The Downside of More: Choice Overload in Health Insurance Markets

Chenyuan Liu[1]
Tsinghua University

Jiacheng Liu[2*]
Beijing Jiaotong University

Li Wei[3]
Renmin University of China

Yi Yao[4]
Peking University

Abstract

Health insurance markets often offer multiple plans, yet consumers frequently choose strictly dominated options. Using two natural experiments in China's private health insurance market, we study how choice set size affects choice quality. An insurer expanded choice menus from two to three plans and later reduced them back to two in an exogenous manner. Exploiting strictly dominated plans and a regression discontinuity design, we find that adding one option increases the probability of choosing a dominated plan by 8.4 percentage points (52%), while removing one reduces the probability by 11.0 percentage points (49%). Overall, the evidence points to choice overload rather than changes in average plan quality as the primary driver of the observed changes in consumer choice quality.
JEL Classification: G22, I13. *Keywords*: Health insurance; Dominated plan; Choice set size; Choice overload.

---

[1] Department of Economics, School of Economics and Management, Tsinghua University, Beijing, China. Email: liuchy3@sem.tsinghua.edu.cn. Liu receives financial support from the National Natural Science Foundation of China, project number 72203120.
[2] Department of Labor Economics, School of Economics and Management, Beijing Jiaotong University, Beijing, China. Email: jcliu@bjtu.edu.cn. *Corresponding author.
[3] Department of Insurance, China Institute of Insurance, School of Finance, Renmin University of China, Beijing, China. Email: weil@ruc.edu.cn and weili_ruc@126.com.
[4] Department of Risk Management and Insurance, School of Economics, Peking University, Beijing, China. Email: yao.yi@pku.edu.cn. Yao receives financial support from the National Social Science Fund of China, Grant/Award Number: 23&ZD178, and the Young Faculty Research Seed Fund at SOE PKU.
All authors contributed equally to this work.

1. Introduction

Health insurance choice is important because it is a point of entry into decision-making in the health care sector (Chandra et al., 2019). In many markets, consumers must choose from health insurance menus with multiple options. However, the literature has documented that consumers often choose suboptimal plans or even dominated plans (Abaluck and Gruber, 2011, 2023; Bhargava et al., 2017; Handel, 2013; Handel et al., 2024; Liu and Sydnor, 2022; Samek and Sydnor, 2025).[1]

There remains a debate about ways to design effective policies to enhance consumers' decision quality in this context. For example, prior literature has found mixed evidence on whether consumers learn over time (Abaluck and Gruber, 2016a, 2016b; Ketcham et al., 2012, 2016) and on whether information interventions improve choice consistency (Bundorf et al., 2019; Ericson et al., 2017; Gruber et al., 2020; Kling et al., 2012; Samek and Sydnor, 2025). The literature has also documented that changing choice architecture, including standardizing plan options (Ericson and Starc, 2016) or reducing choice set size (Abaluck and Gruber, 2023), can enhance choice quality.

In this paper, we build on this literature and explore how choice set size affects consumers' choice quality, which has received limited scrutiny, as reported by Abaluck and Gruber (2023). We leverage a setting in which individuals purchase a private health insurance plan with multiple options in China. Our data are from a leading private health insurance company. In an exogenous manner, the company expanded the menu size for renewing customers from two to three plans, and then reduced from three to two consecutively. These constitute two natural experiments in our setting. Another important feature that enables us to draw a clear comparison of consumer choice quality is the prevalence of strictly dominated plans within each menu, i.e., plans with less coverage (higher deductibles, less covered benefits, and/or narrower networks) but charge a higher premium. This allows us to derive an unambiguous and consistent measure of consumer choice quality across different menus without imposing additional assumptions on loss distributions and risk preferences.

[1] Dominated plans generate higher total spending (plan premiums plus out-of-pocket medical spending) at any level of medical spending, compared to their alternatives (see Liu and Sydnor, 2022, for a detailed discussion).

We apply a regression discontinuity in time design to identify the causal impact of changing the choice set size on consumer choice quality. We provide evidence of identification validity by showing that there is no discontinuous change in the density of the running variable (McCrary, 2008) or in the values of the covariates at the cutoff dates.

We find that increasing the choice set size leads to a decline in consumer choice quality, and conversely, a smaller choice set improves choice quality. Specifically, we show that adding one insurance plan to a two-plan menu significantly increases the probability of choosing a strictly dominated plan by around 8.4 percentage points in absolute terms, or 52% relative to the baseline mean before the menu change. In contrast, reducing the choice set size from three to two plans leads to a decrease of 11.0 percentage points in choosing dominated plans, or 49% in relative terms. The two estimates are statistically close to each other, suggesting the effects are likely to be symmetric.

To rule out unobserved contemporaneous shocks, we additionally conduct a placebo test using a separate sample of enrollees, whose plan menu remains unchanged at the cutoff date of the second natural experiment and includes a strictly dominated plan. For this group, we find no evidence of a discontinuous change in the probability of choosing a dominated plan on the same day.

We further investigate the potential mechanism behind the choice of dominated plans and how the changes in the choice set size impact it. The first potential mechanism is choice overload, that is, consumers' choice functions perform worse in the face of larger choice sets (e.g., Besedeš et al., 2015). Several stylized facts in our data support this mechanism. First, we find that more-experienced consumers make better choices before the menu changes and are less affected by the changes in choice set size. Second, we leverage another natural experiment in the firm, in which some insurance agents are trained with artificial intelligence (AI) tools to help their customers with plan selection. Although the program improves choice quality in general, the improvement is only noticeable among consumers choosing from three-plan menus, that is, the group featuring the lower baseline choice quality. This result suggests that information interventions can mitigate the performance gap in choice functions between the two-plan and three-plan menu settings. Third, we find that consumers seem to be more satisfied with

their chosen plans when selecting from a smaller menu.[2] This result is consistent with the previous literature on choice overload, which documents that larger choice sets can reduce satisfaction with the chosen options (e.g., Iyengar and Lepper, 2000).

Next, we explore another possible mechanism, i.e., whether providing more options leads to worse choices because the average plan quality declines as the plan menu grows larger (Abaluck and Gruber, 2023). We measure the average plan quality using the fraction of plans that are dominated. For more than half of the consumers, the proportion of dominated plans is 50% in a two-plan menu (one dominated plan) and 67% in a three-plan menu (two dominated plans). Thus, worse choice quality under a three-plan menu is consistent with the corresponding deterioration in average plan quality. However, there are other consumers (around 30%) who face only one dominated option in both the two-plan and three-plan menus (so the average quality is lower under the two-plan menu). We find that these consumers are still more likely to choose dominated options under the three-plan menu when the average plan quality improves, suggesting that variation in the average plan quality is not the driving force leading to worse choice quality.

We contribute to a growing strand of literature on choice frictions in health insurance markets, in particular, the impact of choice set size on health insurance choice quality (e.g., Abaluck and Gruber, 2023; Bhargava et al., 2017). The findings from the prior literature have been somewhat mixed. Some survey experiments have found no significant link between menu size and consumers choice quality when plans vary along fewer dimensions (e.g., Bhargava et al., 2017). However, recent field studies have found that consumer choices are improved under a reduced number of plans (e.g., Abaluck and Gruber, 2023). We connect these findings by demonstrating that plan menu reduction can improve choice quality even when the original choice set is small, especially when the number of plan attributes for comparison is large. Within our setting, consumers

[2] Following the spirit of the existing literature, our choice satisfaction measure is derived from the enrollee's decision to sign up for automatic renewal for the following year. Gruber et al. (2020) and Ketcham et al. (2012, 2015) utilize plan switching as an ex-post measure of plan satisfaction. In contrast, our measure captures the complementary dimension: the intention to remain in the chosen plan.

choose from a small choice set of no more than three plans, yet they must evaluate around 20 plan characteristics—in both financial and nonfinancial terms.[3]

Second, our findings add to the discussion on choice overload in the context of health insurance (e.g., Abaluck and Gruber, 2023; Ketcham et al., 2015; Li et al., 2025). Our study differs from prior work by documenting choice overload as the primary driver of consumer choice quality. We also show that other explanations, such as changes in average plan quality (Abaluck and Gruber, 2023), are of secondary importance in our context.

Third, we contribute to the strand of literature on health insurance choice frictions in the developing countries, including China, thus supplement existing studies based mostly on experiences in developed countries. This extension is meaningful, as the previous literature find that health insurance choice patterns vary significantly with socioeconomic factors (Handel et al., 2024). We show that in developing countries like China, where consumers of private insurance plans are not experienced in navigating health insurance plans given the limited choice in the public sector, choice inconsistency can happen even with a small choice set. Our findings also highlight that it is necessary for regulators to monitor the design and changes in private health insurance plan menus cautiously.

The rest of the paper is organized as follows. Section 2 outlines the institutional setting, defines the choice quality measure, and describes the empirical data. Section 3 describes the identification strategy and discusses identification validity. Section 4 presents the main results and robustness checks. Section 5 discusses the potential mechanisms. Section 6 concludes with policy implications.

[3] On the one hand, prior studies consider a larger number of available plans than we do. For example, the number of plans varies from four to twelve in Bhargava et al. (2017). Abaluck and Gruber (2023) study menus containing two to seven plans, and the average number of plans in their main sample is around four. In contrast, the size of our choice set varies from two to three. On the other hand, prior studies consider fewer plan attributes than we do. For example, in Bhargava et al. (2017), the number of plan attributes vary from two (the deductible and maximum out-of-pocket spending) to one (the deductible). Abaluck and Gruber (2023) hold all nonfinancial characteristics of plans constant. In comparison, within our setting, consumers compare a dozen plan characteristics—both financial (e.g., the deductible and the premium) and nonfinancial (e.g., items covered for treatment).

2. Setting and data

2.1 The sampled health insurance product

In China, most of the population is covered by the public basic medical insurance program. Some individuals may purchase additional coverage from private health insurance company. We utilize a popular inpatient care insurance product from a leading private health insurance company in China. Consumers in this market are predominantly young (Zhu et al., 2024) and are less experienced in navigating health insurance plans, given the limited choice in the public sector.[4] Since the product is in the form of one-year contract, policyholders must renew annually to maintain coverage. At renewal, an insured individual is granted the right to make an active choice among the multiple plans offered by the insurer.

This setting provides a unique opportunity to study choice frictions among relatively inexperienced young market participants in developing countries, a population that stands in sharp contrast to the frequently studied consumers in the existing literature, such as Medicare beneficiaries.

2.2 The natural experiments

We study the effect of changing the choice set size on consumer choice quality. We leverage two natural experiments that altered the renewal menus for existing enrollees. These menu changes were exogenous for the following reasons. The menu changes stemmed from supply-side adjustments and were orthogonal to localized or individual-level demand shocks. The changes took place without any preannouncement and took effect for all affected enrollees at the same time; thus, the changes should not have been anticipated by the consumers. As policyholders could only renew through personalized renewal links that directed them to predetermined renewal menus, consumers' strategic manipulation of the choice sets was also unlikely.

The timeline of the natural experiments and detailed choice menus during our observation period (2020–21) are summarized in Figure 1: (i) Menu 1: before July 2020, individuals faced two renewal options—Plan A and Plan B; (ii) Menu 2: since July 1, 2020, Plan C was included in the menu, and the other two plans remained the same; (iii) Menu 3: since January 29, 2021, Plan B and Plan C were

[4] Public basic medical insurance program in China provides basic coverage with very limited choices. Most employed residents are assigned a single plan, while others typically face only a couple options (typically no more than two plans).

replaced by Plan D. These transitions resulted in two exogenous variations in menu size, allowing us to identify the causal effect of changing choice set size on consumer choice quality.

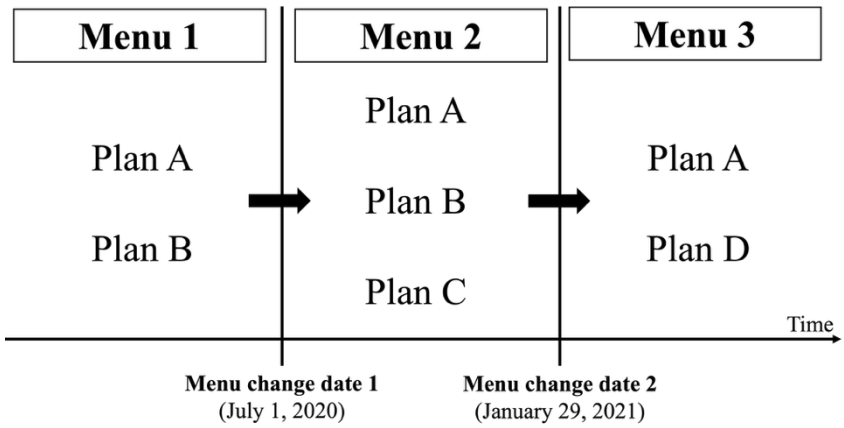

Figure. 1: Timeline of natural experiments and plan menu changes

The four plans mainly covered medical expenses for inpatient and critical-illness-related treatments that exceed the coverage cap for public basic medical insurance. Plan A was the initial plan choice for all sampled enrollees in the years preceding each natural experiment. Relative to Plan A, the subsequent plans offered expanded coverage.[5] Specifically, Plans B, C and D provided additional coverage for several rare diseases. Beyond this, Plan C introduced a shared deductible for enrolled family members, while Plan D extended the coverage window for pre-hospitalization outpatient services. Despite offering better insurance protection, Plans B, C, and D were less expensive than Plan A for some enrollees due to promotional premium discounts. The base premium and premium discounts also varied by enrollee age and household size, resulting in variation in net premiums across different enrollees. See Appendix A.1 for plan details.

Despite the small menu size, the online renewal choice interface displayed high information density, including approximately 20 plan characteristics—in financial terms (e.g., base premium, premium discount, and deductible) and nonfinancial terms (e.g., covered treatment items). Consequently, the overall complicated plan

[5] Enrollees may add additional coverage for certain diseases or services on top of the basic coverage for Plans A, B, and D.

features may pose a significant challenge for consumers in accurately evaluating and comparing their options.

2.3 Consumer choice quality measure

We leverage the presence of strictly dominated plans to construct a consistent measure of choice quality across menus, defined by whether an individual chooses such a strictly dominated plan. Formally, for each option within a menu, we compare its coverage and premium against all alternatives. A plan is deemed strictly dominated if there exists at least one alternative that (i) offers the same or better insurance protection (i.e., a broader network, a lower deductible, or coverage for more diseases and benefits) and (ii) charges an equal or lower premium, with at least one of these two conditions being strictly favorable to the consumer. In other words, a strictly dominated option entails strictly higher enrollee spending (out-of-pocket spending plus premiums) in any scenario (Bhargava et al., 2017; Liu and Sydnor, 2022). In our context, for many enrollees, Plan A is strictly dominated by other options because it offers less coverage and is more expensive. Moreover, sometimes Plan B or C is also strictly dominated, depending on enrollees' household size and age. We detail the logic of identifying strictly dominated options in Appendix A.2.

Our measure of choice quality avoids strong assumptions regarding out-of-pocket expenditures and does not require individual-level medical spending data, compared to existing approaches (e.g., Abaluck and Gruber, 2023).[6]

2.4 Sample construction and summary statistics

Our empirical analysis is based on a dataset acquired from the private health insurance product provider. Our raw dataset consists of individuals who were originally enrolled in 2017 or 2018 (the insurance product was first introduced to the market at the end of 2016), who held Plan A in the years preceding each intervention. We have information on their choice sets (as shown in Figure 1) and plan choices throughout both natural experiments.

To ensure that the choice set contains strictly dominated options regardless of the menu provided, we restrict the sample to enrollees with no claim history and exclude those aged 16–20 or over 60. To obtain cleaner estimates, we impose two

[6] For instance, the perfect foresight assumption in Abaluck and Gruber (2023) implies that enrollees know exactly what their out-of-pocket costs will be in the coming year.

additional sample restrictions. First, we exclude individuals who placed their renewal orders on the exact dates of the menu changes. Because these changes took place in the middle of the day and we only observe daily-level data, removing these observations avoids a downward bias in the estimated treatment effects. Thus, the days immediately following the menu changes are defined as the cutoff dates. Second, to ensure comparability, we only keep those who placed renewal orders within a 60-day bandwidth around each cutoff date.

For this final study sample, we observe significant differences across several observed characteristics for individuals who made their choices under the different menus (see Appendix Table B.1 for details). This indicates that a direct comparison between individuals below and above the cutoffs may lead to biased estimates due to confounders. This fact motivates our regression discontinuity design in time (RDiT) design, which we discuss next.

3. Empirical design

3.1 Main identification strategy

We employ a RDiT design to estimate the effect of changing the menu size on consumer choice quality. Two natural experiments, as discussed in section 2.2, create exogenous variations in the choice sets. We compare consumers' choices immediately prior to and following the menu changes. Specifically, we estimate the following equation:

$$Y_{it} = \alpha + \beta D_{it} + f(X_{it}) + \varepsilon_{it}, \quad (1)$$

where i and t index policyholders and time, respectively. $Y_{i,t}$ is the outcome variable of interest, indicating whether an individual selects a strictly dominated plan at renewal. D_{it} is a dummy variable indicating whether consumer i makes the plan choice under the new menu (Menu 2 and Menu 3 in the first and second natural experiments, respectively). X_{it} is the running variable representing the plan choice date (renewal date) relative to the cutoff date.

We employ a local linear specification: $f(X_{it}) = \gamma_1 X_{it} + \gamma_2 D_{it} X_{it}$. We impose a triangular kernel to assign lower weights to observations further from the cutoffs. For bandwidth selection, we follow the approach proposed by Calonico et al. (2014) for each regression and calculate that the optimal bandwidth is around 14

days. To enable comparisons across regressions, we follow the literature (e.g., Ang et al., 2024; Giuntella and Mazzonna, 2019; Lee and Vabson, 2024) and set a uniform bandwidth of 14 days across all outcomes in the main regressions. Standard errors are clustered at the calendar day level to account for potential correlation in the errors among enrollees who make plan choices on the same day.

3.2 Validity tests

The validity of our identification strategy relies on the assumption that individuals cannot precisely manipulate the running variable around the threshold. Since both menu changes took place without any announcement and plan choices were made via personalized web links, strategic sorting is unlikely. To provide empirical evidence, we test for discontinuity in the density of the running variable (McCrary, 2008) and find no evidence of excess bunching at either cutoff (see Appendix Figure C.1). We also check for covariate balance and find no significant differences in observed characteristics across either cutoff (see Appendix Table C.1, Figures C.1 and C.2).

4. Main results and robustness checks

4.1 Main results

This section presents the main results on the effect of changing choice set size on consumer choice quality. As shown in column (1) of Table 1, expanding the menu from two to three plans increases the probability of choosing a dominated plan by 8.4 percentage points. The estimated effect is highly significant at the 1% level and large in relative terms, at around 52% relative to the baseline mean of 16.1 percentage points. Mirroring this, we estimate that reducing the choice set size from a menu of three plans to a two-plan menu lowers the probability of choosing a dominated plan by 11.0 percentage points, or approximately 49% compared to the baseline mean (column (2)).

Consistently, Figure 2 documents a sharp discontinuity in consumer choice quality measure on the days when the choice set sizes were changed. As panel A shows, the probability of making choice mistakes, measured by the probability of choosing a strictly dominated plan at renewal, increases sharply after the menu size increases from two to three plans. Consistent with our expectation, panel B shows that reducing the menu size from three to two plans reduces the probability of choosing a dominated plan (i.e., a consumer choice quality improvement).

In sum, the findings from the two natural experiments yield a consistent pattern in how the changes in menu size affects consumer choice quality, and the estimated sizes of the effect are significant and generally comparable.

Table 1: Effect of changing choice set size on consumer choice quality

	Outcome: Choosing a strictly dominated plan	
	(1)	(2)
	Increasing menu size from two to three	Reducing menu size from three to two
RDiT estimate	0.084***	-0.110***
	(0.023)	(0.031)
Eff. number of obs.	7,699	5,433
Below-cutoff mean	0.161	0.225

Notes: "Below-cutoff mean" is the unconditional sample mean of the outcome for policyholders who renewed within a 14-day bandwidth below the cutoff.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

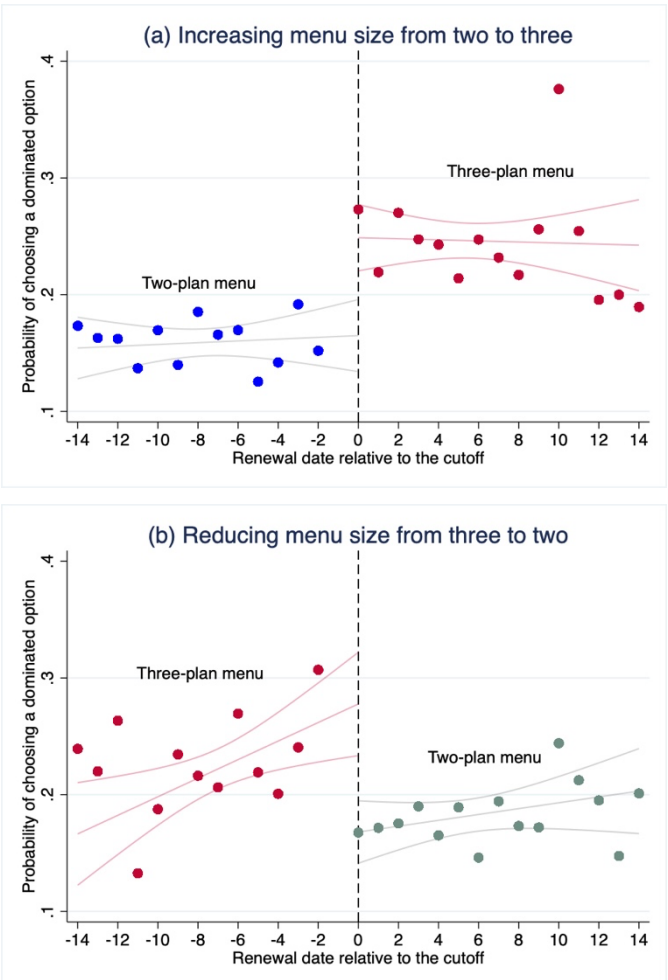

Figure 2: Changes in menu size and consumer choice quality

Notes: Each panel plots the probability of choosing a strictly dominated plan at renewal for each standardized renewal-date bin, together with fitted lines and 95% confidence intervals. The fitted lines are produced using regressions with a triangular kernel. The first, second, and third choice menus are plotted in blue, red, and green, respectively.

4.2 Robustness and validity

Specification checks. We find highly consistent results under alternative specifications: implementing a donut-hole RDiT (excluding the nearest observations within a one-day radius around the cutoff), using a quadratic polynomial, applying a uniform kernel, changing bandwidth, and adding covariates (see Appendix Table C.2).

Placebo test to rule out unobserved contemporaneous shocks. We replicate our baseline analysis based on a separate sample of enrollees of the same health insurance product, whose choice menu remains unchanged at the cutoff date of the second natural experiment and includes a strictly dominated plan. We estimate no significant change in the choice quality measure for this group (see Appendix C.4).

Testing for changes in the renewal rate. We find no significant change in the renewal rate at the cutoffs, suggesting minimal extensive-margin responses (see Appendix D). This finding is noteworthy in its own right; furthermore, it confirms that our baseline findings are not driven by extensive-margin selection effects in the population of consumers who choose to renew at the cutoffs.

5. Discussion of mechanisms

The results in Section 4 demonstrate that larger choice sets reduce consumer choice quality, whereas reducing choice set size improves it. This section explores two potential mechanisms discussed in recent literature: (i) choice overload (e.g., Besedeš et al., 2015) and (ii) changes in average plan quality (e.g., Abaluck and Gruber, 2023).

5.1 Choice overload

5.1.1 Heterogeneous effects by enrollment experience

A plausible explanation behind our main findings is choice overload, i.e., larger choice sets lead to lower performing choice functions (e.g., Besedeš et al., 2015). If this mechanism holds, the negative treatment effect of expanding menu size should be smaller for individuals with better baseline choice performance. The underlying assumption is that these individuals are more likely to choose wisely, even if their cognitive capacity is somewhat restricted by a larger menu, compared to those with worse baseline choice performance.

We split the estimation sample into "more experienced" and "less experienced" consumers based on the median number of historical enrollment records.[7] In both natural experiments, the more experienced group exhibits a lower pre-treatment probability of choice mistakes (see summary statistics in Appendix Table E.1).

Furthermore, consistent with choice overload, Figure 3 reveals substantial heterogeneity in treatment effects by consumers' enrollment experience. Specifically, the more experienced group is not affected by the changes in the menu size. The estimated coefficients are insignificant and close to zero (0.030 and -0.028). In contrast, the less experienced customers make more choice mistakes when choose from a larger menu and fewer mistakes choosing from a smaller one. The estimates for this group are highly significant and substantially larger (0.141 and -0.171). See Appendix E.1 for detailed RDiT estimates and figures.

These results suggest that consumers do not choose solely based on heuristics (e.g., Abaluck and Gruber, 2023) but can learn from their own enrollment experience (e.g., Ketcham et al., 2012), providing evidence supporting the choice overload mechanism.

[7] As a robustness check, we use an alternative definition of "more experienced" and "less experienced" consumers, based on the initial enrollment year. Specifically, we assign consumers who initially enrolled in 2017 to the "more experienced" group and those who enrolled in 2018 to the "less experienced" group. The results remain qualitatively similar to those reported and are available upon request.

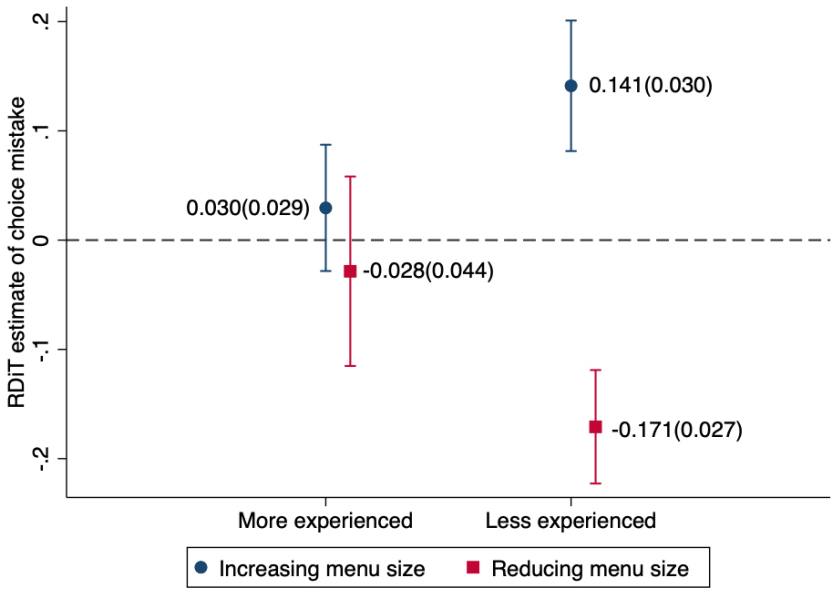

Figure 3: Heterogeneous treatment effects by baseline enrollment experience
Notes: This figure reports the RDiT estimates of choice mistake and 95% confidence intervals by natural experiment and consumers' enrollment experience.

5.1.2 Decision support program

Another piece of evidence point to the choice overload mechanism is provided by exploring the effect of an AI-based decision support program provided to the insurance agent, which was aimed at improving consumer decision-making. In November 2020, the insurer launched the decision support program. An AI-based decision support tool was provided to agents in the program to help them analyze clients' needs and identify the best plans. The program also included training sessions taught by lecturers from the leading universities on how to identify better plans for clients. Treatment turned on for insured enrollees who made plan choices after their agents participating in the program.[8]

Leveraging the staggered implementation of the program, we employ a difference-in-differences (DID) design to investigate its impact on consumer choice quality.

[8] For the sampled individuals, client-agent matching occurred during their initial enrollment in 2017 and 2018, years prior to the implementation of the decision support program.

More importantly, we focus on how this effect varies along the changes in renewal menu size.[9]

To obtain accurate estimates in the presence of multiple treatments (changing choice menu and providing decision support), we follow the literature (e.g., Callaway et al., 2024; de Chaisemartin and d'Haultfoeuille，2023, 2024) and estimate the treatment effects of the decision support program on choice quality separately for different menus. Accordingly, we focus on those who renewed under Menu 2 (from July 10, 2020, to January 25, 2021) and Menu 3 (from February 1 to March 31, 2021). To ensure the existence of strictly dominated plans, we apply the same sample selection criteria as in Section 3. We also exclude enrollees with missing information on the treatment status.

We estimate the following model:

$$Y_{igt} = \alpha_0 + \beta Treat_g \times Post_t + \lambda_g + \Gamma_t + \varepsilon_{igt}, \qquad (2)$$

where i, g, and t index policyholders, treatment groups, and time, respectively. Y_{igt} is the measure of choice mistakes, an indicator of whether an individual chooses a strictly dominated plan at renewal. $Treat_g$ is the treatment group indicator, which equals 1 only for units affected by the support program within the sample period. $Post_t$ is the post-treatment period indicator. λ_g denotes initial treatment month fixed effects, representing the calendar months when the agents first joined the program. Γ_t denotes calendar month fixed effects.

We use the estimator of Callaway and Sant'Anna (2021) to address treatment-effect heterogeneity. We use the never-treated individuals as the control group. Standard errors are clustered at the calendar date level as in the baseline RDiT specification.

Table 2 reports the estimates. Specifically, under the three-plan menu (Menu 2), the decision support program significantly lowers the probability of making choice mistakes by 22.1 percentage points (column (1)). In contrast, for the group

[9] Alternatively, one could test how the treatment effect of changing choice set size varies between individuals with and without decision support in natural experiment 2. However, comparing RDiT estimates across these subsamples is precluded by a lack of observations near the cutoff. In the main estimation sample—those who renewed within 14 days around the cutoff—fewer than 50 individuals were affected by the program, rendering RDiT estimates for this group statistically unreliable.

choosing from the smaller two-plan menu (Menu 3), the treatment effect is insignificant and relatively small in magnitude (3.8 percentage points, column (2)). These findings are further supported by event-study results (Appendix Figure E.2) and robustness checks using the not-yet-treated individuals as control and employing a two-way fixed effects estimator (Appendix Table E.2).

The findings indicate that the decision support program is only effective in reducing choice mistakes when a larger menu is given, where individual choice functions are more likely to be weakened.

Table 2: Effect of decision support on choice quality by plan menu

	Outcome: Choosing a strictly dominated option	
	(1)	(2)
	Choosing under Menu 2 (three-plan menu)	Choosing under Menu 3 (two-plan menu)
DID estimate	-0.221**	0.038
	(0.109)	(0.062)
Obs.	47,046	15,079
Pre-treatment mean	0.294	0.112

Notes: "Pre-treatment mean" reports the unconditional sample mean of the outcome for the treatment group policyholders during the pre-treatment period. The estimation sample here departs from that in the baseline RDiT analysis by: (i) restricting the scope to individuals choosing under Menu 2 and Menu 3; (ii) excluding observations with missing data on agents' enrollment in the decision-aid program; and (iii) extending the sampling window to maximize the inclusion of ever-treated individuals.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5.1.3 Satisfaction with the chosen option

Previous literature on choice overload also documents that consumers may be more satisfied with their chosen options when choosing from a smaller choice set (e.g., Iyengar and Lepper, 2000). We test this prediction by utilizing additional information collected in the second natural experiment.

We measure consumer satisfaction based on the preference to remain enrolled with the current chosen option.[10] In the second natural experiment, immediately after enrollees chose a plan, the insurer sought consent for automatic renewal of the chosen options in the following year, allowing enrollees to modify this

[10] Gruber et al. (2020) and Ketcham et al. (2012, 2015) utilized plan switching as an ex-post measure of plan satisfaction. Our approach aligns with this literature by focusing on the intention to remain in the chosen plan.

decision throughout the year. Our data capture their final choice recorded at the end of the year. We find that reducing menu size increases the probability that the consumer signs up for automatic renewal (see Appendix Table E.3 and Figure E.3).

In sum, this section provides evidence consistent with choice overload, showing that the effects of changing choice set sizes can be mitigated by consumer experience and information interventions, and a smaller menu is accompanied by a higher level of consumer satisfaction.

5.2 Changes in average plan quality

Abaluck and Gruber (2023) document that more options lead to worse choices mainly because the average plan quality declines as the plan menu expands. To test for this mechanism, we examine how menu changes affect average plan quality and then verify whether this trend aligns with the observed changes in consumer choice quality.

We measure average plan quality within a menu as the proportion of baseline plans that are strictly dominated (see Appendix Table A.1 for details). For more than half of the consumers, the proportion of dominated plans is 50% in a two-plan menu (one dominated plan) and 67% in a three-plan menu (two dominated plans). For these individuals, we find that the direction of the changes in plan quality is generally consistent with the changes in consumer choice quality (see Panel A of Appendix Table F.1).

However, a significant proportion of individuals (households with two members enrolled) face only one dominated plan in both the two- and three- plan menus, so the average quality is lower under the two-plan menu. It's noteworthy that these enrollees are nonetheless more likely to make choice mistakes after the menu expands (i.e., the average plan quality improves), indicating a significant decline in consumer choice quality (see Panel B of Appendix Table F.1). This violation indicates that the average menu quality driven mechanism alone is not sufficient to explain the results of changing menu size on consumer choice quality in our setting.[11]

[11] As a robustness check, we utilize a more granular measure by defining average plan quality as the proportion of strictly dominated plan and supplementary coverage combinations (see Appendix A.3 for details). This adjustment yields similar findings, and the results are available upon request.

6. Conclusion

This paper studies how choice set size affects consumer choice quality in private health insurance markets. We leverage a novel dataset and two natural experiments involving a leading health insurance product in China. We find that expanding the plan menu significantly increases the probability of choosing a dominated plan, and shrinking the menu size improves choice quality in a consistent manner. The estimated treatment effects are symmetric in magnitude. We also provide evidence showing that these effects are primarily driven by choice overload.

Overall, our results demonstrate that meaningful choice frictions can arise even with relatively small menus in developing country contexts, where consumers are less familiar with making active health insurance choices when facing many complicated plan attributes. Presented with a menu of two plans, a significant proportion of our sampled consumers (over 15%) still choose strictly dominated options, although they should have recognized such dominance.

These results highlight the need for careful menu design and well-targeted decision support interventions to improve consumer choice quality in health insurance markets. Based on our findings, one straightforward way to deal with such choice inconsistency would be to reduce the choice set size. However, this approach may have undesirable consequences, as certain consumers may not be able to obtain their preferred options, as recently discussed by Brown and Jeon (2024). Our further findings on an AI-based decision support intervention suggest a potential pathway to resolve the regulatory trade-off between product diversity and consumer choice quality. The AI-based tool can bridge the performance gap in consumer decision-making as menu sizes grow larger, thus allowing for a richer variety of plan options without the associated welfare losses from choice overload.

Our findings also shed light on the potential long-term benefits of improved consumer choice quality. We show that a smaller menu not only results in higher choice quality, but also a higher level of consumer satisfaction, supported by the increased probability of signing up for automatic renewal. This suggests that policies that aim to improve choice quality may also foster long-term consumer retention. For developing countries like China, where private health insurance is an optional entry point into the broader health care sector, interventions that aim

to improve consumer decision quality may also contribute to sustainable growth of the industry.

Several limitations in our study point to promising directions for future research. First, our evidence on the choice overload mechanism is indirect. Further analysis would benefit from more granular data on consumers' decision-making processes, such as browsing behavior and measures of consumer attention. Second, the prevalence of strictly dominated plans raises broader questions about product design and regulation in developing countries. Further research is needed to help understand why such plans persist, how they intersect with firms' pricing and marketing strategies, and their implications for regulatory oversight.

References

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.

Abaluck, Jason, and Jonathan Gruber. 2016a. "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program: Reply." *American Economic Review* 106 (12): 3962–87.

Abaluck, Jason, and Jonathan Gruber. 2016b. "Evolving Choice Inconsistencies in Choice of Prescription Drug Insurance." *American Economic Review* 106 (8): 2145–84.

Abaluck, Jason, and Jonathan Gruber. 2023. "When Less Is More: Improving Choices in Health Insurance Markets." *Review of Economic Studies* 90 (3): 1011–40.

Ang, Geer, Ya Tan, Yingjia Zhai, Fan Zhang, and Qinghua Zhang. 2024. "Housing Wealth, Fertility and Children's Health in China: A Regression Discontinuity Design." *Journal of Health Economics* 97: 102915.

Besedeš, Tibor, Cary Deck, Sudipta Sarangi, and Mikhael Shor. 2015. "Reducing Choice Overload without Reducing Choices." *Review of Economics and Statistics* 97 (4): 793–802.

Bhargava, Saurabh, George Loewenstein, and Justin Sydnor. 2017. "Choose to Lose: Health Plan Choices from a Menu with Dominated Option." *Quarterly Journal of Economics* 132 (3): 1319–72.

Brown, Zach Y., and Jihye Jeon. 2024. "Endogenous Information and Simplifying Insurance Choice." *Econometrica* 92 (3): 881–911.

Bundorf, M. Kate, Maria Polyakova, and Ming Tai-Seale. 2019. "How Do Humans Interact with Algorithms? Experimental Evidence from Health Insurance." NBER Working Paper 25976. National Bureau of Economic Research, Cambridge, MA.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2024. "Event Studies with a Continuous Treatment." *AEA Papers and Proceedings* 114: 601–05.

Callaway, Brantly, and Pedro HC Sant'Anna. 2021. "Difference-in-differences with multiple time periods." *Journal of econometrics* 225 (2): 200–30.

Calonico, Sebastian, Matlas D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression‐Discontinuity Designs." *Econometrica* 82 (6): 2295–2326.

Chandra, Amitabh, Benjamin Handel, and Joshua Schwartzstein. 2019. "Behavioral Economics and Health-Care Markets." In *Handbook of Behavioral Economics: Applications and Foundations 1*, Vol. 2, 459–502. North-Holland.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2023. "Two-way Fixed Effects and Differences-in-Differences Estimators with Several Treatments." *Journal of Econometrics* 236 (2): 105480.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2024. "Difference-in-Differences Estimators of Intertemporal Treatment Effects." *Review of Economics and Statistics* Forthcoming.

Ericson, Keith M. Marzilli, Jon Kingsdale, Tim Layton, and Adam Sacarny. 2017. "Nudging Leads Consumers in Colorado to Shop but Not Switch ACA Marketplace Plans." *Health Affairs* 36 (2): 311–19.

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.

Giuntella, Osea, and Fabrizio Mazzonna. 2019. "Sunset Time and the Economic Effects of Social Jetlag: Evidence from US Time Zone Borders." *Journal of Health Economics* 65: 210–26.

Gruber, Jonathan, Benjamin R. Handel, Samuel H. Kina, and Jonathan T. Kolstad. 2020. "Managing Intelligence: Skilled Experts and AI in Markets for Complex Products." NBER Working Paper 27038. National Bureau of Economic Research, Cambridge, MA.

Handel, Benjamin R. 2013. "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts." *American Economic Review* 103 (7): 2643–82.

Handel, Benjamin, Jonathan Kolstad, Thomas Minten, and Johannes Spinnewijn. 2024. "The Socioeconomic Distribution of Choice Quality: Evidence from Health Insurance in the Netherlands." *American Economic Review: Insights* 6 (3): 395–412.

Iyengar, Sheena S., and Mark R. Lepper. 2000. "When Choice Is Demotivating: Can One Desire Too Much of a Good Thing?" *Journal of Personality and Social Psychology* 79 (6): 995.

Ketcham, Jonathan D., Nicolai V. Kuminoff, and Christopher A. Powers. 2016. "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program: Comment." *American Economic Review* 106 (12): 3932–61.

Ketcham, Jonathan D., Claudio Lucarelli, Eugenio J. Miravete, and M. Christopher Roebuck. 2012. "Sinking, Swimming, or Learning to Swim in Medicare Part D." *American Economic Review* 102 (6): 2639–73.

Ketcham, Jonathan D., Claudio Lucarelli, and Christopher A. Powers. 2015. "Paying Attention or Paying Too Much in Medicare Part D." *American Economic Review* 105 (1): 204–33.

Kling, Jeffrey R., Sendhil Mullainathan, Eldar Shafir, Lee C. Vermeulen, and Marian V. Wrobel. 2012. "Comparison Friction: Experimental Evidence from Medicare Drug Plans." *Quarterly Journal of Economics* 127 (1): 199–235.

Lee, Ajin, and Boris Vabson. 2024. "The Value of Improving Insurance Quality: Evidence from Long-run Medicaid Attrition." *Journal of Health Economics* 94: 102865.

Li, Honglin, Chenyuan Liu, and Justin Sydnor. 2025. "Is Choice Overload a Problem for Health-Plan Choice? Evidence from the ACA Marketplace." Working paper.

Liu, Chenyuan, and Justin Sydnor. 2022. "Dominated Options in Health Insurance Plans." *American Economic Journal: Economic Policy* 14 (1): 277–300.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2): 698–714.

Samek, Anya, and Justin Sydnor. 2025. "The Impact of Consequence Information on Insurance Choice." *Review of Economics and Statistics* Forthcoming.

Zhu, Minglai, Youji Lyu, Benke Wang, Jingjing Ou, Rui Zhu, Shanshan Wang, Zihan Diao, et al. 2024. "Blue Paper on the Development of China's Commercial Medical Insurance." Accessed December 13, 2025, https://mp.weixin.qq.com/s/DButxC1wabe1fejwWDrVBg?scene=25&sessionid=#wechat_redirect (in Chinese).

Appendix (For Online Publication Only)

A. Plan details and strictly dominated options

A.1 Plan details

Plan A featured a 30-day waiting period, a 10,000 yuan deductible (roughly $1,450) for inpatient expenses and a zero deductible for expenses related to critical illness. The plan has a 100% compensation ratio, a 300 million yuan (or roughly $435,000) upper limit for inpatient expenses, and another 300 million yuan coverage limit for critical illness. Outpatient services provided within 7 days preceding the claimed hospitalization were covered by the plan.

Compared to Plan A, all subsequent plans provided strictly expanded coverage. Specifically, Plans B, C and D offered additional coverage for 121 rare diseases defined by the government (e.g., albinism).[1] Plan C also introduced a shared deductible for enrolled family members. Furthermore, Plan D extended outpatient coverage from 7 to 30 days preceding the claimed inpatient visit.

The net premium equaled the baseline premium minus the premium discount. For a given insurance plan, the baseline premium was mainly determined by an enrollee's age at renewal. Even though Plan A's base premium was slightly lower, Plan A's enrollees were ineligible for premium discounts, while enrollees in Plan B, Plan C, and Plan D could benefit from a premium discount if they had no claim history. Moreover, on top of this discount, Plans C and D provided a further discount for households with three or more members. As a result, Plan A's net premium was higher than that of the other options for many enrollees.

A.2 Strictly dominated options

We use the following steps to determine dominance:

1. Compare the coverage for the baseline plans. This includes determining whether one plan has a (weakly) higher deductible, less covered benefits, or narrower networks.
2. Compare the choice of supplemental coverage. Plans A, B, and D allow enrollees to purchase certain supplemental coverage for specific medical catalogs and services. If such supplementary coverage is purchased, we

[1] See the official website of the National Health Commission of the People's Republic of China for details: https://www.gov.cn/zhengce/zhengceku/2018-12/31/content_5435167.htm.

then search whether the other options also offer such supplementary coverage.

3. Compare the premiums. The total premium includes both the net premium of the baseline plan (baseline premium minus the premium discount) and the premium of the supplementary coverage (if purchased).

Table A.1 reports the presence of strictly dominated plan options across different menus and household sizes. Note that due to age differences and the purchase of supplementary coverage, the dominance measure may vary even after holding fixed menu, plan, and household size. (Thus, "Yes" in the table indicates that this baseline plan is strictly dominated by another available option for at least some enrollees of that category. In most cases, "Yes" represents over 99% of enrollees in that category with a strictly dominated option.) We find that Plan A is associated with strictly dominated options in all menus, because for a given choice of supplemental coverage, Plan A packages provide strictly less protection at weakly higher premiums than do the corresponding packages from Plans B, C, or D. Further, Plan C features a shared deductible with enrolled family members, offers additional discounts for households with three or more members, and does not allow for the purchase of supplemental coverage. For a single-person household, Plan C is strictly dominated by Plan B with no supplemental coverage, as it offers the same protection at a higher price. Conversely, for households with three or more members, Plan B is strictly dominated by Plan C because it carries a higher deductible at a higher price.

Table A.1: Baseline plans with strictly dominated options

Menu	Plan	Household size = 1	Household size = 2	Household size = 3	Household size > 3
Menu 1	Plan A	Yes	Yes	Yes	Yes
	Plan B	No	No	No	No
Menu 2	Plan A	Yes	Yes	Yes	Yes
	Plan B	No	No	Yes	Yes
	Plan C	Yes	No	No	No
Menu 3	Plan A	Yes	Yes	Yes	Yes
	Plan D	No	No	No	No

Notes: The table shows the choice set within each menu by household size. Each row corresponds to a base plan. "Yes" indicates that at least one plan package associated with this base plan (i.e., the base plan combined with specific supplemental coverage) is strictly dominated by another available option.

A.3 Alternative measure of average plan quality

The availability of supplementary coverage may complicate the measure of choice set size. In the baseline analysis, we measure the total number of choices using the baseline plans and calculate the average plan quality as the fraction of "Yes" for each menu and household size in Table A.1. One alternative is to measure both using the combination of the baseline plan and the availability of supplementary coverage, i.e., treat each unique combination of the baseline plan and supplementary coverage as one option (we call these "plan packages"). The total number of available plan packages is given by: $N_{total} = \sum_m 2^{n_m}$, where m indexes the baseline plans within a menu and n_m is the total number of available supplemental coverage options for plan m. Accordingly, the proportion of strictly dominated plan packages within a menu is: $Quality = N_{dominated}/N_{total}$, where $N_{dominated}$ denotes the total number of strictly dominated packages within that menu.

B. Summary statistics

Table B.1: Summary statistics of the main covariates

	(1) Before menu change	(2) After menu change	(3) Difference (after – before)
Panel A. Natural experiment 1: Increasing choice set size from two to three			
Age	37.540	37.239	-0.301*
Male	0.484	0.475	-0.010*
Household size (#)	1.691	1.755	0.064***
Any rider (previous year)	0.594	0.604	0.009*
Historical policies (#)	2.536	2.724	0.189***
Initial enrollment in 2017	0.532	0.719	0.187***
Obs.	24,211	13,677	-
Panel B. Natural experiment 2: Reducing choice set size from three to two			
Age	37.750	38.164	0.414**
Male	0.480	0.478	-0.002
Household size	1.903	1.854	-0.049***
Any rider (previous year)	0.659	0.591	-0.067***
Historical policies (#)	3.023	3.504	0.481***
Initial enrollment in 2017	0.435	0.500	0.065***
Obs.	13,434	16,180	-

Notes: This table reports the means of the baseline covariates by natural experiment and choice menu. "Historical policies (#)" is the total number of policies held before renewal in the natural experiments. "Initial enrollment in 2017" is a dummy variable that equals 1 for policyholders who initially enrolled in 2017 and 0 for those who enrolled in 2018.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C. Identification validity and robustness tests

C.1 Density test

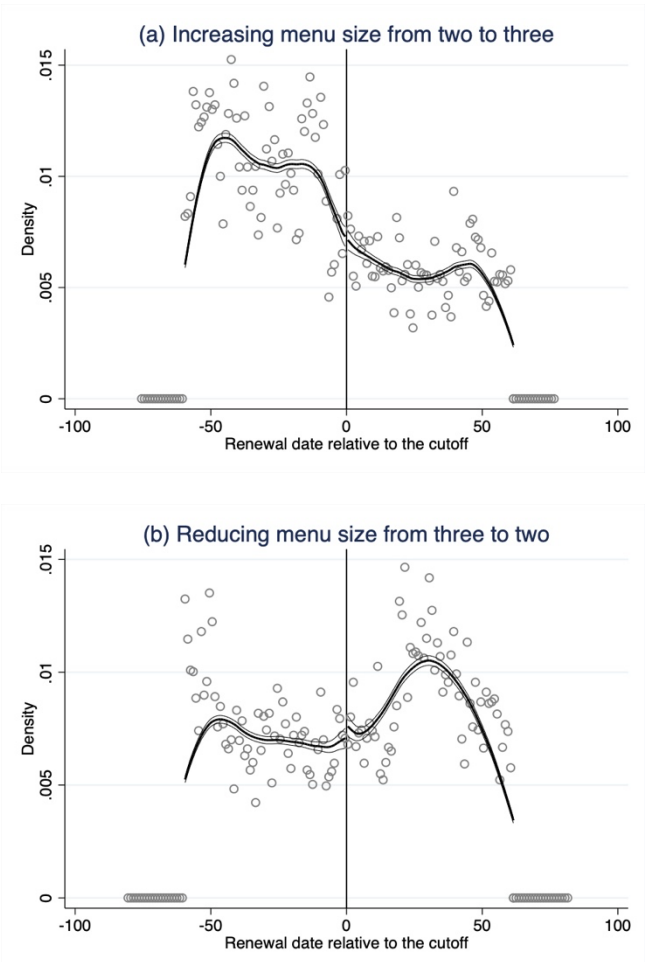

Figure. C.1: Density tests

Notes: We test for discontinuities in the density of the running variable at the cutoffs, following McCrary (2008). The corresponding estimates (standard errors) for the discontinuities are 0.001 (0.047) and 0.075 (0.046) in panels A and B, respectively.

C.2 Covariate balance

Table C.1: Covariate balance

	(1)	(2)
Variable	Baseline mean	RDiT estimate
Panel A. Natural experiment 1: Increasing choice set size from two to three		
Age	37.801	0.856
		(0.550)
Male	0.490	-0.011
		(0.013)
Household size (#)	1.680	-0.024
		(0.083)
Any rider previous year	0.598	-0.001
		(0.030)
Historical policies (#)	2.598	0.036
		(0.027)
Initial enrollment in 2017	0.598	0.034
		(0.027)
Eff. number of obs.	7,699	
Panel B. Natural experiment 2: Reducing choice set size from three to two		
Age	37.733	0.274
		(0.892)
Male	0.481	-0.027
		(0.022)
Household size (#)	1.960	-0.143
		(0.095)
Any rider previous year	0.629	-0.042
		(0.034)
Historical policies (#)	3.447	0.058
		(0.055)
Initial enrollment in 2017	0.468	0.070
		(0.046)
Eff. number of obs.	5,433	

Notes: "Baseline mean" is the unconditional sample mean of the outcome for policyholders who renewed within a 14-day bandwidth below the cutoff. "RDiT estimate" denotes the baseline RDiT estimates.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

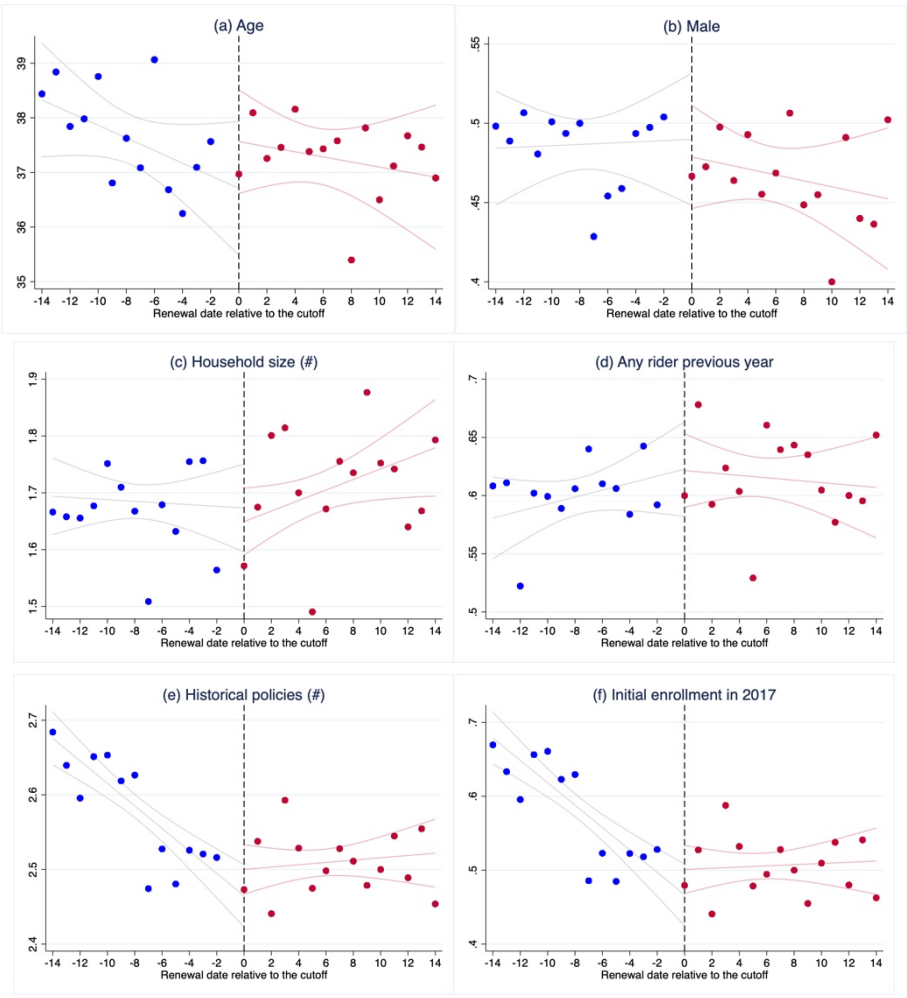

Figure. C.2: Covariate balance (increasing menu size from two to three)

Notes: Each figure plots the means of the outcomes at each standardized renewal date bin, together with fitted lines and 95% confidence intervals. The fitted lines are produced using regressions with a triangular kernel. The first and second choice menus are indicated in blue and red, respectively.

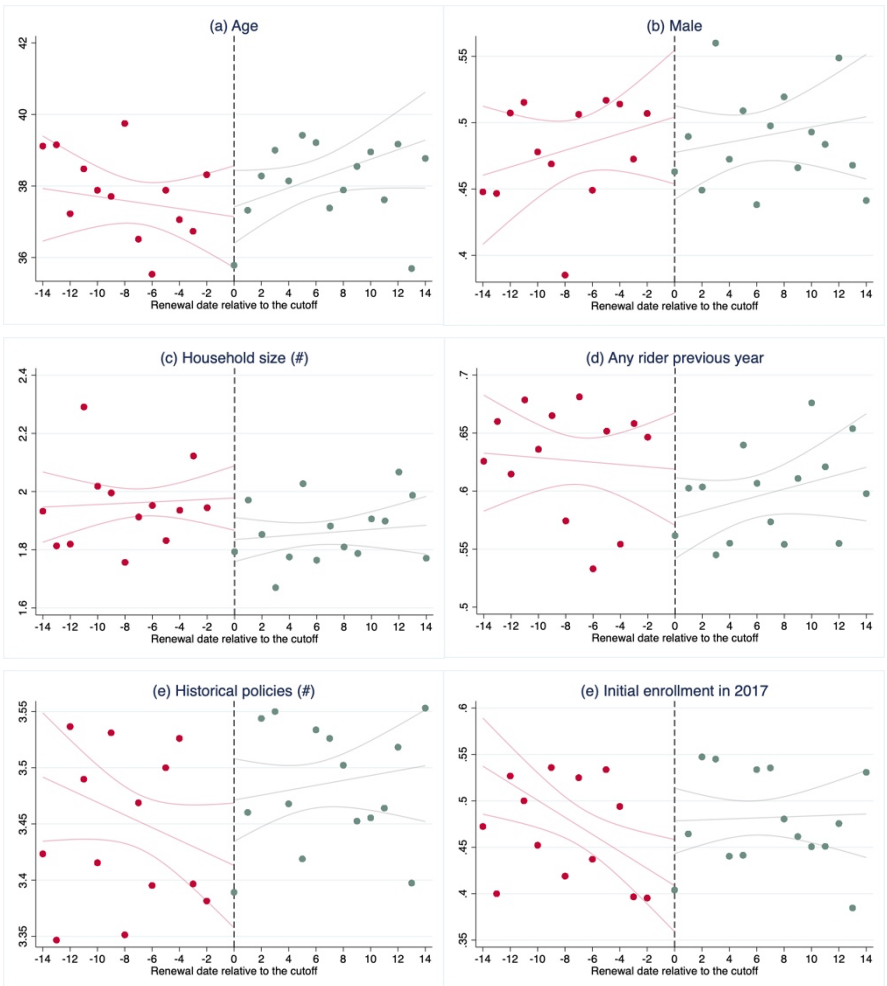

Figure. C.3: Covariate balance (reducing menu size from three to two)

Notes: Each figure plots the means of the outcomes at each standardized renewal date bin, together with fitted lines and 95% confidence intervals. The fitted lines are produced using regressions with a triangular kernel. The second and third choice menus are indicated in red and green, respectively.

C.3 Sensitivity of main results

Table C.2: Sensitivity of main results

	Outcome: Choosing a strictly dominated plan					
	(1) Baseline	(2) Donut hole (radius = 1)	(3) Quadratic polynomial	(4) Uniform kernel	(5) Optimal bandwidth	(6) Add covariates
Panel A. Increasing menu size from two to three						
RDiT estimate	0.084***	0.081***	0.087**	0.097***	0.087***	0.084***
	(0.023)	(0.027)	(0.034)	(0.020)	(0.020)	(0.024)
Eff. number of obs.	7,699	7,092	7,699	8,480	10,491	7,699
Bandwidth	14	14	14	14	17.892	14
Kernel	Triangular	Triangular	Triangular	Uniform	Triangular	Triangular
Polynomial order	1	1	2	1	1	1
Covariates	N	N	N	N	N	Y
Below-cutoff mean	0.161					
Panel B. Reducing menu size from three to two						
RDiT estimate	-0.110***	-0.110***	-0.159***	-0.086***	-0.117***	-0.114***
	(0.031)	(0.032)	(0.046)	(0.028)	(0.032)	(0.028)
Eff. number of obs.	5,433	4,991	5,433	5,775	4,758	5,433
Bandwidth	14	14	14	14	11.368	14
Kernel	Triangular	Triangular	Triangular	Uniform	Triangular	Triangular
Polynomial order	1	1	2	1	1	1
Covariates	N	N	N	N	N	Y
Below-cutoff mean	0.225					

Notes: Panels A and B report the results from the natural experiments of increasing and reducing choice set size, respectively. Column (1) reports the baseline estimates. Columns (2) to (6) further report the results of implementing a donut-hole RDiT (excluding the nearest observations within a one-day radius around the cutoff), using a quadratic polynomial, utilizing a uniform kernel, employing the optimal bandwidth as suggested by Calonico et al. (2014) for each regression, and adding covariates. "Below-cutoff mean" is the unconditional sample mean of the outcome for policyholders who renewed within a 14-day bandwidth below the cutoff.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C.4 Placebo test for contemporaneous shocks

To rule out unobserved contemporaneous shocks, we conduct a placebo test using a separate sample of new enrollees of the same health insurance product. For these consumers, the choice menu does not change around the cutoff date of the second natural experiment (January 30, 2021). Specifically, since January 6, 2021, they choose from a stable menu consisting of two plans (Plan B and Plan D in Figure 1). For a subset of these individuals, Plan D strictly dominates Plan B, offering more generous coverage at the same net premium. Because the plan menu remains unchanged at the cutoff, this setting provides a clean placebo test for detecting any concurrent shocks unrelated to the menu change in the second natural experiment.

We implement the placebo test by replicating the baseline RDiT analysis on this sample. To ensure the existence of strictly dominated plans, we impose sample restrictions consistent with those in the baseline analysis.[2] We then test for any discontinuity at the cutoff in the probability of choosing a dominated plan.

The main results are reported in Column (1) of Table C.3. We find no evidence of a discontinuous change in choice quality at the cutoff in this placebo test. The estimated effect is small in magnitude (0.6 percentage points) and statistically insignificant. Consistent with this finding, the graphical analysis in Figure C.4 shows no visible discontinuity at the cutoff. Columns (2) to (6) in Table C.3 further shows consistent results from alternative RDiT specifications.

[2] The sample restriction here differs from the baseline in two aspects: (i) For new enrollees, eligibility for premium discounts, and hence the existence of strictly dominated options, depends on smoking history rather than claims history, so we exclude individuals with a smoking history. (ii) To ensure that all individuals face the same choice menu, we restrict observations to the left of the cutoff date to policyholders who enrolled since January 6, 2021.

Table C.3: Placebo test based on new enrollees

	Outcome: Choosing a strictly dominated plan					
	(1) Baseline	(2) Donut hole (radius = 1)	(3) Quadratic polynomial	(4) Uniform kernel	(5) Optimal bandwidth	(6) Add covariates
RDiT estimate	0.006	0.007	0.010	0.004	0.009	0.006
	(0.013)	(0.014)	(0.019)	(0.013)	(0.015)	(0.013)
Eff. number of obs.	10,362	7,807	10,362	10,732	8,994	10,362
Bandwidth	14	14	14	14	10.351	14
Kernel	Triangular	Triangular	Triangular	Uniform	Triangular	Triangular
Polynomial order	1	1	2	1	1	1
Covariates	N	N	N	N	N	Y
Below-cutoff mean	0.071					

Notes: Column (1) reports the baseline estimates. Columns (2) to (6) further report the results of implementing a donut-hole RDiT (excluding the nearest observations within a one-day radius around the cutoff), using a quadratic polynomial, utilizing a uniform kernel, employing the optimal bandwidth as suggested by Calonico et al. (2014), and adding covariates (age, gender and household size). "Below-cutoff mean" is the unconditional sample mean of the outcome for policyholders who enrolled within a 14-day bandwidth below the cutoff.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

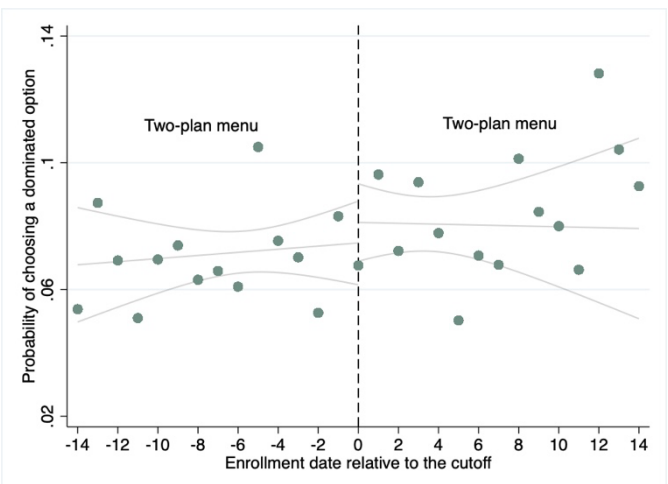

Figure C.4: Placebo test based on new enrollees

Notes: Each panel plots the probability of choosing a strictly dominated plan for each standardized enrollment-date bin, together with fitted lines and 95% confidence intervals. The fitted lines are produced using regressions with a triangular kernel.

D. Testing for changes in the renewal rate

This section tests for changes in the renewal rate around the cutoff dates. We employ an auxiliary data set from the same data source, which contains information on the Plan A policies purchased a year before the natural experiments. These policies were scheduled for renewal around the cutoff dates. We implement the same sample restrictions as in our baseline empirical analysis to ensure the existence of strictly dominated options within each menu and to rule out potential confounders.

For this analysis, we specify the running variable as the individual's policy expiration date relative to the cutoff dates of the menu changes.[3] To ensure comparability, we restrict the sample to individuals whose policy expiration dates were within the 60-day bandwidth before and after the cutoff dates.

For estimation, we use the baseline RDiT strategy presented in section 4 and replace the outcome variable with a dummy indicator for plan renewal. Optimal bandwidth selection method of Calonico et al. (2014) yields optimal bandwidths of around 14 days. Thus, we set a uniform bandwidth of 14.

Table D.1 reports the estimates. The results indicate that the renewal rate barely changes at either cutoff. The estimated changes are 0.019 and 0.004 for the first and second natural experiments, respectively. Both estimates are insignificant and of small magnitude compared to the baseline means below the cutoffs. The graphical evidence presented in Figure D.1 shows consistent findings, with no obvious discontinuities at either cutoff.

[3] In our main study sample of renewed enrollees in the two natural experiments, around 55% of them renewed within 10 days around the expiration date.

Table D.1: Testing for changes in the renewal rate

	Outcome: Choosing to renew	
	(1)	(2)
	Increasing menu size from two to three	Reducing menu size from three to two
RDiT estimate	0.019	0.004
	(0.017)	(0.027)
Eff. number of obs.	14,234	5,415
Below-cutoff mean	0.785	0.788

Notes: "Below-cutoff mean" is the unconditional sample mean of the outcome for policyholders who renewed within a 14-day bandwidth below the cutoff.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

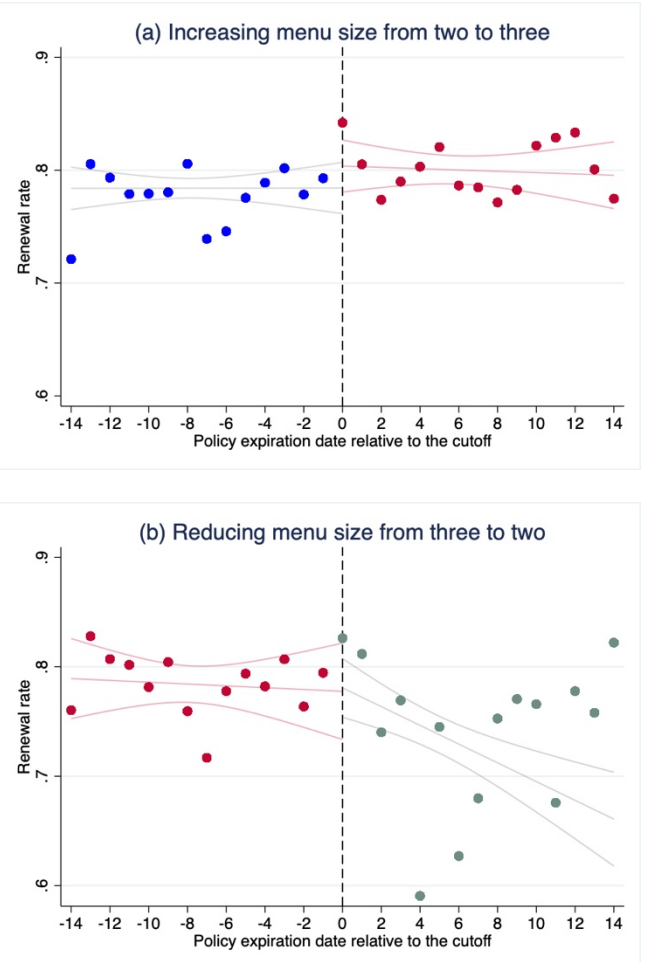

Figure. D.1: Insurance renewal rates around the cutoffs

Notes: Each figure plots the renewal rate at each standardized policy expiration date bin, together with fitted lines and 95% confidence intervals. The fitted lines are produced using regressions with a triangular kernel. The first, second, and third choice menus are indicated in blue, red, and green, respectively.

E. Results for the "choice overload" mechanism

E.1 Heterogeneous effects by enrollment experience

Table E.1: Heterogeneous treatment effects by enrollment experience

	Outcome: Choosing a dominated option	
	(1)	(2)
	Increasing menu size from two to three	Reducing menu size from three to two
Panel A. More-experienced consumers		
RDiT estimate	0.030	-0.028
	(0.029)	(0.044)
Eff. number of obs.	4,217	3,482
Below-cutoff mean	0.151	0.176
Panel B. Less experienced consumers		
RDiT estimate	0.141***	-0.171***
	(0.030)	(0.027)
Eff. number of obs.	2,520	2,913
Below-cutoff mean	0.200	0.246

Notes: "Below-cutoff mean" is the unconditional sample mean of the outcome for policyholders who renewed within a 14-day bandwidth below the cutoff.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

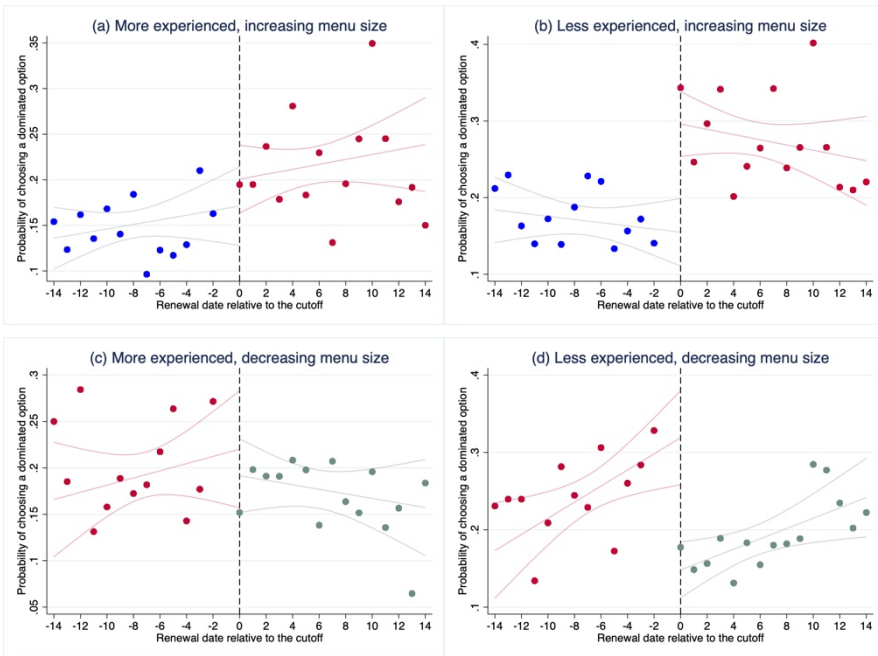

Figure E.1: Heterogeneous treatment effects by enrollment experience

Notes: Each figure plots the probability of choosing a dominated plan at renewal at each standardized renewal-date bin, together with fitted lines and 95% confidence intervals. The fitted lines are produced using regressions with a triangular kernel. The first, second, and third choice menus are indicated in blue, red, and green, respectively.

E.2. Decision support program

Table E.2: Effect of decision support on choice quality by plan menu

	Outcome: Choosing a dominated option		
	(1) Baseline	(2) Not-yet-treated group as control	(3) TWFE
Panel A. Choosing under Menu 2 (three-plan menu)			
DID estimate	-0.221**	-0.252**	-0.223***
	(0.109)	(0.112)	(0.069)
Obs.	47,046	2,084	47,046
Pre-treatment mean		0.294	
Panel B. Choosing under Menu 3 (two-plan menu)			
DID estimate	0.038	0.037	0.038
	(0.062)	(0.088)	(0.063)
Obs.	15,079	614	15,079
Pre-treatment mean		0.112	

Notes: Panels A and B report results based on the subsamples of individuals choosing under Menu 2 and Menu 3, respectively. Column (1) reports the baseline estimate. Columns (2) and (3) report results using the not-yet-treated group as control and employing a two-way fixed effects (TWFE) estimator. "Pre-treatment mean" reports the unconditional sample mean of the outcome for treatment-group policyholders during the pre-treatment period.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

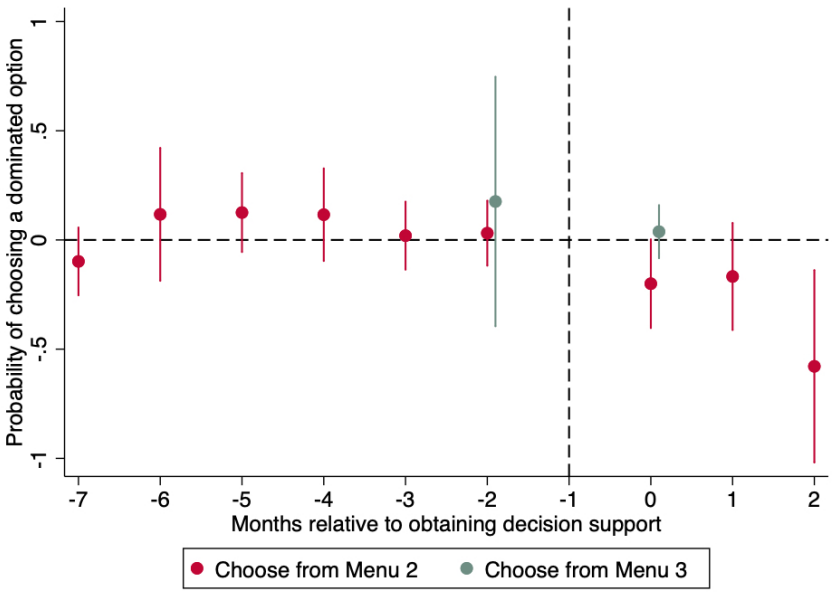

Figure E.2: Event study estimates of the impact of the decision support program on the probability of making choice mistakes

Notes: The figure shows 95% confidence intervals.

E.3 Satisfaction with the chosen option

Table E.3: Effect of reducing the size of the choice set on consumer satisfaction with the chosen option

	Outcome: Automatic renewal
	(1)
	Reducing menu size from three to two
RDiT estimate	0.044*
	(0.025)
Eff. number of obs.	5,433
Below-cutoff mean	0.123

Notes: "Below-cutoff mean" is the unconditional sample mean of the outcome for policyholders who renewed within a 14-day bandwidth below the cutoff.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

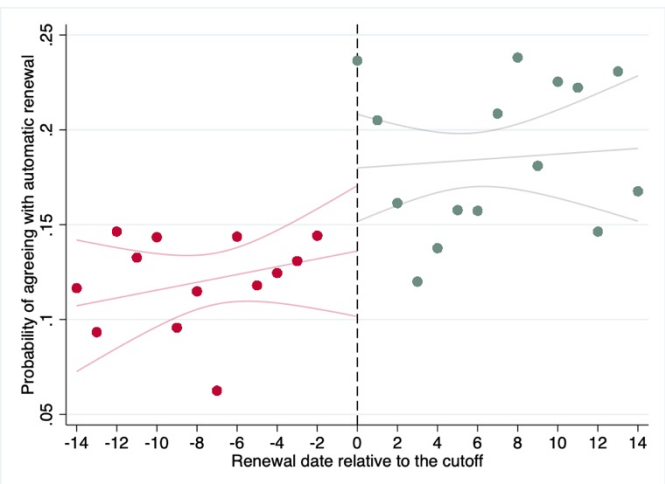

Figure E.3: Reducing menu size and consumer satisfaction

Notes: This figure plots the probability of agreeing with automatic renewal of the currently chosen options at each standardized renewal-date bin, together with fitted lines and 95% confidence intervals. The fitted lines are produced using regressions with a triangular kernel. The second and third choice menus are indicated in red and green, respectively.

F. Results for the "average plan quality change" mechanism

Table F.1: Heterogeneous treatment effects by enrollee type

	Outcome: Choosing a strictly dominated plan	
	(1)	(2)
	Increasing menu size from two to three	Reducing menu size from three to two
Panel A. Other households (plan quality deteriorates as menu expands)		
RDiT estimate	0.086***	-0.095***
	(0.029)	(0.025)
Eff. number of obs.	5,968	4,186
Below-cutoff mean	0.163	0.212
Panel B. Two-member households (plan quality improves as menu expands)		
RDiT estimate	0.077	-0.153**
	(0.074)	(0.069)
Eff. number of obs.	1,731	1,247
Below-cutoff mean	0.155	0.264

Notes: Panel A represents the case where average plan quality deteriorates when the choice menu expands, as the proportion of strictly dominated plans increases. Conversely, Panel B corresponds to the scenario where average plan quality improves when the menu expands, as this proportion decreases. "Below-cutoff mean" is the unconditional sample mean of the outcome for policyholders who renewed within a 14-day bandwidth below the cutoff. In Panel B, the point estimate of increasing menu size (Column (1)) is 0.099 (0.052; $p < 0.10$) when employing a larger bandwidth of 21.419, as suggested by Calonico et al. (2014).
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.