Review of Economic Studies (2023) **90**, 1011–1040 https://doi.org/10.1093/restud/rdac050

Advance access publication 4 August 2022

When Less is More: Improving Choices in Health Insurance Markets

JASON ABALUCK
Yale University and NBER

and

JONATHAN GRUBER
MIT and NBER

First version received December 2019*; Editorial decision October* 2020*; Accepted July* 2022 (*Eds.*)

We study the impact of changing choice set size on the quality of choices in health insurance markets. Using novel data on enrolment and medical claims for school district employees in the state of Oregon, we document that the average employee could save $600 by switching to a lower cost plan. Structural modelling reveals large "choice inconsistencies" such as non-equalization of the dollar spent on premiums and out of pocket, and a novel form of "approximate inertia" where enrolees are excessively likely to switch to other plans that are close to the current plan on the plan design spreadsheet. Variation in the number of plan choices across districts and over time shows that enrolees make lower-cost choices when the choice set is smaller. We show that a curated restriction of choice set size improves choices more than the best available information intervention, partly because approximate inertia lowers gains from new information. We explicitly test and reject the assumption that this is because individuals choose worse from larger choice sets, or "choice overload". Rather, we show that this feature arises from the fact that larger choice sets feature worse choices on average that are not offset by individual re-optimization.

Key words: Health insurance, Choice, Decision-support, Choice set size, Discrete choice, Choice overload

JEL Codes: I1

1. INTRODUCTION

Insurance product choice is a central feature of health insurance markets in the US. Approximately 50% of US residents get their coverage from an employer, and 58% of those offered employer-sponsored insurance (ESI) have a choice of insurance plans.[1] Those who buy private insurance under the state and federal exchanges established by the Affordable Care Act (ACA) had an

1. Data from Kaiser Family Foundation at http://kff.org/other/state-indicator/total-population/ and Kaiser Family Foundation at http://kff.org/health-costs/report/2017-employer-health-benefits-survey/.

The editor in charge of this paper was Adam Szeidl.

average of 20 plans per county being offered on the exchanges in 2016.[2] The Medicare program that provides insurance coverage to approximately 42 million elderly and disabled Americans provides a choice between the traditional Medicare program and an average of 21 "Medicare Advantage" plans that provide a private alternative; the prescription drug plan that was added to Medicare in 2006 offers most beneficiaries a choice of more than 40 private prescription drug insurance plans.[3] The lowest income Americans who are insured through Medicaid typically have a choice of a variety of managed care plans for their coverage, with 275 total managed organizations operating in 38 states and DC (an average choice set of 8 plans per state).[4]

This expansion of choice raises a number of issues.[5] Foremost among them is the question of whether consumers can adequately choose from a variety of complicated health insurance options. Several studies show choice inconsistencies in health insurance markets whereby consumers do not appear to choose plans that maximize their own long-run utility. Most of this work has been focused on the case of prescription drug plan choice in the Medicare program. Abaluck and Gruber (2011) present reduced form facts and structural analysis consistent with large choice inconsistencies and foregone welfare for seniors choosing prescription drug plans; Ketcham, Kuminoff and Powers (2016) and Abaluck and Gruber (2016b) clarify the sensitivity of these results to normative assumptions about the role of omitted characteristics and debate various specification checks of models of choice inconsistencies. Work by Handel (2013), Bhargava, Loewenstein and Sydnor (2017), and Handel and Kolstad (2015) a document choice inconsistencies in the broader insurance context as well.

Importantly, however, there is little work on solutions to this problem of choice inconsistency. Information interventions which have been assessed in practice and typically find small or zero impacts on choice (Kling, Mullainathan, Shafir, Vermeulen and Wrobel, 2012; Ericson, Kingsdale, Layton and Sacarny, 2017, 2019), although a recent study suggests that combining information interventions with skilled agents can improve choice (Gruber, Handel, Kina and Kolstad, 2020). Learning over time does not seem to lead to a reduction in choice inconsistencies (Abaluck and Gruber, 2016a).[6] One potential solution is to aggressively shift the choice architecture to potentially limit the "damage" from inconsistent choices. For example, Ericson and Starc (2016) show that consumer welfare was improved by moving to a more standardized set of choices on the Massachusetts health insurance exchange.

Even more radical than standardizing choices is limiting choices. A smaller choice set reduces the potential for choice inconsistencies, but at the same time does not allow heterogeneous consumers to match to the plan which best fits their tastes. Kamenica (2008) develops a theory where consumers become less likely to participate in larger choice sets if there is some mechanism by which smaller choice sets leave out bad options, and past literature from outside health insurance shows that an individual's willingness to participate in a market decreases as the choice set size increases (Sethi-Iyengar, Huberman and Jiang, 2004; Iyengar and Kamenica, 2010; Bertrand, Karlan, Mullainathan, Shafir and Zinman, 2010). But

2. Calculated from HealthCare.gov (2017) individual QHP landscape data at https://www.healthcare.gov/health-plan-information-2017/.

3. Data from Kaiser Family Foundation at http://kff.org/report-section/whats-in-and-whats-out-medicare-advantage-market-entries-and-exits-for-2016-appendix/, Kaiser Family Foundation at http://kff.org/medicare/issue-brief/medicare-advantage-2016-data-spotlight-overview-of-plan-changes/, and Kaiser Family Foundation at http://kff.org/medicare/fact-sheet/the-medicare-prescription-drug-benefit-fact-sheet/.

4. Data from Kaiser Family Foundation at http://kff.org/other/state-indicator/total-medicaid-mcos/.

5. See Gruber (2017) for a review of these issues.

6. Ketcham, Lucarelli, Miravete and Roebuck (2012) argue that there is learning, but Abaluck and Gruber (2016a) show a lack of learning in higher quality data. See Ketcham *et al.* (2016) and Abaluck and Gruber (2016b) for a further debate over these issues.

there is no field work which goes further to explore the impact of choice set size on the *nature of choices within the set*.[7]

In this article, we provide the first thorough analysis of whether a curated reduction in choice set size for health insurance improves choice quality by reducing mistakes or makes consumers worse off by removing valuable alternatives. We do so using a novel dataset across school districts in the state of Oregon. Beginning in October 2008, each of the roughly 240 school districts in Oregon selected a subset of plans to offer their employees from a menu of 9–13 plans that were made available to them at prices centrally negotiated by the state, including up to 7 from the single largest insurer. Prior to 2012, districts could select up to four plans which employees could then choose from as well as the district contribution towards each option. After 2012, districts were free to choose any number of plans to offer. The result is wide variation in choice sets and premiums available to roughly 63,000 school district employees in Oregon each year.

We have gathered data from 2008 to 2012 on the complete enrolment and medical claims information for school district employees.[8] We matched enrolment and claims data with data carefully collected from school district surveys and union contracts on the number of options and the district contributions towards those options over this period. We use these data to show that the type of choice inconsistencies documented by Abaluck and Gruber (2011, 2016b) extend from the narrower area of prescription drug plan choice to the broader health insurance plan choice environment. In fact, we find comparable results to the Part D context, with less than half of employees making the cost minimizing choice (despite often choosing from 2 to 4 plans) and an average foregone savings of over \$500 (about 20% of total out-of-pocket costs of \$2,600, inclusive of premiums). Structural models document significant choice inconsistencies even when controlling for other aspects of plan choice.

We then make three new contributions. The first is to document an important new driver of choices, "approximate inertia". Not only is there the striking inertia in plan choice documented by previous studies, but when enrolees do move, they tend to move to nearby plans in product space with somewhat lower premiums. The effect of this approximate inertia is quite large; for example, it is one-half as large as the well-documented inertia effect of enrolment in a given plan. It is also largely insensitive to the magnitude of the premium savings—switchers move over time to lower premium plans without regard for the actual dollars that they save by doing so. Remarkably, allowing for approximate inertia explains a greater share of choices in our choice model than allowing choices to be sensitive to premiums. Indeed, adding this single parameter to our previous model reduces the mean-squared error of predicted market shares conditional on the prior year plan by almost 60%.

Our second contribution is to provide the first evidence that consumer choices improve when facing a reduced set of choices. We study this using natural variation in choice set size within districts over time in a setting where premiums are set at the state level, allowing us to isolate demand-side effects. We know of no previous attempt to separate the demand-side impact of the number of choices on how well consumers choose from the supply-side impact on premiums via concentration. We find that limiting plan options, in a curated way, leads to both lower foregone savings and, more strikingly, reduced total costs paid by enrolees. While work like

7. Within the context of health insurance, the only paper to explore the impacts of choice set size issue is Ketcham, Lucarelli and Powers (2015), who finds that larger choice sets tended to increase switching behaviour in the Medicare Part D market. But the variation in choice set sizes ranges from 46 to 55 in their sample, which may not be a relevant range for many insurance option sets, and does not speak to whether the increased switching led to improved choices.

8. Our data also include 2013 claims and choices, although most of our analysis focuses in 2008–12 since a restructuring of plan offerings in 2013 makes it more challenging to determine default options. Our main results are qualitatively unchanged if we include 2013 data as well.

Bhargava *et al.* (2017) document that individuals make dominated choices, here, we show that an actual policy explicitly designed to limit choice set size by giving local officials discretion about what plans to include made consumers better off. Undoing the reform that limited choice sets to four plans prior to 2012 cost consumers about $150 per beneficiary. In our earlier work (Abaluck and Gruber, 2009), we found that randomly excluding plans from choice sets would not make consumers better off, so we illustrate empirically that when policy-makers are given discretion to exclude plans, they do much better than random and actually succeed in making beneficiaries better off.

We also investigate the interaction between approximate inertia and policies designed to improve choices. We use our model to compare policies that limit choices to policies that would induce equivalent savings among new consumers via information provision. We find that the impact of information provision among returning consumers is sharply curtailed because they are subject to both inertia and approximate inertia. The greater the impact of the intervention among new consumers, the more this impact is muted by inertia of both types among returning consumers. In contrast, policies which limit choices are much less impacted by inertia and approximate inertia because unsuitable plans which are removed cannot be chosen.

Our third contribution is to provide the first empirical test of choice overload by separating demand from supply and assessing whether consumers use a different "choice function" when more options are available. This is a major departure from the previous literature which has argued that the factor driving choice size effects was that "decision-making improves when fewer options are considered concurrently" (Besedeš, Deck, Sarangi and Shor, 2015). In fact, we find that "choice overload" is of secondary importance relative to the average quality of the options available in determining the impact of adding more plans. This suggests that the problem of "too many choices" in insurance markets may not be that consumers are especially confused when the number of options is large, but that consumers always choose based on heuristics and this leads them to go awry when some options are unsuitable for most beneficiaries. We conclude that individuals are sufficiently homogenous in their response to plan characteristics that plan administrators can assess whether adding more plans to a choice set can improve welfare simply by evaluating the effect on the average plan enrolee in their district.

Our article proceeds as follows. Section 2 discusses the health insurance choices across Oregon school districts which provides the context for our study. Section 3 describes our data while Section 4 introduces our empirical strategy. Section 5 provides the results on choice inconsistency. Section 6 then estimates the role of choice set limitations on choice quality. Section 6 concludes.

2. HEALTH INSURANCE BENEFITS AMONG OREGON TEACHERS

The state of Oregon was divided into between 226 and 244 school districts, education service districts, or community colleges during our study period, with small variation from year to year. Districts had several classes of workers; a given employee is categorized as one of: administrator licensed, administrator non-licensed, classified, community college non-instructional, community college faculty, confidential, licensed, substitute, or superintendent. Within each type are both part-time and full-time employees. Most workers employed by school districts are a member of one of three unions, either the Oregon School Employees Association (OSEA), the Oregon Education Association (OEA), or the American Federation of Teachers-Oregon (AFT).

Prior to 2008, districts and community colleges independently purchased plans for employees through the Oregon School Employees Association or one of two health plan trusts. Beginning in 2008, health insurance benefits, as well as life and disability coverage, long-term care insurance, an employee assistance program, and pre-tax savings accounts for each of these districts or community colleges are provided by the Oregon Educational Benefit Board (OEBB). The OEBB

TABLE 1
Plans available by year

Plan	2008	2009	2010	2011	2012
Kaiser Medical Plan 1	Y	Y	Y	Y	Y
Kaiser Medical Plan 2	Y	Y	–	–	–
Kaiser Medical Plan 1A	–	Y	Y	Y	Y
ODS Medical Plan 3	Y	Y	Y	Y	Y
ODS Medical Plan 4	Y	Y	Y	Y	Y
ODS Medical Plan 5	Y	Y	Y	Y	Y
ODS Medical Plan 6	Y	Y	Y	Y	Y
ODS Medical Plan 7	Y	Y	Y	Y	Y
ODS Medical Plan 8	Y	Y	Y	Y	Y
ODS Medical Plan 9	Y	Y	Y	Y	Y
Providence Medical Plan 1	Y	Y	–	–	–
Providence Medical Plan 2	Y	Y	Y	Y	–
Providence Medical Plan 1A	–	Y	–	–	–
Providence Medical Plan 2A	–	–	Y	Y	–

Notes: 1. "Y" indicates that a plan was offered in at least one district and "–" indicates that a plan was not offered in any districts in a given year. 2. Between 2012 and 2013 enrolment, ODS changed their name to MODA Health. Plans offered prior to 2013 and in 2013 are listed alphabetically, plans listed in the same row before and after the name change are not necessarily equivalent. In our analysis of plan choices in 2013, we consider whether plans introduced in 2013 were simply renamed or offered qualitatively different benefits.

negotiates rates for the state for a variety of plans from several insurers. These plans are listed over time in Table 1. From 2008 to 2012, there were between 9 and 12 options available from three insurers: Kaiser Permanente, a closed panel Health Maintenance Organization (HMO) that restricts patients to go to Kaiser hospitals and physicians; OMED (later MODA), a Preferred Provider Organization (PPO) plan that allows free choice of providers within a fairly broad statewide network; and Providence, a competing PPO plan.

As the table shows, the set of options a district could offer changed over time: Kaiser added one plan in 2009 and removed one plan in 2010, and Providence eventually withdrew from the choice set in 2012. The options available to a given beneficiary also changed from year to year based on statewide regulations and individual district choices. In addition to these medical plan options, OEBB also offers a choice of prescription drug plans, dental plans, and vision coverage. Supplementary Table E1 summarizes the benefits structures of each of these options.

Each district was then given the option to offer up to four of those plans to their employees for 2008–11. In 2012, there was no cap on the number of medical plans a district could offer. Across all years, Kaiser plans were only offered in a subset of regions.[9]

The district has other tools at its disposal that can impact insurance plan choices as well. One such tool is the rate at which the district will contribute towards plans. These contributions are negotiated with the unions representing workers in each district and are made public to employees before they enrol in health insurance for the upcoming year. Contribution structures differ substantially across districts. For most districts, for each employee type there is one flat contribution for all coverage tiers (employee, employee and child, employee and spouse, family). Districts could also vary the fixed contribution amount by coverage tier and could offer either prorated or full contribution amounts to part-time employees. Some districts offer a percentage contribution in which the district paid some percentage of the chosen plan premium, rather than a fixed dollar amount. This percentage contribution could either be constant for all employees or vary similarly to fixed contributions, by coverage tier and full time vs. part time status. Districts

9. A single district in our sample—comprising 0.02% of all beneficiary years—offered only a single plan. We exclude this district from our analysis because the sample is too small to draw meaningful conclusions.

could also establish a fixed employee contribution to the premium, in which case the district contribution would be the raw premium minus this fixed employee contribution and would vary based on the cost of the chosen plan. This fixed employee contribution could vary similarly to the district contributions above.

In addition, districts could make a contribution directly to accounts that employees could use to pay their out of pocket medical costs. OMED9 was a qualified “high deductible plan” (based on high level of patient deductibles) that allowed enrolees to set up “health savings accounts” (HSA) to which they and the district could contribute funds on a pre-tax basis—and then draw down these funds to pay medical expenses without tax penalty. For the other plans, districts could set up Health Reimbursement Accounts (HRA), which operate in a similar fashion with the important difference that HSAs are owned by the employee (and are therefore portable to other employers) and HRAs are owned by the employer (and if the employee leaves the remaining balance therefore reverts to the employer). If there was an excess district premium contribution (e.g. the negotiated amount a district contributes to an employee’s premiums was greater than the raw premium) unions negotiated that either the complete excess, a percentage of the excess, or the excess up to a maximum value would be contributed by the district to an employee’s HRA. Some districts made a fixed contribution to an HRA regardless of excess contributions.

A key issue for our analysis is the fact that there are very meaningful differences in provider networks across plans. Provider networks are identical within insurer with one exception; in 2011 and 2012 plan, OMED4 was a limited network plan with a narrower network than other OMED plans. But the networks differ across insurers, particularly for Kaiser which has a very limited network of providers. Including insurer fixed effects in the model are not sufficient to capture the overall effect of these differences if there is individual-specific variation in the value of broader networks that may be correlated with our other parameters of interest in ways difficult to capture in any parametric specification. Therefore, we focus only on plans offered by the largest single insurer, OMED/MODA.[10] In doing so, we hold all non-financial characteristics of plans constant, including features such as customer support or network breadth—the differences across plans are purely financial, arising from variation in premium and out of pocket costs. In Abaluck and Gruber (2016c), we show that the results are very similar when using the larger set of choices.

Enrolment in health insurance plans takes place during an open enrolment period that runs from 15th August to 15th September each fall. We have data on choices made by enrolees in open enrolments from fall 2008 through fall 2013. The default option for employees not making an active choice vary by district and year; unfortunately, these defaults are not observable to us.

3. DATA

We have collected data from a variety of sources for this analysis.

3.1. *Institutional details on district plan structure and contributions*

We received complete data on the plans offered by each district in each year to each employee type from OEBB (OEBB, 2013a,d, 2014). We then collected detailed data on the district contributions to employee premiums as well as district policies on HSAs and HRAs from two sources. First, with OEBB’s assistance, we collected detailed surveys from each district (OEBB, 2013e). Surveys regarding district HSA and HRA policies were sent to each district’s benefits manager. Second, we received union contracts from OEBB. These contracts contain the negotiated district contributions

10. The results for OMED/MODA only are identical if OMED4 is excluded from the choice set for 2011 and 2012.

TABLE 2
Choice set by year

Choice set size	2009	2010	2011	2012	All years
2	3,620	2,395	2,768	498	9,282
3	11,724	12,464	9,938	4,394	38,520
4	7,441	8,194	10,844	6,840	33,319
5	0	0	0	1,750	1,750
6	0	0	0	3,712	3,712
7	0	0	0	5,550	5,550

Notes: Table shows the total number of policy holders (including both individuals and families) enrolled in choice sets with the listed number of options in each year among beneficiaries who choose MODA plans. Prior to 2011, choice sets were limited to four plans and benefit managers chose up to four to offer to their employees. From 2012 on, benefit managers could choose between 1 and 10 plans to offer to their employees (only one small district offered a single plan, so we drop that from the analysis), including up to seven MODA plans.

to represented employees as well as whether an HSA or HRA was available. We carefully combined these two sources of data, with priority to directly collected surveys because of the possibility that a contract had been amended, but the amendment was not publically available. We drop district, year, employee type observations for which we do not have data on the district contribution. We also drop observations for employees whose choice sets include only one plan option (a single district).

Table 2 shows the number of beneficiaries with each choice set size in each year. This table is tabulated among the final sample of policy holders included in our perfect foresight analysis.

In addition to variation in the plans available to a beneficiary (choice set), districts vary widely in their contribution policy. While about 96% of policy holders in each year were in districts with a fixed district contribution,[11] the value of that contribution varied because the contribution is sufficient to fully cover premiums for some plans but not others (and beneficiaries cannot keep the residual difference). For more than 90% of policy holders, there is at least one plan with zero premiums after the contribution, meaning that the contribution alters the relative value of plans; that is, since the contribution is worthless once premiums have reached zero (in most cases), a fixed contribution changes the relative prices of plans. Therefore, even among this 96%, the district contribution does change the relative value of plans. As noted above, another source of differentiation across districts is their contributions towards accounts employees can use to pay out of pocket medical costs, either through the HSA that is associated with plan OMED 9 or the HRA that could be offered to policy holders with other OMED plans. There is substantial variation across districts in both whether there are contributions to the HSA/HRA, and the form of those contributions (fixed, percentage of excess contribution, or other).

3.2. *Enrolment and claims data on OEBB employees*

To analyse choice of plan, we gathered a complete universe of enrolment and claims data for OEBB employees over the 2008–12 period (OEBB, 2013b,c). We sometimes rely on a beneficiary's year $t-1$ claims to model year t plan choice, so we keep only beneficiaries with prior year claims (meaning that 2009 is the first year of our analysis sample). Supplementary Table E2 shows the impact of each sample restriction we impose on the total sample size. Our final analysis sample contains 92,133 beneficiaries.

11. Between 2.3% and 5.4% of policy holders in each year were in districts with a percentage district contribution, and the remaining policy holders were in districts with a fixed policy holder contribution.

3.3. *Measuring premium and out of pocket costs*

There are a variety of issues that arise in the measurement of premiums and out of pocket costs. Direct premium payments are determined by the difference between plan cost and district contributions, as discussed above. But districts often also make separate contributions to the HSA account included in the OMED9/MODAH plan, as well as the HRA that could be offered with other OMED/MODA plans. As described above, we have carefully collected data on district policies for how any excess district contribution to premiums is deposited into these savings accounts. We then apply federal legal maximum amounts to these accounts, to arrive at a final dollar value for the amount (which varies by the raw premium of the plan selected and district specific policies) by which a district could fund a savings account. Beneficiaries can use this amount to offset out of pocket costs, so after calculating the raw out of pocket costs faced by a beneficiary in each plan in their choice set, we subtract the district contributed amount in a beneficiary's savings account to arrive at the net out of pocket costs to a beneficiary.

Another issue is treatment of dental and vision premiums. Given the small premium relative to medical and prescription drug, we assume that dental and vision plans are of secondary importance in the choice of a health insurance plan. Therefore, we assume that an individual will enrol in the same dental and vision plan, regardless of the medical plan in which they enrol.[12]

The major issue with measuring out of pocket costs is determining the proper model of expectations. We consider three different models of expectations: perfect foresight, perfect backcast, and rational expectations. In the perfect foresight model, we assume that enrolees know exactly what their out of pocket costs will be in the coming year and run their realized claims through a calculator in each plan to determine out of pocket costs. In the perfect backcast model, we assume that enrolees believe that the coming year's claims will be identical to the prior year; to determine out of pocket costs, we run the prior year claims through each of the plans in the enrolee's choice set.

The rational expectations model assumes that enrolees forecast a distribution of possible out of pocket costs for each plan given the information available at the time when they choose. To create this distribution we use a software program developed by Johns Hopkins Medical School that predicts individual risk for future medical expenditures using past expenditure and demographics, as in Handel (2013).[13] This software develops individual risk scores for future health care expenditure. By creating groups of individuals who are similarly at risk based on the Johns Hopkins software predicted risk score, and using our calculator to model costs in all available plans for randomly selected individuals from each group, we can create distributions of expected expenditures for each group of similarly at risk individuals. We use three methods and three draw sizes, resulting in nine versions of these distributions to test sensitivity; all yield very similar results (see Supplementary Table E3).

More specifically, the ACG software creates three scores based on predicted medical expenditures, predicted drug expenditures and all other expenditures. For the results in this

12. We do not observe chosen dental and vision plans prior to 2010. To calculate this premium cost prior to 2010, we calculate enrolment weighted average dental and vision premiums in each district and employee type for all dental and vision plans and for all non-Kaiser dental and vision plans with plan selection weights based on observed 2010 enrolment. If a beneficiary is enrolled in a Kaiser medical plan, we apply the all-plan weighted average dental or vision premium, and if a beneficiary is enrolled in a non-Kaiser medical plan, we apply the non-Kaiser weighted average dental or vision premium. After 2010, we apply the chosen dental or vision premium to all counterfactual plans, unless the counterfactual medical plan is non-Kaiser and the chosen medical plan was Kaiser—in which case the Kaiser dental or vision plan would not be available. In these cases we apply the mean enrolment weighted average of available non-Kaiser dental or vision plan premiums.

13. Johns Hopkins ACG (Adjusted Clinical Groups) Case-Mix System. See http://acg.jhsph.org/.

article, we create deciles of each of the three dimensions of risk and add an eleventh category in each dimension for zero costs. We then regress year t costs on these three categorical variables (calculated based on year $t-1$ claims) and generate a predicted cost in year t. Next we create deciles of this predicted cost variable to yield 10 groups of similarly at risk individuals. We then randomly sample with replacement 2,000 individuals from each cell. These randomly drawn individuals are all modelled as if they were individual policy holders in all available plans to create 2,000 estimates of out of pocket costs in every plan $\times$ cell combination. An observation for every plan in a beneficiary's choice set is then matched to the 2,000 estimates of out of pocket costs for their cell. Costs are then summed across families for each draw. Finally, family out of pocket and deductible maximums are imposed on total family costs. With these constructed rational expectations measures of out of pocket costs, we can then assess choices against the mean and variance of the distribution of expected costs to investigate preferences for risk protection. In some robustness checks below, we use 20,000 individuals per cell in order to calculate extreme quantiles of the risk distribution with more precision.[14]

4. MODELLING CHOICE INCONSISTENCIES

4.1. *Descriptive facts: the role of "approximate inertia"*

We begin by presenting the basic facts on foregone savings, defined as the total costs to the beneficiary in the chosen plan minus the total cost to the beneficiary in the cost minimizing plan. For each policy holder in our data, we use our calculator to assign the net premium plus out of pocket costs of each option in their choice set. We then compare this quantity for the chosen plan to the lowest cost plan in their choice set, and compute the difference, for each model of expectations.

As Figure 1 shows, we find substantial foregone savings across all methods of modelling expected out of pocket costs. Each bar in the figure represents average foregone savings for an alternative model of expectations. Assessing the chosen plan relative to lowest cost plan yields mean foregone savings between \$511 and \$522.[15] Supplementary Figure E.2 shows the distribution of foregone savings according to the rational expectations measure—10% of beneficiaries could save more than \$2,000. There are meaningful foregone savings from choices regardless of how expectations are modelled.

If people are not cost minimizing, how are they choosing? As in most insurance markets, we find high rates of inertia: 71% of consumers are inertial when their prior year plan continues to be available.[16] Among active choosers, we find that most non-inertial beneficiaries follow a simple heuristic: they move to one of the two most similar plans with lower premiums. We call this phenomenon, "approximate inertia". Consumers are uncertain about what will happen if they choose a radically different plan, so they choose a similar plan with lower premiums. They do

14. Implicitly, constructing out of pocket costs using a fixed basket of claims assumes that consumers would not adjust their consumption for moral hazard reasons. Provided consumers do not have sufficient foresight to take into account expected differences in moral hazard in their plan choices, ignoring moral hazard in our analysis should not bias our positive model. We show in Appendix A of Abaluck and Gruber (2009) that the additional impact of moral hazard on utility for plan j relative to a reference plan is given by: $-1/2\, S\cdot\epsilon\cdot \mathrm{CI}_j\ (CI_j - CI_0)$, where S is out of pocket spending, ϵ is the utilization elasticity, and CI is the fraction of expenditures paid out of pocket. Assuming a utilization elasticity of -0.20 and taking the median plan as the reference plan, this term would typically be less than \$30.

15. Supplementary Figure E.1 shows that savings are even larger if we consider all alternative plans, but these plans may have different provider networks.

16. While high, the degree of inertia is lower than in other markets. We think this is due to two factors: first, there are large changes in premiums from year to year over the period we study (both positive and negative). Second, OEBB reaches out to beneficiaries and encourages them to actively choose plans in each year.

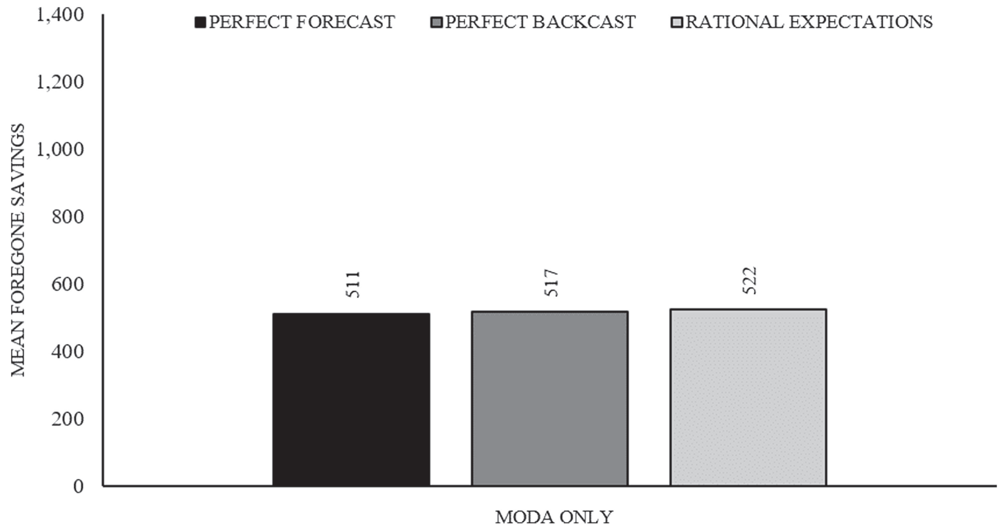

FIGURE 1

Mean foregone savings, MODA

Notes: Figure shows foregone savings among MODA plans using a variety of metrics to compute out of pocket costs. This is computed as the difference in costs between the chosen MODA plan and the lowest cost available MODA plan for beneficiaries who chose MODA plans. The perfect forecast, perfect backcast, and rational expectations models of out of pocket costs are described in the text.

so largely without regard for whether this plan saves them money or even would have saved them money in the previous year. “Approximate inertia” is thus distinct from the usual premium responsiveness we estimate because it occurs regardless of the magnitude of the actual premium savings achievable from switching plans.

The underlying pattern is illustrated in Figure 2. MODA plans are numbered in decreasing order of coverage comprehensiveness. Higher plans have lower premiums, but higher deductibles and lower out of pocket maxes. This means that lower numbered plans cost more and have better risk protection for the average beneficiary. The black bars in the figure represent the actual choices made by those switching plans (i.e. if you switch from OMED3 to OMED5, you have a difference of 2). We can see that conditional on switching, more than 71% of beneficiaries move to a plan which is one or two numbers higher. If beneficiaries chose randomly (the white bars), we would expect this number to be 44%. If beneficiaries minimized costs, we would expect this number to be 59% (the grey bars). These plans are thus substantially overrepresented relative to what we would expect if beneficiaries minimized costs. Cost minimizing switchers would both choose more plans with more generous coverage, and choose more plans with even lower premiums but less generous coverage.

Could this be explained by heterogeneous risk preferences? The fact that consumers who choose OMED3 move to OMED4 and those who choose OMED6 move to OMED7 is consistent with consumers with different risk preferences preferring different levels of coverage. However, the movements we observe cannot be rationalized by risk preference alone: Figure 2 suggests that if beneficiaries minimized costs, 16% would choose plans with more coverage rather than the 10% we observe. Plans with more coverage also offer better risk protection—thus, the fact that beneficiaries overwhelmingly move to less comprehensive, lower premium plans (even absent improved health) does not appear consistent with any level of risk aversion. In our structural model, we explicitly assess the degree to which heterogeneous risk preferences can explain the “approximate inertia” we observe.

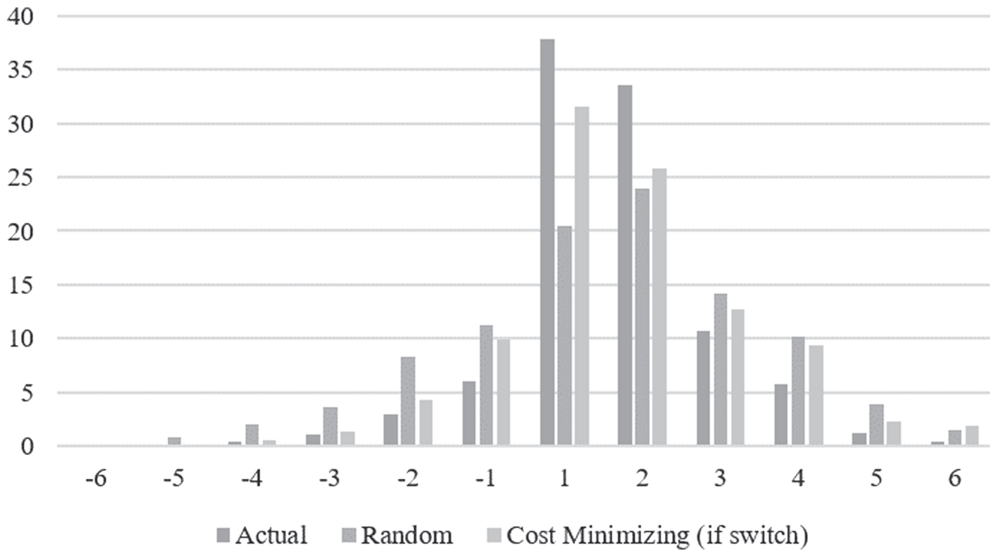

FIGURE 2

Plan number changes for switchers

Notes: Among beneficiaries who chose a MODA plan this year and the previous year but switched plans, "Actual" shows the distribution of the difference of the "number" of the plan chosen this year relative to the previous year. For example, a beneficiary who switched from OMED3 to OMED4 would have a difference of 1. "Random" shows what this distribution would be if we randomly assigned beneficiaries to alternative plans in their choice sets. "Cost Minimizing" shows what this difference would be if we assigned everyone to the cost minimizing plan (conditional on switching).

In our previous work, we document that consumers weight premiums more heavily than out of pocket costs (Abaluck and Gruber, 2011, 2016b). Is this a manifestation of the same phenomenon? Our results in the next section imply that "approximate inertia" is something new: we do find that consumers overweight premiums relative to out of pocket costs, but the estimated response to changes in premiums is far too small to explain the phenomenon we document in Figure 2—and in fact, we see shifts towards lower premium plans even in years when the relative premium savings of those plans decrease.

4.2. *Choice model*

Of course, these facts on foregone savings are not by themselves dispositive that consumers are making errors because of other differences across plans. Even within the MODA plans, the variance in outcomes may be lower for plans with higher measured foregone savings. Additionally, we want to evaluate the degree to which apparent heuristics like "approximate inertia" might be explained by persistent risk preferences and changes in premiums over time. To do so, we turn to a structural model of plan choice.

Define the Gross Premium as the premium listed on the plan design document and define $Net\,Premium = max\{0, Gross\,Premium - Employer\,Contribution\}$ as the amount the beneficiary pays. Additionally, define $Residual\,Premium = Net\,Premium - Gross\,Premium$. We will allow consumer utility to vary as a function of both the gross premium and the residual premium. If these have equal coefficients, then we could equivalently write utility as a function only of the net premium.

Positive utility in our model is given by:

$$\begin{aligned} u_{ijt} = \beta_{0i} Gross\ Premium_{ijt} + \beta_{1i} Residual\ Premium_{ijt} \\ + \beta_{2i} E(OOP)_{ijt} + \beta_{3i} Var_{ijt} + d_{jt} + \xi_{ij} + \theta_{ij} + \epsilon_{ij} \end{aligned} \tag{1}$$

Utility depends firstly on the gross and residual premium terms, both of which vary by plan and tier. Utility additionally depends on the individual's mean and variance of out of pocket costs (or in the case of the perfect backcast or perfect foresight model, just the mean since there is no uncertainty), on plan-year fixed effects d_{jt}, on the inertia dummies ξ_{ij} which are 1 for the plan chosen last year and zero otherwise, on the approximate inertial dummies θ_{ij} which is 1 if plan j is one or two columns higher than the inertial plan on the plan design spreadsheet and zero otherwise, and on the idiosyncratic error terms ϵ_{ij}, which are assumed i.i.d. type I extreme value. In some specifications, we further decompose $d_{jt} = x_{jt}\gamma + e_{jt}$ where x_{jt} are plan financial characteristics.

To motivate the dependence of utility on only the mean and variance of costs, we can assume CARA utility and normally distributed costs. That is, if $U(C) = -exp(-\gamma(W-C))$ where W is wealth and C is total costs (premiums plus out of pocket costs) with $C \sim N(\mu, \sigma^2)$, then expected utility is given by: $EU(C) = -\alpha exp(\gamma\mu + \frac{1}{2}\gamma^2\sigma^2)$, where $\alpha = -exp(\gamma W)$, a constant. Taylor-expanding this gives the indirect utility function in equation (1). In some models, we allow β_{0i} and γ_i to be random coefficients to allow for heterogeneous risk preferences as well as heterogeneous preferences for costs relative to other unobservables. Below, we also consider models where consumers value higher quantiles of the distribution of out of pocket costs to check whether our results could be driven by differences across plans not well-accounted for by the variance of out of pocket cost.

Table 3 shows the structural coefficients from logit estimation of the positive utility equation (1) assuming homogeneous coefficients (we report random coefficient results below). For plans with small market share, these coefficients can be interpreted as the percentage change in choice probabilities induced by a change in the "x" variable. For example, a \$100 increase in gross annual premiums leads to a 7.3% reduction in the probability that a plan is chosen. We report results in our perfect backcast, perfect foresight, and rational expectations models. We include the variance term only for the rational expectations measure since it is not computed for the other measures. All specifications include plan $\times$ tier and plan $\times$ year fixed effects.[17]

The responsiveness to gross premiums is thus identified by differential premium changes over time across tiers: if premiums increase more for families than individual policy holders for a given plan, do we see a reduction in the probability that families choose that plan? The responsiveness to residual premiums is identified principally based on whether, when district contributions "zero-out" some plans but not others, do individuals become more likely to choose the plans which benefit more from this contribution and thus have a lower relative price? The responsiveness to individual variables like out of pocket costs is identified based on whether individuals whose specific set of claims make particular plans expensive are less likely to choose those plans.

The main results in all specifications replicate the choice inconsistencies from our previous work as well as providing evidence of new choice inconsistencies.[18] First, we find a large gap

17. The coefficients on plan characteristics such as the deductible, out of pocket max and copay are recovered by fixing the coefficients other than plan characteristics at their estimated value reported in the table, omitting the plan dummies and including these characteristics. This is equivalent to a weighted regression of the fixed effects on auxiliary plan characteristics.

18. Supplementary Table E6 shows that similar results hold if we estimate our model using the full choice set instead of only MODA plans.

TABLE 3
Logit models of plan choice (MODA only)

	Perfect backcast	Perfect forecast	Rational expectations[1]
Gross premium	−0.073***	−0.073***	−0.073***
(*hundreds*)	(0.006)	(0.006)	(0.006)
Residual premium	−0.026***	−0.026***	−0.026***
(*hundreds*)	(0.004)	(0.004)	(0.004)
Mean OOP costs	−0.028***	−0.035***	−0.035***
(*hundreds*)	(0.003)	(0.004)	(0.007)
Variance OOP costs	–	–	−0.006
(*times* 10^6)			(0.034)
Inertia	3700***	3702***	3706***
	(82)	(82)	(82)
Approximate inertia	1957***	1958***	1962***
	(64)	(65)	(65)
Deductible, in network	−0.048***	−0.046***	−0.046***
(*hundreds*)	(0.001)	(0.001)	(0.001)
Max OOP, in network	−0.035***	−0.034***	−0.037***
(*hundreds*)	(0.0005)	(0.0005)	(0.0005)
PCP copay, in network	−0.733***	−0.475***	−0.501***
(*hundreds*)	(0.177)	(0.178)	(0.1)
Foregone welfare	517	511	522
(*Mean (SD)*)	(968)	(966)	(968)
Foregone welfare	–	–	522
(*no variance*)	–	–	(970)
Percent selecting cost minimizing plan	47.6	47.6	47.6

Notes: 1. Rational expectations using regression predicted approach with 2,000 draws. 2. We drop beneficiaries with only 1 plan in their choice set. However, we do not drop beneficiaries with 1 MODA and 1 or more non-MODA plan, and thus a choice set of 1 when restricting to MODA plans. In Figure 1, we do drop such individuals to avoid having observations with mechanically 0 foregone savings. 3. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

between the coefficient on premium and out of pocket costs; in every specification, the coefficient on the gross premiums is significantly larger than the coefficient on out of pocket costs, and is often several times larger. In addition, while consumers are fairly responsive to gross premiums (the number listed on plan design spreadsheets next to a plan), they are fairly insensitive to residual premiums, which vary conditional on gross premiums to the degree that premium contributions from employers “zero out” some plans and not others. As with out of pocket costs, this suggests that consumers are either not fully informed about what these contributions are or are failing to properly compute their consequences. Second, there is a large willingness to pay for financial plan characteristics even after controlling for the out of pocket consequences of those characteristics. In particular, we find that all of the coefficients on fixed plan cost-sharing characteristics are significant and right signed. For example, turning to the rational-expectations specification in the last column, consumers respond to a \$10 increase in primary care copays seven times more than a \$10 increase in premiums *after conditioning on the individualized out of pocket cost consequences*. In other words, beneficiaries are extremely responsive to the copays listed for primary care visits, but their responsiveness does not vary much with the number of primary care visits they are predicted to make (and thus the individualized out of pocket cost coefficient is very small).

Conditional on the factors in our earlier model, we also find that “approximate inertia” is important in explaining choices conditional on the other effects included in our model. In the rational expectations specification, the increase in choice probability for “approximately inertial” plans in the sense defined above is equivalent to what we would expect from a \$1,960 increase in premiums. For comparison, the degree of inertia is what we would expect if inertial plans were \$3,700 better. Approximate inertia is about $\frac{1}{2}$ as powerful as inertia itself in our context.

Indeed, we in fact find that approximate inertia explains a *greater share of choices than premiums*. To measure this, we re-estimate our models with and without premiums and approximate inertia, and consider both the average absolute error and mean squared error of the prediction of each model relative to the observed shares choosing plan j in each year conditional on the plan chosen in the prior year. We find that adding approximate inertia to the model has a much larger impact on these measures than does adding premiums. This result holds whether we include plan fixed effects (in which case the premium coefficients allow for differential responses by tier and district contribution) or if we omit plan fixed effects (in which case premium coefficients also permit consumers to on average be less likely to choose plans with higher premiums). Indeed, adding a single parameter to our model to allow for approximate inertia reduces the mean-squared error of predicted market shares conditional on prior year plan by almost 60%![19]

In other words, adding a single parameter to utility which allows consumers to give special weight to plans which are adjacent (a difference of "1" or "2" on the plan design spreadsheet) explains more choices than letting consumers choose plans on the basis of premium differences (and this is true regardless of whether we include plan fixed effects in the model). Consumers move over time to lower premium plans, but largely without regard for the actual savings they will realize by doing so. This is an analogue of our finding above that consumers prefer plans with desirable characteristics (small deductibles), but largely without regard for the actual out of pocket cost savings these features will produce.

One explanation for approximate inertia could be that healthy consumers wrongly believe they will remain healthy forever. If they realize they would have saved money in year $t-1$ by choosing a plan with lower premiums, they choose such a plan in year t even if it is not cost minimizing in a rational expectations sense. In fact, this does not appear to be what is going on. If we look specifically at beneficiaries who previously received a shock of more than \$5,000 to their medical expenses in the previous year, these beneficiaries are nearly equally likely to exhibit approximate inertia; the implicit willingness to pay for an approximately inertial plan among shocked consumers is \$1,897, compared to \$1,980 among unshocked consumers.[20] In other words, approximate inertia is not just a phenomenon of people wrongly believing they will remain healthy forever; even beneficiaries who recently had high spending seek out savings by moving to lower premium plans even if they end up losing money by doing so.

To investigate whether these results could be explained by heterogeneous risk preferences, we again estimate equation (1), but add random coefficients on the premium, out of pocket cost and variance terms. These results are reported in Supplementary Table E5. While the random coefficients are often statistically significant indicating that the model with heterogeneity fits the data better, we find no evidence that accounting for heterogeneity explains an appreciable fraction of approximate inertia. Risk preferences are identified because the same plans generate very different risk profiles for different individuals. If consumers are risk averse, we should see that consumers who face substantial risk given their prior year claims make different choices than consumers who face little risk. In fact, this is not what we see, and so our models suggest minimal levels of risk aversion (and relatively little heterogeneity).

Ericson and Sydnor (2018) pose a novel alternative explanation for findings such as this: the presence of liquidity constraints might cause individuals to overvalue regular premiums over

19. These results are reported in Supplementary Table E4.

20. These numbers are computed by adding a separate "shock $\times$ approximate inertia" term to the behavioural logit model and dividing the estimated coefficient by the coefficient on gross premiums. We report here the numbers for the rational expectations model; the numbers in other models are similar.

irregular and large deductibles. To test this, in Supplementary Table E7, we show that our results are comparable if we restrict to individuals of at least 50 years of age who are less likely to be liquidity constrained. Thus, liquidity constraints seem unlikely to explain our results.[21]

Our results could be consistent with a world in which consumers fail to accurately model the distribution of costs they will face in each plan, but still have determinate risk preferences which they express via their choice of coverage comprehensiveness. Nevertheless, our model implies that the patterns generated by heuristics—such as choosing a plan based on a vague sense of coverage comprehensiveness and then "staying close" to that plan—fit the data better than modelling choices based on the actual distribution of costs with persistent risk preferences.

4.3. *Evidence of choice inconsistencies and welfare consequences*

In our baseline specifications, we assume that normative utility in money-metric terms is given by:

$$u^N_{ijt} = Net\ Premium_{ij} + E(OOP)_{ijt} - \frac{\beta_{3i}}{\beta_{0i}} Var_{ijt}. \tag{2}$$

This embeds four normative assumptions:

Assumption 1. *A dollar of premiums has the same normative utility impact as a dollar of out of pocket costs. Additionally, a dollar of gross premiums has the same normative utility impact as a dollar of residual premiums.*

A dollar is a dollar regardless of its provenance (at least once we control for risk). This assumption follows (Abaluck and Gruber, 2011, 2016b).[22] Even if preferences for risk or the marginal utility of income are heterogeneous, we should see that premiums and expected out of pocket costs are given equal weight once we appropriately control for risk.

Assumption 2. *Financial characteristics of plans are not relevant for utility, except via out of pocket costs.*

Conditional on the individualized mean and variance of out of pocket exposure for a given plan, consumers should not independently value features such as deductibles and copayments; they should only care about such plan characteristics to the extent that they affect the consumer. This restriction is natural provided we have correctly specified how consumers value different cost distributions. As noted in the previous section, CARA utility and normal costs imply that utility depends only on the mean and variance of costs. Below, we consider alternative models where utility depends on other moments of the cost distribution and show this makes little difference to our results.

Assumption 3. *The inertia terms are normatively irrelevant.*

This is assumption follows our previous work as well as other papers in this literature (Handel and Kolstad, 2015). If inertia reflects factors like inattention or "satisficing" (Schwartz, Ward, Monterosso, Lyubomirsky, White and Lehman, 2002), then consumers may be

21. Olafsson and Pagel (2018) find that consumers may behave *as if* liquidity constrained even in settings where there are no such constraints. This pattern of behaviour is consistent with the evidence documented here.

22. Appendix C of Abaluck and Gruber (2009) considers one microfoundation.

better off were they enrolled in an alternative plan even if empirically, we see that they rarely switch. When switching between insurers with different networks, individuals may have to pay substantial adjustment costs to switch providers. However, in our context, looking within MODA plans, any adjustment costs should be negligible.[23]

Assumption 4. *The other omitted characteristics—including the plan fixed effects and idiosyncratic error term—are not relevant to utility.*

Conditional on choosing a MODA plan, physician network and other non-financial characteristics of plans are held constant—so the only meaningful source of differentiation are the financial characteristics which we observe.

The second panel of Table 3 computes foregone welfare given the normative utility function specified in the normative utility equation (equation 2) as well as the case where the variance term is assumed to be zero (in the perfect backcast or perfect foresight models, there is no variance). We find that including the variance makes essentially no difference—the measured degree of risk aversion is extremely small—and we find foregone welfare of \$510–\$520. These figures are very close to the results shown in Figure 1, indicating that foregone savings is a good summary measure of welfare loss.

Supplementary Table E8 computes the same number under alternative normative assumptions. Accounting for heterogeneity in risk preferences and the marginal utility of income (and computing expected utility by integrating over the estimated distributions) makes little difference to our estimates because the estimated degree of risk aversion is always small. We also consider whether our findings are driven by misspecification of risk magnitudes or preferences. In fact, we find little evidence of sensitivity to the 90th, 95th, or 99th percentiles of the distribution of out of pocket costs conditional on the mean that we include in our model. If rather than use revealed preference we impute the degree of risk aversion based on values commonly used in the literature (CARA = 0.0001, 0.0003, 0.0005, 0.001, 0.002), we find that this makes little difference to our conclusions about foregone welfare. In fact, our welfare conclusions are not monotonic in the imputed level of risk aversion. This reflects the fact that chosen plans are sometimes inferior both in terms of average costs and variance of costs (meaning that higher risk aversion increases foregone welfare), while sometimes they are higher cost but lower variance (meaning higher risk aversion decreases foregone welfare).

5. IMPROVING CHOICES: INFORMATION OR CHOICE SET CHANGES?

Given these large foregone savings, and given that consumers use recognizable heuristics, are there mechanisms that could improve consumer choice of health care plans? A natural such mechanism is decision support, providing clearer information and presentation on the implications of plan choice for total enrolee costs. As noted earlier, most studies of decision support find only very modest benefits if any in terms of improving choice consistency (Kling *et al.*, 2012; Ericson *et al.*, 2017, 2019), although a recent study suggests that combining information interventions with skilled agents may lead to larger improvements (Gruber *et al.*, 2020).

23. One way of quantifying this is to estimate the “default-specific consideration” model of Abaluck and Adams (2017) in order to recover the fraction of inertia which is due to inattention. Estimating this model in our context implies that roughly 2/3 of inertia is driven by inattention, supporting our normative assumption. The remaining “utility-relevant” components of inertia need not be normatively relevant either—they may reflect factors like consumers paying attention but requiring a threshold benefit before switching to compensate for “unknown unknowns” that might go wrong if they switch. In our view, while adjustment costs across plans or networks may be large and non-trivial in other settings, large adjustment costs within MODA plans are unlikely.

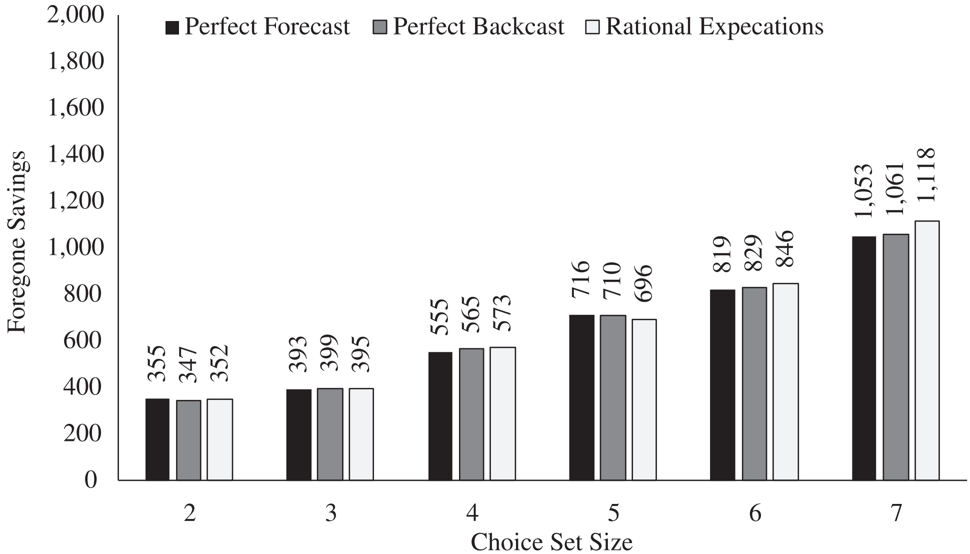

FIGURE 3

Foregone savings by choice set size (MODA only)

Notes: Figure shows the distribution of foregone savings by choice set size, counting only MODA plans.

A more radical alternative is to limit the plans available to consumers. That is, do beneficiaries in smaller choice sets leave less money on the table relative to the best plan, and ultimately, do they spend less on premiums and out of pocket costs? And how does this depend on how the smaller plans are curated? In this section, we investigate this more radical alternative and compare the benefits to potential information interventions. We will also investigate how the presence of inertia and approximate inertia impacts the benefits of both types of intervention.

As noted in Section 1, a common problem with analysing the welfare implications of choice set variation is that it can affect both the supply and demand sides of the market. Reducing the choice set can reduce competitive pressures on premium setting as well as the nature of consumer choices across plans. This is not a problem in our context, however, because premiums are set at the state level. This means that the supply side is fixed with respect to any individual district. When districts vary the number of choices facing enrolees, this has no impact on the prices that will be paid by these enrolees, allowing us to isolate the demand side impacts of choice set size variation in our context. Doing so is important because tools such as auctions could be used to maintain competition between providers while still limiting the options ultimately made available to consumers.

5.1. *Reduced form impact of choice set size and the 2012 reform*

Figure 3 shows foregone savings by choice set size for all three models of expectations. While we find substantial foregone savings in each case, there is a significant increase with choice set size: foregone savings rises from about $350 in choice sets with two MODA plans to $1,050–$1,100 in those with seven MODA plans.

Foregone savings measures the quality of choices relative to the best possible option. For welfare purposes, however, we are principally concerned with whether consumers are made better off by larger choice sets. Changing choice set size impacts foregone savings via *both* the chosen plan and the best available plan—foregone savings might increase if the best available plan is better in larger choice sets even if the plan in which beneficiaries are enrolled is no less appropriate. Therefore, we may want to examine effects not only on foregone savings, but on total costs.

Additionally, many other factors may differ across districts with different numbers of plan choices. For example, districts with healthier or sicker enrolees may offer more or fewer choices. To evaluate whether larger choice sets make consumers worse off, we therefore turn to a reduced form model of total costs in the chosen plan as a function of choice set size and a rich set of covariates. We start by constructing a dataset which consists of just the chosen plan for each beneficiary. We then estimate the coefficients on dummies for the number of plans, controlling for other covariates that might be correlated with districts' decisions to offer more or fewer choices.

Specifically, we estimate the equation:

$$u_{itr}^{N} = \xi_J + x_{it}\gamma + \xi_{d,r} + \xi_{t,e(i)} + \epsilon_{itr}, \tag{3}$$

where i indexes beneficiaries, t indexes time (in years), and r indexes tiers (individual/couple/family). u_{itr}^{N} are either the foregone savings or the total costs in the chosen plan and ξ_J are the coefficients of interest (the dummies for the number of plans). The remaining controls in the model capture other differences in districts. In particular, $\xi_{d,r}$ includes choice set (d) $\times$ tier (r) fixed effects. With these included, we are only identifying the effect of choice set size from within district/tier changes in the number of choices offered to employees; that is, any fixed differences across district/tier combinations (such as taste for variety) are controlled for.

Of course, it is possible that even changes in the number of choices offered are correlated with underlying employee health. To address this point, we also include in the model $\xi_{t,e(i)}$, a set of year $\times$ decile of individual expenditure fixed effects. That is, we control for how individual expenditures impact choices, so that any choice set variation within districts is independent of enrolee health. Finally, x_{it} are controls which include the employer contribution and the number of years the beneficiary has been present in the sample.

The resulting coefficient measure how foregone savings and total costs vary as the number of plans vary within a choice set over time—holding fixed district characteristics, individual expenditure, employer contribution, and other factors which might impact total costs and foregone savings. We omit the fixed effect for choice sets with two plans (the smallest observed choice sets); thus, the estimated plan size effects are all defined relative to choice sets with two plans.

Column 1 of Table 4 shows the results of this regression with foregone savings on the left hand side. Each coefficient shows the foregone savings associated with choice sets of each size, relative to a choice set with only two options. The results are quite close to Figure 3; the levels are somewhat lower, but the slope by choice set size is steeper.[24] The results in Table 4 thus show that the raw patterns in Figure 3 persist once we use panel variation and control for differences in beneficiary characteristics.

Column 2 reports the results with total costs (rather than foregone savings) on the left hand side. While not monotonic, after partialling out covariates, the basic pattern in Figure 3 remains—choice sets with more plans not only lead to higher foregone savings, but higher total costs in the plan in which beneficiaries enrol. Therefore, smaller choice sets do appear to be associated with higher quality choices. The differences are quite large; in a choice set of 7 relative to a choice set of 2, foregone savings are higher by \$742 and total costs are higher by \$450.

24. While they show qualitatively the same pattern, the quantitative results in the Table differ somewhat from Figure 3. This is because the figure shows the raw cross-sectional variation ....

TABLE 4
Total costs vs. number of plans

Sample	MODA only	
Metric	Foregone savings	Total costs
2	0	0
3	264	105
4	408	128
5	437	302
6	567	224
7	742	450
N	92,133	92,133

Notes: 1. Table shows results from a regression of foregone savings or total costs on dummies for the number of plans in the choice set controlling for choice set × tier × rate structure fixed effects, year× decile of expenditure fixed effects, and for the subsidy amount and the number of years in which the beneficiary appears in the data. 2. MODA only shows results from the regression considering only beneficiaries who chose MODA plans, the number of plans offered by MODA, and savings relative to the best MODA plan in the foregone savings column.

In Supplementary Table E9, we replicate these results with alternative control sets—the results are extremely robust to the inclusion of different controls, and barely shift whether we include year fixed effects, year × expenditure quantile fixed effects, experience fixed effects, or subsidy controls.

While we estimate zero risk aversion in practice, we might worry that this reflects a lack of information on the part of consumers and that a more normatively appropriate standard is to allow for some risk aversion. Supplementary Table E10 demonstrates that these results are robust to a range of values for risk aversion commonly seen in the literature—larger choice sets are not simply leading consumers to pay more for better coverage.[25] Therefore, limiting choice set size, in our context, appears to hold the promise to substantially improve choice quality and lower total enrolee costs.

5.2. *The impact of the 2012 reform*

In 2012, OEBB relaxed a requirement that districts offer at most four plans per choice set. The reform itself increased the average number of plans from 3.5 to 5.4. In this section, we use the 2012 reform to assess whether our structural model can replicate the reduced form effects of this policy change.

Based on the results from equation (3), this suggests that the reform increased total costs by $157; that is, expanding choice set sizes from 3.5 to 5.4 raised the total annual spending by enrolees by $157. In Supplementary Appendix C, we show that our structural model implies a similar value. Specifically, when we estimate a decision rule for which plans are included as a function of their total cost and apply this decision rule to remove plans from choice sets with more than 4 plans in 2012, we find that the average beneficiary saves $139. Note the difference between these two estimates: the reduced form estimate uses the empirical (panel) variation in choice set size over time and asks, do we see that consumers enrol in lower cost plans when choice sets are smaller? Even absent any variation in choice set size, our structural model makes a prediction about how consumers would choose as we vary the choice set. The alignment of the structural and

25. Abaluck and Gruber (2011) demonstrate the assumptions under which we can derive a coefficient on the variance of costs from assumptions about the degree of CARA risk aversion. We consider the same range of values for risk aversion used in the simulations there.

TABLE 5
Simulated foregone welfare and total costs

Sample	MODA only	
Metric	Foregone welfare	Total costs
Simulated—2 plans	550	2,785
Simulated—3 plans	535	2,770
Simulated—4 plans	502	2,736
Simulated—5 plans	479	2,713
Simulated—6 plans	506	2,740
Simulated—7 plans	555	2,790

Notes: The first row of Table 6 shows mean foregone savings and total costs in the "All Plans" and "MODA Only" samples. The remaining rows show results from a simulation constructed as follows. First, we estimate equation (1) allowing β, d_{jt}, ξ, and θ to all vary flexibly with the number of available plans. This gives us a choice function that tells us how people choose from a given choice set. Each row of Table 6 uses the choice function associated with the listed number of plans to simulate choices from all choice sets.

reduced form numbers suggests that these predictions perform well. We explore this issue more directly in Section 5.3, where we ask if a more general model in which our structural parameters vary with choice set size fits the data better. As we might expect from the results reported here, we find that the same parameters predict choices well regardless of choice set size.

5.3. *Choice overload?*

As noted earlier, the existing literature argues that more options may be worse because "decision-making improves when fewer options are considered concurrently" (Besedeš *et al.*, 2015). In other words, larger choice sets lead to lower performing choice functions—which, if applied to any observed choice set, would lead to worse choices. But, this assumption has never been tested against alternatives, such as that larger choice sets impact choice quality via the scope for bad choices. Our model allows us to test this hypothesis more explicitly.

To investigate "choice overload", we re-estimate equation (1) allowing the coefficients β, ξ and θ to vary flexibly with the number of plans. This allows for the possibility that consumers choose worse from larger choice sets. For example, perhaps with just two choices individuals weight premiums and out of pocket costs equally and do not need to resort to heuristics like choosing plans with low deductibles, but, as the choice set size increases, individuals begin to weight out of pocket costs less than premiums. Were this the case, we would find that the choice function estimated in choice sets with fewer plans would lead to better choices than the choice function estimated on larger choice sets, when both functions are used to simulated choices on a given choice set.

In Table 5, we perform such an exercise. Each row of Table 5 considers all choice sets regardless of the actual number of options. In each row, we simulate the results for a given choice function estimated from a specific number of plans; that is, we are varying the demand side parameters to show what costs would be if beneficiaries chose "as if" there were the listed number of plans for that row. In other words, fix a given choice set size (say four choices). If beneficiaries chose from that choice set given the structural parameters estimated with seven choices, do we see worse outcomes than if beneficiaries chose given the structural parameters estimated with two choices? One way we might is if, for example, consumers in smaller choice sets tended not to overweight premiums relative to out of pocket costs or exhibited less approximate inertia.

TABLE 6
Logit model of plan choice (MODA only) by number of plans

	Number of plans					
	2	3	4	5	6	7
Gross premiums	−0.065***	−0.070***	−0.078***	−0.089***	−0.080***	−0.079***
(*hundreds*)	(0.010)	(0.006)	(0.007)	(0.010)	(0.007)	(0.007)
Residual premium	−0.028***	−0.025***	−0.027***	−0.017	−0.030***	−0.029***
(*hundreds*)	(0.011)	(0.005)	(0.006)	(0.012)	(0.006)	(0.006)
Mean OOP costs	−0.082***	−0.044***	−0.032***	−0.007	−0.018	−0.024*
(*hundreds*)	(0.030)	(0.016)	(0.006)	(0.028)	(0.014)	(0.013)
Inertia	3,973***	3,816***	3,432***	3,192***	−3,805***	−4,397***
	(219)	(112)	(135)	(285)	(120)	(126)
Approximate inertia	1,782***	2,061***	1,865***	1,546***	1,776***	1,911***
	(353)	(92)	(87)	(255)	(154)	(175)
Deductible, in network	−0.024***	−0.046***	−0.042***	−0.074***	−0.067***	−0.056***
(*hundreds*)	(0.006)	(0.001)	(0.001)	(0.005)	(0.004)	(0.003)

Notes: 1. This table reports the logit coefficients which are used to generate the simulation results in Table 10. To obtain this coefficients, we estimate equation (1) separately by choice set size. This table is thus identical to column 3 of Table 3, with results broken out by choice set size. 2. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

The result of this exercise is that we see little variation in choice quality—we find no evidence that choices are systematically worse as the number of plans increase. Compared to Figure 3, we see relatively little variation, and we see no systematic trend. That is, allowing the choice function to vary with number of choices does not seem to explain the pattern we see in Figure 3.

Table 6 shows how the key coefficients on different attributes vary by number of plans. In other words, this table replicates Column 3 of Table 3 (the conditional logit model using the rational expectations measure), but interacts all observables with the number of plans. The qualitative choice inconsistencies documented in Section 4 are generally present, with one interesting exception. In almost all cases, we see that consumers are more responsive to gross premiums than out of pocket costs, more responsive to gross premiums than residual premiums, and consistently prefer plans with lower deductibles and lower copays after controlling for the fact that they are likely to pay less out of pocket given their claims. The one exception is that in choice sets with two plans, we see that premium and out of pocket cost sensitivity is comparable, perhaps suggesting that with only two options, consumers are better able to appraise the out of pocket cost consequences of alternative plans. Table 5 shows that this alone is not sufficient to choose well, as these consumers in choice sets with two plans still underweight residual premiums, overweight average plan attributes and demonstrate substantial inertial and approximate inertia.

Table 6 shows some suggestive evidence of trends by choice set size, but these trends go in different directions for different choice inconsistencies. As noted, the relative sensitivity to premiums and out of pocket costs and the excess weight given to deductibles is smallest in smaller choice sets. Inertia appears “u-shaped” and the variation in the degree of inertia drives the small amount of variation we see in Table 5. On net, while our estimated choice function is not constant by choice set size, we see similar qualitative patterns in all cases with similar results; choice overload cannot explain why larger choice sets lead to worse choices in our setting.

5.4. *Why are large choice sets worse?*

If the reason for higher costs for larger choice sets is not choice overload, then it must be that the choice set is leading to worse choices as it grows larger. This finding is driven by the behaviour of administrators—if administrators chose randomly which plans to include in the choice set,

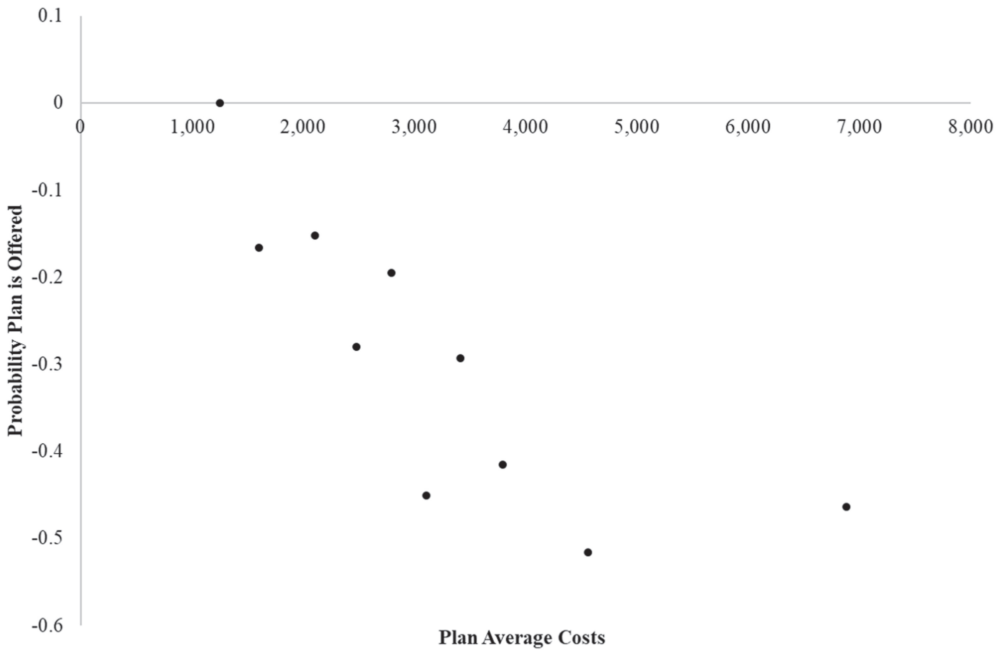

FIGURE 4

Probability offered vs. average costs (MODA only)

Notes: Figure shows plan offer probabilities by quantile of average costs. To construct this graph, we create an observation for each (year, district, plan, tier) and regress whether the plan is offered on average costs of that plan in that (year, plan, tier) as well as (year, tier) fixed effects. In other words, the regression is asking whether plans which are low cost relative to other plans available for that (year, tier) tend to be offered in more districts. The regression is normalized relative to the offer probability for (year, plan, tiers) in the first decile of average costs.

absent choice overload, the quality of choices would be independent of the size of the choice set. That is clearly not the case in the OEBB.

Specifically, district benefit managers appear able to identify better plans and include them in choice sets. In Figure 4, we plot the relationship between the average cost of a plan and the probability that the plan is offered. To construct this graph, we create an observation for each (year, district, plan, tier) and regress whether the plan is offered on average costs of that plan in that (year, plan, tier) as well as (year, tier) fixed effects. In other words, the regression is asking whether plans which are low cost relative to other plans available for that (year, tier) tend to be offered in more districts. The regression is normalized relative to the offer probability for (year, plan, tiers) in the first decile of average costs. The resulting trend is stark: higher average cost plans are substantially less likely be included in choice sets. When average costs increase by \$1,000, the probability that a plan is offered decreases by 10 percentage points. That is, the marginal plan that is added to choice sets as they get larger is systematically more expensive than the plans included when choice sets are smaller.

In what respect are bad plans worse? Table 7 replicates the total cost results from Table 4, replacing total costs on the left-hand side of the regression with gross premiums, net premiums and out of pocket costs. The result is that we see that districts with larger choice sets have worse total costs because they have higher premiums. While in general, higher premium plans offer better coverage to compensate, districts that offer all plans include plans which have higher premiums but whose additional out of pocket coverage is not valuable for the beneficiaries in those districts in an ex ante or ex post sense. Districts which offer fewer plans remove high premium plans whose

TABLE 7
What varies by choice set size?

	Net premiums	Gross premiums	OOP
2	0	0	0
3	89	155	18
4	138	145	−14
5	187	21	107
6	229	243	−11
7	425	546	23

Notes: Table 7 parallels Table 4 but replaces total costs on the left-hand side of the number of plans regression with different components of total costs. These components are regressed on dummies for the number of plans in the choice set controlling for choice set × tier × rate structure fixed effects, year × decile of expenditure fixed effects, and for the subsidy amount and the number of years in which the beneficiary appears in the data.

TABLE 8
Expenditures and demographics by district restrictions

	Restricted	Not restricted
Expenditures	11,294	10,582
Sex (% male)	28.5%	26.4%
Individual	0.283	0.250
Individual and child	0.229	0.233
Individual and spouse	0.120	0.120
Family	0.368	0.397

Notes: Comparison of average expenditures and beneficiary demographics for districts that did not offer (restricted) or did offer (not restricted) every available plan in 2012 and 2013 (pooled average), when benefit managers have the option of offering all available plans.

specific coverage characteristics are not beneficial for enrolees in those districts. This suggests then that district benefit managers are playing an important role in helping consumers choose better by exercising discretion.

Why is it that some district benefit managers prune choice sets while others do not? One reason could be different tastes for broad or narrow networks—but once again we restrict here only to MODA plans so that the only differences are plan financial characteristics. Another could be that there are omitted characteristics across districts that are correlated with both choice quality and the number of plans offered. To address this possibility, Table 8 compares expenditures and demographics in districts with and without any restriction in 2012 and 2013, when districts had the option of offering all plans. Districts that offer all plans tend to have slightly higher expenditures, a few percentage points more employees in single employee plans and a few percentage points less in family plans. Demographics are otherwise quite similar. In Table 9, we report results from a panel regression of the number of plans in each district on ranges of enrolment size, quantiles of the district subsidy amount, quantiles of district medical expenditures, as well as fixed effects for district, year, and tiers. The upshot is that none of these district level observables explain an appreciable (or even statistically significant) fraction of the variation in the number of plans offered.

These analyses raise the further question of whether our results are likely to generalize to other settings. Is the fact that larger choice sets lead to higher costs due to distinctive features of how OEBB plans are differentiated, or is it likely to apply in other health insurance exchanges?

In Supplementary Appendix C, we develop a model to investigate these issues. We find that the marginal benefit of adding another option to a choice set depends on the cost of that plan on average for all plan beneficiaries, plus the product of the sensitivity to individual heterogeneity times a function which determines the degree to which individuals can take advantage of that heterogeneity. The first is clearly context specific, so we consider a simulation where we randomly

TABLE 9
Correlates of number of plans

Quartile	Enrolment size	Medical expenditure	Subsidy amount	% Families
2	−0.158	0.150	−0.033	0.011
	(0.329)	(0.256)	(0.107)	(0.012)
3	−0.554	0.402	−0.120	0.004
	(0.418)	(0.279)	(0.121)	(0.013)
4	−0.016	0.268	−0.119	0.008
	(1.057)	(0.268)	(0.123)	(0.017)

Notes: Table reports regression coefficients from a single regression of the number of plans in each district on quartiles of number of beneficiaries enrolled, average expenditures in the district, subsidy amount and the percent of beneficiaries which are families as well as year, tier, and district fixed effects. Quartile 1 is the omitted category in each case.

replace two plans in each Oregon choice set with a plan that provides full coverage and a plan with the largest deductible allowed under the Affordable Care Act, \$7,000. This allows us to consider a much more heterogeneous choice sets. But even with this diverse choice set, the benefits from heterogeneity remain small relative to the variation in average plan cost. If we imagine a hypothetical information intervention that increased the sensitivity of beneficiaries to out of pocket costs, this intervention would need to make individuals *10 times* more sensitive to out of pocket costs before the benefits of heterogeneity would outweigh the fact that choice sets that include high average cost plans are worse for the average beneficiary.

Taken together, our findings show that smaller, curated choice sets lead to beneficiaries being enrolled in lower cost plans because larger choice sets have more plans which are higher cost on average. Individual choices do little to offset this "more dangerous" choice environment. The very small benefits of heterogeneity, relative to the large increase in average costs with choice set size, suggest that a large reduction in choice frictions would be required before heterogeneity could offset the effects of poorer choices when we remove curation.

A few important caveats to this finding are in order. First, we are concerned principally with choices among plans provided by a single insurer. Thus, while our results suggest that consumer choices can be improved by restricting the choice set to leave out bad options, the finding that the left out options have little value may change if the left-out options have different provider networks. Our empirical results have little to say about this, but they do suggest that restricting the number of offerings from a given insurer can lead to better choices. Additionally, our analysis is partial equilibrium—we are concerned with how choice architecture can help enrolees make better choices. This is only one part of the more general question of the optimal design of health insurance menus. If one were to consider total surplus divided among insurers, employers, and employees, expected costs would be a transfer, and total surplus would depend on whether consumers were matched to plans with desirable attributes such as risk protection and how this relates to moral hazard (Marone and Sabety, 2019). Even more generally, a full analysis must consider how the degree of competition induced by the choice environment would impact equilibrium prices (via competition and adverse selection) as well as the choices of insurers about how much coverage to offer. We believe the partial equilibrium analysis is nonetheless informative as a necessary piece of the larger puzzle of optimal insurance market design.

5.5. *Information vs. choice set restrictions*

In this section, we simulate two types of policies that might impact choices: information interventions and choice set size restrictions. By doing so we illustrate the power of choice set size restrictions—in particular given the importance of both inertia and approximate inertia found in our models. For this exercise, we consider choice sets with more than two plans, since some of our simulations involve removing two plans from each choice set.

TABLE 10
Simulations: changing choice set size and providing information

	Year: best plan	District: best plan	Individual best plan	Year: remove two worst	District: remove two worst	Individual: remove two worst
Choice set intervention						
Status quo impact	215	332	537	75	245	377
Only approximate inertia	–	–	–	57	242	387
No inertia	–	–	–	19	227	378
Equivalent information intervention						
Status quo impact	164	192	204	95	175	197
Only approximate inertia	200	255	303	96	221	268
No inertia	211	329	510	71	248	375

Notes: The top panel shows simulations of the impact of changing choice sets on total costs in the chosen plan. The first row shows the estimated impact given the current degree of inertia, the second row the estimated impact given only approximate inertia, and the third row the estimated impact removing all inertia. The first column shows the impact of keeping only the best plan based on average costs among all beneficiaries in that year, the second column keeping only the best plan among all beneficiaries in that district, and the third keeping only the best plan for that individual (the degree of inertia is irrelevant with only one plan). The fourth column shows the impact of removing the two worst plans by average costs in that year, the fifth column the impact of removing the two worst plans by average cost in that district, and the sixth column the impact of removing the two worst plans for that individual. The bottom panel shows the impact of an information intervention (simulated as described in the Section 5.5. In this case, no plans are removed, but the information intervention is simulated so that, with no inertia, it is as effect as the choice set intervention in the status quo (this is done using a grid search for the scale parameter, and the sixth and first row do not match exactly due to the coarseness of this grid).

We first consider restricting choices to the lowest total cost plans (premium plus expected out of pocket costs) in the choice set. In the first three columns of the top panel of Table 10, we consider restricting choices to either: (a) the lowest cost plan based on average costs among all beneficiaries in that year, (b) the lowest cost plan based on average costs in each district, or (c) the lowest cost plan for each individual.[26] Restricting to the lowest cost plan for each individual saves \$537.[27] Substantial savings remain available if we offer all beneficiaries only the lowest cost plan in their district available to them (\$332), or the lowest cost plan based on average costs in a given year (\$215). This exercise shows that, while there is heterogeneity in which plans are best, the most inexpensive plans tend to save money across the board.

We show in Supplementary Table E11 that replicating these analyses with foregone welfare does not change the result. Assigning beneficiaries to lower cost plans saves money. It sometimes affords greater risk protection and sometimes less, but on net, these differences are negligible relative to the monetary savings. This exercise is partial equilibrium—if employers offered only a single plan, this would change the nature of their bargaining with insurers and change premiums. Restricting the set of offered plans should increase the bargaining power of employers and thus lower premiums for plans that would be offered anyway (Dafny, Ho and Lee, 2019). Our results suggest that, if this plan were chosen wisely, it might yield large benefits over the status quo. In fact, if we allow benefits managers to choose any number of plans ranked based on average cost in their district, choosing only a single plan is optimal more than half the time, and the mean number of plans is 1.4.

26. For simulations (a) and (b), to keep the beneficiaries we are comparing comparable across all simulations, we restrict to the plan in each individual's choice set with the lowest average cost in that year (in a) and the plan in each individual's choice set with the lowest average cost in their district (in b).

27. This number differs slightly from the \$522 reported in Figure 1 because we are now restricting to choice sets with more than two plans.

In the next three columns of Table 10, we consider the alternative policy of removing the worst two plans by each criterion. Simulating the impact of this policy requires using the choice model estimated in Section 4 to evaluate counterfactual choices. We can do so allowing for the status quo choices, and also removing various types of inertia. These simulations therefore allow us to understand the role that both inertia and, separately, approximate inertia play in driving our results.

In the first row of Columns 4, 5, and 6 of Table 10 we simulate status quo choices. Removing the two plans which are worst in a given year achieves only small benefits, while removing the two worst plans in each district saves \$245 per person, and removing the two worst plans for each individual saves almost \$400 per person. The plans which are worst in some districts are often pretty good in others, so removing them from all choice sets produces small benefits. By contrast, the plans which are best are rarely bad in our specific context, so restricting only to the best plans on average produces large benefits. Prior to 2012, benefits managers were empowered in each district to decide how to meet the four plan limit. They removed two plans per district on average and saved about \$150 per person. Since \$150 is not far from the \$245 limit of what they could conceivably have saved, district benefits managers did a reasonable job of identifying plans which were unsuitable for beneficiaries in their district. We provide more direct evidence on this point in Section 5.3 below when we investigate which plans are offered as a function of their cost characteristics.

The second and third rows of the first panel of Table 10 show that the benefits of removing choices are not too sensitive to the degree of inertia. If the plan a beneficiary was going to choose is removed, they are forced to choose a new plan regardless of their degree of inertia. If an alternative plan is removed, they do not alter their choices very much. Inertia has a non-zero impact because the presence of inertia alters the likelihood that they choose an alternative plan if their preferred option is not removed.

In contrast to choice set policies, policies which seek to inform consumers are very sensitive to the degree of inertia at least if they would be effective without inertia, as we now show. To model strong information interventions, we simulate changes which, for new beneficiaries (e.g. without inertia or approximate inertia) are structured to be as effective as each choice set restriction in the status quo. In this way, the information interventions can be readily compared to choice set restrictions.

In particular, we assume that the information intervention sets positive utility equal to normative utility plus an idiosyncratic error term, where the variance of the remaining error term determines the effectiveness of the information intervention. We choose this variance so that, with no inertia (i.e. among new beneficiaries), the information intervention has the same effectiveness as each choice set restriction. In most cases, this is already more effective than any intervention in the literature. For example, Gruber *et al.* (2020) find that an information intervention targeting insurance brokers reduced foregone savings among new beneficiaries by 22%, corresponding to a \$120 per capita reduction in total costs. That is about half the size of the benefits of removing the two worst plans in each district. Importantly, we assume that the information intervention itself does not alter the utility impact of approximate inertia. If approximate inertia arises because consumers are concerned that dissimilar plans will have higher out of pocket costs, then informing them about out of pocket costs could reduce the degree of approximate inertia.

In Rows 1 and 2 of the second panel of Table 10, we show how adding inertia to the model impacts the effectiveness of these interventions.[28] The larger the impact of the intervention, the more that inertia reduces its effectiveness. This is intuitive: if consumers would actively choose

28. Implicitly, our modelling assumes that the estimated inertia parameters are structural and that while information reduces the variance of the idiosyncratic error term, it does not impact the inertia coefficients. Our results in Section 4.2

poorly, their inertial plan will not be much worse than their active choice and may even be better. But if consumers are induced to make better choices by an information intervention, inertial tendencies mitigate these gains. For example, if we simulate an information intervention that saves \$329 among new beneficiaries, these benefits are reduced by 42% to \$192 due to the presence of both approximate and direct inertia.

These results have two important implications. First, they highlight the value of our introduction of approximate inertia into choice models. Across the information interventions we consider, the gains from removing inertia but retaining approximate inertia are on average 44% as large as those from removing both sources of inertia (and approximate inertia is more important the more consequential the information intervention we consider).[29] For example, if we introduce an information intervention as consequential as choosing the best plan in each district (with no inertia), it produces savings of \$329 per year for new enrolees who do not suffer from inertia, while the savings for existing enrolees are only \$192. But only 46% of the \$137 loss in savings to existing enrolees is achieved by removing standard inertia (i.e. forcing an active choice), while the remaining 54% is achieved by also removing approximate inertia. The fact that consumers, even with switching plans, switch only to "similar" plans greatly undercuts the value of information interventions.

Second, these results highlight an important challenge to using information to help people make better choices. Information interventions are inherently limited in their effectiveness because of the presence of both direct inertia and approximate inertia. This limitation is clearly visible in existing studies. For example, in Medicare Part D, Kling *et al.* (2012) provide information to returning beneficiaries and find that switching rates increase by 11 percentage points, from 17% to 28%. While this is a substantial increase, the striking fact is that 72% of beneficiaries remain inertial. In Abaluck and Gruber (2016b), we show that removing inertia alone is not sufficient to produce good choices. Removing bad options more directly deals with the complementary problems of inertia and choice inconsistencies by preventing any consumers from choosing those options.

6. CONCLUSIONS

Debates over the role of choice in health insurance markets are likely to grow in the coming years. The exchanges that form the backbone of the ACA are under political attack, and the Republican majority in the Senate has stated its preferences for further promoting choice through "premium support" programs for Medicare. As a result, it is critical to understand the implications of choice over insurance products, and one of the most important elements of such understanding is how choice impacts the quality of consumer insurance plan enrolment.

The setting explored in this article has a number of unique advantages for addressing this question. We have sizeable variation in the nature of choice sets facing otherwise similar individuals, with variation in the number of insurance options, the relative prices of these options, and the individual cost sharing implications of these options.

We use these data to first document sizeable choice inconsistencies. Our finding confirms evidence from Handel (2013) and Bhargava *et al.* (2017) that choices are inconsistent in the broader insurance context, as well as a series of studies which document such inconsistencies in the choice over prescription drug plans. The dollars at stake are sizeable; even among plans

lend some support to this assumption by showing that we see similar degrees of approximate inertia and inertia even as choice sets change. This is also consistent with the finding that information interventions in the literature among returning beneficiaries tend to find modest impacts on switching behaviour (Kling *et al.*, 2012; Ericson *et al.*, 2017, 2019).

29. Note that this does not imply that approximate inertia is *more* important than inertia. Removing approximate inertia but keeping inertia also generates only small benefits.

which are identical in all aspects other than financial coverage characteristics, foregone savings is in the range of \$500–\$600 per year on average and exceeds \$2000 for a non-trivial fraction of consumers.

We also extend the choice inconsistencies literature in an important way to documenting a novel inconsistency, "approximate inertia". Individuals are not only inertial in their choice of plans, but when they do change plans they disproportionately move to "nearby" plans, regardless of cost savings. This compounds the uphill battle faced by tools such as decision support interventions; not only must they overcome inertia, but even when they do, choices are still influenced by factors such as plan placement.

Critically, we find that insurance costs are much lower in smaller than in larger choice sets; since pricing is set at the state and not district level, this effect arises solely through choice differences and not competitive effects. This is the first evidence that curated choice sets improve choices in practice. Strikingly, this effect does not arise from choice overload, as is often assumed in the literature on choice set size. Rather, it appears to arise from variation in the quality of choices that are offered by plan administrators as choice set sizes grow, and the fact that poorer choices on average are not offset by individuals through better decision making.

A key question raised by our results is the generalizability of the finding that marginal plans in larger choice sets are worse. It is useful to contrast two scenarios. Consider first cases like the Affordable Care Act exchanges, where entry of plans is determined endogenously by market forces and not by the active curation of a manager. In these cases, insurers are incentivized to exploit the types of choice inconsistencies documented here, for example, by offering plans with low premiums and obscured high cost sharing. In such a situation, standardizing plan options (as in Ericson and Starc (2016)) might be a more important tool than limiting choices, since there is no clear argument that limited choices would be curated to be best for (inconsistent) consumers.

Consider next cases like our setting, where managers choose a set of plan options to offer their enrolees. In this case, managers who are choosing a limited set of plans that are best for the average enrolee can do better than allowing more plans that are tailored for smaller groups of heterogeneous consumers. The question then arises of how managers will choose in practice. The evidence from Oregon is encouraging, in that managers tended to choose suitable plans for their districts. Our results suggest that making this choice "locally", at the district level rather than state-wide, produces additional benefits. This also reduces concerns of regulatory capture by limiting the stakes for any given choice.

In many settings, choice set construction will also have supply side consequences. Understanding better how choice set curation impacts equilibrium price-setting, especially in a dynamic context with a high-degree of inertia, is an important topic for future research.

Acknowledgments. We are grateful to Christopher Behrer, Emily Crawford, Adrienne Sabety, and Sean Sall for exceptional research assistance, and to NIA grant number R01 AG031270 and the MODA Foundation for financial support. We are extremely grateful to Joan Kapowich for making this study possible, and to Mary French and Heidi Williams of the Oregon Educators Benefits Board (OEBB) for valuable information on program administration. We also thank seminar participants at the Behavioral Economics Annual Meeting, NBER, the University of Georgia, and the Canadian Public Economics Group for helpful comments.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online. And the replication packages are available at https://dx.doi.org/10.5281/zenodo.6829021.

Data Availability Statement

The data underlying this study were provided by the Oregon Educators Benefit Board (OEBB) by permission. The code is provided in our online repository: https://doi.org/10.5281/zenodo.6829021.

REFERENCES

ABALUCK, J. and ADAMS, A. (2017), "What Do Consumers Consider Before They Choose? Identification from Asymmetric Demand Responses" (NBER Working Paper 23566).

ABALUCK, J. and GRUBER, J. (2009), "Choice Inconsistencies Among the Elderly: Evidence from Plan Choice in the Medicare Part D Program" (NBER Working Paper 14759) https://www.nber.org/papers/w14759.

ABALUCK, J. and GRUBER, J. (2011), "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program", *American Economic Review*, **101**, 1180–1210.

ABALUCK, J. and GRUBER, J. (2016a), "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program: Reply", *American Economic Review*, **106**, 3962–3987.

ABALUCK, J. and GRUBER, J. (2016b), "Evolving Choice Inconsistencies in Choice of Prescription Drug Insurance", *American Economic Review*, **106**, 2145–2184.

ABALUCK, J. and J. GRUBER (2016c), "Improving the Quality of Choices in Health Insurance Markets" (NBER Working Paper 22917).

BERTRAND, M., KARLAN, D., MULLAINATHAN, S., SHAFIR, E. and ZINMAN, J. (2010), "What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment", *The Quarterly Journal of Economics*, **125**, 263–306.

BESEDEŠ, T., DECK, C., SARANGI, S. and SHOR, M. (2015), Reducing Choice Overload without Reducing Choices", *Review of Economics and Statistics*, **97**, 793–802.

BHARGAVA, S., LOEWENSTEIN, G. and SYDNOR, J. (2017), "Choose to Lose: Health Plan Choices from a Menu with Dominated Option", *The Quarterly Journal of Economics*, **132**, 1319–1372.

DAFNY, L., HO, K. and LEE, R. S. (2019), "The Price Effects of Cross-Market Mergers: Theory and Evidence from the Hospital Industry", *The RAND Journal of Economics*, **50**, 286–325.

ERICSON, K. and SYDNOR, J. (2018), "Liquidity Constraints and the Value of Insurance" (NBER Working Paper w24993).

ERICSON, K. M. M., KINGSDALE, J., LAYTON, T. and SACARNY, A. (2017), " Nudging Leads Consumers in Colorado to Shop but not Switch ACA Marketplace Plans", *Health Affairs*, **36**, 311–319.

ERICSON, K. M. M., KINGSDALE, J., LAYTON, T. and SACARNY, A. (2019), "How do Humans Interact with Algorithms? Experimental Evidence from Health Insurance" (NBER Working Paper 25976).

ERICSON, K. M. M. and STARC, A. (2016), "How Product Standardization Affects Choice: Evidence from the Massachusetts Health Insurance Exchange", *Journal of Health Economics*, **50**, 71–85.

GRUBER, J. (2017), "Delivering Public Health Insurance through Private Plan Choice in the United States", *Journal of Economic Perspectives*, **31**, 3–22.

GRUBER, J., HANDEL, B., KINA, S. and KOLSTAD, J. (2020), "Managing Intelligence: Skilled Experts and AI in Markets for Complex Products" (NBER Working Paper 27038).

HANDEL, B. R. (2013), "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts", *American Economic Review*, **103**, 2643–2682.

HANDEL, B. R. and KOLSTAD, J. T. (2015), "Health Insurance for 'Humans': Information Frictions, Plan Choice, and Consumer Welfare", *American Economic Review*, **105**, 2449–2500.

HEALTHCARE.GOV (2017), "2017 QHP Landscape Data" https://www.healthcare.gov/health-plan-information-2017/. Last accessed 16 January 2018.

IYENGAR, S. S. and KAMENICA, E. (2010), "Choice Proliferation, Simplicity Seeking, and Asset Allocation", *Journal of Public Economics*, **94**, 530–539.

KAISER FAMILY FOUNDATION (2015a), "Medicare Advantage 2016 Data Spotlight: Overview of Plan Changes" https://www.kff.org/medicare/issue-brief/medicare-advantage-2016-data-spotlight-overview-of-plan-changes/. Last accessed.

KAISER FAMILY FOUNDATION (2015b), "What's In and What's Out? Medicare Advantage Market Entries and Exits for 2016" https://www.kff.org/report-section/whats-in-and-whats-out-medicare-advantage-market-entries-and-exits-for-2016-appendix/. Last accessed.

KAISER FAMILY FOUNDATION (2017a), "2017 Employer Health Benefits Survey" https://kff.org/health-costs/report/2017-employer-health-benefits-survey/. Last accessed 19 September 2017.

KAISER FAMILY FOUNDATION (2017b), "Total Medicaid MCOs" http://kff.org/other/state-indicator/total-medicaid-mcos/. Last accessed.

KAISER FAMILY FOUNDATION (2018), "Health Insurance Coverage of the Total Population" http://kff.org/other/state-indicator/total-population/. Last accessed.

KAISER FAMILY FOUNDATION (2019), "An Overview of the Medicare Part D Prescription Drug Benefit" https://www.kff.org/medicare/fact-sheet/an-overview-of-the-medicare-part-d-prescription-drug-benefit/. Last accessed.

KAMENICA, E. (2008), "Contextual Inference in Markets: On the Informational Content of Product Lines", *American Economic Review*, **98**, 2127–2149.

KETCHAM, J. D., KUMINOFF, N. V. and POWERS, C. A. (2016), "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program: Comment", *American Economic Review*, **106**, 3932–3961.

KETCHAM, J. D., LUCARELLI, C., MIRAVETE, E. J. and ROEBUCK, M. C. (2012), Sinking, Swimming, or Learning to Swim in Medicare Part D", *American Economic Review*, **102**, 2639–2673.

KETCHAM, J. D., LUCARELLI, C. and POWERS, C. A. (2015), "Paying Attention or Paying Too Much in Medicare Part D", *American Economic Review*, **105**, 204–233.

KLING, J. R., MULLAINATHAN, S., SHAFIR, E., VERMEULEN, L. C. and WROBEL, M. V. (2012), “Comparison Friction: Experimental Evidence from Medicare Drug Plans”, *The Quarterly Journal of Economics*, **127**, 199–235.

MARONE, V. R. and SABETY, A. (2019), “Should There Be Vertical Choice in Health Insurance Markets?” (Technical Report, Working Paper).

OEBB (2008–2013a), “Choice Set Data”.

OEBB (2008–2013b), “Claims and Eligibility Data”.

OEBB (2008–2013d), “Premiums Data”.

OEBB (2008–2014), “Plan Characteristics Data”.

OEBB (2012–2013c), “Informed Enrollment Data”.

OEBB (2013e), “School District Surveys and Union Contracts”.

OLAFSSON, A. and PAGEL, M. (2018), “The Liquid Hand-to-Mouth: Evidence from Personal Finance Management Software”, *The Review of Financial Studies*, **31**, 4398–4446.

SCHWARTZ, B., WARD, A., MONTEROSSO, J., LYUBOMIRSKY, S., WHITE, K. and LEHMAN, D. R. (2002), “Maximizing versus Satisficing: Happiness is a Matter of Choice”, *Journal of Personality and Social Psychology*, **83**, 1178.

SETHI-IYENGAR, S., HUBERMAN, G. and JIANG, W. (2004), “How Much Choice Is Too Much? Contributions to 401(K) Retirement Plans”, *Pension Design and Structure: New Lessons from Behavioral Finance*, **83**, 84–87.