Design Document and Analysis Plan

Project Name: The Impact of Personalized Telephone Outreach on Health Insurance Choices: Evidence from a Randomized Controlled Trial

Date Finalized: September 24, 2020

This document serves as a basis for distinguishing between planned (confirmatory) analysis and any unplanned (exploratory) analysis that might be conducted on project data. Documenting these planned analyses is crucial to ensuring that the results of statistical tests will be properly interpreted and reported. For the Analysis Plan to fulfill this purpose, it is essential that it be finalized and date-stamped before we begin looking at outcome data.

Contents

1. Project Objectives

The goal of the project is to assess whether a personalized phone-based information intervention can help health insurance marketplaces improve take-up among consumers who may face barriers to enrollment and improve marketplace stability, both issues of direct relevance to policy makers. To achieve this goal, we propose to shed light on the following questions of interest to Covered California and policymakers broadly:

I. **What is the effect of personalized phone outreach on enrollment (take-up) and market risk?** The rate of take-up is a critical measure of well-being for the individual consumers served by the marketplace, and marketplace risk is an important indicator of the stability of the health insurance marketplace. Unlike previous RCTs (such as a 2016 effort in California based on a low touch letter intervention), this phone call intervention is more intensive and personalized. A personalized intervention may allow consumers with greater health care needs - and hence, greater need to understand plan options - to enroll.

II. **How do personalized phone calls affect the consumers' decision of which plan to choose?** Plan choice has important implications for sorting of consumer risk across plans, and consumer financial protection from insurance (avoiding choice errors and choosing the "right" benefit design for them).

III. **Are personalized outbound calls particularly effective with certain hard-to-reach populations?** Equitable take-up among the eligible population is an important ingredient in Marketplace success. Certain disadvantaged groups may face higher barriers to plan choice and enrollment. In principle, the intervention lowers these barriers, and may have resulted in increased enrollment and better plan choice for otherwise disadvantaged groups.

IV. **Phone calls are resource intensive; what are the financial implications for the Marketplace and for the insurance market generally of a phone-based intervention to encourage take-up?** There are two components to this question. The first is whether, by reducing enrollment frictions, a phone-based intervention may lead to changes in plan premiums and subsidies. Second, Covered California bore the cost of this intervention;

what are the implications for the agency of taking this intervention to scale? In addition to this budgetary impact analysis, we will also quantify the dollar value equivalent of the intervention in terms of its impact on take-up and market risk and compare the financial and market benefits to the cost of implementation.

2. Evaluation Design

To understand the effect of personalized information intervention on health insurance enrollment and plan choice, Covered California implemented phone-based consultations during the open enrollment period for 2019 coverage to help inform consumers of their plan options, costs, plan benefits, and assist in enrollment. In this intervention, eligible consumers who had not yet picked a plan by a date close to the deadline were randomly assigned to receive no phone call, or a call with a Service Center Representative (SCR) who would discuss their health insurance options with them. In contrast to low-touch interventions, which passively inform consumers about plan options, this intervention assisted consumers more intensively and interactively.

Because call-center time is a scarce resource during the open enrollment period, and it was known in advance that not all eligible consumers could be reached, Covered California implemented the outbound call campaign with a randomized control group. We propose to use administrative data on health insurance choices and phone calls from this randomized controlled trial to understand how personalized assistance informs take-up, plans selected, and costs. The data gathered from the trial include consumers' random treatment assignment, personal characteristics, and subsequent health insurance decisions. See section 3.1 for a detailed description.

2.1 Introduction to the marketplace and consumer enrollment process

A major provision of the ACA was the establishment of regulated insurance marketplaces, or "Exchanges", for the non-group and small group markets made up of individuals without health insurance coverage through a large employer or another public program. Each spring, insurers announce their intention to enter a region in the subsequent calendar year and undergo a state certification process. After certification, insurers offer coverage options. The plans and prices are set and publicized at the end of every summer. Premiums are not allowed to vary based on buyers' observable characteristics except for the buyer's age (smoking is a factor that could be priced on in some states but not in California). The adjustment factor for the premiums is pre-specified, and monotone in age starts from 1 for age 21 to 3 for age 64.

In Covered California, private insurers offer a set of five standardized coverage options. The five coverage levels are Catastrophic, Bronze, Silver, Gold and Platinum. The latter four are ordered by an estimate of the actuarial value of their coverage: 60% for Bronze, 70% for Silver, 80% for Gold, and 90% or more for Platinum plans. Actuarial value of a plan summarizes the deductible, out-of-pocket maxima, and co-pays. Consumers compare and purchase insurance plans during an open enrollment period at the end of the year. Coverage then lasts for the subsequent calendar year. Low-income (less than 400% of FPL) households receive premium subsidies. Households with lower income (less than 250% of FPL) are also eligible to receive cost-sharing subsidies.

2.2 Study Sample

The sampling population for the study is the "funnel" into Covered California. Funnel consumers are those who submitted an application for the 2019 coverage year, were found Covered California-eligible but had not yet selected a plan. This a self-selected sample, and these consumers' actions are informative about their barriers and benefits to enrollment.[1] On the one hand, by the initiation of the search process suggests these consumers had a certain level of interest in obtaining marketplace coverage via Covered California. On the other hand, the fact that these consumers had not chosen a plan and completed the enrollment process is suggestive of the costs, benefits, and barriers that they face. Some consumers may not have completed the enrollment process because, upon reflection, they decided that the expected financial costs of enrollment exceeded the expected benefits from enrollment. Other consumers may not have completed enrollment because they faced behavioral barriers to enrollment, such as confusion about their options or logistical barriers. For example, some consumers might have had difficulty choosing between plan options, been unaware of the subsidy for which they are eligible, had difficulty with the website, or may have planned to complete enrollment but forgot to do so before the deadline.

[1] In addition to consumers who apply through Covered California directly, there is a large subset of consumers who may have been referred to Covered California automatically through county social services agencies, likely through a re-determination of eligibility for the Medicaid program. These consumers may or may not have actively explored Covered California coverage, and less is known about their entire journey in administrative data. For this reason, many analyses will be stratified by this key characteristic.

Still other consumers may not have completed the enrollment process because, in the time elapsed between being found eligibility for Covered California and the end of the Open Enrollment period, they have become eligible for another source of coverage (such as job-based coverage under new employment).

Total Number of Observations

During the 2019 Open Enrollment period, Covered California identified 79,522 individuals who had initiated the paperwork to obtain Covered California health insurance coverage for the 2019 coverage year but had yet to select a plan and enroll. This group comprises our study sample.

2.3 Treatment Arms

This is a two-arm research design where approximately 30 percent of cases were assigned to a control group and 70 percent of cases were assigned to the treatment group. A description of these two arms follows:

Arm 1. Treatment arm (N_1=55,519): These consumers were placed in a list to receive a call from an SCR. The goal of the call was to provide information about their likely eligibility for coverage through Covered California, provide personalized information about plan options, and provide live assistance in choosing a plan including answering on-the-spot questions. If the call went to voicemail, the SCR left a voicemail instructing the recipient to call the service center hotline if they would like further assistance. In total, 39,309 members of the treatment group had their file reviewed to make sure they were still eligible according to the administrative data available to Covered California on the date of the call (i.e., were "worked" by the service center). Those who were eligible received an outbound call from a SCR if staff were available. In the end, 27,123 received an outbound call and 28,396 did not receive a call before the end of open enrollment.

Arm 2. Control arm (N_2=24,003): No phone call was placed to these consumers.

Assignment Process

Consumers were randomly assigned to treatment arms. Randomization occurred at the household level and was based on the last digit of the individual case ID. Cases that end in 1, 2, or 3 were assigned to the control group, and cases that end in any other number were assigned to the treatment group. Preliminary balance checks support the validity of randomization, and further checks are proposed as part of this analysis plan.

2.4 Power and Effect Size

Minimum Detectable Effect Size

To arrive at an estimate for the minimum detectable effect (MDE), we assume a baseline plan selection rate of 10 percent based on prior data (Domurat, Menashe, and Yin 2019). With that assumed base rate and 79,500 cases allocated to treatment and control arms in a 70-30 split, the analysis is powered at the 80% level to detect a 0.7 percentage point increase in plan selection rates.

Meaningful Effect Size

We will consider the intervention to have a meaningful effect if we detect a statistically significant difference in our key outcomes (e.g. take-up, premium changes) between the treatment and control arms of the RCT. From Covered California's perspective, the intervention will be considered "successful" under various criteria if the dollarized benefit from the increases in take-up (or plan choices decisions) exceed the cost of implementing the intervention, or if the intervention is particularly efficient for harder-to-reach segments of the population.

Likely Effect Size

In prior outbound call interventions seeking to increase health insurance take-up and plan switching, Covered California has observed intent-to-treat effects between 1-4 percentage points. As such, we would expect to see a comparable effect in this campaign.

3. Data and Data Structure

This section describes data and variables that will be analyzed, as well as changes that will be made to the raw data with respect to data structure and variables.

3.1 Data Sources

We will combine multiple data files for the project. Our core analytic dataset will be a table which includes consumer demographics, plan selection information and Service Center tracker information. We will also complement this core dataset with other administrative and survey data.

Data File 1: Covered California Outbound Call Extract

Summary of Information: This is a comprehensive extract that includes every consumer that was part of the intervention along with attributes such as:

- demographic detail (age, race/ethnicity, etc.)
- eligibility details (income, Covered California aid codes)
- residence information (5-digit zip code and county)
- enrollment information (metal tier chosen)
- application information (whether consumer submitted application)
- CDPS Risk Scores

Universe of File and Unit of Analysis / Record Granularity: One record for each household who was part of the intervention.

Data File 2: Covered California Inbound Call Extract

Summary of Information: An extract of all inbound calls made by the intervention population to the Service Center during the 2019 Open Enrollment period. This dataset allows us to track inbound calls to the service center for consumers in both treatment arms.

Universe of File and Unit of Analysis / Record Granularity: One record for each call received during the Open Enrollment period.

Data File 3: Covered California Member Survey

Summary of Information: Covered California's 2019 Member Survey oversampled the group of consumers targeted in this study, i.e., consumers who had initiated the enrollment process but did not complete it. These survey data include information about the consumers' knowledge about health insurance, perceived barriers to enrollment in Covered California coverage, and decisions about enrolling in coverage from Covered California vs. another source.

Universe of File and Unit of Analysis / Record Granularity: One survey respondent per household.

Data File 4: Healthcare Evidence Initiative Utilization Data

Summary of Information: A summary of utilization and cost data from Covered California's Healthcare Evidence Initiative dataset (as managed by IBM Watson Health) for consumers who ultimately enrolled in Covered California coverage. This dataset provides an alternate measure of health spending risk to be used in robustness checks.

Universe of File and Unit of Analysis / Record Granularity: One record per enrollee, with summarized utilization data, derived risk scores, and allowed cost financial details such as total allowed charges, member cost-share, or similar.

Data File 5: Covered California Product Extract

Summary of Information: A comprehensive extract of all products and their prices, representing the "product shelf" through CoveredCA.com, designed to allow the research team to identify the set of plan options (and their prices) available to each customer during the open enrollment period. These data include:

- Product features
- Product HIOS code
- Product premium rate
- Price for each age (or "base rates" for a 21-year-old, plus age inflation factor) in each zipcode
- Region (for the price)
- Zip code / service area in which the product is available
- Flags for the relevant crosswalk plan in the next year

Universe of File and Unit of Analysis / Record Granularity: One record for each product option for each region, county, zip code combination, since 2014.

Data File 6: Covered California Provider Directory

Summary of Information: These are the raw input files from Qualified Health Plan issuers that Covered California uses to build the provider directory. These data provide a comprehensive picture of the provider networks available in each product option for each consumer.

Universe of File and Unit of Analysis / Record Granularity: One record for each provider listed for each issuer's directory for each year (or equivalent).

Data File 7: Covered California Enrollee Extract

Summary of Information: A comprehensive enrollment extract, summarizing individual and household characteristics related to each plan selection event, including:

- demographic detail (age, race/ethnicity, etc.)
- eligibility details (income, CC aid codes)
- residence information (not specific addresses, but 5-digit zipcode and county)
- enrollment information (product chosen, prices, whether enrollment was effectuated)
- subsidy information (amount of federal tax credits, or enrollment into CSR Silver plans)
- application information (method applied, whether county submitted application, creation of user accounts, etc.)
- encrypted (masked) identifiers for longitudinal tracking of individuals and households
- CDPS Risk Scores

These data can be used to study health insurance decisions of the broader Covered California enrollee population and the influence on these decisions of factors such as premiums, subsidies, and member risk. This will be used to compare the choice behaviors of the treatment arm of the RCT against prior-year choices of Covered California enrollees, and to project choices and financial implications from a scaled-up intervention.

Universe of File and Unit of Analysis / Record Granularity: One record for each enrollee in the RCT population since 2014, for each plan selection event.

3.2 Key Variables

Outcomes of Interest

Enrollment: The primary outcome of interest is an indicator for whether a consumer selected a plan during Open Enrollment. A related outcome will be an indicator for effectuated enrollment, i.e., whether the consumer paid for their coverage.

Choice quality: A growing literature has documented that many consumers in the United States may purchase the "wrong" plan for them based on the options available (so-called "dominated plans," i.e., plans that cost more but provide equivalent or worse coverage than other available options) due to their lack of understanding of health insurance options (Hoerl et al. 2017; Bhargava, Loewenstein, and Sydnor 2017; Liu and Sydnor 2018; B. J. Abaluck and Gruber 2016; Wang et al. 2017b; Baicker, Congdon, and Mullainathan 2012; Loewenstein et al. 2017; J. Abaluck and Gruber 2011; Wang et al. 2017a; K. M. M. Ericson and Starc 2016). Consumers facing challenges in plan choice may be more likely to choose such plans. It is plausible that an intensive interactive information intervention could help consumers weigh their options more appropriately and steer away from these choices, with direct implications for consumer expenditure risk and well-being.

Our key outcome related to choice quality will be an indicator for whether a consumer made a choice error by selecting a dominated plan. We considered a given consumer to have selected a dominated plan under any of the following conditions: a) the consumer was eligible for a Silver 87 plan but chose a Gold plan; b) the consumer was eligible for a Silver 94 plan but chose a Gold or Platinum plan; c) the consumer was eligible for a Silver 94, \$1-premium plan but chose a \$1-premium Bronze plan; d) the consumer had household income greater than 200% of FPL and chose a Silver plan when eligible for a lower-premium Gold plan offered by the same insurer.

The risk mix of enrollees: We will examine the effect of the intervention on risk levels of consumers who purchased a plan in the marketplace, using the CDPS risk scores to capture consumers' recent health spending. In particular, we will study two measures of risk mix in the marketplace: a) average market risk, driven by the health risk of consumers *brought into the market* by the intervention; and b) the *sorting* of consumers, by risk, across plans with higher vs. lower actuarial value (e.g., Silver vs. Gold tier plans). The former captures the effect of the intervention on total market risk, while the latter can impact relative profits of plans across tiers.

It is possible that the intervention, like a low-touch letter intervention, encouraged healthier people to enroll in a plan. Alternately, the more intensive nature of the intervention may have lowered the barriers in making difficult plan choice decisions, which sicker patients find more valuable. By jointly examining the enrollment and plan choice decisions, we will be able to assess which frictions in the enrollment process the intervention is likely targeting and assess its implications for risk selection and sorting across plans.[2]

Engagement with the intervention: If the outbound call intervention has a larger impact than low-touch interventions, such as reminder letters, in-depth conversations with the SCR could account for this difference. Yet, not all consumers who received an outbound call ultimately engaged in such a conversation. Some consumers who received an outbound call hung up briefly after taking the call; others let the call from the SCR go to voicemail and did not return the call.

To better understand this issue, we will examine the factors predicting consumers' engagement with the intervention, measured by having a conversation with an SCR. Differing levels of engagement with the intervention can help to explain heterogeneous effects of the intervention across different groups of consumers, and inform future targeting of the intervention. The outcome of interest for this analysis will be determined by assessing the typical call length for a sample of calls that were denoted as voicemail only in the tracker in which the SCRs recorded call outcomes, believed to be roughly whether a consumer has ever engaged with an SCR in a conversation lasting 1 minute or longer; other cutoffs such as 30 seconds, 2 minutes, or above-mean or above-median call length will be used in robustness checks.

The dollar value of the intervention and cost-effectiveness: We will use a simple demand model to quantify the subsidy equivalent dollar value of the caller intervention, i.e., how much additional

[2] Moreover, taken together, we can also forecast the joint effect of a) extensive margin risk effects and b) risk sorting effects on equilibrium enrollment, premiums, government subsidies, and adverse selection welfare, if the intervention were to be scaled up. This would involve a structural model similar to our modeling for the state subsidies and AB1810 work, calibrated by the experimental estimates from the RCT.

premium subsidy would be required to achieve the impact on take-up achieved by the phone calls. We will then compare the value to the cost of implementing the intervention to characterize the program's social cost-benefit. We will also do an internal financial cost-benefit analysis, comparing the increase in marketplace fees induced by the intervention to program costs to help inform policymakers and administrators of health insurance marketplaces who may be considering similar interventions.

Variables Used in Heterogeneity Checks

For the analyses of enrollment and plan choice quality, we will conduct heterogeneity analyses that stratify the data into pre-specified groups. The goal of this analysis is to assess whether the impact of the intervention varies for consumers who may face different barriers to enrollment or differ in other observable characteristics, to support improved targeting of the intervention in the future. These groups include:

- Whether the consumer's initial eligibility for Covered California occurred through CoveredCA.com (whether by the consumer's own application or through a broker, navigator, SCR, or renewal process), or was triggered by a county referral process (from the SAWS Medicaid eligibility system)
- People who prefer to use a language other than English, and received service in their preferred language[3]
- People who prefer to use a language other than English, but received service in English
- Low-income (subsidy-eligible) individuals
- Younger age (e.g., above vs. below median age; additional age brackets such as under 30 and 50+ will be examined in supplemental analyses)
- Racial or ethnic minorities

Covariates Used in Multivariable Modeling

Due to the randomization, adjustment for confounders is not required for our statistical models to obtain unbiased treatment effects. To improve power, however, we will include covariates in

[3] We assume that consumers received in-language assistance if they worked with an SCR who had also served over 100 other people with the same preferred language.

multivariable modeling that are predictive of our outcomes of interest. We propose to include the following set of variables as covariates in the main model:

- Location (zip code) fixed effects will be used to adjust for the set of insurance plans available in the location of residence and other factors specific to the location of residence
- *Factors affecting premiums:* age, household income (as percent of federal poverty level), Covered California's age-based community-rating premium ratio, and household size
- *Factors correlated with potential barriers to insurance take-up:* preferred language, race/ethnicity

We will systematically drop or add covariates in robustness checks, as follows. In a first robustness check, we will eliminate covariates, first dropping location fixed effects and then dropping all other covariates.

3.3 Transformations of Variables

When working with inbound call data from the telephony database, we will collapse all calls at the household level to create our call indicator. We will use exchange aid code to create indicators for CSR eligibility and subsidy ineligibility, and we will use the preferred spoken language variable to create language preference indicators.

3.4 Imported Variables

We will merge inbound call data from the telephony database into the core analytic dataset to enable analysis of the impact of having a conversation with an SCR.

3.5 Treatment of Missing Data

In the baseline model, we will use listwise deletion to eliminate missing data on covariates of interest. In a robustness check, we will model missingness using an indicator variable to avoid listwise deletion.

4. Quality Control Checks

4.1 Balance check

Since random assignment is a key feature of our study, we will take great care in verifying the random assignment. We checked for covariate balance across several observable variables, which indicated the randomization was successfully implemented.

Covariate	Treatment	Control
Subsidy FPL %	221%	223%
Spanish Speaker	19%	19%
English Speaker	77%	77%
SAWS-initiated application	62%	62%
CSR Eligible	62%	62%
Subsidy Ineligible	18%	18%
Female HOH	63%	63%
Age	39	39
N	55,519	24,003
Note: columns report the mean value of the covariate by treatment group.		

In the full analysis, we will assess the balance of several additional variables:

- Number of insurers operating in the region
- Median premium of health plans in each tier of the ACA marketplace in the zipcode/region
- Number of agents or certified enrollment counselors in the zipcode/region
- Household size
- Race/ethnicity
- Risk score
- Any prior recorded enrollment in Covered California

Tables 1 and 2 provide sample table shells for the balance checks using administrative and survey data, respectively.

4.2 Simulation check

We will additionally use computer simulations to assess whether we can replicate the results of the Covered California randomization procedure (to within sampling error). The Covered California randomization was conducted by using the last digit of the user ID (1, 2, or 3 vs all others). We will write our own program to implement a similar randomized procedure - randomly selecting three final digits and assigning people with those final digits to a simulated treatment group – and run it 500 times. For each of the key characteristics of eligible individuals that we can observe (i.e. age, gender, preferred language, geographic location, etc.), we will assess whether the sample mean in the actual selected sample is within two standard deviations of the sample means from the 500 simulations. Table S1 provides a sample table shell for the simulation check using administrative data.

5. Anticipated Limitations

Our study has three main limitations. First, Covered California implemented multiple randomized evaluations concurrently during the Open Enrollment period. As such, there is some possibility that, by chance alone, funnel consumers are in the Control group in one RCT but are in a Treatment group in another RCT. However, random assignment helps ensure that unrelated outreach communications are likely to be evenly distributed across treatment arms. Second, the data reported to Covered California on income, household size, and age is self-reported and therefore the standard limitations for self-reported data apply. Third, we study the impact of the intervention for the Covered California funnel population only – which is specific to a set of consumers who did not immediately pick a plan after becoming eligible. Nonetheless, given that the intervention is designed to address common barriers to health insurance enrollment and risk factors for choice errors, our findings can be informative for other marketplaces in which consumers must actively choose their own health insurance plans.

6. Statistical Models & Hypothesis Tests

We will analyze the effects of being randomized to the treatment group and the effects of receiving a call from an SCR on our outcomes of interest using the following model specifications.

6.1 Effects of randomization to the treatment group on enrollment, the health risk of enrollees, and choice errors

We will first examine the effects of assignment to the treatment group, that is, the Intent to Treat (ITT) effects. This will be achieved by using ordinary least squares (OLS) regression in which the outcome of interest (e.g. plan selection) for consumer *i* is regressed on an indicator representing assignment to the treatment group:

$$Y_i = \beta_0 + \beta_1 1(Randomized\ to\ Treatment)_i + \beta_2 X_i + \varepsilon_i$$

The coefficient on the treatment indicator variable, β_1, will be the estimate of the causal effect of being randomized to the treatment group on our outcomes of interest. When modeling risk scores, we will only include people who enrolled in Covered California insurance in the model, and the coefficient β_1 will capture the impact of randomization to the treatment group rather than the control group on the average risk of the study sample. For example, a positive and significant β_1 would suggest that the marginal enrollee added to the market because of randomization to the treatment group rather than the control group had a higher risk score than prior enrollees.

The main analysis will use linear models with robust standard errors for all outcomes to account for heteroscedasticity. In robustness checks, we will use logit or probit models to model binary outcomes. The reduced form estimates will be reported in a table such as Table 4 (see Column 2), where each row corresponds to a dependent variable of interest.

6.2 Effects of randomization to the treatment group on sorting across market tiers

It is helpful to understand how the intervention changed the sorting of consumers across different market tiers by their health spending risk. To study this, we propose to use a multinomial logit model. The probability of consumer *i* choosing a certain tier *j* is:

$$Prob(i\ choosing\ \text{Tier j}) = \exp\left(U_{\text{ij}}\right) / \left(1 + \sum_{k \in T/None} exp(U_{\text{ik}})\right)$$

where

$$U_{\text{ij}} = \alpha_0^j + \alpha_1^j 1(Randomized\ to\ Treatment)_i \times \text{Risk Score}_i + \alpha_2^j\, 1(Randomized\ to\ Treatment)_i + \alpha_3^j \text{Risk Score}_i + \alpha_4^j X_i$$

where the dependent variable is whether consumer i chooses Tier j plan or not selecting a Covered California plan, and $j \in T \equiv \{Platinum, Gold, Silver, HSA\ Bronze, non-HSA\ Bronze, and\ None\}$.[4] In a supplemental analysis, we will limit the data to only those people who are eligible for an enhanced silver plan, and include enhanced silver in the choice set.

U_{ij} can be thought of as consumer i's utility from choosing plan j. X_i includes covariates noted in section 3.2. The parameter of interest is α_1^j, the sign of which indicates whether the intervention resulted in an increase or decrease in the average health spending risk of people in tier j compared to the households not selecting a Covered California plan. To aid in interpretation of the model, we will calculate two average marginal effects that capture key quantities of interest, which will be reported in a table such as Table 5 (see Panel A).

First, to capture the average effect of randomization to the treatment group on enrollment in tier j, we will compute and report average marginal effects defined as follows:

$$t_j \equiv \frac{1}{N}\left(\sum_{i=1}^{N} Prob(i\ choosing \text{Tier } j | 1(RandomizedtoTreatment)_i = 1, X_i, Risk\ Score_i) - \sum_{i=1}^{N} Prob(i\ choosing \text{Tier } j | 1(RandomizedtoTreatment)_i = 0, X_i, Risk\ Score_i)\right)$$

If t_j is positive and statistically significant, this would suggest that the outbound call intervention resulted in higher probability of enrolling in tier j.

[4] In addition to these categories, Covered California offers a "minimum coverage plan," also known as a "catastrophic plan." These plans are only available to consumers younger than 30 or those who prove they are without affordable coverage options or are experiencing financial hardship. Among all Covered California enrollees, only 2.8% enrolled in Catastrophic plans in 2018. Thus, we conjecture that there will be small number of consumers in the funnel sample enrolled in catastrophic plans. If usage of these plans is higher than anticipated in the sample, we can also include these plans in the tier-selection model as a robustness check.

Additionally, there are large variations in the catastrophic enrollment rate among all Covered California enrollees across rating areas, e.g., ranging from 0.3% in rating area 13 to 3.8% in rating area 4. Therefore, as a second robustness check, we propose to estimate the multinomial logit model separately for each rating area by stratifying the data. This will allow us to precisely understand how the intervention shapes sorting of enrollees in regions with a larger vs. smaller share of enrollees selecting Catastrophic plans.

Second, to capture the average effect of randomization to the treatment group on *sorting by risk* into tier j, we will use the following average marginal effect:

$$\chi_j \equiv \frac{1}{N}\left(\sum_{i=1}^{N} \frac{\partial Prob\left(i\ choosing \text{Tier } j \middle| 1(RandomizedtoTreatment)_i = 1, X_i, Risk\ Score_i\right)}{\partial\ Risk\ Score_i} - \sum_{i=1}^{N} \frac{\partial Prob\left(i\ choosing \text{Tier } j \middle| 1(RandomizedtoTreatment)_i = 0, X_i, Risk\ Score_i\right)}{\partial\ Risk\ Score_i}\right)$$

χ_j measures the average effect of randomization to the treatment group on the link between a 1-unit increase in risk score and the probability of choosing tier j. If χ_j is positive and statistically significant, this would suggest that the outbound call intervention increased the concentration of high-risk people in tier j. To aid in interpretation, we will also use a graph to display and compare the effects of randomization to the treatment group rather than control group on tier choice for consumers with risk scores at the 10^{th}, 25^{th}, 50^{th}, 75^{th}, or 90^{th} percentile (see Figure 1).

Alternate specifications

The multinomial models specified above require the assumptions of the independence of irrelevant alternatives (IIA). This assumption implies that, for example, the availability of a bronze plan does not affect a consumer's preference for a platinum plan over a gold plan. To relax the IIA assumption in an alternate analysis, we will use a nested logit model. This model will have two nests, representing two layers of consumer choice: the first choice is whether to select a Covered California plan at all, and the second choice is which type of Covered California plan to select (i.e., choice of tier.). We will report the treatment effects for each dependent variable in a table such as panel B of Table 5.

6.3 Effects of receiving a call from an SCR on enrollment, the health risk of enrollees, and choice errors

Because of the scarcity of call-center time, not all consumers assigned to the treatment arm were reached by phone by an SCR before the end of the open enrollment period. Members of the treatment group had their file reviewed by a SCR to make sure they were still eligible (i.e., were "worked" by the service center). Those who were eligible would then receive an outbound call from a SCR if staff were available.

This outbound call could have changed health insurance decisions through two mechanisms. First, consumers who could not be reached were left a voicemail, which could have changed their health insurance decisions through a "reminder effect." Second, those who were reached, or who called the number in their voicemail, could have changed their health insurance decisions based on the information learned during a conversation with an SCR.

It is therefore helpful to ask: what is impact of a consumer receiving an outbound SCR call on take-up and plan choice on the individual-level, and on risk mix on the group-level? These estimates should be different from the ITT estimates (β_1) because some consumers randomized to treatment did not receive an outbound call (i.e., one-sided non-compliance).

We will use two-stage least squares regression (2SLS) to estimate the causal effects of being reached out to by a SCR on take-up of Covered California insurance, plan choice, and risk mix in the marketplace. The variable $1(Outbound\ Call)_i = 1$ when case of consumer i in the treatment group is reached out to by a SCR. Receiving a call from a SCR is endogenous because of the selection process used by the SCR, which selectively reached out to households who were eligible and therefore more likely to enroll in Covered California insurance. The structural equation to be estimated is:

$$Y_i = \gamma_0 + \gamma_1 1(Outbound\ Call)_i + \gamma_2 X_i + \omega_i,$$

and the first stage is:

$$1(Outbound\ Call)_i = \theta_0 + \theta_1 1(Randomized\ to\ Treatment)_i + \theta_2 X_i + \vartheta_i$$

In the model above, Y_i is the outcome of interest, e.g., take-up, effectuated enrollment, choice errors, or risk score. γ_1, the coefficient of interest, captures the impact of an outbound call on this outcome, or the Local Average Treatment Effect (LATE).

When modeling risk scores, we will only include people who enrolled in Covered California insurance in the model, and the coefficient γ_1 will capture the impact of the intervention on the average risk of the study sample. For example, a positive and significant γ_1 would suggest that the marginal enrollee added to the market because of randomization to the treatment group rather than the control group had a higher risk score than prior enrollees. Along with the estimated treatment effect on take-up, the impact on average risk can be used to estimate the average risk across respondents to the treatment.

A key identifying assumption behind the 2SLS LATE estimator is the exclusion restriction, that is, that assignment to the treatment arm only influenced the outcomes through being worked by a SCR. This assumption is plausible given the random assignment of individuals to the treatment arm, and the fact that only people in the treatment arm received an outbound call. The first stage and 2SLS estimates will be reported in tables such as Table 3 and Table 4.

6.4 Consumers' engagement with the intervention: predictors of having a conversation with an SCR

If the outbound call intervention has a larger impact than low-touch interventions, such as reminder letters, in-depth conversations with the SCR could account for this difference. Yet not all consumers who received an outbound call ultimately engaged in such a conversation. Some consumers who received an outbound call hung up briefly after taking the call; others let the call from the SCR go to voicemail and did not return the call. While we cannot directly test the causal effect of conversation,[5] it is of interest to understand who responded to the intervention by engaging with an SCR. The model we consider is an OLS model:

$$\text{Conversation with SCR}_i = \varphi_0 + \varphi_1 1(Randomized\ to\ Treatment)_i + \varphi_2 X_i + \varepsilon_i,$$

The dependent variable is whether a consumer have a conversation with an SCR, defined as 1 minute of discussion or longer between consumer and an SCR; other cutoffs such as 30 seconds, 2 minutes, or above-median or above-mean call length will be used in robustness checks. φ_1 captures how being randomized into treatment affects the probability of consumer having conversation with an SCR, X_i is the vector of covariates used in previous models, which capture how consumers' characteristics predict their probability of engaging in a conversation with an SCR. The results will be reported in a table such as Table 6.

[5] This issue arises because there are two endogenous variables on the path to having a conversation – first, the consumer has to appear eligible for insurance according to Covered California SCRs and therefore be selected for an outbound call; and second, the consumer must decide to engage in a conversation rather than hanging up the phone or letting the call go to voicemail. While the randomization provides one instrument, we lack the second instrument required to study the impact of having a telephone conversation.

6.5 Follow-up analyses: Additional heterogeneity checks

After conducting our main analysis, we will examine heterogeneity in treatment effects by stratifying the data. To be specific, we will estimate ITT and LATE models by aid code (subsidy-eligible vs. subsidy-ineligible), application source (county-referred vs. CoveredCA.com), language preference (English vs. any other language), service language for people with non-English language preference (people who prefer to use a language other than English, and received service in their preferred language vs. people who prefer to use a language other than English, but received service in English), race/ethnicity (minority vs. non-minority), and by age (above vs. below median age; additional age brackets such as under 30 and 50+ will be examined in supplemental analyses). The summary of heterogeneity checks for each outcome of interest and heterogeneity variable will be reported in a table such as Table 7.

6.6 Inference criteria, including any adjustments for multiple comparisons

We will not perform any corrections for multiple hypothesis tests, and we will use two-tailed tests with p-values <= .05 to denote statistically significant effects.

6.7 Exploratory analysis: Combining administrative and survey data

Covered California's Member Survey oversampled funnel consumers in 2019. We will join the core analytic dataset to Covered California's Member Survey in order to explore questions related to ease of enrollment and whether consumers had another source of coverage, among other topics. Given the smaller sample size of this joined sample, power may be low for these analyses and we therefore consider them to be exploratory.

First, we will conduct additional heterogeneity analyses to investigate potential predictors of intervention impact. Specifically, we will assess whether the following factors predicted the impact of the intervention:

- Factors related to the consumer's motivation to find health insurance:
    - Lack employer sponsored health insurance[6]

[6] Defined as answering "No" to the question, "Does an employer offer you any additional pay or benefits to help pay for health insurance? This could be your employer or a family member's employer."

- Have a chronic condition[7]
- In fair or poor health[8]
- Currently take any prescriptions[9]
- Factors related to possible information barriers experienced by the consumer:
  - Low English proficiency[10]
  - Less than high school education
  - Thought there was no penalty for being uninsured in 2018[11]

The heterogeneity effects of each variable will be reported in a table such as Table 9 (see Panel B).

Second, we will estimate the causal effects of the intervention on other outcome variables. We will examine the following factors:

- Satisfaction with the application process:
  - Likelihood of recommending Covered California to a friend or colleague[12]
  - Had difficulty with entering details in the Covered California application[13]
- Potential mechanisms underlying the effect of the intervention (e.g., closing information gaps or facilitating shopping/comparing plans):
  - Unaware of penalty for being uninsured in 2019
  - Had difficulty shopping and comparing plans[14]
  - Had difficulty getting needed information during the enrollment process[15]
  - Had difficulty finding if a doctor or hospital was covered by a plan

---

[7] Defined as answering “Yes” to the question “Do you currently have a health condition that has lasted for a year or more or is expected to last for a year or more? This could be a physical health condition (such as arthritis, asthma, cancer, dementia, diabetes, heart disease, high cholesterol, hypertension or stroke), a behavioral health or mental health condition, or a developmental disability.”

[8] Defined as answering “fair” or “poor” to the question, “In general, would you say your health status is..”)

[9] Defined as answering “yes” to the question “Do you currently need or take any medicine prescribed by a doctor? Do not include birth control.”

[10] Defined as rating that they speak English “not well” or “not at all”; question was asked of people who did not speak English at home

[11] Defined as answering “No” to the question “Think back to last year (2018). As far as you know, was it required by law for most people to have health insurance in 2018 or else pay a fine?

[12] Rated on a Likert scale from 0 to 10, with 0 as “not at all likely” and 10 as “extremely likely.”

[13] Defined as answering “somewhat difficult” or “very difficult” to the question, “How easy or difficult was it to fill out or update the Covered California application for 2019? As a reminder, the Covered California application asks for information like social security number, household size, address, and income.”

[14] Defined as answering “Somewhat difficult” or “Very difficult” to the question “Overall, how easy or difficult was it to shop and compare health plans through Covered California?”

[15] Defined as answering “Somewhat difficult” or “Very difficult” to the question “In general, how easy or difficult was it to get the help or information you needed during the most recent Open Enrollment period?”

  - Had difficulty finding what their monthly premium would be for a Covered California plan[16]
  - Had difficulty understanding total cost-sharing[17]
- Health insurance enrollment decisions:
  - Have health coverage
  - Main source of health coverage is via current or former employer or union
  - Main source of health coverage is via Medi-Cal or Medicaid

The first stage results, reduced form, and 2SLS results will be reported tables such as Table 3 and Table 4 (see Panel B).

[16] Defined as answering "somewhat difficult" or "very difficult" to the question, "How easy or difficult was it to find out how much your monthly premium would be for a plan through Covered California for 2019?"
[17] Defined as answering "Somewhat difficult" or "Very difficult" to the question "Overall, how easy or difficult was it to understand your total expense estimate for plans through Covered California? Your total expense estimate for a plan reflects the monthly premium and your estimated out of pocket costs such as the deductible and copays from using health care services."

References

Abaluck, By Jason, and Jonathan Gruber. 2016. “Evolving Choice Inconsistencies in Choice of Prescription Drug Insurance.” *The American Economic Review* 106 (8): 2145–2184.

Abaluck, Jason, and Jonathan Gruber. 2011. “Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program.” *American Economic Review* 101 (4): 1180–1210. https://doi.org/10.1257/aer.101.4.1180.

Baicker, Katherine, William J. Congdon, and Sendhil Mullainathan. 2012. “Health Insurance Coverage and Take-up: Lessons from Behavioral Economics.” *The Milbank Quarterly* 90 (1): 107–34. https://doi.org/10.1111/j.1468-0009.2011.00656.x.

Bhargava, S, G Loewenstein, and J Sydnor. 2017. “Choose to Lose: Health Plan Choices from a Menu with Dominated Options.” *Quarterly Journal of Economics* 132: 1319–1372.

Domurat, Richard, Isaac Menashe, and Wesley Yin. 2019. “The Role of Behavioral Frictions in Health Insurance Marketplace Enrollment and Risk: Evidence from a Field Experiment.” Working Paper 26153. National Bureau of Economic Research. https://doi.org/10.3386/w26153.

Ericson, Keith M Marzilli, and Amanda Starc. 2016. “How Product Standardization Affects Choice : Evidence from the Massachusetts Health Insurance Exchange.” *Journal of Health Economics* 50: 71–85. https://doi.org/10.1016/j.jhealeco.2016.09.005.

Ericson, Keith Marzilli, and Amanda Starc. 2015. “Measuring Consumer Valuation of Limited Provider Networks.” *American Economic Review* 105 (5): 115–19. https://doi.org/10.1257/aer.p20151082.

Gruber, Jonathan, and Robin McKnight. 2016. “Controlling Health Care Costs through Limited Network Insurance Plans: Evidence from Massachusetts State Employees.” *American Economic Journal: Economic Policy* 8 (2): 219–50. https://doi.org/10.1257/pol.20140335.

Hoerl, Maximiliane, Amelie Wuppermann, Silvia H Barcellos, Sebastian Bauhoff, Joachim K Winter, and Katherine G Carman. 2017. “Knowledge as a Predictor of Insurance Coverage Under the Affordable Care Act.” *Medical Care* 55 (4).

Liu, Chenyan, and Justin Sydnor. 2018. “Dominated Options in Health-Insurance Plans.” Working Paper. National Bureau of Economic Research.

Loewenstein, George, David Hagmann, Janet Schwartz, Keith Ericson, Judd Kessler, Saurabh Bhargava, Jennifer Blumenthal-Barby, et al. 2017. “A Behavioral Blueprint For Improving Health Care Policy.” *Behavioral Science & Policy* 3 (January): 53–66. https://doi.org/10.1353/bsp.2017.0005.

Wang, Annabel Z., Karen A. Scherr, Charlene A. Wong, and Peter A. Ubel. 2017a. “Poor Consumer Comprehension and Plan Selection Inconsistencies Under the 2016 HealthCare.Gov Choice Architecture.” *MDM Policy & Practice* 2 (1): 2381468317716441. https://doi.org/10.1177/2381468317716441.

———. 2017b. “Poor Consumer Comprehension and Plan Selection Inconsistencies Under the 2016 HealthCare.Gov Choice Architecture.” *MDM Policy & Practice* 2 (1): 2381468317716441. https://doi.org/10.1177/2381468317716441.

Appendix: Sample Table Shells

Table 1: Balance tests: Administrative data

	Control Mean	Treatment Mean	*T-test p-value*
	(1)	(2)	(3)
Subsidy FPL % *SE*			
Prefer a language other than English			
SAWS-initiated application			
Subsidy Ineligible			
Female HOH			
Age			
Race/ethnicity			
Non-Hispanic White			
Hispanic			
Non-Hispanic Black			
Races besides Black, Hispanic, White			
Number of insurers operating in the region			
Median premium of health plans in each tier of the ACA marketplace in the zipcode/region			
Number of agents or certified enrollment counselors in the zipcode/region			

Household size			
Any prior recorded enrollment in Covered California			
Risk score			
Pooled F – stat ***p-value*** ***N***			

Table 2: Balance tests: Survey data

	Control Mean	Treatment Mean	*T-test p-value*
	(1)	(2)	(3)
Match Rate with Admin Sample			NA
Subsidy FPL % *SE*			
Prefer a language other than English			
SAWS-initiated application			
CSR Eligible			
Subsidy Ineligible			
Female HOH			
Age			
Race/ethnicity Non-Hispanic White			
Hispanic			
Non-Hispanic Black			
Races besides Black, Hispanic, White			
Number of insurers operating in the region			
Median premium of health plans in each tier of the ACA marketplace in the zipcode/region			
Number of agents or certified enrollment counselors in the zipcode/region			

Household size			
Any prior recorded enrollment in Covered California			
Risk score			
Factors related to the consumer's motivation to find health insurance			
Lack employer sponsored health insurance			
Have a chronic condition			
In fair or poor health			
Currently take any prescriptions			
Factors related to possible information barriers experienced by the consumer			
Less than high school education			
Thought there was no penalty for being uninsured in 2018			
Pooled F – stat ***p-value*** ***N***	. . .		

Table 3: First stage: Effect of randomization into the treatment group on receiving an outbound call

	Entire List (Main Sample)		Survey Respondent Sub-sample	
	Control mean (1)	Estimated First Stage (2)	Control Mean (3)	Estimated First Stage (4)
SE *p*				
Sample Size				

Table 4: Impacts of randomization to the treatment group (reduced form) and of receiving an outbound call from an SCR (2SLS)

	Control mean	Reduced form	2SLS
	(1)	(2)	(3)
Panel A: CC Administrative Data			
Plan selection/Effectuated enrollment *SE* *p*			
Selecting a dominated plan			
CDPS risk scores			
Panel B: CC Survey Data			
Satisfaction with the application process			
Likelihood of recommending Covered California to a friend or colleague			
Had difficulty with entering details in the Covered California application			
Potential mechanisms underlying the effect of the intervention			
Unaware of penalty for being uninsured in 2019			
Had difficulty shopping and comparing plans			
Had difficulty getting needed information during the enrollment process			
Had difficulty finding if a doctor or hospital was covered by a plan			
Had difficulty finding what their monthly premium would be for a Covered California plan			
Had difficulty understanding total cost-sharing			

Health insurance enrollment decisions			
Have health coverage			
Main source of health coverage is via current or former employer or union			
Main source of health coverage is via Medi-Cal or Medicaid			

Table 5: Effects of randomization to the treatment group on sorting of consumers across market tiers

Panel A. Average marginal effects from multinomial logit model

All households

	Bronze - HSA	Bronze– non-HSA	Silver	Gold	Platinum	None
	(1)	(2)	(3)	(4)	(5)	(6)
Randomization into treatment x CDPS risk scores (χ) SE p						
Randomization into treatment						
CDPS risk scores						

Households eligible for enhanced silver plans

	Bronze - HSA	Bronze– non-HSA	Silver	Enhanced Silver	Gold	Platinum	None
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Randomization into treatment x CDPS risk scores (χ) SE p							
Randomization into treatment							
CDPS risk scores							

Panel B. Average marginal effects from nested logit model

All households

	Bronze - HSA	Bronze–non-HSA	Silver	Gold	Platinum	None
	(1)	(2)	(3)	(4)	(5)	(6)
Randomization into treatment x CDPS risk scores (χ) SE p						
Randomization into treatment						
CDPS risk scores						

Households eligible for enhanced silver plans

	Bronze - HSA	Bronze–non-HSA	Silver	Enhanced Silver	Gold	Platinum	None
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Randomization into treatment x CDPS risk scores (χ) SE p							
Randomization into treatment							
CDPS risk scores							

Table 6: Variation in consumers' engagement with the intervention: Predictors of having a conversation with an SCR

	OLS model	Logit model	Probit model
Randomization to the treatment arm			
Subsidy FPL %			
Prefer language other than English, and spoke with bilingual SCR			
Prefer language other than English, and spoke with non-bilingual SCR			
SAWS-initiated application			
CSR Eligible			
Subsidy Ineligible			
Female HOH			
Age			
Race/ethnicity			
Non-Hispanic white			
Hispanic			
Non-Hispanic black			
Races besides black, Hispanic, white			
Number of insurers operating in the region			
Median premium of health plans in each tier of the ACA marketplace in the zipcode/region			
Number of agents or certified enrollment counselors in the zipcode/region			
Household size			
Risk score			
Any prior recorded enrollment in Covered California			

Note: Average marginal effects are presented in all three columns.

Table 7: Heterogeneity tests

	N	First stage	Enrollment - Reduced form	Enrollment – 2SLS	Choice error – Reduced form	Choice error – 2SLS	Health risk of enrollees – Reduced form	Health risk of enrollees – 2SLS	Having a conversation with SCR
Full sample									
Panel A: CC Admin Data									
Aid code									
Subsidy-eligible									
Subsidy-ineligible									
p-value of difference									
Language preference									
English									
Any other language									
p-value of difference									
Service language									
Prefer language other than English, and spoke with bilingual SCR									
Prefer language other than English, and spoke with non-bilingual SCR									
p-value of difference									
Race/ethnicity									
Minority									
Non-minority									
p-value of difference									
Age									
Above median age									
Below median age									
p-value of difference									
Panel B: CC Survey Data									
Factors related to the consumer's motivation to find health insurance									
Feel it is important to have health insurance									
Have a chronic condition									
In fair or poor health									
Currently take any prescriptions									
Factors related to possible information barriers experienced by the consumer									
Less than high school education									
Thought there was no penalty for being uninsured in 2018									

Table S1: Comparison of actual and simulated randomization to the treatment group

	Mean in those selected into randomization	Mean of mean in simulations	SD of mean in simulations	*((1)-(2))/(3)*
Subsidy FPL %				
Spanish speaker				
English speaker				
SAWS-initiated application				
CSR Eligible				
Subsidy Ineligible				
Female HOH				
Age				
Race/ethnicity				
Non-Hispanic white				
Hispanic				
Non-Hispanic black				
Races besides black, Hispanic, white				
Number of insurers operating in the region				
Median premium of health plans in each tier of the ACA marketplace in the zipcode/region				
Household size				
Risk score				
Any prior recorded enrollment in Covered California				

Table S2: Sensitivity of treatment effects to inclusion or omission of covariates

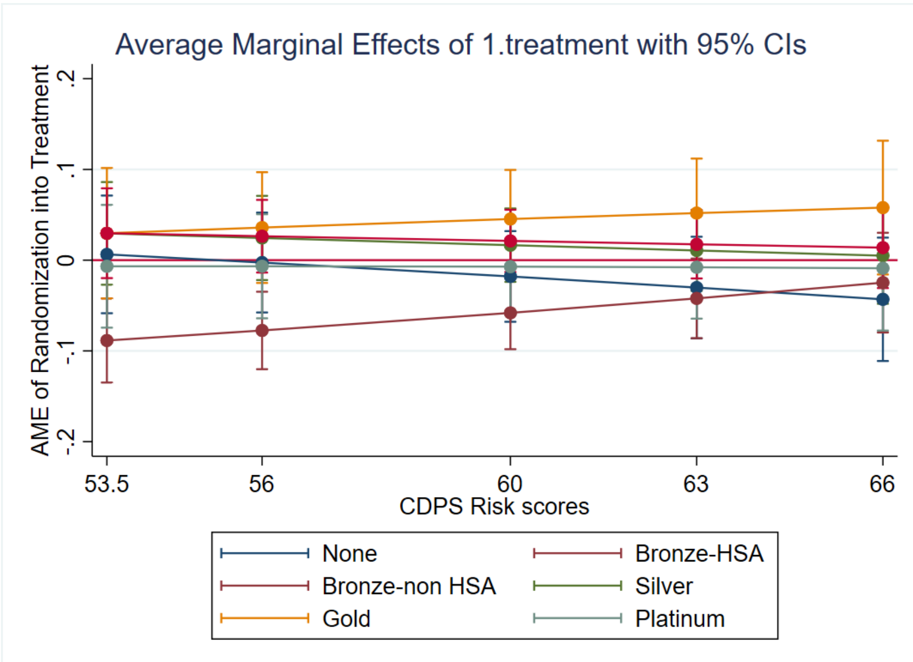
	Reduced Form				2SLS			
	Baseline	No location fixed effects	No covariates	No listwise deletion	Baseline	No location fixed effects	No covariates	No listwise deletion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: CC Administrative Data								
Plan selection/Effectuated enrollment								
Selecting a dominated plan								
CDPS risk scores								
Conversation with an SCR					NA	NA	NA	NA
Panel B: CC Survey Data								
Satisfaction with the application process								
Likelihood of recommending Covered California to a friend or colleague								
Had difficulty with entering details in the Covered California application								
Potential mechanisms underlying the effect of the intervention								
Unaware of penalty for being uninsured in 2019								
Had difficulty shopping and comparing plans								
Had difficulty getting needed information during the enrollment process								
Had difficulty finding if a doctor or hospital was covered by a plan								
Had difficulty finding what their monthly premium would be for a Covered								

California plan								
Had difficulty understanding total cost-sharing								
Health insurance enrollment decisions								
Have health coverage								
Main source of health coverage is via current or former employer or union								
Main source of health coverage is via Medi-Cal or Medicaid								

Figure 1: Impacts of randomization to the treatment group on tier choice: Heterogeneity by risk score

Note: *These data are simulated and used for expositional purposes only.*

Panel A: All households



Panel B: Households eligible for enhanced silver plans