Simplifying Health Insurance Choice with Consequence Graphs

Anya Samek, University of Southern California

Justin Sydnor, University of Wisconsin – Madison & NBER

Abstract

Standard theories of insurance demand are based on the idea that people select plans that maximize expected utility over the distribution of final wealth outcomes determined by the plan choice. However, a number of recent empirical studies in health insurance markets document patterns of sub-optimal choices that cannot be rationalized by standard models. These seemingly inefficient choices may be linked to consumers' poor understanding of insurance and the complexity involved in mapping the cost-sharing features of plans to their distribution of financial consequences. We develop an approach that we call "consequence graphs" to presenting health insurance options that combines information about the ex-ante distribution of medical spending needs with plan cost-sharing rules to present each plan option graphically. The resulting consequence graphs show the distribution of total spending generated by each plan option. We use an incentivized laboratory experiment to compare consequence graphs to providing information on cost-sharing features (e.g., deductibles and coinsurance rates) of plans. We find that when plan menus have options that are financially dominated, the majority of people violate financial dominance when choosing with typical feature-based information displays, but very few do so with consequence graphs. We interpret this result as clear evidence of the role that poor understanding of financial consequences plays in health-insurance choice patterns. The consequence-graph approach also provides a practical way to simplify and clarify health insurance options.

We thank the BRITE Lab at the University of Wisconsin for financial support for this project. We also thank participants at the University of Wisconsin Center for Financial Security Household Finance Seminar and the CEAR/MRIC Behavioral Insurance Workshop for useful feedback. Special thanks to Christa Gibbs for invaluable support as a research assistant on this project.

One of the defining features of health insurance in the United States in recent years has been an expansion of choices for individuals about their insurance coverage. This has occurred across a range of health insurance settings, including public health insurance marketplaces and employer-sponsored insurance coverage. With the rise of this focus on plan choices, there has been a wave of economic studies exploring the nature of choice patterns. These studies have documented patterns of decisions that are difficult to reconcile with standard economic theories of insurance demand that are based on the idea that health-insurance consumers are well-informed expected utility maximizers whose utility is defined over final wealth states.[1]

A number of these studies also provide some evidence that part of the reason people deviate from economic predictions of efficiency is that they have difficulty understanding the distribution of financial consequences they face with different plan options. One of the starkest demonstrations of these patterns comes from recent work by Bhargava, Lowenstein, and Sydnor (forthcoming), which shows that the majority of employees at a large U.S. firm selected health plans that were strictly state-wise dominated by alternative plans available to them in their option set. These authors also replicate those choice patterns with a separate sample in hypothetical-choice experiments and show that among experimental subjects the likelihood of selecting a dominated plan is correlated with measures of insurance literacy. Their hypothetical-choice experiments also include one condition that shows the maximum and minimum spending a person could have with each plan and they find that when shown that information people are less likely to select a dominated option. That finding suggests that providing more information about the distribution

[1] See for example: Frank and Lamiraud, 2009; Heiss, McFadden, and Winter 2010; Abaluck and Gruber 2011; Schram and Sonnemans, 2011; Sinaiko and Hirth, 2011; Ericson and Starc, 2012; Kling et al. 2012; Zhou and Zhang 2012; Handel 2013; Johnson et al. 2013; Loewenstein et al., 2013; Anastov and Baker, 2014; Erickson, 2014; Handel and Kohlstad 2015; Bhargava, Loewenstein, and Benartzi forthcoming; Bhargava, Loewenstein and Sydnor, forthcoming

of consequences from different plan options could improve the efficiency of consumer choices. Our study builds on that idea and seeks to test it in a direct and comprehensive way using an incentivized experimental framework.

In this study, we examine how patterns of health insurance choices change when people receive information about the distribution of financial consequences for different plan options directly rather than the typical information about the features of plans. In standard economic models of insurance demand, the utility of different coverage options is determined by the distribution of final wealth outcomes a person faces under different plans (Rothschild and Stiglitz 1976) . Typical health insurance choice environments are "feature based" in that they provide people with tables that show different cost-sharing features, such as premiums, deductibles, co-insurance rates and maximum-out-of-pocket costs. Standard theory would suggest then that to make a utility-maximizing choice people need to know the distribution of final wealth consequences for each plan, which requires mapping plan features to out-of-pocket costs for some subjective distribution of beliefs about medical spending needs. This is clearly a cognitively challenging exercise. To test the importance of this issue and provide a possible resolutions to the problems posed by this complexity, we developed an approach that we call "consequence graphs" in which instead of giving people information about plan features, we do this mapping for them. The consequence graphs we construct for each plan show the individual's total spending, combining premiums and out-of-pocket costs, for different quantiles of an ex-ante distribution of potential medical spending the person faces. This approach allows us to represent each plan as a single line in a simple graph in a way that conveys the distribution of potential financial consequences a person would face if they select that plan.

We then use an incentivized laboratory experiment to directly compare the effects of information delivered via consequence graphs relative to typical feature-based cost-sharing information. Subjects in the experiment are asked to make plan choices from a series of menu options. In order to allow for as natural of a presentation of health insurance options as possible in an incentivized lab experiment, we frame these choices by asking them to choose plans for a fictional young adult who is selecting a health plan at a new job. Importantly, the decisions are incentive compatible since subjects' payments are directly and proportionally related to the realized wealth minus medical spending of the fictional person. Subjects first see information on the distribution of medical spending possibilities for this individual, which ensures that all participants have a common set of beliefs about the health-risk distribution. They then make a series of choices from four plan menus and in three of those menus there is an option that under standard theory should dominate the alternatives. Each menu is displayed in both a traditional feature-table format and the new consequence-graph format. We randomize the ordering of these presentations to facilitate the possibility for both between-subject and within-subject analysis.

We find that consequence graphs substantially change how people select plan options, causing their choices to be more in line with predictions of standard economic theory. When subjects make choices from feature-table displays for menus with dominant options, we find that only 40% to 60% of subjects select the dominant option. With consequence graphs, the fraction selecting the dominant option rises to around 90% both for menus with state-wise dominant options and for menus where the dominant option is second-order stochastically dominant but not state-wise dominant. Importantly, because all participants received information about the distribution of potential medical spending, there was, in theory, no new information contained in the consequence graph that a well-informed and insurance-savvy participant could not have calculated themselves in the standard table display. The dramatic difference in choice patterns, however, reveals that many subjects were not, in fact, creating this mapping from plan features to the distribution of consequences when they selected options using the standard choice platform.

Displaying the consequences of insurance choices also interacts in important ways with underlying insurance literacy. When people select plan options from standard table displays there is a strong correlation between measured insurance literacy and the propensity for the plan choice to violate dominance. However, we find that with the consequence graphs this difference is completely eliminated and the likelihood of selecting a dominant plan (which is very high with consequence graphs) is unrelated to one's baseline understanding of how to calculate spending from insurance features. This could be an important feature of the consequence-graph approach, given that many surveys have documented low insurance literacy in the population (e.g., Winter et al., 2006; Johnson et al. 2013; Loewenstein et al. 2013)

We also test a menu of options with no financially dominant plan but where plans can be rank-ordered and all can be rationalized by some level of risk aversion. We do not find significant differences in overall choice patterns between consequence graph and standard presentations; however, within-subject analysis shows that half the subjects select different plans from the same menu when the display changes. We also find that the choices people make when using the consequence graph have a modest and statistically significant correlation with an independent measure of risk aversion from a common Eckel-Grossman (2002) style gamble-choice task. In contrast, there is zero correlation between measured risk aversion and plan choice when subjects choose using the standard feature-table display. This suggests that even when the consequence-graph display does not affect the overall distribution of choices, it may cause people to select plans in ways that are more consistent with models of expected-utility-of-wealth maximization.

The results of this study are closest to prior work by Johnson et al. (2013) and the previously discussed study by Bhargava et al. (forthcoming). Johnson et al. (2013) used plan options from menus modeled on the Affordable Care Act exchanges and asked experimental subjects to select the plan that would result in the minimum spending for the year given a specific expectation about the number of doctor's visits and overall medical spending. Similar to our findings, Johnson et al. (2013) find that subjects had difficulty identifying the most cost-effective option, including in some treatments where the selection was incentivized, and that errors tended to be in the direction of people selecting plans that had a lower deductible and high premiums,. Johnson et al. find that the proportion selecting the cost-effective plan is significantly improved by giving subjects a calculator or the correct answer as a default option. Our work builds on these previous studies and contributes by systematically studying the role that improving the ease of understanding the distribution of consequences has for changing health insurance choices. Importantly, the consequence-graph approach is consistent with underlying economic theory about insurance demand in a way that simpler decision aids (e.g., showing only the expected level of spending but not variance) are not.

The results of this study have academic and practical implications. For the academic literature on insurance demand, this experiment provides clear evidence of how choice patterns that are inefficient relative to standard economic benchmarks can be improved by making it simpler to map health plan designs to their consequences. Our results also drive home an argument that has been made in a number of recent studies (e.g., Sydnor, 2010; Barseghyan et al., 2013; Handel, 2013; Handel and Kohlstad, 2015; Baicker, Mullainathan and Schwartzstein, 2015; Spinnewijn, forthcoming; Bhargava et al, forthcoming) that observations of insurance choices are not directly informative about underlying preferences over financial risk when people find insurance too complex to understand the consequences of their choices. The consequence-graph approach also has practical applications because these graphs could be included easily in any choice platform where people are asked to make health insurance decisions over different levels of financial coverage. The results here suggest that displaying consequences in this way could simplify health insurance choice and move decisions in directions that are more in line with standard notions of economic efficiency. We also find that 75% of subjects prefer to make choices using the consequence graphs.

2. Data and Experimental Design

The study took place in May of 2016 at the Behavioral Research Insights through Experiments (BRITE) Lab at the University of Wisconsin, Madison, which like most economic laboratories has a policy of no deception. We recruited 201 undergraduate student participants using the lab's subject-pool recruitment system. Subjects made their decisions individually at a private computer terminal. Subjects earned $24 on average for sessions that lasted approximately 30 minutes.

The experiment consisted of a simulation of health insurance choices programmed using Qualtrics. Subject instructions (available in Appendix A) were displayed on the screen and subjects moved through the screens of the experiment at their own pace. Subjects were told to consider a fictional young adult named Jamie, who just started a new job and needed to select a health insurance plan from a menu of options provided by Jamie's employer. We used the framing of a young adult choosing health insurance so that the choice task would be naturalistic and we could use menus denoted in realistic dollar amounts for insurance contract features like deductibles. To exclude any complexities related to differences health network quality, subjects were explicitly instructed that all plan options provided the same access to doctors and hospitals. We further abstracted from any complications where out-of-pocket costs for health spending can depend not only on total medical bills but also on the types of medical spending (e.g., drug vs emergency room) by simply giving overall cost-sharing parameters for each plan that did not differentiate by service.

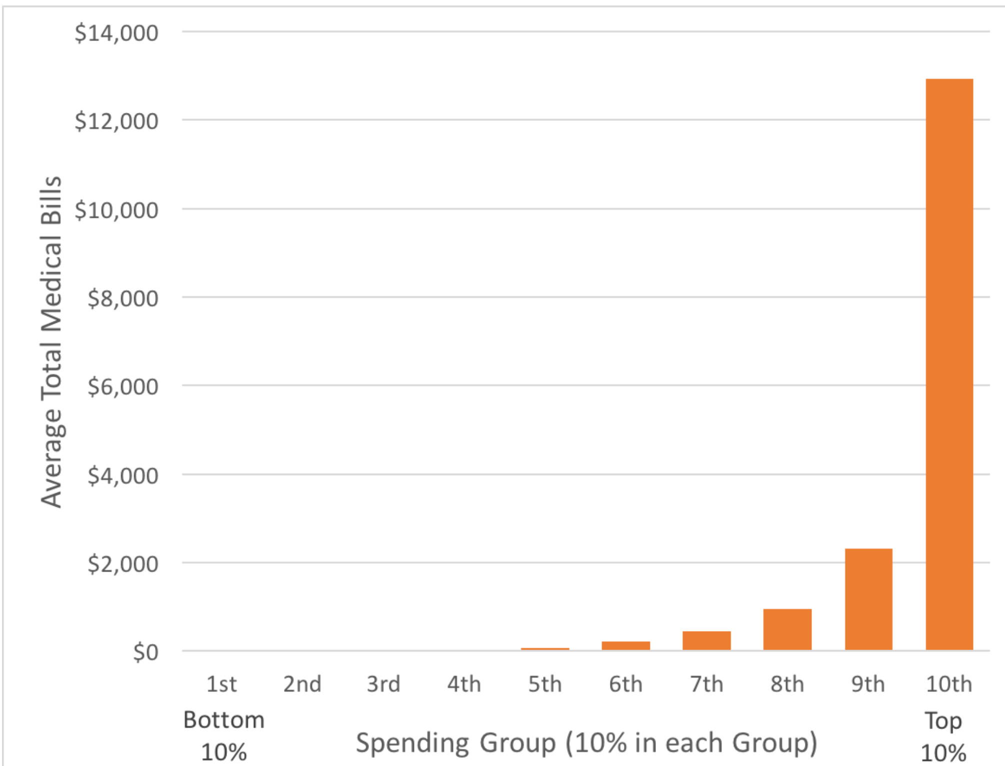
Subjects were incentivized based on Jamie's wealth level at the end of the year, which consisted of Jamie's salary (randomized between-subjects as either $40,000 and $60,000) minus the health insurance premium paid by Jamie (a direct consequence of the subject's decision) minus any out-of-pocket medical spending Jamie had to pay for the year (determined via a random draw from Jamie's distribution of possible medical spending).[2] Subjects made a series of 8 choices about health insurance for Jamie, but only one of the menus was randomly selected to determine the payout. The computer randomly selected a medical spending amount (i.e., total doctor and hospital bills for the year) for Jaimie from actual spending amounts for the sample of young adults (both men and women) between the ages of 21 and 25 in the Medical Expenditure Panel Survey from 2012 and 2013. That random draw of medical spending resulted in an out-of-pocket spending amount for Jamie based on the insurance contract chosen by the participant in the selected menu. Jamie's wealth was translated to the subjects' payment as $2 for every $1,000 of Jamie's wealth.

After reading about this information on their computer screens, all subjects (regardless of the order in which they would later see the insurance menus) were shown Figure 1, which gives the distribution of total medical bills for the sample of young adults from which Jamie's medical

[2] We randomized the size of the salary because ex-ante it was not clear whether the salary level would substantially affect how people viewed the insurance options and the level of effective risk aversion the subjects would display. In practice we detect zero difference in choice patterns across the randomly-selected salary level for any choice. As such, throughout we simply present results pooling across the two salary levels.

spending would be drawn. The figure shows spending levels in deciles, which maps to how we constructed the consequence graphs (described in the next section). Subjects were given a paper copy of this figure so that they could refer to it throughout the session. Before proceeding with the choice task, subjects answered a comprehension question about reading this figure. All but 3 of the 201 subjects correctly answered this question.

Figure 1. Distribution of Underlying Medical Bills



All subjects made choices from 4 different plan menus in two different display scenarios: standard table display or consequence-graph display, for 8 choices in total. Each subject chose from the same menu of options twice, once in each display format. The tables showed the level of contract features for different plan options: premium, deductible, coinsurance rate, and maximum out-of-pocket. These terms were described using natural language both in the instructions and at the top of each of the table menu screens. Subjects could sort the tables on two of these features, so as to facilitate as natural a comparison process as was feasible using the Qualtrics survey platform. The graphical display showed subjects the total health spending Jamie would have, defined as premium plus out-of-pocket, for different realizations of overall medical bills coming from the distribution shown in Figure 1. In this way, each plan is represented by a line in the consequence graph. Section 3 shows examples of all of the table and consequence graphs subjects saw and discusses the consequence-graph format in more detail.

We randomized the order of whether subjects saw the menus in feature-table format or consequence-graph format first, which allows for a between-subjects analysis based on the first display format they saw. Within each display block the order of the menus presented was also randomized. However, the most direct test of whether people are calculating consequences directly from the standard display is to look at insurance choices within-subject. Hence, we asked subjects to make the decisions using both display formats, and compare the choice in each menu in the first display to the choice in the same menu in the second display. To decrease the probability of order effects, we labeled the options with either the names of colors (e.g., plan "Red") or shapes (e.g. plan "Square") and we randomized this across the two display formats. In this way, a subject who selected say the "Black" plan in Menu 1 in table format may not be aware that this was the same plan as the "Diamond" plan in Menu 1 in the graph format.[3] We detect no order effects for any menus. As such, we present within-subjects analysis (using both their choices from the table display and menu display) throughout. We present the details of these menus in the next section.

After making their insurance selections from the 8 different option sets (4 menus in each of two display formats), subjects answered a question about which display format they preferred

[3] We further randomized the order of the menus the subjects saw within each of the display formats.

and an open-ended question about how they made their selections. Subjects further made a selection of a gamble from an incentivized Eckel-Grossman risk elicitation task (Eckel and Grossman, 2002), which we use as a separate measure of risk aversion. We discuss this measure in Section 3.4. As a secondary risk-aversion metric, we asked 4 questions from the Domain Specific Risk Taking (DOSPERT) scale to measure subjective attitudes toward risk. Finally, subjects were also given an incentivized insurance literacy task in which they were shown a menu with plan options with different insurance-contract features, e.g., deductible co-insurance. They were then asked to calculate for a specific level of total medical spending what the total out-of-pocket spending would for one of the insurance contracts in the table. We asked two of these questions and use these as a measure of whether subjects possess the insurance literacy to calculate spending consequences given a description of the contract features. Finally, subjects answered the short questionnaire of demographic questions.

Table 1 shows summary statistics on the limited set of demographic characteristics we collected from all subjects. A total of 201 subjects participated. The average age was 21 and 73% of subjects were female, though we did no special recruitment by gender. Just over half of the subjects stated that they themselves anticipated using some level of healthcare next year, defined as anticipating to have at least a couple of doctor visits during the year. The table shows that these demographics were balanced across the randomized order of whether the subject saw tables or graph displays first.

Table 1. Summary Statistics

Variable	(1) Full Sample	Display type seen first: (2) Tables	(3) Graphs	t-test p-value (2) vs (3)
age	21.3	21.2	21.4	0.21
female	0.73	0.77	0.68	0.15
gpa	3.4	3.4	3.4	0.4
expects to use some healthcare next year	0.54	0.54	0.55	0.85
N	201	110	91	

Notes: Table shows means. The last column shows the p-value on a t-test of the difference in means for those who saw the standard table display first versus those who saw the consequence-graph display first. Order of first display type was randomized.

3. Results

We describe the results of the experiment broken down by the different menus of options we presented subjects.

3.1 Simple 4-plan menu with a strictly dominant option

Menu 1 gave subjects a choice between 4 plans that differed in premium and deductible. Each of these plans had the same 10% coinsurance rate (the fraction of bills the individual would have to pay for total bills in excess of the deductible) and the same $2,500 in out-of-pocket spending above the deductible. The maximum out of pocket varied across plans because the deductible was different.

Table 2. Menu 1 Table Display

Plan Name	Annual Premium	Annual Deductible	Coinsurance Rate	Maximum Out of Pocket
Purple*	$817	$1,000	10%	$3,500
Blue	$1,321	$750	10%	$3,250
Red	$1,419	$500	10%	$3,000
Black	$1,957	$250	10%	$2,750

Notes: * Denotes dominant option (not shown to subjects)

The key defining feature of Menu 1 is that one of the options, option Purple in Table 2 state-wise dominated all of the other options. This dominant option had a deductible of $1,000 and premium of $817 for the year. The other three options with lower deductibles required premium increases that were greater than the reduction in deductible. Because of this, no matter the level of medical spending, the total spending for Jamie (and hence the subjects) would be lowest with the Purple plan. This menu was chosen because it was one of the sets of options in a larger menu seen by employees in the study by Bhargava et al. (2016), where a majority of employees at a large U.S. firm selected a dominated health plan. Bhargava et al., found using hypothetical-choice experiments with a separate population that the majority of subjects selected a dominated option from this menu. Using this menu gives us a first baseline for investigating the effect of consequence graphs.

Figure 2 shows the consequence graph for this menu based on the distribution of potential total medical bills for Jamie. Each plan from Table 2 is represented by a line in the consequence graph. The graph shows how total spending for Jamie (and hence the subject) varies with each plan option depending on the medical-spending realization. The graph starts at the premium level (the far left dot) and then rises for each of the 10 equally-likely spending deciles as those increased spending levels translate into higher out-of-pocket costs based on the cost-sharing parameters of the plan. At the far right of each line, we show the maximum possible total spending for Jamie with that plan. The state-wise dominance of the Purple plan in this menu can be seen in the consequence graph easily by simply noting that at all points the Purple line lie strictly below the lines of all of the other plan options.

Figure 2. Menu 1 Consequence Graph Display

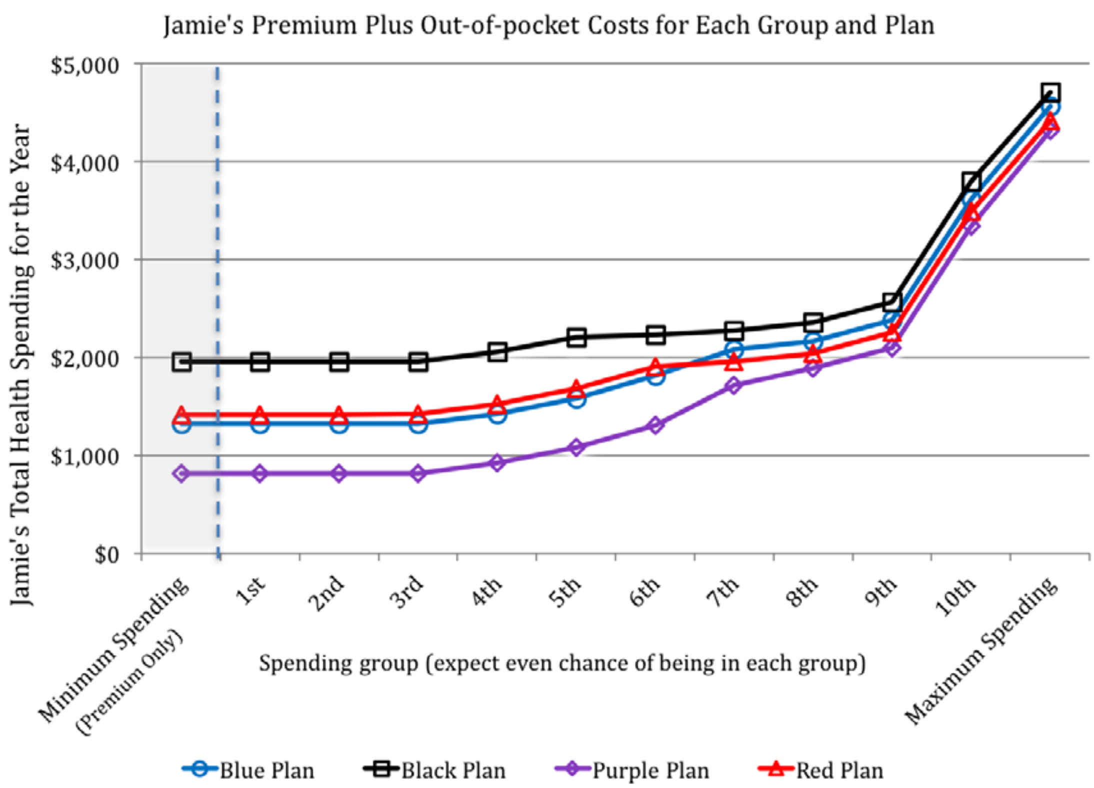


Figure 3 shows the distribution of choices the 201 subjects made across these four plan options in both the table-display format and the consequence-graph-display format. When choosing from the table, only 39% of subjects selected the dominant Purple plan with the lowest premium and highest deductible. The second most popular option in the table format was the Red plan with a $500 deductible, which was chosen by 36% of subjects. The choice shares we see here are very similar to those reported in the Bhargava et al study. In striking contrast, 93% of subjects selected the dominant Purple plan when they chose plans using the consequence-graph format. Figure 3 shows 95% confidence intervals around the choice shares and it is clear to see that the large difference in the share of subjects selecting the dominant plans is statistically significant. The p-value on a Mann-Whitney two-sample test for equality of distributions between the two choice formats is well below 0.01.[4]

It is also worth noting that in this menu there is not only a dominant option but also an option that is strictly dominated by all other plan options – namely the Black plan. To the extent that one might think of selecting that plan as an especially large deviation from standard economic benchmarks of insurance choice, it is notable that in the table format 14% of people select this highest-premium dominated plan, while 0.5% (i.e., 1 person) did so with the consequence graph.

Figure 3. Menu 1 Choice Shares

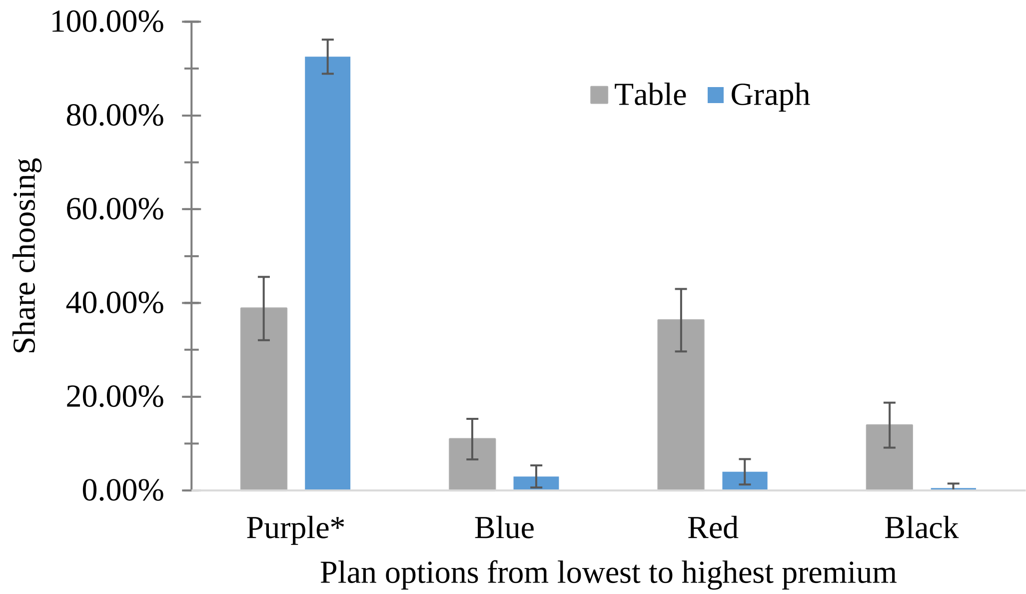


Figure note: * Denotes dominant option. Errors bars give 95% confidence intervals

[4] For this test we use the rank-ordering along the dimension of premium used in Figure 3 for the Mann-Whitney rank-sum test.

3.2 Complex 6-plan menu with a strictly dominant option

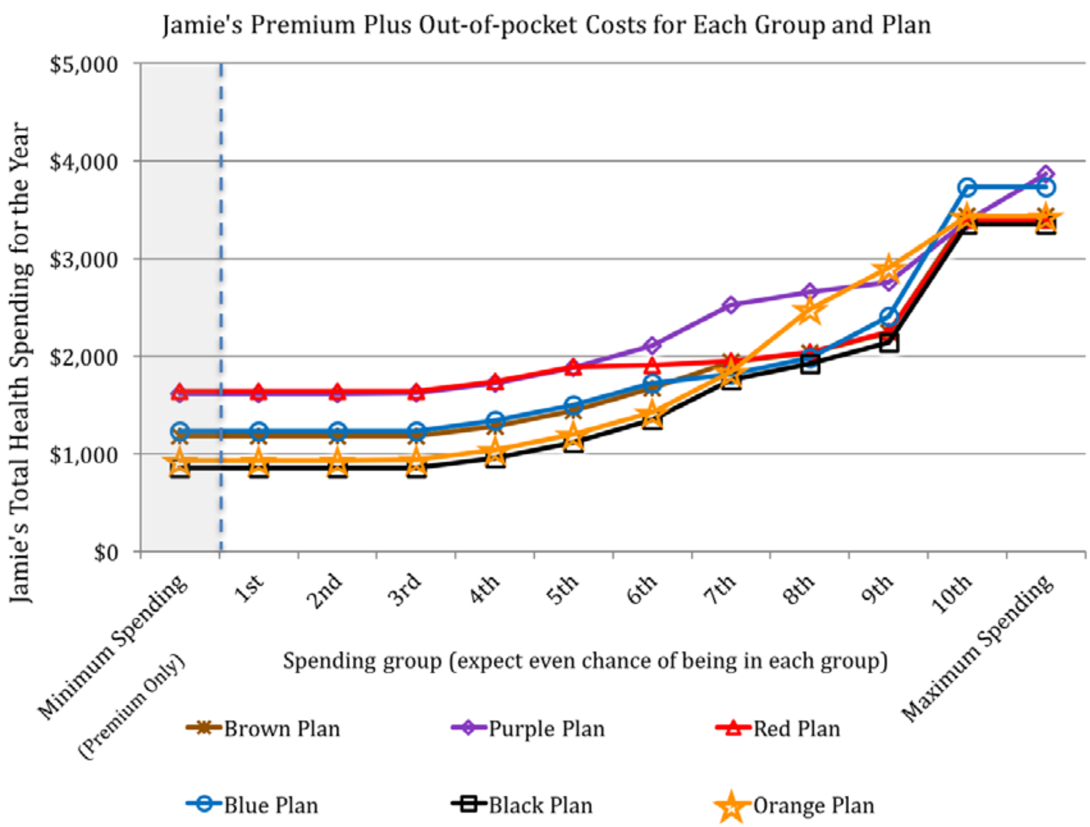
Table 3. Menu 2 Table Display

Plan Name	Annual Premium	Annual Deductible	Coinsurance Rate	Maximum Out of Pocket
Black *	$851	$1,000	10%	$2,500
Orange	$932	$1,500	20%	$2,500
Brown	$1,177	$750	10%	$2,250
Blue	$1,231	$500	20%	$2,500
Purple	$1,616	$1,000	5%	$2,250
Red	$1,635	$250	10%	$1,750

Notes: * Denotes dominant option (not shown to subjects)

The second menu we consider, displayed in Table 3, also had a state-wise dominant plan, but this menu was significantly more complex in that it had 6 options instead of 4 and the options varied by a range of combinations of deductible, coinsurance rate and out-of-pocket maximums.

Figure 4. Menu 2 Consequence Graph Display



As in the prior menu, the dominant option, the Black plan in this case, is visible as the lowest line in the consequence graph (Figure 4). Unlike the first menu, however, for many spots on the graph (i.e., for many realizations of total medical bills) there are other plans with similar total spending.

Figure 5 shows the choice shares for this menu. We again see a striking difference in choice patterns between the two display formats. When choosing from the table display, 38% of subjects selected the dominant option. Using consequence graphs, 88% selected the dominant option. Again a Mann-Whitney two-sample test for equality of distributions between the two choice formats has a p-value well below 0.01.

Figure 5. Menu 2 Choice Shares

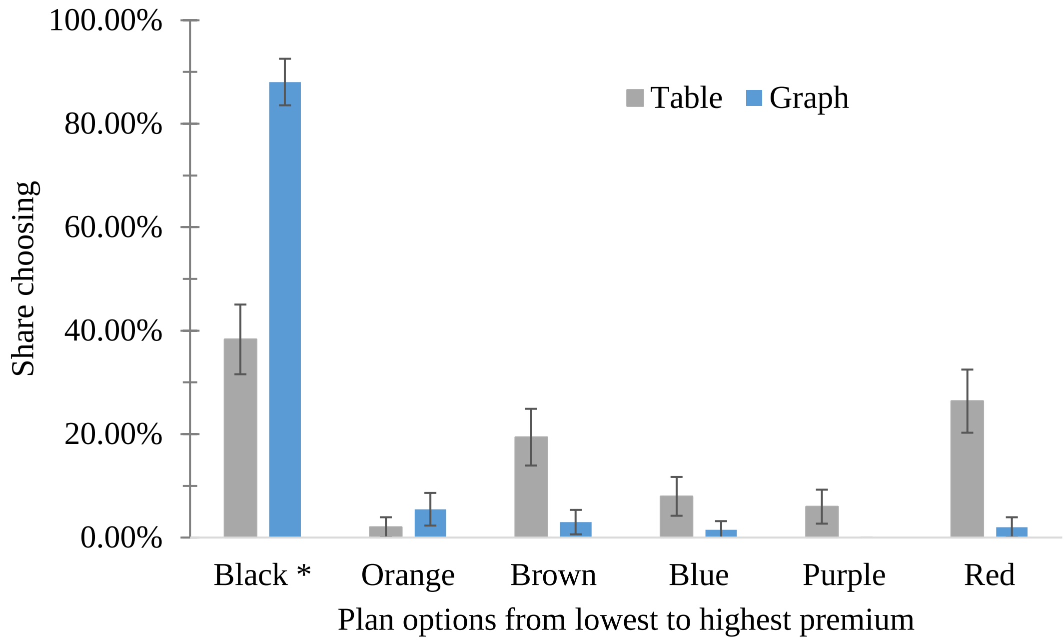


Figure note: * Denotes dominant option. Errors bars give 95% confidence intervals

3.3 Complex 6-plan menu with a second-order stochastically dominant option

The first two menus provide settings where we might expect the consequence-graph approach to perform well since the dominant options are clear to see for subjects who understand how to read the graph. In the third menu, we maintain the idea of having a dominant plan that allows for a clear litmus test of choice relative to standard economic models of insurance choice. However, here we use an option that rather than being state-wise dominant, instead only second-order stochastically dominates the other options. Under second-order stochastic domination, all risk-averse expected utility maximizers should select the dominant plan.

Table 4 shows this third menu, where there are again 6 options with a range of variations in plan features. In this menu, the Brown plan second-order stochastically dominates the other options. This plan is in the middle of the distribution in terms of premium and deductible level.

Table 4. Menu 3 Table Display

Plan Name	Annual Premium	Annual Deductible	Coinsurance Rate	Maximum Out of Pocket
Red	$863	$1,500	20%	$4,000
Blue	$913	$2,750	10%	$3,750
Brown *	$988	$1,250	10%	$2,750
Orange	$1,317	$1,000	20%	$3,500
Black	$1,589	$750	10%	$2,250
Purple	$2,113	$500	10%	$2,000

Notes: * Denotes dominant option (not shown to subjects)

Figure 6 shows the consequence graph for this menu. One can see the second-order stochastic dominance for the Brown plan by noting that it has only slightly higher total costs than the lowest-spending alternatives (Red and Blue plans) for low levels of medical spending but dramatically lower spending in the worst case scenarios than these alternatives. We also can see in the graph that the Brown plan strictly dominates three of the other options – the Orange, Black and Purple plans. This menu provides a stronger test of the consequence-graph approach, as it requires that subjects will successfully notice the attractive risk-reducing properties of the Brown plan even when it is not a strictly dominant option that is lowest at all point in the graph.

Figure 6. Menu 3 Consequence Graph Display

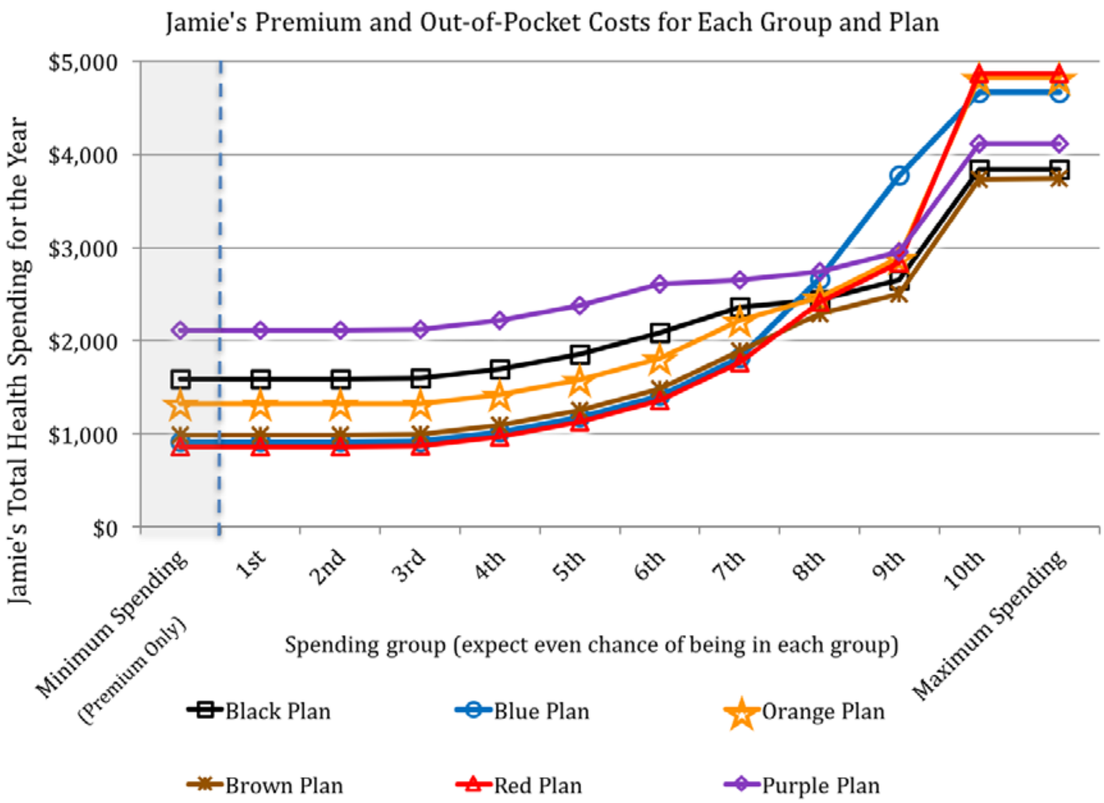


Figure 7 shows the distribution of choices for this menu. We again see a substantial increase in the fraction selecting the dominant option when the menu is presented with a consequence graphs relative to the table. We find that 88% of subjects selected the second-order stochastically dominant plan with the consequence graph, while 63% of subjects selected it with the table. Again a Mann-Whitney two-sample test for equality of distributions between the two choice formats has a p-value well below 0.01.

Figure 7. Menu 3 Choice Shares

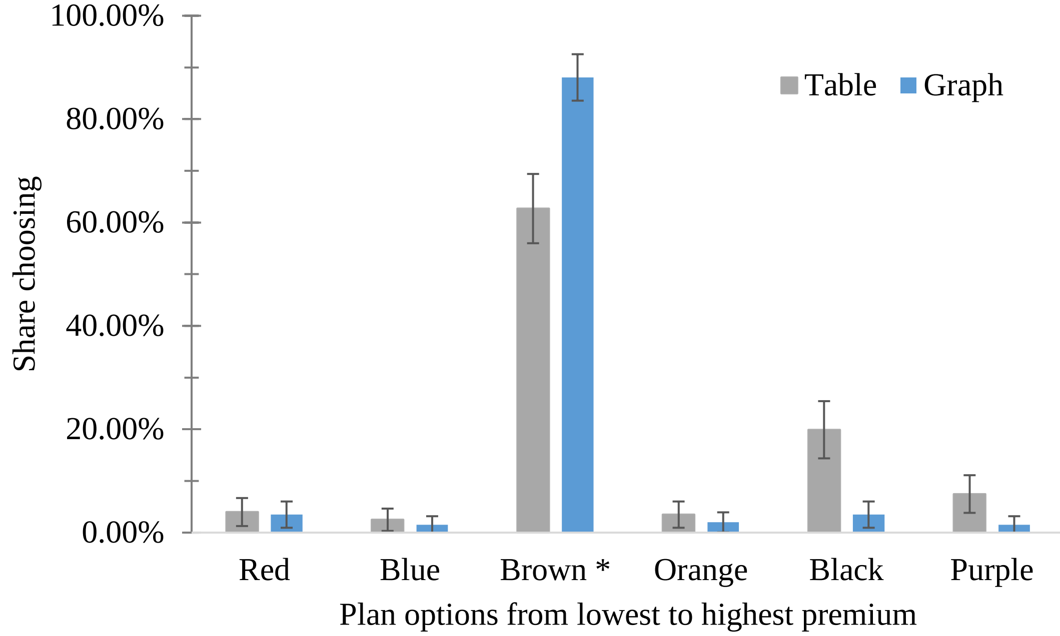


Figure note: * Denotes dominant option. Errors bars give 95% confidence intervals

Interestingly, in this menu unlike the previous menus, the second-order stochastically dominant option is chosen by the majority of people when they use the table format. This could be because the choice process when people select using tables involves some sort of trading off of plan features looking for “balance” that leads them to this second-order stochastically dominant option but did not so easily lead them to select state-wise dominant plans in the prior menu. It is also interesting to note that the second most popular plan in the table format was the Black plan, which is strictly dominated by the Brown plan, but which does have the second lowest maximum out of pocket spending. This could also be because this option was a middle option in the table for most sort orders that a subject could select and perhaps people gravitated to middle options, which we did not anticipate when we designed the interface.

3.4 Regression analysis for menus with dominant options

In this subsection, we present regression analysis to quantify the magnitude and statistical significance of the differences in the choice patterns across display format in these three menus with dominant options. These regressions also allow us to explore interactions of the display format with our measures of insurance literacy and display-format preference.

Table 5 presents our main regression results. The dependent variable is an indicator for selecting the dominant option from the menu versus the alternative of selecting any one of the other plans. We run simple ordinary least squares regressions separately for each menu and regress the indicator for dominant plan choice on a dummy variable for whether the choice was made in the consequence-graph display format. Each subject made two choices for each menu (one in each format) for a total of 402 choices for 201 subjects. We account for the repeated measures by clustering the standard errors in the regression at the subject level.

Table 5. Regression Results for Graphical Display in Menus with Dominant Options

OLS Regression; Dependent variable = indicator for chose the dominant plan.

	Menu 1 4 plans, FOSD (1)	Menu 2 6 plans, FOSD (3)	Menu 3 6 plans, SOSD (5)
Graph display	0.54*** (0.04)	0.50*** (0.04)	0.25*** (0.04)
Constant	0.39*** (0.03)	0.38*** (0.03)	0.63*** (0.03)
Number choices	402	402	402
Number participants	201	201	201

Note: Standard errors clustered at the participant level.

The regression results confirm the patterns seen in the choice-distribution figures in the preceding subsections. We see increases in dominant-plan choice of 54 percentage points, 50 percentage points, and 25 percentage points across menus 1 through 3 respectively. All of these results are highly statistically significant.

We also collected an incentivized measure of insurance literacy based on whether subjects could correctly calculate the total spending a person would have with a specific total medical bill level ($10,000) under two specific plans given the cost-sharing parameters for those plans. The plans we used for this literacy test were not the same as any of the plans in the choice menus and were selected so that in one case the $10,000 in medical bills did not generate maximum out of pocket spending (due to a low coinsurance rate) and another case where it did lead to hitting the maximum out of pocket. We define those who answered both questions wrong (35% of subjects) as having low insurance literacy and create an indicator variable to denote that low insurance literacy.

Table 6 shows regression results analogous to the main regression results but now including the indicator for low insurance literacy and an interaction between graph display and low literacy. The coefficient on the low-literacy measure is negative in all regressions, implying that in the table format subjects with low insurance literacy were less likely to select the dominant option. In particular, in the more complex menus we see that those with low literacy were 19 percentage points (Menu 2) and 22 percentage points (Menu 3) less likely to choose the dominant option, both highly statistically significant results. These results are consistent with a number of recent studies that have document correlations between measured insurance literacy and competence and choice patterns, such as avoiding dominance.

The interaction between low insurance literacy and graphical display in these regressions shows that the graphical display erased the literacy gap in dominant plan choice. The interaction coefficient fully offsets the low-literacy disadvantage in menus 2 and 3 and nearly does for Menu 1, where the difference between low literacy and higher-literacy subjects is lower to begin with. This result suggests that using consequence graphs that make it easier to see the consequences of choices may have an especially big effect on those with the lowest insurance literacy.

Table 6. Regression Results for Interaction of Graph Display and Insurance Literacy

OLS Regression; Dependent variable = indicator for chose the dominant plan.

	Menu 1 4 plans, FOSD (1)	Menu 2 6 plans, FOSD (3)	Menu 3 6 plans, SOSD (5)
Graph display	0.51***	0.43***	0.18***
	(0.06)	(0.04)	(0.04)
Low literacy	-0.11*	-0.19***	-0.22***
	(0.06)	(0.06)	(0.07)
Low ins. literacy * Graph display	0.07	0.20**	0.22***
	(0.07)	(0.08)	(0.08)
Constant	0.43***	0.45***	0.70***
	(0.05)	(0.03)	(0.03)
Number choices	402	402	402
Number participants	201	201	201

Note: Standard errors clustered at the participant level. Definition of low insurance literacy: participants were given two incentivized questions asking them to calculate out-of-pocket spending for a given plan design and total medical-bill scenario. We define low insurance literacy as getting neither question right (35% of subjects).

We also asked subjects after they had made all their choices, but before they found out which plan was selected for payment, whether they preferred to make choices using the table form or the graph format. Seventy-five percent of subjects stated that they preferred to make choices using the graphs, while the other 25% preferred the tables.

Table 7. Regression Results for Interaction of Graph Display and Preference for Tables

OLS Regression; Dependent variable = indicator for chose the dominant plan.

	Menu 1 4 plans, FOSD (1)	Menu 2 6 plans, FOSD (3)	Menu 3 6 plans, SOSD (5)
Graph display	0.56***	0.55***	0.30***
	(0.05)	(0.02)	(0.04)
Prefers table display	-0.06	0.08	0.02
	(0.08)	(0.07)	(0.09)
Prefers table * graph display	-0.08	-0.21***	-0.20**
	(0.10)	(0.07)	(0.10)
Constant	0.4***	0.36***	0.62***
	(0.05)	(0.02)	(0.04)
Number choices	402	402	402
Number participants	201	201	201

Note: Standard errors clustered at the participant level. We asked participants after they made all decisions whether they preferred making choices using the table display or graph display. The variable "prefers table display" in this table is an indicator for stating a preference for tables, which was true for 25% of subjects.

Table 7 presents a regression including an indicator for preferring the table display and the interaction between stating a preference for the table display and seeing the graphical display. The small coefficients on the indicator for preferring tables imply that people who stated a preference for using tables were not systematically more likely to select dominant plans when they made choices with tables. So, based on a decision metric of selecting the dominant plan, we would not conclude that those who prefer tables are better at using them. However, the interaction term between preferring tables and the graphical-display format reveals that those who preferred tables benefited less (in terms of being more likely to select a dominant plan) by seeing the graph. For example, in Menu 2 seeing the graph increased the share selecting the dominant play by 55 percentage points for those who preferred seeing graphs. However, the effect of the graph was 21 percentage points lower for those who preferred tables. The net effect is still that those who state they prefer tables are substantially more likely to select a dominant option when they see the graphical display (34 percentage points more likely in Menu 2, for example). However, it does appear that the preference for tables might be reflecting some greater difficulty in using the graphs.

3.4 Choice when all options are rationalizable by standard theory

For our final menu of options, we move away from having a natural benchmark for choice based on economic theory and instead present a set of options where all plans could potentially be rationalized by some level of risk aversion for an expected utility maximizer. This menu was designed so that each option would be preferred by an expected-utility-of-wealth maximizer with a constant absolute risk aversion (CARA) utility function for some range of risk aversion given the distribution of medical bills Jamie faced.

Table 8. Menu 4 Table Display

Plan Name	Annual Premium	Annual Deductible	Coinsurance Rate	Maximum Out of Pocket
Orange	$1,000	$75	15%	$2,525
Red	$1,059	$100	12%	$2,050
Purple	$1,119	$125	9%	$1,575
Black	$1,179	$150	6%	$1,125
Blue	$1,238	$175	3%	$675
Brown	$1,295	$200	0%	$200

Table 8 shows the table-format of Menu 4. The plans now have a natural ordering. The consequence graph for this menu is shown in Figure 8. One can see that the lines for the plans all cross between the 8th and 9th decile of medical spending. A completely risk-neutral expected-utility-of-wealth maximizer would be predicted to select the orange plan from this menu, while the most risk averse expected-utility maximizers would be expected to select the brown plan.

Figure 8. Menu 4 Consequence Graph Display

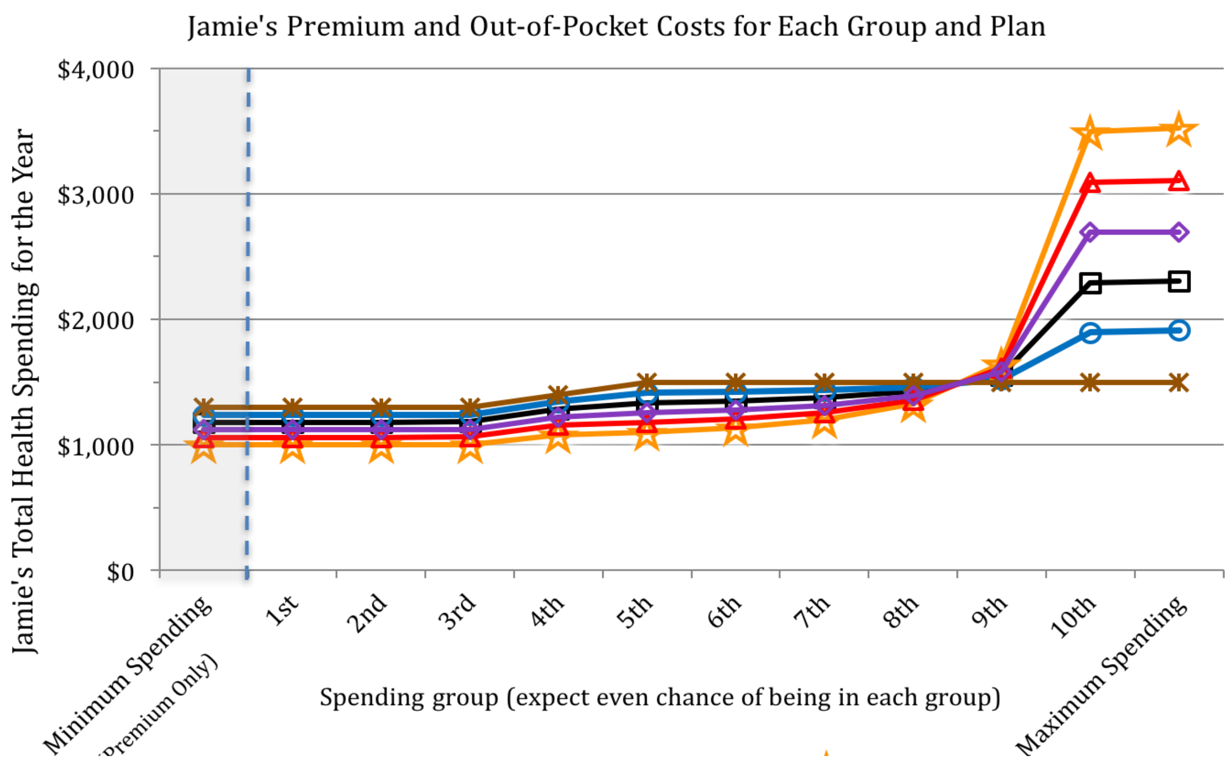
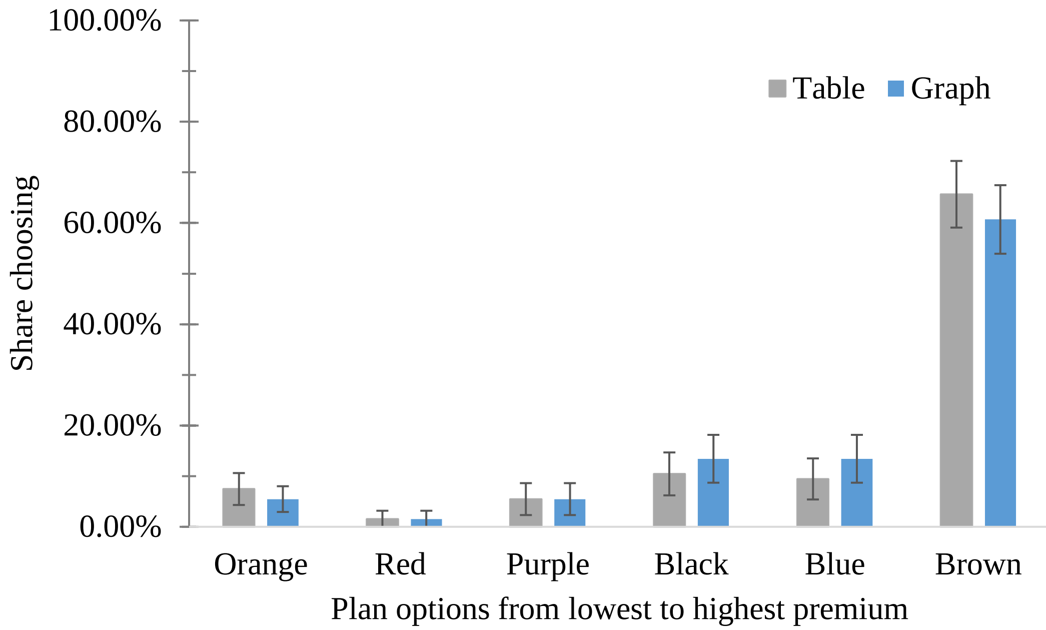


Figure 8 shows the choice shares for this menu. We see that in both display formats, the most popular choice was the Brown plan, with the highest premium and lowest variation in medical spending. In stark contrast to the menus with dominated options, we find that the choice patterns in this well-ordered menu are nearly identical between the table and graphical displays. A Mann-Whitney two-sample test for the equality of distributions has a p-value of 0.51 so that we cannot reject that the distribution of choices is the same between the formats.

Figure 9. Menu 3 Choice Shares



Although the overall choice shares are very similar for the two display formats, when we analyze the choice patterns at the individual level we find that there is actually some meaningful variation in what people chose under the different menus. Only half of subjects selected the exact same plan in Menu 4 in the two display formats. Most of the others selected a plan that was one or two spots up or down the coverage-level ranking in this menu when comparing across display formats. These variations are symmetric and may simply represent noise in the decision process.

However, the fact that there are some differences in choices across the menu formats at the individual level, raises the question of whether either of the two display formats corresponds more closely to some stable underlying preference for risk coverage. There is no obvious way to assess whether choices in one format or the other better reflect some "true preference". Yet one potentially instructive exercise is to investigate how choices from this menu correlate with a separate measure of risk aversion.

Table 9. Eckel-Grossman Risk-Elicitation Task

Gamble	Outcomes (each with 50% chance) Better	Worse	EV	Choosing
1	\$7	\$7	\$7.00	13%
2	\$6	\$9	\$7.50	18%
3	\$5	\$11	\$8.00	29%
4	\$4	\$13	\$8.50	7%
5	\$3	\$15	\$9.00	21%
6	\$1	\$18	\$9.00	12%

Notes: All subjects selected one of the six gambles in this table to be played out. The expected values are shown here but were not shown to subjects.

To that end, we elicited a measure of risk aversion for all subjects by asking them to make a choice of a gamble from a menu of options using the Eckel-Grossman task (Eckel and Grossman, 2002). Subjects selected one of 6 gambles that each had a 50% chance of having a better or worse outcome from two possible outcomes. Table 9 shows the menu of options, along with their expected value and subjects choice shares. Gamble 1 presents a sure option of \$7. The remaining gambles increase in expected value but also in risk. The sixth gamble is a mean-preserving spread of gamble 5. Subjects were informed that one out of every four subjects would be randomly selected by the computer to have this gamble choice played out for real money, in which case the realization of the 50/50 gamble outcome was randomly generated by the computer. We see here that there is substantial variation in the level of risk aversion subjects display in this gamble choice. While choices from this menu can, in theory, be mapped to a specific level of risk aversion under an assumption about the underlying utility, we prefer to remain agnostic about the specific utility function that best represents choice patterns. Instead, we use the choice from this table as a way of rank-ordering subjects by risk aversion.

Table 10. Correlation between Amount of Coverage Chosen in Menu 4 Choice and Risk Aversion Level in Eckel-Grossman Choice

Display format	Pearson correlation	p-value Pearson	Spearman correlation	p-value Spearman
Table display	0.002	0.98	-0.02	0.78
Graph display	0.16	0.02	0.12	0.08

Notes: Table shows correlation coefficients and p-values on test of correlation coefficient equal to zero. Correlations between choice in Menu 4 with Eckel-Grossman Gamble Selection where higher numbers for both imply more risk-neutral choices and lower numbers imply more risk-averse choices.

Table 10 shows how the level of risk aversion as measured in the Eckel-Grossman gamble choice (with lower gamble implying higher risk aversion) correlates with the level of coverage selected in the Menu 4 choice task (where options with higher premiums have more coverage). We see that choices subjects make from the table display show zero correlation with their risk aversion as measured by the Eckel-Grossman task. The correlation coefficients from both Person and Spearman correlations are close to zero and have high p-values. In contrast, we detect a modest and statistically significant correlation between risk aversion measured with the gamble-choice task and coverage level for the choices these same subjects made with the graphical-display format. This implies that the difference in choices subjects made when choosing from the graph display versus the table display moved them in the direction of greater correspondence with their risk aversion as measured by the gamble choice. In all cases, we see low correlation between coverage and the risk aversion measure, which suggests that perhaps the lab measure of risk aversion is not that relevant for understanding preferences over health insurance options. Yet, taken together with the evidence in the preceding sections, we believe the greater correlation between coverage choices and risk aversion in the graphical display is consistent with an interpretation that graphical display formats facilitate insurance choices that are more in line with standard expected utility-of-wealth models.

4. Conclusion

We showed that displaying health insurance options using consequence graphs instead of standard feature-based tables substantially changes how people select plan options. People are substantially less likely to violate dominance, both state-wise and second-order-stochastic dominance, when they use consequence graphs. Consequence graphs eliminate the gaps in dominant-plan choices between those with low measured insurance competence and higher insurance competence that we see when people select from tables describing plan features. We also find that 75% of subjects say they prefer to use consequence graphs and that even for those who prefer tables, they are more likely to select a dominant plan when using the consequence graph. Finally, we find a greater correlation between coverage choices and risk aversion measured with an abstract gamble when people see consequence graphs.

These results contribute to a growing body of literature exploring how poor understanding of insurance affects choice patterns in health insurance. Relative to the prior literature, we think this study provides the clearest evidence yet that insurance choices in standard display formats – where options are almost always presented as a menu of plan features – are affected by subjects having a low ability to map plan features to a distribution of final wealth consequences. If people were actually able to do that mapping, there is no reason to expect to see differences in choice patterns between the table and graphical display formats. Because standard models of insurance choice used for estimating risk aversion and selection patterns from observations of health insurance markets rest of the premise that utility is defined over the distribution of final wealth states, our findings help to illustrate a fundamental flaw in these empirical studies. Namely, unless we help people better understand the consequences of their health insurance choices, they are not making selections based on utility over the distribution of final wealth states.

On a practical level, the consequence-graph approach we introduce in this paper may provide a way of simplifying and clarifying health insurance options in settings where people are asked to select between plans with different cost-sharing tradeoffs. The primary challenge to using this approach in practice is that it requires one to use a distribution of expected medical spending, and hence the consequence graph will be specific to the individual. However, choice platforms for health insurance are increasingly providing information based on an expected distribution of spending. For example, the federal health exchanges operated as part of the Affordable Care Act in the United States provide participants with estimates of their expected out-of-pocket spending for different plan options based on distributions of medical spending linked to a few simple questions related to the person's expected medical-spending needs. The benefit of the consequence graph approach relative to approaches that provide only expected spending levels is that it allows the decision-maker to see the distribution of spending consequences they face, and hence retains and clarifies information about the amount of variance in spending and not just the average spending level.

Another potential issue with consequence graphs in real health insurance choice is that they clarify the distribution of final wealth states people face with different insurance options, but not the flow of spending they face. Although most models of insurance choice in the economics literature use the expected-utility-of-final-wealth formulation, it is unclear at this point how important the flow of spending implied by different insurance choices are to decision patterns and welfare in health insurance. If consequence graphs push people toward decisions based on final wealth states, they may actually be detrimental to welfare if their true preferences incorporate features that depend on not just the overall level of spending for the year but also the flow of that spending. This may be especially important for people with liquidity constraints, behavioral hazard that causes them to under-utilize medical services that require out-of-pocket costs (Baicker, Mullainathan, and Schwartzstein, 2015), or whose experienced utility incorporates reference dependence (Köszegi and Rabin, 2007). This is an important area for research generally on decision aids in health insurance

REFERENCES

Abaluck, J., and J. Gruber. (2011). "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program." *American Economic Review* 101 (4): 1180-1210.

Atanasov, P and T. Baker. (2014). "Putting Health Back Into Health Insurance Choice." *Medical Care Research and Review*, 71(4): 337-355.

Baicker, K., Mullainathan, S., and J. Schwartzstein. (2015). "Behavioral Hazard in Health Insurance." *Quarterly Journal of Economics*, 130(4): 1623-1667.

Bhargava, S., G. Loewenstein, and S. Benartzi. "The Costs of Poor Health Plan Choice and Prescriptions for Reform," *Behavioral Science & Policy*, forthcoming.

Bhargava, S., G. Loewenstein, and J. Sydnor (forthcoming). "Evaluating health insurance decisions: health plan choices from a menu with dominated options" *Quarterly Journal of Economics,* forthcoming.

Ericson, K. (2014). "Consumer Inertia and Firm Pricing in the Medicare Part D Prescription Drug Insurance Exchange." *American Economic Journal: Economic Policy* 6(1): 38-64.

Ericson, K., and A. Starc, (2012). "Heuristics and Heterogeneity in Health Insurance Exchanges: Evidence from the Massachusetts Connector." *American Economic Review Papers and Proceedings* 102(3): 493-497.

Frank, R.G., and K. Lamiraud. (2009). "Choice, Price Competition and Complexity in Markets for Health Insurance." *Journal of Economic Behavior and Organizaion,* 71 (2), 550-562.

Handel, B. (2013). "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts." *American Economic Review* 103(7): 2643-2682.

Handel, B., and J. Kolstad. (2015). “Health Insurance for “Humans”: Information Frictions, Plan Choice, and Consumer Welfare,” *American Economic Review* 105(8):2449-2500.

Heiss, F., D. L. McFadden, and J. K. Winter. (2010). "Mind the Gap! Consumer Perceptions and Choices of Medicare Part D Prescription Drug Plans." David A. Wise (Ed.), Research Findings in the Economics of Aging, The University of Chicago Press, 413-481.

Johnson E., R. Hassin, T. Baker, A. Bajger, and G. Treuer. (2013). “Can Consumers Make Affordable Care Affordable? The Value of Choice Architecture.” *PLoS ONE* 8(12): e81521. doi:10.1371/journal.pone.0081521.

Kling, J., S. Mullainathan, E. Shafir, L. Vermeulen, and M. Wrobel. (2012). “Comparison Friction: Experimental Evidence from Medicare Drug Plans.” *Quarterly Journal of Economics* 127 (1): 199-235.

Koszegi, B., and M. Rabin. (2007). “Reference-Dependent Risk Attitudes.” *American Economic Review* 97(4): 1047-73.

Loewenstein G, Friedman JY, McGill B, Ahmad S, Linck S, Sinkula S, Beshears J, Choi JJ, Kolstad J, Laibson D, Madrian BC, List JA, Volpp KG. (2013). “Consumers' misunderstanding of health insurance.” *Journal of Health Economics*, 32(5):850-862.

Loewenstein, G., J.Y. Friedman, B. McGill, S. Ahmad, S. Linck, S.Sinkula, J. Beshears, J.J. Choi, J. Kolstad, D. Laibson, B.C. Madrian, J. A. List, and K. G. Volpp. (2013). “Consumers’ Misunderstanding of Health Insurance.” *Journal of Health Economics* 32(5): 850-862.

Rothschild, M., and J. Stiglitz, (1976). “Equilibrium in competitive insurance markets: An essay on the economics of imperfect information.” In *Foundations of Insurance Economics* (pp. 355-375). Springer Netherlands.

Schram, A., and J. Sonnemans. (2011.) “How individuals choose health insurance: An experimental analysis.” *European Economic Review* 55(6): 799-819.

Sinaiko, A., and R. A. Hirth (2011), “Consumers, health insurance and dominated choices,” *Journal of Health Economics* 30(2): 450-457.

Winter, J., R. Balza, F. Caro, F. Heiss, B. Jun, R. Matzkin, and D. McFadden. (2006). “Medicare Prescription Drug Coverage: Consumer Information and Preference,” *Proceedings of the National Academy of Sciences*, 103(20): 7929–7934.

Zhou, C., and Y. Zhang. (2012). “The Vast Majority of Medicare Part D Beneficiaries Still Don’t Choose the Cheapest Plans That Meet Their Medication Needs.” *Health Affairs* 31(10): 2259-65.