$$\begin{aligned}\frac{d\log(e)}{dT} &= \frac{1}{E_H}\frac{dE_H}{dT} - \frac{1}{E_H + E_L}\left(\frac{dE_H}{dT} + \frac{dE_L}{dT}\right) \\ &= \frac{(E_H + E_L)\frac{dE_H}{dT} + E_H\left(\frac{dE_H}{dT} + \frac{dE_L}{dT}\right)}{E_H(E_H + E_L)}\end{aligned}$$

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

$$\begin{aligned}\frac{dW}{dT} &= \mu_l \frac{dA_l}{dT} + \mu_h \frac{dA_h}{dT} - \left[(\pi_l B_l + g_l)\frac{dA_l}{dT} + (\pi_h B_h + g_h)\frac{dA_h}{dT}\right] \\ &= (\mu_l - \pi_l B_l - g_l)\frac{dA_l}{dT} + (\mu_h - \pi_h B_h - g_h)\frac{dA_h}{dT} \\ &= (\mu_l - \pi_l B_l - g_l)\left(\frac{dA_l}{dT} + \frac{dA_h}{dT}\right) + (\mu_h - \pi_h B_h - g_h - \mu_l + \pi_l B_l + g_l)\frac{dA_h}{dT} \\ &= (\mu_l - \pi_l B_l - g_l)\frac{dA}{dT} + \left[(\mu_h - \mu_l) - (\pi_h B_h + g_h) - (\pi_l B_l + g_l)\right]\frac{dA_h}{dT}.\end{aligned}$$

Now, since dA/dT is held constant (so that $dA_l/dT + dA_h/dT = 0$), we just need to see how the right-hand term varies with $dlog(e)/dT$. We begin with the following derivatives given model assumptions (all derivatives dX implicitly refer to dX/dT):

$$
\begin{aligned}
dE_H &= \pi_{hH} dA_h + \pi_{lH} dA_l \\
dE_L &= \pi_{hL} dA_h + \pi_{lL} dA_l
\end{aligned}
$$

These equations can be used to solve for dA_h in terms of dE_H and dE_L:

$$
\begin{aligned}
dA_h &= (\pi_{lL} dE_H - \pi_{lH} dE_L)/\Delta \\
dA_l &= (-\pi_{hL} dE_H + \pi_{hH} dE_L)/\Delta
\end{aligned}
$$

where $\Delta = \pi_{hH}\pi_{lL} - \pi_{lH}\pi_{hL} > 0$. Since $dA_l = -dA_h$, this can be used to solve for dE_H in terms of dE_L:

$$
\begin{aligned}
(\pi_{hL} dE_H - \pi_{hH} dE_L)/\Delta &= (\pi_{lL} dE_H - \pi_{lH} dE_L)/\Delta \\
(\pi_{lL} - \pi_{hL}) dE_H &= -(\pi_{hH} - \pi_{lH}) dE_L \\
dE_H &= -\frac{(\pi_{hH} - \pi_{lH})}{(\pi_{lL} - \pi_{hL})} dE_L = -\Pi dE_L
\end{aligned}
$$

Now, we can write dA_h in terms of dE_L:

$$
\begin{aligned}
dA_h &= (-\pi_{lL}\Pi dE_L - \pi_{lH} dE_L)/\Delta \\
&= ((-\pi_{lL}\Pi - \pi_{lH})/\Delta) dE_L
\end{aligned}
$$

Thus, we can solve for dA_h in terms of $dlog(e)/dT$:

$$\begin{aligned}
\frac{d\log(e)}{dT} &= \frac{1}{E_H}\frac{dE_H}{dT} - \frac{1}{E_H+E_L}\left(\frac{dE_H}{dT}+\frac{dE_L}{dT}\right)\\
\frac{d\log(e)}{dT} &= -\frac{1}{E_H}\Pi dE_L - \frac{1}{E_H+E_L}\left(-\Pi dE_L + dE_L\right)\\
\frac{d\log(e)}{dT} &= (-\frac{1}{E_H}\Pi - \frac{1}{E_H+E_L}\left(1-\Pi\right))dE_L\\
\frac{d\log(e)}{dT} &= \frac{(\Pi-1)E_H-(E_H+E_L)\Pi}{E_H(E_H+E_L)}dE_L\\
\frac{d\log(e)}{dT} &= \frac{(\Pi-1)E_H-(E_H+E_L)\Pi}{E_H(E_H+E_L)}\frac{\Delta}{-\pi_{lL}\Pi-\pi_{lH}}dA_h\\
\frac{d\log(e)}{dT} &= \frac{E_H+E_L\Pi}{E_H(E_H+E_L)}\frac{\Delta}{\pi_{lL}\Pi+\pi_{lH}}\frac{dA_h}{dT}\\
\frac{d\log(e)}{dT} &= \frac{E_H(\pi_{lL}-\pi_{hL})+E_L(\pi_{hH}-\pi_{lH})}{E_H(E_H+E_L)}\frac{dA_h}{dT}\\
\frac{dA_h}{dT} &= \frac{d\log(e)}{dT}\left(\frac{E_H(E_H+E_L)}{E_H(\pi_{lL}-\pi_{hL})+E_L(\pi_{hH}-\pi_{lH})}\right)
\end{aligned}$$

which can then be substituted back into the dW/dT expression above. Then, taking the partial derivative with respect to $dlog(e)/dT$ gives the expression in Proposition 3:

$$\begin{aligned}
\frac{\partial}{\partial(\frac{dlog(e)}{dT})}\left(\frac{dW}{dT}\right)\bigg|_{\frac{dA}{dT}} &= \frac{\partial}{\partial(dlog(e)/dT)}\left[\left[(\mu_h-\mu_l)-(\pi_hB_h+g_h)-(\pi_lB_l+g_l)\right]\frac{dA_h}{dT}\right]\\
&= \frac{\partial}{\partial(dlog(e)/dT)}*\\
&\left[\left[(\mu_h-\mu_l)-(\pi_hB_h+g_h)-(\pi_lB_l+g_l)\right]\frac{d\log(e)}{dT}\left(\frac{E_H(E_H+E_L)}{E_H(\pi_{lL}-\pi_{hL})+E_L(\pi_{hH}-\pi_{lH})}\right)\right]\\
&= \left[(\mu_h-\mu_l)-(\pi_hB_h+g_h)-(\pi_lB_l+g_l)\right]\frac{E_H(E_H+E_L)}{E_H(\pi_{lL}-\pi_{hL})+E_L(\pi_{hH}-\pi_{lH})}.
\end{aligned}$$

E2. Extensions: Propositions and Proofs

E2.1 Extensions to the neoclassical model (mean unbiased beliefs)

Welfare Effects of Non-Marginal Changes without Misperceptions (Away From Envelope Theorem)

We have focused discussion of model on marginal changes in perceptions and costs in response to interventions. However, non-marginal changes can be analyzed in a similar fashion. In particular, with non-marginal changes, the envelope theorem will no longer apply even in the neoclassical benchmark case with accurate beliefs. However, we can still show the same lack of a general relationship between targeting properties of the intervention and changes in welfare. This is illustrated in the following proposition, which focuses on the case of accurate beliefs for simplicity:

Proposition 3. *If $\epsilon_l = \epsilon_h = 0$, then the effect of the non-marginal Information Only treatment ($\Delta T = \Delta\epsilon$) on welfare is given by:*

$$\Delta W = u'(y_l) \int_{(1+\epsilon_l)\pi_l B_l}^{(1+\epsilon_l+\Delta\epsilon)\pi_l B_l} (\pi_l B_l - c) f_l(c) dc + u'(y_h) \int_{(1+\epsilon_h)\pi_h B_h}^{(1+\epsilon_h+\Delta\epsilon)\pi_h B_h} (\pi_h B_h - c) f_h(c) dc$$
$$- [(\pi_l B_l + g_l)\Delta A_l + (\pi_h B_h + g_h)\Delta A_h] \quad (6)$$

and the effect of the non-marginal Information Plus Assistance treatment ($\Delta T = \Delta\epsilon,\ -\Delta c$) on welfare is given by:

$$\Delta W = u'(y_l) \int_{\pi_l B_l}^{(1+\Delta\epsilon)\pi_l B_l} (\pi_l B_l - c) f_l(c) dc + u'(y_h) \int_{\pi_h B_h}^{(1+\Delta\epsilon)\pi_h B_h} (\pi_h B_h - c) f_h(c) dc$$
$$+ u'(y_l) \int_{\pi_l B_l - \Delta c}^{\pi_l B_l} (\pi_l B_l - c) f_l(c) dc + u'(y_h) \int_{\pi_h B_h - \Delta c}^{\pi_h B_h} (\pi_h B_h - c) f_h(c) dc \quad (7)$$
$$+ u'(y_l)\Delta c A_l + u'(y_h)\Delta c A_h - [(\pi_l B_l + g_l)\Delta A_l + (\pi_h B_h + g_h)\Delta A_h]$$

Alternatively, if we define expected benefits net of costs as $z_i = \pi_j B_j - c_i$, the effect of the non-marginal Information Only treatment on welfare is given by

$$\Delta W = u'(y_l) \int_{-\Delta\epsilon\pi_l B_l}^{0} z k_l(z) dz + u'(y_h) \int_{-\Delta\epsilon\pi_h B_h}^{0} z k_h(z) dz - [(\pi_l B_l + g_l)\Delta A_l + (\pi_h B_h + g_h)\Delta A_h] \quad (8)$$

and the effect of the non-marginal Information Plus Assistance treatment on welfare is given by:

$$\Delta W = u'(y_l) \int_{-\Delta\epsilon\pi_l B_l}^{0} z k_l(z) dz + u'(y_h) \int_{-\Delta\epsilon\pi_h B_h}^{0} z k_h(z) dz + u'(y_l) \int_0^{\Delta c} z k_l(z) dz + u'(y_h) \int_0^{\Delta c} z k_h(z) dz \quad (9)$$
$$+ u'(y_l)\Delta c A_l + u'(y_h)\Delta c A_h - [(\pi_l B_l + g_l)\Delta A_l + (\pi_h B_h + g_h)\Delta A_h]$$

where $k_j(z)$ is the distribution of the z's.

To see why these terms are not obviously signed in the case of high-benefit targeting, note that in the non-marginal case the increase in private welfare from a given change in enrollees 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 private welfare, 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 ($\Delta T = \Delta\epsilon$): If ϵ_j increases to $\epsilon_j + \Delta\epsilon$, then W changes by:

$$\begin{aligned}
\Delta W = {} & u'(y_l) \int\limits_{(1+\epsilon_l)\pi_l B_l}^{(1+\epsilon_l+\Delta\epsilon)\pi_l B_l} (\pi_l B_l - c) f_l(c) dc + u'(y_h) \int\limits_{(1+\epsilon_h)\pi_h B_h}^{(1+\epsilon_h+\Delta\epsilon)\pi_h B_h} (\pi_h B_h - c) f_h(c) dc \\
& - \left[(\pi_l B_l + g_l) \int\limits_{(1+\epsilon_l)\pi_l B_l}^{(1+\epsilon_l+\Delta\epsilon)\pi_l B_l} f_l(c) dc + (\pi_h B_h + g_h) \int\limits_{(1+\epsilon_h)\pi_h B_h}^{(1+\epsilon_h+\Delta\epsilon)\pi_h B_h} f_h(c) dc \right] \\
= {} & u'(y_l) \int\limits_{(1+\epsilon_l)\pi_l B_l}^{(1+\epsilon_l+\Delta\epsilon)\pi_l B_l} (\pi_l B_l - c) f_l(c) dc + u'(y_h) \int\limits_{(1+\epsilon_h)\pi_h B_h}^{(1+\epsilon_h+\Delta\epsilon)\pi_h B_h} (\pi_h B_h - c) f_h(c) dc \\
& - \left[(\pi_l B_l + g_l) \Delta A_l + (\pi_h B_h + g_h) \Delta A_h \right]
\end{aligned}$$

If $\epsilon_h = \epsilon_l = 0$, this simplifies to

$$\begin{aligned}
\Delta W = {} & u'(y_l) \int\limits_{\pi_l B_l}^{(1+\Delta\epsilon)\pi_l B_l} (\pi_l B_l - c) f_l(c) dc + u'(y_h) \int\limits_{\pi_h B_h}^{(1+\Delta\epsilon)\pi_h B_h} (\pi_h B_h - c) f_h(c) dc \\
& - \left[(\pi_l B_l + g_l) \Delta A_l + (\pi_h B_h + g_h) \Delta A_h \right].
\end{aligned}$$

To simplify notation, we can define the expected benefits net of costs as $z_i = \pi_j B_j - c_i$. Using a change of variables, we have $\pi_j B_j - c_i = z_i$, $f_j(c) = f_j(\pi_j b_j - z) = k_j(z)$, and $dc = -dz$. Therefore,

$$\begin{aligned}
\Delta W = {} & -u'(y_l) \int\limits_{0}^{-\Delta\epsilon\pi_l B_l} z f_l(\pi_l B_l - z) dz - u'(y_h) \int\limits_{0}^{-\Delta\epsilon\pi_h B_h} z f_h(\pi_h b_H - z) dz \\
& - \left[(\pi_l B_l + g_l) \Delta A_l + (\pi_h B_h + g_h) \Delta A_h \right] \\
= {} & u'(y_l) \int\limits_{-\Delta\epsilon\pi_l B_l}^{0} z k_l(z) dz + u'(y_h) \int\limits_{-\Delta\epsilon\pi_h B_h}^{0} z k_h(z) dz - \left[(\pi_l B_l + g_l) \Delta A_l + (\pi_h B_h + g_h) \Delta A_h \right]
\end{aligned}$$

where $k_j(z)$ is the distribution of the expected benefits net of costs. This completes the first part of the proof.

Assistance Only ($\Delta T = -\Delta c$): If c falls to $c - \Delta c$, then W changes by:

$$\Delta W = \int_{(1+\epsilon_l)\pi_l B_l - \Delta c}^{(1+\epsilon_l)\pi_l B_l} u'(y_l)(\pi_l B_l - c) f_l(c) dc + \int_{(1+\epsilon_h)\pi_h B_h - \Delta c}^{(1+\epsilon_h)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc$$
$$+ u'(y_l)\Delta c A_l + u'(y_h)\Delta c A_h - [(\pi_l B_l + g_l)\Delta A_l + (\pi_h B_h + g_h)\Delta A_h]$$

If $\epsilon_j = 0$, this simplifies to

$$\Delta W = \int_{\pi_l B_l - \Delta c}^{\pi_l B_l} u'(y_l)(\pi_l B_l - c) f_l(c) dc + \int_{\pi_h B_h - \Delta c}^{\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc$$
$$+ u'(y_l)\Delta c A_l + u'(y_h)\Delta c A_h - [(\pi_l B_l + g_l)\Delta A_l + (\pi_h B_h + g_h)\Delta A_h]$$

Information and Assistance ($\Delta T = \Delta \epsilon, -\Delta c$): Combining the information and assistance effects, we have

$$\Delta W = u'(y_l) \int_{\pi_l B_l}^{(1+\Delta\epsilon)\pi_l B_l} (\pi_l B_l - c) f_l(c) dc + u'(y_h) \int_{\pi_h B_h}^{(1+\Delta\epsilon)\pi_h B_h} (\pi_h B_h - c) f_h(c) dc$$
$$+ u'(y_l) \int_{\pi_l B_l - \Delta c}^{\pi_l B_l} (\pi_l B_l - c) f_l(c) dc + u'(y_h) \int_{\pi_h B_h - \Delta c}^{\pi_h B_h} (\pi_h B_h - c) f_h(c) dc$$
$$+ u'(y_l)\Delta c A_l + u'(y_h)\Delta c A_h - [(\pi_l B_l + g_l)\Delta A_l + (\pi_h B_h + g_h)\Delta A_h]$$

where ΔA_j is the change in total applications from both $\Delta\epsilon$ and $-\Delta c$. To simplify notation, we can define the expected benefits net of costs as $z_i = \pi_j B_j - c_i$. Again, using change of variables, we have $\pi_j B_j - c_i = z_i$,$f_j(c) = f_j(\pi_j B_j - z) = k_j(z)$, and $dc = -dz$. Therefore,

$$\Delta W = -u'(y_l) \int_{0}^{-\Delta\epsilon\pi_l B_l} z k_l(z) dz - u'(y_h) \int_{0}^{-\Delta\epsilon\pi_h B_h} z k_h(z) dz - u'(y_l) \int_{\Delta c}^{0} z k_l(z) dz - u'(y_h) \int_{\Delta c}^{0} z k_h(z) dz$$
$$+ u'(y_l)\Delta c A_l + u'(y_h)\Delta c A_h - [(\pi_l B_l + g_l)\Delta A_l + (\pi_h B_h + g_h)\Delta A_h]$$
$$= u'(y_l) \int_{-\Delta\epsilon\pi_l B_l}^{0} z k_l(z) dz + u'(y_h) \int_{-\Delta\epsilon\pi_h B_h}^{0} z k_h(z) dz + u'(y_l) \int_{0}^{\Delta c} z k_l(z) dz + u'(y_h) \int_{0}^{\Delta c} z k_h(z) dz$$
$$+ u'(y_l)\Delta c A_l + u'(y_h)\Delta c A_h - [(\pi_l B_l + g_l)\Delta A_l + (\pi_h B_h + g_h)\Delta A_h] .$$

This completes the proof.

Heterogeneity in ϵ (Shifting the Mean)

Proposition 4. *Let ϵ_j be distributed according to the distribution $e_j(\cdot)$ which is symmetric and centered around zero with support $[\underline{\epsilon}_j, \overline{\epsilon}_j]$. Let costs be distributed according to the distribution $f_j(\cdot)$ which is symmetric around $\pi_j B_j$ with support $[\underline{c}_j, \overline{c}_j]$. Further, let $e_j(\cdot)$ and $f_j(\cdot)$ be independent.*

Define $d\epsilon$ as a small shift in each type j agent's value of ϵ_j. If $g_j \geq 0$, then $\frac{dW}{d\epsilon} = \frac{dW}{dT} < 0$.

Proof Sketch: We start with some intuition and a rough sketch before the formal proof:

- Select an arbitrary $\epsilon_j < 0$ from the support of $e_j(\cdot)$.
  - There is mass $e_j(\epsilon_j)$ of these agents.
  - Within this subgroup, any agent who satisfies

    $$(1+\epsilon_j)\pi_j B_j - c_{j,i} = 0 \Rightarrow c_{j,i} = (1+\epsilon_j)\pi_j B_j = \pi_j B_j + \epsilon_j \pi_j B_j$$

    is indifferent between applying and not applying.
  - Since ϵ_j and c are independent, there is a mass $e_j(\epsilon_j)f_j((1+\epsilon_j)\pi_j B_j)$ of these agents.
  - These indifferent agents will be induced to apply by a small increase in $d\epsilon_j$. Each agent's expected utility gain is $-\epsilon_j \pi_j b_j$ (these are agents who should have been applying, and recall that $\epsilon_j < 0$ so $-\epsilon_j \pi_j B_j$ is positive). Therefore, the total welfare increase is

    $$\underbrace{[-\epsilon_j \pi_j B_j]}_{expected\ utility\ gain\ per\ agent} \times \underbrace{e_j(\epsilon_j)f_j((1+\epsilon_j)\pi_j B_j)}_{mass\ of\ indiff't\ agents}$$

- Now consider $-\epsilon_j$.
  - There is mass $e_j(-\epsilon_j) = e_j(\epsilon_j)$ of these agents due to the symmetry of $g_j(\cdot)$ around 0.
  - Within this subgroup, any agent who satisfies

    $$(1-\epsilon_j)\pi_j B_j - c_{j,i} = 0 \Rightarrow c_{j,i} = (1-\epsilon_j)\pi_j B_j = \pi_j b_j - \epsilon_j \pi_j B_j$$

    is indifferent between applying and not applying.
  - Since ϵ_j and c are independent, there is a mass $e_j(-\epsilon_j)f_j((1-\epsilon_j)\pi_j B_j)$ of these agents. Since $f_j(\cdot)$ is symmetric around $\pi_j B_j$, $f_j((1-\epsilon_j)\pi_j B_j) = f_j((1+\epsilon_j)\pi_j B_j)$. Therefore, $e_j(-\epsilon_j)f_j((1-\epsilon_j)\pi_j B_j) = e_j(\epsilon_j)f_j((1+\epsilon_j)\pi_j B_j)$.
  - These indifferent agents will be induced to apply by a small increase $d\epsilon$. Each agent's expected utility loss is $\epsilon_j \pi_j b_j$ (these are agents who should not have been applying, and recall that $\epsilon_j < 0$). Therefore, the total welfare decrease is

    $$\underbrace{[\epsilon_j \pi_j B_j]}_{utility\ loss\ per\ agent} \times \underbrace{e_j(-\epsilon_j)f_j((1-\epsilon_j)\pi_j B_j)}_{mass\ of\ indiff't\ agents} = [\epsilon_j \pi_j B_j] \times e_j(\epsilon_j)f_j((1+\epsilon_j)\pi_j B_j)$$

- Since the net change in private welfare for any $(\epsilon_j, -\epsilon_j)$ pair is zero (the two cancel), integrating over the full support $[\underline{\epsilon}_j, \overline{\epsilon}_j] = [-\overline{\epsilon}_j, \overline{\epsilon}_j]$will yield zero change in private welfare.

- However, we have encouraged a mass of $2e_j(\epsilon_j)f_j((1+\epsilon_j)\pi_j B_j)$ to apply, at a cost of $\pi_j B_j + g_j$ to apply. Since $g_j \geq 0$, we know that we incurred an additional cost.

- Therefore, total welfare (private welfare plus public costs) has fallen.

Formal Proof: In this case, social welfare can be defined as

$$
\begin{aligned}
W &= \iint E_l[u()]f_l(c)e_l(\epsilon)dcd\epsilon + \iint E_h[u()]f_h(c)e_h(\epsilon)dcd\epsilon - [(\pi_l B_l + g_l)A_l + (\pi_h B_h + g_h)A_h] \\
&= \iint_{apply} E_l[u()]f_l(c)e_l(\epsilon)dcd\epsilon + \iint_{\neg apply} u(y_l)f_l(c)e_l(\epsilon)dcd\epsilon \\
&+ \iint_{apply} E_h[u()]f_h(c)e_h(\epsilon)dcd\epsilon + \iint_{\neg apply} u(y_h)f_h(c)e_h(\epsilon)dcd\epsilon - [(\pi_l B_l + g_l)A_l + (\pi_h B_h + g_h)A_h].
\end{aligned}
$$

Individuals of type h apply if

$$(1+\epsilon_{h,i})\pi_h u(y_h + B_h - c_{h,i}) + (1 - (1+\epsilon_{h,i})\pi_h)u(y_h - c_{h,i}) > u(y_h).$$

If we take first-order Taylor approximation around $u(y_h)$, this simplifies to

$$(1+\epsilon_{h,i})\pi_l B_h - c_{h,i} > 0 \Rightarrow c_{h,i} < (1+\epsilon_{h,i})\pi_l B_h$$

Since c and ϵ are independently distributed, we can re-write

$$\iint_{apply} E_h[u(\cdot)]f_h(c)e_h(\epsilon)dcd\delta = \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h}\left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h}(\pi_h u(y_h + b_H - c) + (1-\pi_h)u(y_h - c))\, f_h(c)dc\right]e_h(\epsilon)d\epsilon$$

and

$$
\begin{aligned}
\iint_{\neg apply} u(y_h)f_h(c)e_h(\epsilon)dcd\epsilon &= \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h}\left[\int_{(1+\epsilon)\pi_h B_h}^{\bar{c}_h} u(y_h)f(c)dc\right]e_h(\epsilon)d\epsilon \\
&= \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} u(y_h)\left[1 - F((1+\epsilon)\pi_h B_h)\right]e_h(\epsilon)d\epsilon
\end{aligned}
$$

Therefore, again using a first-order Taylor approximation, we can define the private welfare of type h as

$$
\begin{aligned}
V_h &\approx \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} \left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} \big(u(y_h) + u'(y_h)(\pi_h B_h - c)\big) f_h(c)dc \right] e_h(\epsilon)d\epsilon \\
&+ \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} u(y_h) \left[1 - F((1+\epsilon)\pi_h B_h)\right] e_h(\epsilon)d\epsilon \\
&= \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} \left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c)dc + u(y_h)F((1+\epsilon)\pi_h B_h \right] e_h(\epsilon)d\epsilon \\
&+ \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} u(y_h) \left[1 - F((1+\epsilon)\pi_h B_h)\right] e_h(\epsilon)d\epsilon \\
&= \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} \left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c)dc + u(y_h) \right] e_h(\epsilon)d\epsilon \\
&= \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} \left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c)dc \right] e_h(\epsilon)d\epsilon + u(y_h)
\end{aligned}
$$

Doing the same for type l, we get total welfare is

$$
\begin{aligned}
W &\approx \int_{\underline{\epsilon}_l}^{\bar{\epsilon}_l} \left[\int_{\underline{c}_l}^{(1+\epsilon)\pi_l B_l} u'(y_l)(\pi_l B_l - c) f_l(c)dc \right] e_l(\epsilon)d\epsilon + u(y_l) \\
&+ \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} \left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c)dc \right] e_h(\epsilon)d\epsilon + u(y_h) - \left[(\pi_l B_l + g_l)A_l + (\pi_h B_h + g_h)A_h\right].
\end{aligned}
$$

We want to know how welfare changes if every agent experiences a small increase in ϵ. We call this $\frac{dW}{d\epsilon}$.

$$
\begin{aligned}
\frac{dW}{d\epsilon} &= \frac{d}{d\epsilon} \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} \left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c)dc \right] e_h(\epsilon)d\epsilon \\
&+ \frac{d}{d\epsilon} \int_{\underline{\epsilon}_l}^{\bar{\epsilon}_l} \left[\int_{\underline{c}_l}^{(1+\epsilon)\pi_l B_l} u'(y_l)(\pi_l B_l - c) f_l(c)dc \right] e_l(\epsilon)d\epsilon - \left[(\pi_l B_l + g_l)\frac{dA_l}{d\epsilon} + (\pi_h B_h + g_h)\frac{dA_h}{d\epsilon} \right].
\end{aligned}
$$

First we can focus on the h types. Bringing the derivative inside the integral, we have

$$
\begin{aligned}
&\frac{d}{d\epsilon} \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} \left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c)dc \right] e_h(\epsilon)d\epsilon = \\
&\int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} \frac{d}{d\epsilon} \left[\left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c)dc \right] e_h(\epsilon) \right] d\epsilon.
\end{aligned}
$$

To calculate the derivative now inside the integral, imagine that every type h agent experiences a $+\Delta$ increase in his value of ϵ. Now fix a value of $\epsilon = \epsilon'$. All type h agents who previously had

$\epsilon = \epsilon'$ now have $\epsilon = \epsilon' + \Delta$. The change in welfare for the type agents who had $\epsilon = \epsilon'$ is

$$\Delta W_{h,\epsilon'} = \left[\int_{\underline{c}_h}^{(1+\epsilon'+\Delta)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc\right] e_h^{\Delta}(\epsilon+\Delta) - \left[\int_{\underline{c}_h}^{(1+\epsilon')\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc\right] e_h(\epsilon).$$

where $e_h^{\Delta}(\cdot)$ is the new distribution of ϵ's, because the entire distribution shifted up. Therefore, $e_h^{\Delta}(\epsilon + \Delta) = e_h(\epsilon)$, which allows us to re-write

$$\Delta W_{h,\epsilon'} = u'(y_h) e_h(\epsilon) \int_{(1+\epsilon')\pi_h B_h}^{(1+\epsilon'+\Delta)\pi_h B_h} (\pi_h B_h - c) f_h(c) dc.$$

Dividing by this by Δ and taking the limit as $\Delta \to 0$ gives the derivative evaluated at ϵ':

$$\begin{aligned}
\frac{d}{d\epsilon}\left[\left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc\right] e_h(\epsilon)\right]_{\epsilon=\epsilon'} &= \lim_{\Delta \to 0} \frac{u'(y_h) e_h(\epsilon') \int_{(1+\epsilon')\pi_h B_h}^{(1+\epsilon'+\Delta)\pi_h B_h} (\pi_h B_h - c) f_h(c) dc}{\Delta} \\
&= \lim_{\Delta \to 0} \frac{u'(y_h) e_h(\epsilon') \frac{d}{d\Delta} \int_{(1+\epsilon')\pi_h B_h}^{(1+\epsilon'+\Delta)\pi_h B_h} (\pi_h B_h - c) f_h(c) dc}{1} \\
&= \lim_{\Delta \to 0} \frac{-u'(y_h) e_h(\epsilon')((\epsilon' + \Delta)\pi_h B_h) f_h((1 + \epsilon' + \Delta)\pi_h B_h)}{1} \\
&= -u'(y_h) e_h(\epsilon')(\epsilon' \pi_h B_h) f_h((1 + \epsilon')\pi_h B_h)
\end{aligned}$$

where the second line follows from L'Hopital's rule and the third line follows from Leibniz's rule. Now consider agents with $\epsilon = -\epsilon'$. Following the exact same steps as above,

$$\begin{aligned}
\frac{d}{d\epsilon}\left[\left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc\right] e_h(\epsilon)\right]_{\epsilon=-\epsilon'} &= \lim_{\Delta \to 0} \frac{u'(y_h) e_h(-\epsilon') \int_{(1-\epsilon')\pi_h B_h}^{(1-\epsilon'+\Delta)\pi_h B_h} (\pi_h B_h - c) f_h(c) dc}{\Delta} \\
&= \lim_{\Delta \to 0} \frac{u'(y_h) e_h(-\epsilon') \frac{d}{d\Delta} \int_{(1-\epsilon')\pi_h B_h}^{(1-\epsilon'+\Delta)\pi_h B_h} (\pi_h B_h - c) f_h(c) dc}{1} \\
&= \lim_{\Delta \to 0} \frac{u'(y_h) e_h(-\epsilon')((\epsilon' - \Delta)\pi_h B_h) f_h((1 - \epsilon' + \Delta)\pi_h B_h)}{1} \\
&= u'(y_h) e_h(-\epsilon')(\epsilon' \pi_h B_h) f_h((1 - \epsilon')\pi_h B_h) \\
&= u'(y_h) e_h(\epsilon')(\epsilon' \pi_h B_h) f_h((1 + \epsilon')\pi_h B_h)
\end{aligned}$$

where the second line follows from L'Hopital's rule and the third line follows from Leibniz's rule, and the last line follows from symmetry assumptions on $f_h(\cdot)$ and $e_h(\cdot)$. So we have shown that for any $(\epsilon', -\epsilon')$ pair,

$$\frac{d}{d\epsilon}\left[\left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc\right] e_h(\epsilon)\right]_{\epsilon=\epsilon'} + \frac{d}{d\epsilon}\left[\left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc\right] e_h(\epsilon)\right]_{\epsilon=-\epsilon'} = 0.$$

Therefore,

$$\int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} \frac{d}{d\epsilon}\left[\left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc\right] e_h(\epsilon)\right] d\epsilon = 0.$$

The exact same logic implies

$$\int_{\underline{\epsilon}_l}^{\bar{\epsilon}_l} \frac{d}{d\epsilon}\left[\left[\int_{\underline{c}_l}^{(1+\epsilon)\pi_l B_l} u'(y_l)(\pi_l B_l - c) f_l(c) dc\right] e_l(\epsilon)\right] d\epsilon = 0.$$

Therefore,

$$\frac{dW}{d\epsilon} = -\left[(\pi_l B_l + g_l)\frac{dA_l}{d\epsilon} + (\pi_h B_h + g_h)\frac{dA_h}{d\epsilon}\right].$$

Since

$$A \approx \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} F_h((1+\epsilon)\pi_h B_h e_h(\epsilon) d\epsilon$$

we know that

$$\frac{dA_h}{d\epsilon} = \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} \frac{d}{d\epsilon}\left[F_h((1+\epsilon)\pi_h B_h)e_h(\epsilon)\right] d\epsilon.$$

To calculate the derivatives inside the integrals, imagine that every h agent experiences a $+\Delta$ increase to his value of ϵ. Now, fix a value of $\epsilon = \epsilon'$. All type h agents who previously had $\epsilon = \epsilon'$ now have $\epsilon = \epsilon' + \Delta$. The change in applications for the type h agents who have $\epsilon = \epsilon'$ is

$$\Delta A_{h,\epsilon'} = F((1+\epsilon'+\Delta)\pi_h B_h)e_h^{\Delta}(\epsilon'+\Delta) - F_h((1+\epsilon')\pi_h B_h)e_h(\epsilon')$$

where $e_h^{\Delta}(\cdot)$ is the new distribution of ϵ's, because the entire distribution shifted up. Therefore, $e_h^{\Delta}(\epsilon+\Delta) = e_h(\epsilon)$. Therefore, we can re-write

$$\Delta A_{h,\epsilon'} = e_h(\epsilon')\left[F_h((1+\epsilon'+\Delta)\pi_h B_h) - F_h((1+\epsilon')\pi_h B_h)\right].$$

Dividing by Δ and taking the limit as $\Delta \to 0$ gives the derivative evaluated at $\epsilon = \epsilon'$:

$$\begin{aligned}
\frac{d}{d\epsilon}\left[F((1+\epsilon)\pi_h B_h)e_h(\epsilon)\right]_{\epsilon=\epsilon'} &= \lim_{\Delta\to 0}\frac{e_h(\epsilon')\left[F_h((1+\epsilon'+\Delta)\pi_h B_h) - F_h((1+\epsilon')\pi_h B_h)\right]}{\Delta} \\
&= \lim_{\Delta\to 0}\frac{e_h(\epsilon')f_h((1+\epsilon'+\Delta)\pi_h B_h)\pi_h B_h}{1} \\
&= e_h(\epsilon')f_h((1+\epsilon')\pi_h B_h)\pi_h B_h \\
&> 0.
\end{aligned}$$

Similarly, $\frac{dA_l}{d\epsilon} > 0$. Together with the fact that $g_j \geq 0$, this implies

$$\frac{dW}{d\epsilon} = -\left[(\pi_l B_l + g_l)\frac{dA_l}{d\epsilon} + (\pi_h B_h + g_h)\frac{dA_h}{d\epsilon}\right] < 0.$$

Therefore, $\frac{dW}{dT} < 0$. This completes the proof.

Heterogeneity in ϵ (Treatment Shrinks Variance of Beliefs)

There are many ways to reduce the variance of a distribution. Assume that ϵ_j is distributed symmetrically around 0 with support $[\underline{\epsilon}_j, \bar{\epsilon}_j]$. One way to reduce the variance is to take every ϵ_j and move it to $\epsilon_j(1-\Delta)$ where $\Delta \in [0,1]$. If $\Delta = 0$, we haven't reduced the variance at all. If $\Delta = 1$, we have reduced the distribution to a single mass at 0, with no variance.

Proposition 5. *Let ϵ_j be distributed according to the distribution $e_j(\cdot)$ which is symmetric and centered around zero with support $[\underline{\epsilon}_j, \overline{\epsilon}_j]$. Let costs be distributed according to the distribution $f_j(\cdot)$ which is symmetric around $\pi_j B_j$ with support $[\underline{c}_j, \overline{c}_j]$. Further, let $e_j(\cdot)$ and $f_j(\cdot)$ be independent. Let $dT = d\epsilon$ be the variance reduction described above. Then $\frac{dW}{dT} = \frac{dW}{d\epsilon} > 0$ where $\frac{dW}{d\epsilon}$ corresponds to the welfare change from a very small reduction in the variance.*

Proof Sketch: First we start with some intuition and a sketch of the proof. Some agents who have $\epsilon_j < 0$ find their value of ϵ_j increase and are induced to apply. They are made better off, because they should have been applying (they under-estimated the probability of acceptance). On the other hand, some agents who have $\epsilon_j > 0$ find their value of ϵ_j decrease and are induced not to apply. They are made better off, because they should not have been applying (they overestimated the probability of acceptance). So private welfare unambiguously increases. However, under some symmetry assumptions, the number of applications stays constant. Therefore, there is no increase in application costs, and so total welfare increases.

Formal Proof: Just as in Proposition 4, welfare can be written as

$$W \approx + \int_{\underline{\epsilon}_l}^{\overline{\epsilon}_l} \left[\int_{\underline{c}_l}^{(1+\epsilon)\pi_l B_l} u'(y_l)(\pi_l B_l - c) f_l(c) dc \right] e_l(\epsilon) d\epsilon + u(y_l)$$
$$+ \int_{\underline{\epsilon}_h}^{\overline{\epsilon}_h} \left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc \right] e_h(\epsilon) d\epsilon + u(y_h) - \left[(\pi_l B_l + g_l) A_l + (\pi_h B_h + g_h) A_h \right].$$

Therefore,

$$\frac{dW}{d\epsilon} \approx \int_{\underline{\epsilon}_l}^{\overline{\epsilon}_l} \left[\left[\frac{d}{d\epsilon} \int_{\underline{c}_l}^{(1+\epsilon)\pi_l B_l} u'(y_l)(\pi_l B_l - c) f_l(c) dc \right] e_l(\epsilon) \right] d\epsilon$$
$$+ \int_{\underline{\epsilon}_h}^{\overline{\epsilon}_h} \left[\left[\frac{d}{d\epsilon} \int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc \right] e_h(\epsilon) \right] d\epsilon - \left[(\pi_l B_l + g_l) \frac{dA_l}{d\epsilon} + (\pi_h B_h + g_h) \frac{dA_h}{d\epsilon} \right].$$

Focusing on the h types, suppose that everyone has their value of $\epsilon_h \to \epsilon_h(1 - \Delta)$). How does welfare change for these types? Consider all agents who have $\epsilon_h = \epsilon'$. Their collective change in welfare is

$$\Delta W_{H,\epsilon'} = \left[\int_{\underline{c}_h}^{(1+\epsilon'(1-\Delta))\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc \right] e_h^{\Delta}(\epsilon'(1-\Delta)) - \left[\int_{\underline{c}_h}^{(1+\epsilon')\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc \right] e_h(\epsilon')$$

where $e_h^{\Delta}(\cdot)$ is the "compressed" PDF. Note that $e_h^{\Delta}(\epsilon'(1 - \Delta)) = e_h(\epsilon')$ so we can rewrite:

$$\Delta W_{H,\epsilon'} = e_h(\epsilon') \left[\left[\int_{\underline{c}_h}^{(1+\epsilon'(1-\Delta))\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc \right] - \left[\int_{\underline{c}_h}^{(1+\epsilon')\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc \right] \right]$$
$$= u'(y_h) e_h(\epsilon') \int_{(1+\epsilon')\pi_h B_h}^{(1+\epsilon'(1-\Delta))\pi_h B_h} (\pi_h B_h - c) f_h(c) dc$$

Dividing by this by Δ and taking the limit as $\Delta \to 0$ gives the "derivative" (We put derivative in quotations because we've defined it in a very specific and strange way) evaluated at ϵ':

$$\begin{aligned}
\frac{d}{d\epsilon}\left[\left[\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c)f_h(c)dc\right] e_h(\epsilon_h)\right]_{\epsilon_h=\epsilon'} &= \lim_{\Delta\to 0} \frac{u'(y_h)e_h(\epsilon')\int_{(1+\epsilon')\pi_h B_h}^{(1+\epsilon'(1-\Delta))\pi_h B_h}(\pi_h B_h - c)f_h(c)dc}{\Delta} \\
&= \lim_{\Delta\to 0} \frac{u'(y_h)e_h(\epsilon')\frac{d}{d\Delta}\int_{(1+\epsilon')\pi_h B_h}^{(1+\epsilon'(1-\Delta))\pi_h B_h}(\pi_h B_h - c)f_h(c)dc}{1} \\
&= \lim_{\Delta\to 0} u'(y_h)e_h(\epsilon')(-\epsilon'(1-\Delta)\pi_h B_h)f_h((1+\epsilon'(1-\Delta))\pi_h B_h)(-\epsilon'\pi_h B_h) \\
&= u'(y_h)e_h(\epsilon')(\epsilon'\pi_h B_h)^2 f_h((1+\epsilon')\pi_h B_h)
\end{aligned}$$

Note that the sign of this is positive regardless of the sign of ϵ' - in other words, for any value of ϵ, the "derivative" is positive. This implies that when we integrate over the full support, the integral is positive. In other words,

$$\int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h}\left[\left[\frac{d}{d\epsilon}\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c)f_h(c)dc\right] e_h(\epsilon)\right] d\epsilon > 0.$$

By the exact same logic,

$$\int_{\underline{\epsilon}_l}^{\bar{\epsilon}_l}\left[\left[\frac{d}{d\epsilon}\int_{\underline{c}_l}^{(1+\epsilon)\pi_l B_l} u'(\overline{y})(\pi_l B_l - c)f_l(c)dc\right] e_l(\epsilon)\right] d\epsilon > 0.$$

Finally, we know that

$$A_h \approx \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} F_h((1+\epsilon)\pi_h B_h e_h(\epsilon)d\epsilon.$$

Therefore,

$$\frac{dA_h}{d\epsilon} = \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} \frac{d}{d\epsilon}\left[F_h((1+\epsilon)\pi_h B_h)e_h(\epsilon)\right] d\epsilon.$$

To calculate the derivative inside the integrals, again imagine that every H agent has his value of $\epsilon \to \epsilon(1-\Delta)$ Now, fix a value of $\epsilon = \epsilon'$. All type h agents who previously had $\epsilon = \epsilon'$ now have $\epsilon = \epsilon'(1-\Delta)$. The change in applications for the type h agents who have $\epsilon = \epsilon'$ is

$$\Delta A_{h,\epsilon'} = F((1+\epsilon'(1-\Delta))\pi_h B_h)e_h^{\Delta}(\epsilon'(1-\Delta)) - F_h((1+\epsilon')\pi_h B_h)e_h(\epsilon')$$

where $e_h^{\Delta}(\cdot)$ is the "compressed" PDF. Note that $e_h^{\Delta}(\epsilon'(1-\Delta)) = e_h(\epsilon')$ so we can rewrite:

$$\Delta A_{h,\epsilon'} = e_h(\epsilon')\left[F_h((1+\epsilon'(1-\Delta))\pi_h B_h) - F_h((1+\epsilon')\pi_h B_h)\right].$$

Dividing by Δ and taking the limit as $\Delta \to 0$ gives the derivative evaluated at $\epsilon = \epsilon'$:

$$\begin{aligned}\frac{d}{d\epsilon}\left[F((1+\epsilon)\pi_h B_h)e_h(\epsilon)\right]_{\epsilon=\epsilon'} &= \lim_{\Delta \to 0} \frac{e_h(\epsilon')\left[F_h((1+\epsilon'(1-\Delta))\pi_h B_h) - F_h((1+\epsilon')\pi_h B_h)\right]}{\Delta} \\ &= \lim_{\Delta \to 0} \frac{-e_h(\epsilon') f_h((1+\epsilon'(1-\Delta))\pi_h B_h)(\epsilon'\pi_h B_h)}{1} \\ &= -e_h(\epsilon') f_h((1+\epsilon')\pi_h B_h)(\epsilon'\pi_h B_h)\end{aligned}$$

where the second line follows from L'Hopital's rule. Now consider the derivative evaluated at $-\epsilon'$:

$$\begin{aligned}\frac{d}{d\epsilon}\left[F((1+\epsilon)\pi_h B_h)e_h(\epsilon)\right]_{\epsilon=-\epsilon'} &= \lim_{\Delta \to 0} \frac{e_h(-\epsilon')\left[F_h((1+-\epsilon'(1-\Delta))\pi_h B_h) - F_h((1+-\epsilon')\pi_h B_h)\right]}{\Delta} \\ &= \lim_{\Delta \to 0} \frac{e_h(-\epsilon') f_h((1+-\epsilon'(1-\Delta))\pi_h B_h)(\epsilon'\pi_h B_h)}{1} \\ &= e_h(-\epsilon') f_h((1-\epsilon')\pi_h B_h)(\epsilon'\pi_h B_h) \\ &= e_h(\epsilon') f_h((1+\epsilon')\pi_h B_h)(\epsilon'\pi_h B_h)\end{aligned}$$

where the second line follows from L'Hopital's rule and the last line follows from the symmetry assumptions. Therefore, we have shown that

$$\frac{d}{d\epsilon}\left[F((1+\epsilon)\pi_h B_h)e_h(\epsilon)\right]_{\epsilon=\epsilon'} + \frac{d}{d\epsilon}\left[F((1+\epsilon)\pi_h B_h)e_h(\epsilon)\right]_{\epsilon=\epsilon'} = 0$$

for every $(\epsilon', -\epsilon')$ pair. Therefore, integrating over the full support, we have

$$\int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} \frac{d}{d\epsilon}\left[F_h((1+\epsilon)\pi_h B_h)e_h(\epsilon)\right] d\epsilon = 0.$$

Following an identical argument,

$$\frac{dA_l}{d\epsilon} = \int_{\underline{\epsilon}_l}^{\bar{\epsilon}_l} \frac{d}{d\epsilon}\left[F_h((1+\epsilon)\pi_l B_l)e_l(\epsilon)\right] d\epsilon = 0.$$

Putting this all together, we have

$$\begin{aligned}\frac{dW}{d\epsilon} &\approx \int_{\underline{\epsilon}_h}^{\bar{\epsilon}_h} \left[\left[\frac{d}{d\epsilon}\int_{\underline{c}_h}^{(1+\epsilon)\pi_h B_h} u'(y_h)(\pi_h B_h - c) f_h(c) dc\right] e_h(\epsilon)\right] d\epsilon + \int_{\underline{\epsilon}_l}^{\bar{\epsilon}_l} \left[\left[\frac{d}{d\epsilon}\int_{\underline{c}_l}^{(1+\epsilon)\pi_l B_l} u'(y_l)(\pi_l B_l - c) f_l(c) dc\right] e_l(\epsilon)\right] d\epsilon \\ &> 0\end{aligned}$$

which implies $\frac{dW}{dT} > 0$. This completes the proof.

E2.2 Alternative Non-Neoclassical Model: Inattention

Suppose that agents now have correct beliefs, but some fraction $(1-\alpha)$ of agents are inattentive where $\alpha \in [0, 1]$. Assume this fraction is independent of c. 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 low-touch treatment then reduces the fraction of inattentive agents (or increases the fraction of attentive agents). In other words, $dT = d\alpha$.

As before,

$$W = \int E_l[u(\cdot)]f_l(c)dc + \int E_h[u(\cdot)]f_h(c)dc - [(\pi_l B_l + g_l)A_l + (\pi_h B_h + g_h)A_h].$$

Attentive agents apply if

$$\begin{aligned} E_h[u(\cdot)|apply] &> u(y_h) \\ \pi_h u(y_h + b_H - c_i) + (1-\pi_h)u(y_h - c_i) &> u(y_h) \end{aligned}$$

If we take a first-order Taylor approximation around $u(y_h)$, this simplifies to

$$\pi_h B_h - c_i > 0.$$

Inattentive agents don't apply, regardless of their value of c_i. Therefore, we can rewrite:

$$\begin{aligned} \int E_h[u(\cdot)]dc &= \alpha \underbrace{\left[\int_{apply} E_h[u(\cdot)|apply]f_h(c)dc + \int_{\neg apply} u(y_h)f_h(c)dc\right]}_{attentive\ agents} + (1-\alpha)\underbrace{\left[\int u(y_h)f_h(c)dc\right]}_{inattentive\ agents} \\ &= \alpha\left[\int_{\underline{c}_h}^{\pi_h B_h} E_h[u(\cdot)|apply]f_h(c)dc + \int_{\pi_h B_h}^{\bar{c}_h} u(y_h)f_h(c)dc\right] + (1-\alpha)u(y_h) \\ &= \alpha\int_{\underline{c}_h}^{\pi_h B_h} \left[\pi_h u(y_h + b_H - c_i) + (1-\pi_h)u(y_h - c)\right] f_h(c)dc + \alpha(1 - F(\pi_h B_h))u(y_h) + (1-\alpha)u(y_h) \\ &\approx \alpha\int_{\underline{c}_h}^{\pi_h B_h} \left[u(y_h) + u'(y_h)(\pi_h B_h - c)\right] f_h(c)dc + \alpha(1 - F(\pi_h B_h)) + (1-\alpha)u(y_h) \\ &= \alpha\int_{\underline{c}_h}^{\pi_h B_h} u'(y_h)(\pi_h B_h - c)f_h(c)dc + \alpha u(y_h) + (1-\alpha)u(y_h) \\ &= \alpha\int_{\underline{c}_h}^{\pi_h B_h} u'(y_h)(\pi_h B_h - c)f_h(c)dc + u(y_h) \end{aligned}$$

and similarly,

$$\int E_l[u(\cdot)]dc = \alpha\int_{\underline{c}_l}^{\pi_l B_l} u'(y_l)(\pi_l B_l - c)f_l(c)dc + u(y_l).$$

Applicants are given by

$$A_l \approx \alpha F_l(\pi_l B_l)$$

and

$$A_h \approx \alpha F_h(\pi_h B_h)$$

Therefore, welfare is

$$W \approx \alpha u'(y_l) \int_{\underline{c}_l}^{\pi_l B_l} (\pi_l B_l - c) f_l(c) dc + u(y_l) + \alpha u'(y_h) \int_{\underline{c}_h}^{\pi_h B_h} (\pi_h B_h - c) f_h(c) dc + u(y_h)$$
$$- \alpha \left[(\pi_l B_l + g_l) F_l(\pi_l B_l) + (\pi_h B_h + g_h) F_h(\pi_h B_h) \right].$$

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

$$W \approx u'(y_l) \int_{\underline{c}_l}^{\pi_l B_l} (\pi_l B_l - c) f_l(c) dc + u(y_l) + u'(y_h) \int_{\underline{c}_h}^{\pi_h B_h} (\pi_h B_h - c) f_h(c) dc + u(y_h)$$
$$- \left[(\pi_l B_l + g_l) F_l(\pi_l B_l) + (\pi_h B_h + g_h) F_h(\pi_h B_h) \right].$$

We consider the effect of an intervention which increases attention:

Proposition 6. *If $0 \leq \epsilon_l, \epsilon_h < 1$ (i.e., there is at least some inattention), then the effect of Attention treatment ($dT = d\alpha$) on welfare is given by:*

$$\frac{dW}{dT} = u'(y_h) \int_{\underline{c}_h}^{\pi_h B_h} (\pi_h B_h - c) f_h(c) dc + u'(y_l) \int_{\underline{c}_l}^{\pi_l B_l} (\pi_l B_l - c) f_l(c) dc$$
$$- \left[(\pi_l B_l + g_l) F_l(\pi_l B_l) + (\pi_h B_h + g_h) F_h(\pi_h B_h) \right].$$

Proof: Simply differentiating W with respect to α yields

$$\frac{dW}{dT} = u'(y_h) \int_{\underline{c}_h}^{\pi_h B_h} (\pi_h B_h - c) f_h(c) dc + u'(y_l) \int_{\underline{c}_l}^{\pi_l B_l} (\pi_l B_l - c) f_l(c) dc$$
$$- \left[(\pi_l B_l + g_l) F_l(\pi_l B_l) + (\pi_h B_h + g_h) F_h(\pi_h B_h) \right].$$

Cost Misperceptions

Now instead of misperceiving probabilities, agents misperceive costs. Specifically, each type may also misperceive true cost of applying by δ_j, which raises the "perceived" cost of applying for type j. Note that individual of type h applies if expected utility of applying is greater than (certain) utility of not applying. In other words, if following condition holds:

$$E_h[u()|apply] > u(y_h)$$

For individual with private cost c_i this can be defined as follows

$$\pi_h u(y_h + B_h - c_i - \delta_h) + (1-\pi_h)u(y_h - c_i - \delta_h) > u(y_h).$$

If we take first-order Taylor approximation around $u(y_h)$, this simplifies to

$$\pi_h B_h - c_i - \delta_h > 0.$$

As a result, the share of individuals of type h applying is $F_h(\pi_h B_h - \delta_h)$. We can also use same first-order Taylor approximation to define the (private and public) welfare of type h.

$$\begin{aligned}
W_h &= \int_0^{\pi_h B_h - \delta_h} (\pi_h u(y_h + B_h - c) + (1-\pi_h)u(y_h - c)) f_h(c)dc + (1 - F_h(\pi_h B_h - \delta_h))u(y_h) \\
&- (\pi_h B_h + g_h)F_h(\pi_h B_h) \\
&\approx \int_0^{\pi_h B_h - \delta_h} (u(y_h) + u'(y_h)(\pi_h B_h - c)) f_h(c)dc + (1 - F_h(\pi_h B_h - \delta_h))u(y_h) - (\pi_h B_h + g_h)F_h(\pi_h B_h) \\
&= u'(y_h) \int_0^{\pi_h B_h - \delta_h} (\pi_h B_h - c) f_h(c)dc - (\pi_h B_h + g_h)F_h(\pi_h B_h) + u(y_h)
\end{aligned}$$

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

$$\frac{dW}{dT} = u'(y_l)\delta\left(\frac{dA_l}{dT}\right) + u'(y_h)\delta\left(\frac{dA_h}{dT}\right) - \left[(\pi_l B_l + g_l)\frac{dA_l}{dT} + (\pi_h B_h + g_h)\frac{dA_h}{dT}\right]$$

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

$$\frac{dW}{dT} = u'(y_l)A_l + u'(y_h)A_h + u'(y_l)\delta\frac{dA_l}{dT} + u'(y_h)\delta\frac{dA_h}{dT} - \left[(\pi_l B_l + g_l)\frac{dA_l}{dT} + (\pi_h B_h + g_h)\frac{dA_h}{dT}\right].$$

Proof: Welfare (assuming $\delta_h = \delta_l = \delta$) is given by

$$\begin{aligned}
W = &\int_0^{\pi_l B_l - \delta} u'(y_l)(\pi_l B_l - c) f_l(c)dc + u(y_l) + \int_0^{\pi_h B_h - \delta} u'(y_h)(\pi_h B_h - c) f_h(c)dc + u(y_h) \\
&- \left[(\pi_l B_l + g_l)F_l(\pi_l B_l - \delta) + (\pi_h B_h + g_h)F_h(\pi_h B_h - \delta)\right].
\end{aligned}$$

Information Only ($dT = -d\delta$): Taking the derivative with respect to to δ yields

$$\frac{dW}{d\delta} = \frac{d}{d\delta} \int_0^{\pi_l B_l - \delta} u'(y_l)(\pi_l B_l - c) f_l(c) dc + \frac{d}{d\delta} \int_0^{\pi_h B_h - \delta} u'(y_h)(\pi_h B_h - c) f_h(c) dc.$$

$$+ [(\pi_l B_l + g_l) F_l(\pi_l B_l - \delta) + (\pi_h B_h + g_h) F_h(\pi_h B_h - \delta)]$$

Applying Leibniz's Rule, we get

$$\frac{d}{d\delta} \int_0^{\pi_l B_l - \delta} u'(y_l)(\pi_l B_l - c) f_l(c) dc = -u'(y_l) \delta f_l(\pi_l B_l - \delta).$$

Similarly,

$$\frac{d}{d\delta} \int_0^{\pi_h B_h - \delta} u'(y_h)(\pi_h B_h - c) f_h(c) dc = -u'(y_h) \delta f_h(\pi_h B_h - \delta).$$

Since the number of applicants is given by $A_l = F_l(\pi_l B_l - \delta)$ and $A_h = F_h(\pi_h B_h - \delta)$,

$$\frac{dA_l}{d\delta} = -f_l(\pi_l B_l - \delta)$$

and

$$\frac{dA_h}{d\delta} = -f_h(\pi_h B_h - \delta).$$

Therefore, we can re-write

$$\frac{d}{d\delta} \int_0^{\pi_l B_l - \delta} u'(y_l)(\pi_l B_l - c) f_l(c) dc = u'(y_l) \delta \left(\frac{dA_l}{d\delta} \right)$$

and

$$\frac{d}{d\delta} \int_0^{\pi_h B_h - \delta} u'(y_h)(\pi_h B_h - c) f_h(c) dc = u'(y_h) \delta \left(\frac{dA_h}{d\delta} \right).$$

Putting all this together, we have

$$\frac{dW}{d\delta} = u'(y_l) \delta \left(\frac{dA_l}{d\delta} \right) + u'(y_h) \delta \left(\frac{dA_h}{d\delta} \right) - \left[(\pi_l B_l + g_l) \frac{dA_l}{d\delta} + (\pi_h B_h + g_h) \frac{dA_h}{d\delta} \right].$$

Since $dT = -d\delta$, we know that

$$\frac{dW}{dT} = u'(y_l) \delta \left(\frac{dA_l}{dT} \right) + u'(y_h) \delta \left(\frac{dA_h}{dT} \right) - \left[(\pi_l B_l + g_l) \frac{dA_l}{dT} + (\pi_h B_h + g_h) \frac{dA_h}{dT} \right].$$

This is the first part of the proposition.

Assistance Only ($dT = -dc$): Define $-dc$ as a downward shift in every applicant's value of c. First, let's focus on the h types. Suppose the entire distribution $f_h(c)$ shifts down by Δc. Define W_h as the welfare associated with the h types (the sum of private welfare V_h and the public costs). How does W_h change with $-\Delta c$?

$$\Delta W_h = \left[\underbrace{\int_0^{\pi_h B_h - \delta + \Delta c} u'(y_h)(\pi_h B_h - (c - \Delta c)) f_h^{\Delta}(c - \Delta c) dc - (\pi_h B_h + g_h) F_h(\pi_h B_h - \delta + \Delta c)}_{W_h \; after} \right]$$
$$- \left[\underbrace{\int_0^{(1+\epsilon)\pi_h B_h - \delta} u'(y_h)(\pi_h B_h - c) f_h(c) dc - (\pi_h B_h + g_h) F_h(\pi_h B_h - \delta)}_{W_h \; before} \right]$$

where $f_h^{\Delta}(\cdot)$ is the shifted cost PDF. Note that $f_h^{\Delta}(c - \Delta c) = f_h(c)$ at any value of c. Therefore, by combining the integrals on the region $[0, (1 + \epsilon)\pi_h B_h]$, we can re-write

$$\begin{aligned} \Delta W_h = u'(y_h)\Delta c \int_0^{\pi_h B_h - \delta} f(c) dc + u'(y_h) \int_{\pi_h B_h - \delta}^{\pi_h B_h - \delta + \Delta c} (\pi_h B_h - (c - \Delta c)) f_h(c) dc \\ + (\pi_h B_h + g_h) \left[F_h(\pi_h B_h - \delta) - F_h(\pi_h B_h - \delta + \Delta c) \right]. \end{aligned}$$

Next, note that

$$\begin{aligned} F_h(\pi_h B_h - \delta) - F_h(\pi_h B_h - \delta + \Delta c) &= \int_0^{\pi_h B_h - \delta} f_h(c) dc - \int_0^{\pi_h B_h - \delta + \Delta c} f_h(c) dc \\ &= - \int_{\pi_h B_h - \delta}^{\pi_h B_h - \delta + \Delta c} f_h(c) dc \end{aligned}$$

and so we can re-write

$$\begin{aligned} \Delta W_h = u'(y_h)\Delta c \int_0^{\pi_h B_h - \delta} f(c) dc + u'(y_h) \int_{\pi_h B_h - \delta}^{\pi_h B_h - \delta + \Delta c} (\pi_h B_h - (c - \Delta c)) f_h(c) dc \\ - (\pi_h B_h + g_h) \int_{\pi_h B_h - \delta}^{\pi_h B_h - \delta + \Delta c} f_h(c) dc. \end{aligned}$$

Taking this expression, dividing by Δc, and taking the limit as $\Delta c \to 0$ yields $-\frac{dW}{dc}$. However, both the numerator and denominator approach zero (indeterminate), so I apply L'Hopital's Rule. The denominator is 1, so I just need to differentiate each term of the numerator. This yields:

$$\frac{d}{d(\Delta c)} \left[u'(y_h)\Delta c \int_0^{\pi_h B_h - \delta} f_h(c) dc \right] = u'(y_h) \int_{\underline{c}_h}^{\pi_h B_h - \delta} f_h(c) dc$$
$$\frac{d}{d(\Delta c)} \left[u'(y_h) \int_{\pi_h B_h - \delta}^{\pi_h B_h - \delta + \Delta c} (\pi_h B_h - (c - \Delta c)) f_h(c) dc \right] = u'(y_h)(\delta) f_h(\pi_h B_h \delta_h + \Delta c) \text{ (Leibniz)}$$
$$\frac{d}{d(\Delta c)} \left[-(\pi_h B_h + g_h) \int_{\pi_h B_h - \delta}^{\pi_h B_h - \delta + \Delta c} f_h(c) dc \right] = -(\pi_h B_h + g_h) f_h(\pi_h B_h - \delta + \Delta c) \text{ (Leibniz)}$$

Taking each term as $\Delta c \to 0$ and combining yields

$$\begin{aligned}
-\frac{dW_h}{dc} &= u'(y_h)\int_0^{\pi_h B_h-\delta} f_h(c)dc + u'(y_h)(\delta)f_h(\pi_h B_h-\delta) - (\pi_h B_h + g_h)f_h(\pi_h B_h-\delta)\\
&= u'(y_h)F_h(\pi_h B_h-\delta) + u'(y_h)(\delta)f_h(\pi_h B_h-\delta) - (\pi_h B_h + g_h)f_h(\pi_h B_h-\delta)
\end{aligned}$$

Similarly,

$$-\frac{dW_l}{dc} = u'(y_l)F_l(\pi_l B_l-\delta) + u'(y_l)(\delta)f_l(\pi_l B_l-\delta) - (\pi_l B_l + g_l)f_l(\pi_l B_l-\delta).$$

Putting this together implies

$$\begin{aligned}
\frac{dW}{dT} &= -\frac{dW}{dc}\\
&= -\frac{dW_l}{dc} - \frac{dW_h}{dc}\\
&= u'(y_l)F_l(\pi_l B_l-\delta) + \left(u'(y_l)\delta - (\pi_l B_l + g_l)\right) f_l(\pi_l B_l-\delta)\\
&+ u'(y_h)F_h(\pi_h B_h-\delta) + \left(u'(y_h)\delta - (\pi_h B_h + g_l)\right) f_h(\pi_h B_h-\delta)\\
&= u'(y_l)A_l + u'(y_h)A_h + u'(y_l)\delta\frac{dA_l}{dT} + u'(y_h)\delta\frac{dA_h}{dT} - \left[(\pi_l B_l + g_l)\frac{dA_l}{dT} + (\pi_h B_h + g_h)\frac{dA_h}{dT}\right].
\end{aligned}$$

Information + Assistance ($dT = -d\delta, -dc$): Combining information and assistance yields

$$\frac{dW}{dT} = u'(y_l)A_l + u'(y_h)A_h + u'(y_l)\delta\frac{dA_l}{dT} + u'(y_h)\delta\frac{dA_h}{dT} - \left[(\pi_l B_l + g_l)\frac{dA_l}{dT} + (\pi_h B_h + g_h)\frac{dA_h}{dT}\right]$$

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

F: Complier Characteristics

We report the characteristics of enrollees and applicants separately for always takers and compliers. 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. 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

$$\mu_C = \frac{(\pi_{AT} + \pi_C)\mu_T - \pi_{AT}\mu_{AT}}{\pi_C}. \tag{10}$$

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),

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

$$\mu_{AT} = E[\, X \,|\, Z = 1,\, D = 0] \tag{11}$$

(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

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

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.[47] 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.[48] 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)[49] 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 24 months, although there are exceptions that require earlier re-certification.[50]

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.

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

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

[49]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).

[50]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.