

University of Pennsylvania Carey Law School

## Penn Law: Legal Scholarship Repository

---

Faculty Scholarship at Penn Law

---

12-5-2017

### Partitioning Sorted Sets: Overcoming Choice Overload While Maintaining Decision Quality

Benedict C.G. Dellaert  
*Erasmus University Rotterdam*

Tom Baker  
*University of Pennsylvania Carey Law School*

Eric J. Johnson  
*Columbia Business School*

Follow this and additional works at: [https://scholarship.law.upenn.edu/faculty\\_scholarship](https://scholarship.law.upenn.edu/faculty_scholarship)



Part of the [Behavioral Economics Commons](#), [Health Economics Commons](#), [Health Law and Policy Commons](#), [Insurance Commons](#), and the [Insurance Law Commons](#)

---

#### Repository Citation

Dellaert, Benedict C.G.; Baker, Tom; and Johnson, Eric J., "Partitioning Sorted Sets: Overcoming Choice Overload While Maintaining Decision Quality" (2017). *Faculty Scholarship at Penn Law*. 2741.  
[https://scholarship.law.upenn.edu/faculty\\_scholarship/2741](https://scholarship.law.upenn.edu/faculty_scholarship/2741)

This Article is brought to you for free and open access by Penn Law: Legal Scholarship Repository. It has been accepted for inclusion in Faculty Scholarship at Penn Law by an authorized administrator of Penn Law: Legal Scholarship Repository. For more information, please contact [PennlawIR@law.upenn.edu](mailto:PennlawIR@law.upenn.edu).

# **PARTITIONING SORTED SETS: OVERCOMING CHOICE OVERLOAD WHILE MAINTAINING DECISION QUALITY**

*November 22 2017*

*Working paper – ERIM – Erasmus University Rotterdam*

Benedict G.C. Dellaert

Erasmus School of Economics, Erasmus University Rotterdam,  
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands, phone: + 31 10 408 1301,  
email: [dellaert@ese.eur.nl](mailto:dellaert@ese.eur.nl)

Tom Baker

University of Pennsylvania Law School, University of Pennsylvania,  
3501 Sansom Street, Philadelphia, PA 19104, phone: (215) 746-2185,  
email: [tombaker@law.upenn.edu](mailto:tombaker@law.upenn.edu)

Eric J. Johnson

Columbia Business School, Columbia University,  
514 Uris Hall, New York, NY 10027, phone: (212) 854-5068, email: [ejj3@columbia.edu](mailto:ejj3@columbia.edu)

## **Acknowledgments**

We would like to thank Mieke van Os at Independer.nl for providing us with the data used in the field study of this research and for practical insights in the recommendation website's structure, Martijn Willemsen for feedback and support in creating the MouselabWeb iconographs, and Nuno Camacho for valuable feedback. The first author would like to thank Netspar for financial support for part of this research. The second two authors would like to thank the Alfred E. Sloan Foundation for financial support for part of the research.

## **Disclosure**

Benedict Dellaert is a member of the independent board of supervisors of Independer. Tom Baker and Eric Johnson are affiliated with Picwell

## **PARTITIONING SORTED SETS: OVERCOMING CHOICE OVERLOAD WHILE MAINTAINING DECISION QUALITY**

### **Abstract**

We investigate the joint use of partitioning and sorting as a choice architecture to overcome consumer choice overload in large product sets. Partitioning first presents a small initial set of alternatives with the option to click through to see the remaining alternatives. Sorting presents alternatives in order of attractiveness based on a user model that is helpful to the decision-maker. We propose that Sets with Partitioning and Sorting (SPSs) improve consumers' choice outcomes by increasing their focus on the most attractive alternatives and their use of more compensatory decisions. Results from two controlled survey-based experiments and a field study in the domain of health insurance support this positive impact of SPSs when sorting quality is high. However, there is also a potential harmful effect of partitioning when sorting quality is low. We discuss implications of our findings and propose a practical approach to select partitioning size depending on sorting quality.

### **Key words**

Choice architecture, Partitioning, Sorting, Decision Making, Product Recommendations

Consumers and other decision-makers are facing massive increases in the available options in many product categories, such as financial services, travel, consumer electronics, clothing, etc. Firms can easily generate many options, and large online markets make it easy to present these options to consumers. When faced with so many choices, consumers may pick worse options, a phenomenon labeled in the popular media as choice overload or the curse of choice (for excellent reviews see: Chernev, Böckenholt, and Goodman 2015; Scheibehenne, Greifeneder, and Todd 2010). This curse of choice also puts firms in a quandary. When deciding how many and which options to present to consumers, firms face a difficult balance: presenting more options holds out the possibility of a *better match* between what the decision-maker wants and what is available, but adding options also increases the decision-maker's *cost of evaluating* these options. Because decision-makers have limited time and cognitive resources, they may abandon search too quickly and/or adopt a simplified rule or heuristic ignoring information, potentially leading to suboptimal choice. Thus, providing more options may be both a blessing and a curse.

In this paper, we suggest that well-designed choice architecture (Johnson et al 2012; Thaler and Sunstein 2008) can help resolve this challenge, allowing firms and consumers to realize the benefits of increasing the number of options while minimizing the costs. This choice architecture involves an *ensemble* of choice architecture interventions that work together to maximize the quality of decisions. The ensemble concentrates decision-makers' attention on the options most likely to be attractive to them and reduces the risk that they will adopt heuristic choice processes that lead to worse choices.

While most work in choice architecture considers one intervention or another, we propose a synthesis of choice architecture tools to help overcome consumer information overload in large choice sets while preserving consumer autonomy. We know an increasing amount about the effects of single interventions like defaults on decisions, but there is much less known about how multiple interventions interact in ways that can both help or hurt decision-makers. The label for our proposed ensemble of interventions is a *Set with Partitioning and Sorting* (SPS). This ensemble consists of three interventions. The first is to partition the choices into two sets: a primary set of only the most attractive options, and a secondary set consisting of all of the rest. The second intervention is to sort the options in the set. The

third is to apply a user model to predict, stochastically, how well each option fits consumer preferences as a basis for sorting the options in the set. SPSs focus decision-makers attention on a small set of recommended alternatives but still provide them with the possibility of clicking through to see the full set of alternatives. The small number of alternatives in the primary set discourages the use of heuristics that may produce suboptimal choices and encourages compensatory decisions. Thus, SPSs represent an ensemble of choice architecture components that function in a complementary fashion to help individuals cope with information overload.

We present theoretical hypotheses and empirical evidence describing how SPSs improve choice outcomes across a broad and well-delineated set of contexts. We also show why it is important to evaluate these choice architecture interventions as an ensemble. While partitioning and sorting may be helpful, we provide evidence that their combination can also lead to worse outcomes in some circumstances, demonstrating a boundary effect that occurs when user model quality is poor. In this case, SPSs can harm consumer decision outcomes.

Applying user models to sort products in order of predicted attractiveness (based on attributes) can be a very effective way to help individuals' select more attractive products and minimize product search (Dellaert and Häubl 2012; Häubl and Trifts 2000; Xiao and Benbasat 2007). For example, researchers have investigated recommendation systems with user models of various levels of complexity and sophistication to provide consumers with comprehensive recommendations for movies (Ansari, Essegai, and Kohli 2010), music (Chung, Rust, and Wedel 2009), and garments (Lu, Xiao and Ding 2016). However, implementing sorting on its own can produce a problem: Finding the best product may still be daunting for consumers (Chernev, Böckenholt, and Goodman 2015; Iyengar and Lepper 2000; Johnson et al. 2013) leading to the use of potentially suboptimal heuristics. When choice sets are large, decision-makers adopt these simplification procedures that ignore information. For example, eliminating options based on a single attribute can lead to inferior choices. In practice, sorted choice sets can indeed be very large, consisting of 50 different alternatives or more (e.g., [www.ehealthmedicareplans.com](http://www.ehealthmedicareplans.com) or [www.moneywise.co.uk/compare-pension-funds](http://www.moneywise.co.uk/compare-pension-funds)). Research shows that the use of heuristics can,

paradoxically, even cause consumers to search too much in sorted lists, leading to the erroneous selection of products further down the list and lower in quality (Diehl 2005).

SPSs improve on sorting by predicted attractiveness alone, adding a partitioning that focuses the decision-maker on the most highly recommended products (which we call the “recommended set”), while allowing the option to click through to see the complete list of all available alternatives (i.e., the “full set”). The previously studied practice of sorting by predicted attractiveness based on a user model increases the probability that the best alternatives are considered by the decision-maker. Adding the partitioning limits the number of options that are first presented to the decision-maker, encouraging the use of better trade-off based strategies to choose among those options. At the same time, SPSs preserve decision-makers’ autonomy by providing them with the ability to pick any available option.

Decision processes vary in both the effort required to use them and the accuracy that they provide (Payne, Bettman and Johnson 1993). SPSs focus the decision-maker’s limited cognitive resources on the best subset of alternatives, lowering individuals’ search costs when choosing from large sets, while also improving the quality of their decision outcome. This focus may backfire, however, if the sorting quality is too low (i.e., user model quality is low) and the partitioning is too small. In that case, the best alternatives may fall outside the partitioning, and consumers are likely to fail to search beyond the recommended set. The impact of the quality of the user model on the effectiveness of the partitioning illustrates how important it is to evaluate choice architecture tools as an ensemble, whenever there is a possibility that the effectiveness of one choice architecture tool depends on how another tool is implemented. Indeed, identifying and empirically demonstrating the importance of this ensemble approach is one of the main contributions of the research that we report here.

Our paper is structured as follows. We first develop theory and hypotheses describing the impact of SPSs on choice outcomes. We next examine these hypotheses in two online survey based choice experiments and a field study, all in the domain of consumer health insurance product choices. We selected the domain of health insurance choices because it has a strong practical impact on consumer welfare, and recent developments in the U.S. and elsewhere have strongly increased the opportunities

(and corresponding decision-making burden) for consumers to make their own health insurance decisions (Bhargava, Loewenstein, and Sydnor 2017; Ericson and Starc 2012; Johnson, et al. 2013; Ubel, Comerford, and Johnson 2015). Similar market developments can be seen in other product categories where online aggregators provide consumers with access to a huge array of different products (e.g., travel, investments, music, and clothing). We close with a discussion of how partitioning and sorting jointly affect consumer decision outcomes and suggest a practical approach that firms can take to select an optimal partitioning size depending on sorting quality.

### *THE IMPACT OF PARTITIONING AND SORTING ON CHOICE*

SPSs apply a user model to present decision-makers with sets of options that are sorted in order of their predicted attractiveness to the decision-maker and partitioned so that the decision-maker sees a recommended subset that includes only the most highly recommended products, with the option to choose to also see the full set of products. By partitioning the sorted set, the initial inspection for the recommended set (which is available at first) is marginally less costly (in terms of effort) to the decision-maker than inspecting the full set (which requires an additional information request, such as a mouse click). In this section, we theorize and provide hypotheses on why and when SPSs may improve consumer choice outcomes.

Decisions in SPSs represent a special case of the more general product search paradigm. Most normative models of product search (related to the “secretary problem”) have the reservation value property (e.g., Adam 2001; Rosenfield, Shapiro, and Butler 1983; Weitzman 1979). The reservation value property implies that solving the global search optimization problem (how many and which alternatives to inspect) is equivalent to solving a simpler problem. Specifically, decision-makers can optimize their product selection by a sequential procedure: If they determine, as they search alternatives, if the value of the most attractive alternative inspected so far exceeds a certain reservation value, their choice outcome

will be equivalent to a global optimization in which they determine (fully forward-looking) if the cost of inspecting another product is higher than the expected gain of doing so. When search is stopped, the decision-maker chooses the alternative that has the highest value among those already inspected. Note that, because partitioning and sorting affect the distribution of alternatives presented in a set, they will affect the returns to searching options. In the next sections we discuss their impact and propose specific hypotheses on how the joint use of partitioning and sorting affects consumer decision making under different levels of user model quality.

### *The impact of partitioning*

When consumers search for the best alternative in a SPS they need to decide whether to only inspect the initial smaller recommended set, or to also look for products in the full set. Normatively, because the user model ranks the (expected) most attractive alternatives the highest, consumers should first inspect the recommended set that is defined by the partitioning (Moorthy, Ratchford and Talukdar 1993). When it comes to the decision of whether or not to continue the search beyond the recommended set and inspect the full list, it is important to note that while consumers can easily scan the (distribution of) products in the recommended set, they –due to the partitioning– must incur some additional search cost to see the remainder of the full set. The uncertainty of the sorting in the unobserved part of the full set is inherently greater than in the observed partitioned part (because the full set cannot be directly observed). Therefore, normatively, under the plausible condition that the additional costs for continuing the search with one click into the full set are negligible compared to the possible returns, consumers should continue to inspect the full set if no product choice has as yet been made, rather than end the search. The reason is that even when a sorted list is partitioned, consumers should gain from further search beyond the recommended set if they have not found a sufficiently attractive alternative before.

However, in contrast to the normative result, we anticipate that decision-makers are myopic with respect to the expected positive returns to search of drawing from a more uncertain distribution (i.e., the full set) and instead will focus on the local, perceptually prominent, higher cost of having to click through



to the next screen (Häubl, Dellaert and Donkers 2010; Hogarth and Einhorn 1992; Wilson, et. al. 2000). Therefore, we propose that -compared to when consumers see a full set directly- fewer consumers will inspect products in positions further down the list when they initially only see the recommended set. We hypothesize below how this affects consumer choice outcomes.

### *The impact of sorting*

Sorting alternatives in order of predicted attractiveness to the decision-maker differs from the usual search paradigm in which the order of alternatives contains no information about their attractiveness. To see this, imagine that the alternatives in a sorted list are perfectly ordered in line with a decision-maker's preferences. In this case, the decision-maker only needs to inspect the first alternative. As the quality of the user model (in terms of sorting to the decision-maker's underlying utility) decreases, the number of options that the decision-maker should examine increases. More specifically, consumers should continue their search in a sorted list if the expected improvement in utility after inspecting the next alternative is greater than the cost of inspecting the next alternative (i.e., the search cost). This implies that when sorting quality goes down (i.e., as the user model error increases), consumers should be expected to search more, because the probability of finding a better alternative further down the list increases. Thus, the higher the sorting quality, the less consumers should search in the list and also the better their choice outcomes should be (Häubl and Trifts 2000).

Sorting has been shown to influence consumer choice outcomes, but not to entirely determine them. Two important patterns emerge from the research. First, paradoxically, even when sorting quality is high, consumers have been found to search too much in sorted lists, leading to the erroneous selection of products further down the list that are lower in quality (Diehl 2005). Second, when sorting quality is low, consumers may resort to (screening) heuristics in their search. This may be aggravated by specific types of sorting. For example, consider the case of sorting by only one attribute (such as price) that is only partially aligned with consumer preferences. Doing so can make that attribute easier to process, leading to the attribute receiving greater weight, because decision-makers screen on that attribute (Lynch and Ariely

2000). In decisions that entail tradeoffs, these attribute-based heuristics can result in lower quality choices.

*Hypotheses: How SPSs affect decision-making*

Based on these impacts of partitioning and sorting, we now develop hypotheses regarding the effect of SPSs on consumer decision outcomes. First, we propose that partitioning in a sorted list creates a discontinuity at the end of the partitioned set that encourages the consumer to end the search at that point. (Note that SPSs impose only a “soft” limit on the possibility of further search.) This property reduces the probability that consumers will over-search in sorted lists. Therefore, we propose that by lowering over-search, SPSs will lead consumers to *focus* their attention more strongly on the best alternatives in the sorted list.

Second, by reducing the number of options inspected, partitioning also facilitates a more thorough evaluation of the alternatives in the recommended set, which may not be cognitively feasible in a search of the full set (Besedeš et al 2015; Hanoch et al 2011; Hauser and Wernerfelt 1990; Roberts and Lattin 1991). Typically, when consumers are faced with a large choice set, a more in-depth comparison only takes place for a small subset of products after a less detailed evaluation of the larger set of alternatives (Payne 1976; Chakravarti, Janiszewski, and Ülkümen 2006; Ge, Häubl, and Elrod 2012). Therefore, we propose that presenting alternatives directly in a smaller recommended set, such as in a SPS, will increase the probability that consumers will follow a *compensatory decision strategy* (Johnson and Meyer 1984; Payne 1976; Payne, Bettman, and Johnson 1993; Swait and Adamowicz 2001). More specifically, Johnson and Meyer (1984) showed that with increasing complexity, weight is shifted away from the least important attributes toward attributes that are easily used for screening, reflecting a less compensatory choice strategy. In the context of random utility choice models, researchers have observed similar shifts in parameter preferences in discrete choice models due to shifts in decision complexity (Dellaert, Donkers and van Soest 2012; Swait and Adamowicz 2001).

Thus, the combined effects of partitioning and high-quality sorting lead us to propose that SPSs transform consumer product selection decisions in large sets such that consumers will use a richer, more compensatory decision process that is also more focused on the most highly ranked alternatives, compared to when they are presented with the full large and unsorted set directly. Jointly, this increases the probability that consumers will select the best alternative from a large set.

H1: SPSs increase consumer choice of the best alternative in large sets.

However, if the quality of the sorting, is poor, the proposed impacts of partitioning and sorting may also constitute a limiting condition on the beneficial effect of SPSs. In particular, the benefit of focusing consumers' decisions on only the most highly recommended alternatives depends on the quality of the predictive user model. If the user model has a zero or negative correlation with the consumer's actual preferences, partitioning the sorted list may encourage consumers to focus on less attractive alternatives at the top of the list and hence discourage them from inspecting other more attractive alternatives which are not in the recommended set. In this case, consumers will be better off if they are presented with the full set directly. This is the basis of our second hypothesis:

H2: The positive impact of partitioning in SPSs is moderated by sorting quality: When sorting quality is low, partitioning decreases consumer choice of the best alternative.

In addition to the effects of partitioning and sorting on choice, we also examine, in a controlled online study (Study 2), the nature of the changes in the consumer decision process that occur when an SPS choice architecture is introduced. Recall that when large sets are partitioned, the total set of alternatives is divided into two, one small set that is directly available on screen (the recommended set), and the other (the full set) that is just a click away. Even when incentives for more extensive search are present, we argue that the first set will receive the most attention. When sorting quality is low, the best

alternative may well be in the second set, but because of limited search beyond the recommended set, it may not be examined. Because our Study 2 allows us to observe decision-makers' distribution of attention in their search, we offer H3:

H3: Partitioning increases consumers' focus on the best alternatives when sorting quality is high, but reduces it when sorting quality is low.

The other impact of partitioning that we hypothesize is that there will be a shift to the use of compensatory decision strategies making more extensive tradeoffs between attributes. This shift in decision process is triggered by fact that there are fewer alternatives in the recommended set. It mirrors the difference in decision strategies in larger and smaller choice sets that has been found in many studies in decision-making starting with Payne (1976) (see also Payne, Bettman, and Johnson 1993), but this switch is not commonly anticipated in research on search (see Dellaert and Häubl 2012 for an exception).

H4: Partitioning increases consumers' use of compensatory decision strategies.

Our central hypothesis (H1) is that SPSs will help produce better choice outcomes when sorting quality is high, that is when the user model is strongly correlated with the decision-maker's utility. As hypothesized in H3 and H4, we expect this positive impact to occur because of consumers' increased focus on the best alternatives, and a shift to using more compensatory decision strategies. However, if sorting quality is low, we expect partitioning to lead to worse choice outcomes (H2). The reason is that lower sorting quality will lead consumers to focus on alternatives that are inaccurately predicted as being the most attractive. Thus, the moderating effect of sorting quality is particularly important to understanding the impact of partitioning.

We argue that the changes in choice outcomes are due to shifts in decision processes. Therefore we hypothesize that both the direct effect of partitioning and the moderation of this effect by sorting

quality (H3) is mediated by consumers' focus on the best alternatives, while the direct effect of partitioning (H4) is mediated by consumers' use of compensatory decision strategies. In other words, we expect mediated moderation for consumers' focus on the best alternatives, and regular mediation for consumer's use of compensatory decision strategies:

H5a: The effect of partitioning on consumer choice of the best alternative and its moderation by sorting quality, are mediated by consumers' focus on the best alternatives.

H5b: The effect of partitioning on consumer choice of the best alternative is mediated by consumers' use of compensatory decision strategies.

In the following sections, we examine these hypotheses in detail using data from two online survey based experiments and a field study. Study 1 provides support for H1 and H2 in a controlled hypothetical online choice environment. Study 2 replicates the results of Study 1 in a consequential online choice experiment. In addition, it provides a process-level analysis of the proposed impact of SPSs on consumer decision making using a MouselabWEB interface. This allows us to also test H3, H4, and H5a,b. Study 3, finally, presents the results of a quasi-experimental field study in which a combined partitioning and high quality sorting was introduced, and that provides a real-world validation of the hypothesized positive impact of SPSs (H1) on consumer choice outcomes.

### *STUDY 1: DO SPSs IMPROVE CONSUMER CHOICE?*

In Study 1 we conducted an experiment in a controlled hypothetical choice environment to investigate the hypothesized beneficial effect of SPSs on consumer choice (H1) and its reversal in the

case of poor sorting quality (H2). We simulated, in a simplified and more controlled setting, a typical consumer decision task as presented by health insurance exchange websites allowing consumers to compare and purchase health insurance plans. The website we used as the basis for constructing the experiments is based in The Netherlands (where our data was collected); and is similar in many respects to Healthcare.gov and the other health insurance exchange websites established under the Affordable Care Act in the U.S. (Wong et al 2016). We manipulated two key components of SPSs independently from one another: (1) The absence or presence of partitioning and (2) The quality of the sorting criterion on which products were ranked in the list (low, medium and high). Note that in our manipulations we collapsed sorting and user model quality into one dimension. The reason is these two choice architecture tools work so closely in tandem: When user model quality is low, the presence of sorting becomes indiscernible, while when user model quality is high, sorting is clearly present. Participants were from a nationally representative on-line panel in the Netherlands. The experiment employed a principle-agent task that instructed participants which trade-off to optimize in selecting an alternative from the list. This eliminated unobserved preference heterogeneity between participants from the decisions, so that we could determine the quality of the decision that each participant made.

### *Method*

Participants were randomly assigned to one of the six treatment conditions of a two (top three partitioning vs. no partitioning) by three (high quality sorting, medium quality sorting, low quality sorting) between-subjects design. In all conditions participants were presented with a list of the same 10 health insurance plans, but in different partitioning and sorting formats. Participants were asked to make a choice from the health insurance plans presented to them as if they were choosing a plan on behalf of a person they knew well, but that was not part of their own household (e.g., an elderly aunt or uncle). They were told that this person desired a minimum (well-defined) level of coverage, and otherwise to choose

the best buy, defined as the highest ratio of the monthly premium paid and the review ratings that customers gave for the insurance provider. The plans reflected a realistic product set in the participants' health insurance market (The Netherlands). The plans were based on existing options and participants could click on a help link to see a short explanation of each feature. Fictitious brand names were employed. Minor modifications were made to prevent dominance and facilitate understanding.

Two partitioning conditions were used. The partitioned condition showed participants the first three recommended products with a possibility to click through to the full list. The full list condition showed the sorted list of all recommended products directly to participants (see Web Appendix A for screen shots of the two partitioning conditions). We used three sorting conditions: In the high-quality sorting condition, alternatives were sorted almost perfectly aligned with the stated decision criterion. The only deviation was that alternatives three and four in the ranking were inverted. The medium-quality sorting condition was identical to that of the high-quality sorting condition but with the first and third alternatives swapped. This second condition allowed to us to observe if presenting consumers with a top three partitioning could help increase participants' ability to select the best alternative even when sorting was noisy. The low-quality sorting condition randomized the options with two exceptions. First, the third ranked alternative was kept identical to that in high quality sorting condition to be able to more clearly compare consumer choices between the three conditions using this alternative as a common base alternative. Second, the most attractive alternative was placed at the fifth rank position, so that participants would have to search beyond the partitioning to find the best alternative in the list, but did not need to inspect the full list.

For the study participants of over 18 years old were recruited through a large scale online panel run by a marketing research company in the Netherlands. In study 1, we predetermined the sample size to be 250 per treatment condition, which is twice that used in a similar paradigm (Johnson et al. 2013). Similar sample sizes were used throughout. A total of 1,577 valid responses were obtained (assignment to treatment conditions was randomized and data collection was stopped by the marketing research firm after all conditions contained at least 250 observations). The average age of the participants in the sample

was 45.5 years old, 51.1% of the participants were female, 42.2 % had obtained a bachelor or higher degree, and 45.3% had bought health insurance for themselves or for someone else in their household in the past two years.

### *Results*

We tested the effect of SPSs (H1 and H2) using as a dependent variable whether or not participants selected the best product in each condition. We also analyzed a second, complementary, dependent variable that was a measure of how much a participant overpaid compared to the best insurance product by selecting an alternative that was not optimal in terms of the objective decision criterion.<sup>1</sup> Note that the best insurance product was specified by the decision rule communicated to participants.

- INSERT FIGURE 1 ABOUT HERE -

Figure 1 presents both dependent measures, the percentage of consumers choosing the best product in each condition (the upper bars in the figure), and the amount they overpaid in euros (the lower bars in the figure). The figure shows that the quality of sorting clearly made a difference, with lower quality sorting producing worse performance in both the percent of correct choices and in the amount overpaid. Similarly, partitioning helps, but with one very important exception. Partitioning helps when the sorting quality is high or medium, but it hurts performance when the sorting quality is low. Note that also in the medium quality sorting condition many participants selected the best alternative (though fewer than in the high-quality condition), and they were more likely to do so in when the list was partitioned. This pattern of results is consistent with H1 and H2.

---

<sup>1</sup> We obtained this measure by first calculating the ratio of price to consumer review score for each product. Next, using this ratio, we calculated a virtual yearly price for each product corresponding to the consumer review score of the best available product. The difference between this virtual yearly price for the selected product and the actual yearly price for the optimal product was used in our analysis of the amount a consumer overpaid.



We confirmed this pattern using a logistic regression analysis that modelled the probability that participants selected the best product vs. others, analyzing the effects of partitioning, sorting quality and their interaction on consumer choice outcomes. Sorting quality level was represented with an indicator variable with low quality sorting as the base level. In support of H1 and H2, the results support the hypothesized positive impact of SPSs when sorting quality is of high or medium quality (H1), as well as the hypothesized negative impact of partitioning when sorting quality is low (H2). More specifically, we find a positive main effect of greater sorting quality ( $\beta_{\text{HIGH}} = .44$ ,  $p < 0.05$ ;  $\beta_{\text{MEDIUM}} = .10$ , n.s). The negative main effect of partitioning ( $\beta_{\text{PARTITION}} = -2.67$ ,  $p < 0.001$ ) may initially seem surprising but it is driven by the negative effect of partitioning in the low-quality sorting condition, and more than compensated in the high and medium quality sorting conditions by the large and significant interaction effect between sorting quality level and partitioning ( $\beta_{\text{HIGH*PARTITION}} = 2.94$ ,  $p < 0.001$ ;  $\beta_{\text{MEDIUM*PARTITION}} = 3.05$ ,  $p < 0.001$ ). These results reflect the fact that with high and medium quality sorting, partitioning makes it more likely that consumers choose the best product, but with low quality sorting, partitioning makes it less likely that they do so. We also estimated the logistic regression model while controlling for participants' experience with buying health insurance in the past two years (yes/no). The effect of this control variable was significant and positive ( $\beta = .24$ ,  $p < 0.05$ ), but adding it in the model did not affect the significance of the other results.

We also conducted an analysis of the average loss per condition with an ANOVA of the two partitioning levels by the three sorting quality levels and their interactions. The results again strongly support the hypothesized effects. There is a strong significant effect of sorting in the predicted direction ( $M_{\text{HIGH}} = 13.2$ ;  $M_{\text{MEDIUM}} = 14.0$ ;  $M_{\text{LOW}} = 42.6$ ;  $F(2, 1571) = 69.7$ ;  $p < 0.001$ ), with lower losses in case of higher sorting quality. We find an opposing main effect of partitioning ( $M_{\text{NOT-PARTITIONED}} = 20.1$ ;  $M_{\text{PARTITIONED}} = 27.2$ ;  $F(1, 1571) = 6.6$ ;  $p < 0.05$ ), which again is compensated by the hypothesized large interaction effect of sorting and partitioning ( $F(2, 1571) = 44.2$ ;  $p < 0.001$ ). This interaction reflects the

detrimental effect of partitioning in the low quality condition, which in the main effect dominates the positive impact of sorting in the high and medium quality conditions.<sup>2</sup>

An additional analysis of the simple effects of partitioning within each sorting condition reveals the significance of these contrasting effects per condition and in the expected direction. As hypothesized in H1 and H2, we find that the impact of partitioning is significantly positive in the high and medium quality conditions ( $p < 0.05$  for each comparison), but significantly negative in the low quality conditions ( $p < 0.001$ ).

Finally, we explored two process measure proxies to obtain an initial understanding if the underlying consumer decision making was shifting in the expected direction. As a first proxy for how extensive participants' evaluation of the alternatives was, we divided the total time spent in the experiment by the number of alternatives that was visible to participant (taking into account whether or not the participant in the partitioned condition had clicked through to see the full list). We find that, as expected, in the partitioned condition participants spent significantly more time per alternative ( $M_{\text{PARTITIONED}} = 72.9$ ;  $M_{\text{NOT-PARTITIONED}} = 24.1$ ;  $t = -28.9$ ;  $p < 0.001$ ). As a second proxy, we looked at how many participants clicked through to look at products beyond the recommended set in the partitioned conditions. We find that only 10.4% of the participant made use of this possibility, and that there was no significant difference between the three levels of sorting quality ( $M_{\text{HIGH}} = 8.9\%$ ;  $M_{\text{MEDIUM}} = 10.7\%$ ;  $M_{\text{LOW}} = 11.3\%$ ;  $p > .65$ ). This confirms our expectation that consumers have a strong focus on the top three alternatives when partitioning is present, and that this lowers their focus on the best alternatives when sorting quality is low.

---

<sup>2</sup> We also conducted the ANOVA while controlling for participants' experience with buying health insurance in the past two years (yes/no). The effect of this control variable was not significant and adding it in the model did not affect the significance of the other results.

### *Discussion*

The results of Study 1 provide support for the hypothesized impact of SPSs on consumer choice in a normative setting with objectively different quality levels between different choice outcomes. The results as shown in Figure 1 and in the outcome of our analyses strongly support both H1 and H2. We find that with good and medium quality sorting, partitioning is beneficial for consumers, but that with (close to random) low quality sorting, partitioning is not beneficial and may even be harmful to consumer choice outcomes. This finding underlines the fact that, to be helpful to consumers, partitioning requires good quality sorting based on overall product attractiveness to the consumer.

### ***STUDY 2: UNDERSTANDING THE PROCESS UNDERLYING THE IMPACT OF SPSs***

Study 2 expands Study 1 in two important ways. First, in Study 2 we investigated how consumers' decision processes change with partitioning, sorting, and their interaction. We used a MouselabWEB interface that allowed us to directly observe which product attribute levels participants inspected and in which order during their decision process, a technique that has been shown to be useful in the study of individual choice as well as games and time preferences (Gabaix, et al. 2006; Payne Bettman and Johnson 1993; Willemsen, Böckenholt and Johnson 2011; Willemsen and Johnson 2011). Second, we conducted the experiment for Study 2 in a consequential setting. As in Study 1, participants were asked to follow a clearly defined decision rule on behalf of someone else, but now they received a monetary reward based on their performance in the task. This design allowed us to objectively determine the quality of participants' decision outcomes in a task in which participants also experienced monetary consequences of making lower quality choices. In addition, the experiment for Study 2 was conducted in

the U.S. with participants who were presented with an online health insurance choice, which provides for a generalization of the findings to a different cultural and health care system setting.

### *Method*

We employed a principle-agent task that instructed participants which trade-off to optimize on behalf of another individual when selecting a health insurance alternative from a recommendation list. Participants received a monetary reward that depended on how well they performed in this selection task. Participants were randomly assigned to one of four treatment conditions based on a two (top three partitioning vs. no partitioning) by two (high quality sorting vs. low quality sorting) between-subjects design. All conditions presented the same eight health insurance plans. The choice task was repeated once for each participant, within the same experimental condition, but with different attribute values.

Choice sets were either partitioned or not. The partitioned condition showed participants the first three recommended products with the possibility to click through to the full list. The non-partitioned, full set condition showed the sorted list of all recommended products directly to participants (see Web Appendix A for screen shots of the two partitioning conditions). Sorting was manipulated using a high- and low-quality sorting. In the high-quality sorting, alternatives were sorted nearly perfectly in line with the participants' decision criterion. The only deviation was that the first and third alternatives switched. The low-quality sorting presented a randomized order of alternatives to participants with one exception, which was that the best alternative was placed at the fourth ranked position.

Health insurance plans were described in terms of three key characteristics for the U.S. market (monthly premium, doctor visit copayment, and annual deductible) that were also successfully used in earlier research in this domain and market (Johnson, et al. 2013). The plans reflected a realistic product set in the participants' health insurance market (USA). Participants were also given the relevant health care needs of the person for whom they were making the choice. This information allowed them to

calculate the annual expected costs of each of the health insurance products presented in the recommendation list.

To obtain deeper insights into participants' decision process while selecting an alternative from the list, we implemented the experiment in a MouselabWEB online interface (Willemsen and Johnson 2011). MouselabWEB allows researchers to trace the content and sequence of decision processes. Participants viewed a table in which each row represented a different health plan and each column displayed a different attribute of the plans. This format corresponds with the recommendation lists presented to participants in our research. In the MouselabWEB interface, participants use their mouse to move the pointer on the screen across the different boxes in the table representing the different attributes of each alternative. When the pointer hovers over a box, the level (value) of the respective attribute is revealed, and when the pointer is moved, the information is hidden. This process reveals how often, for how long, and in which order, participants examine the information about the attributes of each plan. Past research suggests there is a close correlation between MouselabWEB observations and eye-tracking data, which can be an alternative way of tracing participants' decision processes (Lohse and Johnson 1996; Reisen et al. 2008). These data provide deep insight into the process behind the participant's decision making. We use the MouselabWEB data to investigate differences in participants' decision processes depending on whether or not they are presented with SPSs. These differences can reveal if the proposed shifts in decision making process leading up to our hypotheses explain why partitioning sorted lists may or may not lead to improved consumer decision outcomes.

For Study 2 participants were recruited through Amazon MTurk. A total of 846 responses were obtained, of which 18 outliers were eliminated because they had made extraordinarily few (less than 4) or many (more than 600) information acquisitions during the choice process. The average age of the participants in the sample was 34.5 years old, 43.6% were female, and 50.3% had obtained a bachelors or higher degree.

In the next sections, we discuss the results of Study 2. We first look at the quality of the participant choice outcomes, then we describe the type of processing used by participants, and finally we relate the two by testing for mediated moderation of the decision process on choice outcome. Recall that we hypothesized that SPSs increase consumer choice of the best alternative (H1), but that the positive effect of partitioning is attenuated when sorting quality is low, in which case partitioning reduces consumer choice of the best alternative (H2). In addition we hypothesized that these effects of partitioning and sorting are driven by consumers' focus on the best alternatives (H3) and their use of a more compensatory decision making strategy (H4), and more specifically that these two decision making process shifts mediate the direct and moderated effect (by sorting) of partitioning on consumer choice of the best alternative (H5a,b).

#### *Results – Choice outcomes*

To test the hypothesized direct effects of introducing SPSs on consumer choice outcomes, we again analyzed (1) whether or not the participants selected the best alternative in each condition, and the complementary dependent variable of (2) how much each participant overpaid as a result of selecting a sub-optimal product in each condition. Both variables reflected the best insurance product as specified by the decision rule communicated to participants (Johnson et al. 2013).

Figure 2 presents the percentage of consumers choosing the best product in each condition (the upper bars in the figure), and the amount overpaid in dollars (the lower bars in the figure), both averaged over the two choices per participant. As in Study 1, the quality of sorting clearly made a difference, with lower quality sorting leading to a worse performance in both the probability of a correct choice and in the amount overpaid by participants. We also see that partitioning helps when the sorting quality is high, and that it hurts participants' performance when sorting quality is low.

- INSERT FIGURE 2 ABOUT HERE -

We also conducted a statistical analysis of these effects. First, we estimated a random effects logit model for making the best choice (yes, no) and analyzed the effects of partitioning, sorting quality, and their interaction. The results support the hypothesized positive impact of partitioning when sorting quality is high, but not when sorting quality is low (see Web Appendix B for detailed estimation results). More specifically, we find a positive main effect of greater sorting quality ( $\beta_{\text{HIGH}} = 1.15, p < .001$ ) and a negative main effect of partitioning ( $\beta_{\text{PARTITION}} = -.77, p < .001$ ) which is strongly compensated by the significant interaction between sorting quality level and partitioning ( $\beta_{\text{HIGH*PARTITION}} = 1.09, p < .001$ ). Thus, we find, as in Study 1, that partitioning increases the probability that consumers choose the best product when sorting quality is high and that partitioning decreases this probability when sorting quality is low. This provides further support for H1 and H2. We also estimated a random effects regression model of the (higher) amount participants paid compared to the best insurance product. There is a strong significant effect of sorting in the predicted direction, with high quality sorting lowering the amount that participants overpaid compared to low quality sorting ( $\beta_{\text{HIGH}} = -189.14, p < .001$ ). We find a main effect of partitioning ( $\beta_{\text{PARTITION}} = 94.00, p < .001$ ), which is more than compensated by the hypothesized interaction of sorting and partitioning that reflects that consumers overpay least in the partitioned high sorting quality condition ( $\beta_{\text{HIGH*PARTITION}} = -182.94, p < .001$ ). These results also support H1 and H2.

#### *Process data*

Next, we studied the process level data obtained from the MouselabWEB interface. This interface allowed us to directly observe which product attribute levels participants inspected during their decision process and in which order they inspected them. Figure 3 summarizes these data graphically in four icon graphs (Willemsen and Johnson 2011), each representing one of the four experimental conditions. The icon graphs summarize multiple aspects of participants' information acquisition processes. The boxes in the graphs correspond to the information acquisition for the attribute cells in the table that participants

viewed. The horizontal length of each box represents the average time participants spent inspecting the attribute level in that cell, while the vertical height represents the average number of times participants inspected the cell. The length of the arrows between the boxes represents the number of transitions participants made in their information acquisition process. Results are averaged across the two repeated choice tasks. Following standard practice (Willemsen and Johnson 2011), we eliminated all information acquisition observations of 200ms or less, because they are too short to be read by the participants and probably reflect movements of the pointer from one cell to the next.

It is apparent from the figure that partitioning leads to considerably more inspections of the attributes in the recommended set (i.e., a focus on the top three alternatives presented in the partitioned conditions) as compared to the alternatives that were not in the recommended set. This should be beneficial when sorting quality is high, but not when sorting quality is low, and reflects the earlier estimation results that supported the hypothesized interaction effect between sorting and partitioning. In addition, the figure suggests that compared to the other conditions, the least attribute-based (i.e., vertical) comparisons are made in the partitioned, high quality condition. This is reflected in the lower number of vertical arrows in this condition. As explained in more detail below, this finding suggests that in line with our expectations, partitioning triggered a more compensatory decision process, which can increase decision quality if sorting quality is sufficiently high. In the next section, we conduct a statistical analysis of two key process measures we obtained from this data, and if these process measures mediate the impact of partitioning and sorting on consumer choice outcomes (H3, H4, and H5a,b).

- INSERT FIGURE 3 ABOUT HERE -

### *Mediation analysis*

We hypothesized two complementary decision processes to explain the positive impact of the combination of partitioning and sorting. First, H3 hypothesized that when sorting quality is high, partitioning may produce a decision process that *focuses* participants' attention more strongly on the best



alternatives (but not when sorting quality is low). This increased focus should be reflected in participants spending more time on inspecting the best alternatives in the list (compared to other alternatives). This in turn, should increase participant decision quality. In contrast, when sorting quality is low, partitioning will lead participants to focus away from the best alternatives (as they will not (or only partially) be presented in the recommended set). Thus, we anticipate that the impact of partitioning on choice outcome via focus is moderated by sorting quality (H5a). Second, in H4 we proposed that partitioning may lead to higher quality, more *compensatory decision-making* processes, because consumers tend to focus on a smaller number of alternatives. This shift in the decision-making process should also increase the quality of the decision outcome as participants are better able to evaluate a smaller number of alternatives (H5b). Figure 4 reflects the proposed dual pathway mediated moderation structure.

- INSERT FIGURE 4 ABOUT HERE -

Using the MouselabWEB data, we operationalized the two mediating process variables as follows: 1. *Focus* was operationalized as the total time spent per participant on inspecting the best three alternatives in the list, and 2. *Compensatory decision making* was operationalized using the Payne index measure (Payne 1976). The Payne index captures if a participant's information acquisition process is more alternative-based or attribute-based. Alternative-based information acquisition is reflective of a more compensatory decision making process, while attribute-based information acquisition is reflective of a more screening oriented, heuristic decision making process. The Payne index ( $PI$ ) takes the ratio of the difference between the number of alternative-based (horizontal) acquisition steps ( $NrALT$ ) and the number of attribute-based acquisition steps ( $NrATT$ ) and the sum of these two types of information acquisition steps, all from the MouselabWEB data ( $PI = (NrALT - NrATT)/(NrALT + NrATT)$ ). A score of 1 represents a completely alternative-based decision process, and a score of -1 represents a completely attribute-based search.

To test the hypothesized mediated moderation structure on choice outcome (H3, H4, and H5a,b), we conducted our analyses using Hayes' PROCESS module in SPSS (Hayes 2013). We analyzed as our key dependent variable whether or not participants had made the best choice. Partitioning (yes, no) and sorting quality (high, low) and their interaction were the main independent variables. The new mediating variables *focus* (i.e., time spent on best options) and *compensatory decision making* (i.e., the Payne index) were also included, as was the anticipated moderating effect of sorting on the effect of partitioning on focus. In the model estimation, we used data on the two observations per participant, and included an order variable as covariate.

#### *Results mediation analysis.*

The results of the analyses are presented in Table 1. Following the procedure outlined in Zhao, Lynch and Chen (2010), the table first reports the unstandardized coefficients for the different pathways in the model. These include the anticipated effects of partitioning and sorting on the mediating variables focus and compensatory decision making (H3 and H4). In the table these results are followed by the 95% confidence intervals from the bootstrap analysis. The confidence intervals represent the effects of the mediators and of the moderating effect of sorting quality on the focus mediator with respect to the dependent variables. Note that if these confidence intervals are significantly different from zero, there is support for a mediating process, and hence for H5a and H5b.

The results provide strong support for our hypotheses. When we first look at the effect of *focus*, we find that increased focus significantly increases participants' choice of the best alternative (see "Focus on Choice of Best Alternative). In support of H3, we find support for a significant interaction between partitioning and sorting in their effect on focus (see "Interaction of Partitioning and Sorting Quality on Focus): partitioning has a positive impact on focus when sorting quality is high, but a negative impact when sorting quality is low. As hypothesized in H5a, focus also serves as a mediator of the effect of partitioning on the dependent variable (see the rows for "Indirect effect of Focus"). Furthermore, the

moderating effect of sorting quality on how partitioning affects focus is also mediated (see “Moderating Effect of Sorting Quality on indirect effect of Focus”). This provides further support for H5a.

The results for *compensatory decision-making* support a positive effect of more compensatory decision making on the choice of the best alternative (see “CDM on Choice of Best Alternative”). Here, in support of H4, we find a significant positive effect of partitioning on compensatory decision making (see “Partitioning on CDM”). The results also support H5b and show that compensatory decision-making serves as a mediator of the effect of partitioning on the choice of the best alternative (see the rows for “Indirect Effect of CDM”). For completeness, we also allowed for a moderating effect of sorting quality on how partitioning impacts compensatory decision making. As expected this effect was not significant.

- INSERT TABLE 1 ABOUT HERE -

### *Discussion*

The results of Study 2 provided further outcome as well as process level support for the hypothesized impact of SPSs on consumer choice, in a consequential decision-making setting with objectively defined different quality levels between the different choice outcomes. The results strongly support H1 and H2. As in Study 1, we find that partitioning is beneficial for consumers when sorting quality is high, but not when it is low.

In addition, the MouselabWEB data-based analysis allowed us to test H3, H4, and H5a,b and showed that the impact of SPSs on choice outcomes is mediated by participants’ increased focus on the best alternatives and their use of a more compensatory decision making process. These findings show that partitioning combined with high quality sorting (e.g., sorting based on an overall product attractiveness to the consumer) constitutes a choice architecture ensemble that is beneficial for consumers, but that when sorting quality is low, partitioning may harm consumers, particularly by shifting their attention to less attractive alternatives in the set.

### *STUDY 3: FIELD DATA*

To investigate the external validity of the key predicted beneficial effects of SPSs (H1), Study 3 was a quasi-experiment using field data from a leading financial product comparison website in the Netherlands. The website is similar in many respects to Healthcare.gov and the other health insurance exchange websites established under the Affordable Care Act in the U.S. (Wong et al 2016). The website conducted a major interface redesign that introduced a SPS intervention into their personalized health insurance recommendation lists. The firm serves as an intermediary between consumers and suppliers of health insurance. Health insurance plans can be purchased directly through the website and many consumers go to the site each year to switch insurance providers.

Specifically, the interface redesign partitioned the recommendation lists presented to consumers so that consumers were presented with the top three recommended plans and provided a click-through to view the rest of the list, in contrast to the prior design, which simply presented the entire recommendation list. Simultaneously, the interface redesign changed the sorting criterion from a price-based only sorting to a user modeled approach that balanced price and quality. This was an aggregate ranking that was the same for all consumers, but the site did offer some individual customization of the choice set by screening out products that were not relevant to individual consumers. The website anticipated this price-quality sorting to more closely correspond to user preferences than the prior price-based only sorting. By comparing consumer insurance choice data before and after the interface redesign, we can examine the impact of this particular SPS intervention, and see whether the results in a real-world environment are

consistent with our key hypothesis H1 that SPSs improve consumer choice outcomes (Ericson and Starc 2012).

### *Data*

Users of the website entered their personal characteristics and desired insurance specifications – in particular additional coverage above the legal minimum. Based on this information, they were then presented with a sorted list of recommended health insurance alternatives. In the Year 1 choice architecture, this list was a full recommendation list of health insurance products that met the consumer’s pre-specified criteria, sorted according to the premium (i.e. the price of purchasing the plan) and displayed in groups of ten to fit on a webpage. In the Year 2 choice architecture, the recommendation list was partitioned in two sections. First a top three of the most highly recommended products was presented, sorted according to the price-quality algorithm of the website (see Web Appendix A for screen shots before and after the redesign was implemented). Next, consumers could choose to click through to inspect the full list of health insurance products, which, as before, was displayed in groups of ten. They could choose to sort this full list based on a product criterion that they selected, with a price-based sorting being the default.<sup>3</sup>

For most products, consumers were able to purchase the insurance directly via the recommendation website. For a subset of products consumers had to visit the insurer’s website to make the purchase. In the latter case, we were not able to observe whether or not a contract was in fact closed with the insurance company. Therefore, for the purpose of this study, we only analyzed the visits in which consumers closed a contract directly through the recommendation website. This allows us to track the impact of the recommendation list changes on consumers’ actual market choices.

---

<sup>3</sup> As an intermediate step, consumers were automatically shown the lowest price alternative if this was not already part of the top three price-quality ranking. This somewhat lowers the search cost of moving to the full list for some consumers in the sample.

We obtained data from November-December of Year 1 (before the website choice architecture redesign) and from November-December of Year 2 (after the website choice architecture redesign). Almost all health-insurance purchases are made in the Netherlands during these two months because this is the “open enrollment” period set by law. In the data, a health insurance product was purchased in a total of 8,519 visits in Year 1 and 35,113 visits in Year 2. This represented 2.7% and 1.6% of the total visits in these months respectively, which reflects the strong increase in visitor numbers after the new site design was launched. As socio-demographic background variables, we only obtained consumers’ age for the two years. In Year 1 the average age was 36.2 years, and in Year 2 it was 39.1 years. For both Year 1 and Year 2, the data we obtained captured the first 10 alternatives presented in the recommendation list. Only very few consumers chose outside of the first 10 alternatives and hence no other alternatives were stored by the website. In Year 1 this represented the full first page of recommended products, and in Year 2 this represented the top three products first presented to consumers as well as the subsequent page that they could click through with the first 10 alternatives. We also obtained the rank position of each alternative, as well as the user’s individual characteristics and his/her insurance specifications, and, finally, the alternative that was purchased.

### *Results*

Figure 5 provides an overview of the results from the field study. In the full recommendation list setting before the interface redesign we found that 47.6 % of the visitors (4,053 individuals) who bought health insurance selected the first ranked alternative. After the redesign introduced partitioning, 60.8 % of the visitors (21,345 individuals) who bought health insurance selected the first ranked alternative.

To test the impact of introducing the SPS design on the rank of the products that consumer purchased between the years we estimated an ordered probit model. The ordered probit is most appropriate for this analysis because it takes into account the fact that, although a higher ranked position is superior to a lower ranked position, we cannot observe the metric distance in attractiveness between the ranks.

The results of the ordered probit analysis strongly support H1. We find that there is a significant effect of the new SPS design and in the expected direction ( $\beta=.67$ ;  $p<0.001$ ). This effect shows that after the introduction of the new choice architecture consumers were significantly more likely to choose a higher ranked alternative. As a robustness check, we also estimated the ordered probit model while controlling for participants' age. The effect of this control variable was significant ( $\beta=-.01$ ;  $p<0.05$ ) indicating that overall older consumers were less likely to choose a higher ranked alternative. Adding this control variable in the model did not affect the significance of the other results.

- INSERT FIGURE 5 ABOUT HERE -

#### *Further analysis*

Because the field data reflect real-world constraints, there are potential inherent confounds that may have co-occurred with the SPS introduction. One important such potential effect is the fact that the website introduced a significant marketing campaign along with the redesign, likely bringing a broader, less expert, consumer segment to the redesigned website. Because recommendations were personalized, we therefore cannot rule out the possibility that differences in consumers and in the composition of recommendation lists may also have affected the response to the partitioned recommendations. For example, less experienced consumers may more easily benefit from a recommendation compared to their current insurance product and hence may also more readily accept the most highly recommended product.

To rule out this alternative explanation, we conducted a controlled follow-up lab-study that closely mimicked the conditions of the field study, but in which participants were randomly assigned to different conditions and had access to the exact same set of alternatives (see Web Appendix C for details). More specifically, participants were randomly assigned to one condition in a two-by-two (sorting by partitioning) between-subjects design. Health insurance products were sorted based on one of two measures of attractiveness: a price-based sorting (sorted from most to least attractive price), and a price-quality based sorting (sorted from most to least attractive price-quality score), and we also varied whether

or not the set was partitioned (yes, no). The results of this additional study strongly supported the findings from the field study in the more controlled environment.

### *Discussion*

The results of the field study provide a real-world validation of H1, i.e., the hypothesized positive baseline effect of introducing SPSs on consumer choice outcomes. We found that consumers made different choices in the full recommendation list setting as compared to the SPS setting and the results of the ordered probit model estimates revealed a strong positive effect of introducing SPSs, providing support for H1. Furthermore, we replicated this result in a simplified controlled online lab-study.

## **GENERAL DISCUSSION**

The results from all three studies show that, in line with H1, SPSs have a significant and beneficial impact on consumer choice outcomes when sorting quality is medium or high. Consumers are more likely to choose the best alternative when presented with a list that is sorted on overall product attractiveness and partitioned into a small recommended set with the possibility to click through to see the full list. This is due in part to the fact that consumers make use of more compensatory decision making strategies when they are faced with a partitioned set (H4 and H5b). Another major insight of our research, however, is the interaction between the elements of an ensemble of the choice architecture tools, in our case of partitioning, sorting and user model quality. While both partitioning and sorting can help decision-makers make better choices, this result depends on the quality of the model underlying the sorting. The quality of consumers' choices declines along with the quality of the sorting. Partitioning can magnify this effect. Therefore, as we hypothesized in H2, if the quality of the sorting is too low, introducing a partition leads to a decrease in choice quality. The results from the MouselabWEB data show that this moderation of partitioning by sorting quality is mediated by the shift in consumers' focus on the best alternatives (H3



and H5a). Thus, the results from our research highlight how important it is for firms to develop predictive user models that are in line with consumer preferences. In addition, firms need to think carefully about how to set the partitioning at a level that benefits consumers the most. When sorting quality is high, a large partitioned set may harm decision outcomes because consumers don't focus enough and end up searching too much, making a worse selection from the larger set (Diehl 2005). However, when sorting quality is low, a small partitioned set may unduly limit consumers' choice because they focus on a recommended set that does not include the best alternative, and the partitioning inhibits consumers from searching further down the list.

This result illustrates an important challenge facing choice architects. While some choice architecture tools may be helpful in isolation, their effects may differ when used in combination. We demonstrate the importance of thinking of choice architecture tools as ensembles, by showing that the effect of partitioning depends on the quality of the user model underlying the sorting with which it is combined. We believe such interactions are commonplace, and deserving of future research.

### *Theoretical implications*

In theory, SPSs can lead to welfare losses. Indeed, given the very small cost of clicking through to the full sorted set in an on-line environment, a rational actor who has not yet reached decision based on the recommended set, would almost always click through to consider the full set, at least in the kinds of circumstances presented in the empirical portion of this research. Thus, partitioning that leads a significant number of individuals to not even consider the full set, or to pick less frequently outside of that set, leads to welfare losses in the normative model whenever the sorting fails to adequately reflect the preferences of all the people choosing within the partitioned set. As our research highlights, the size of that welfare loss will be a joint function of the ensemble of choice architecture tools, in our case the quality of the sorting (as measured by the fit between the sorting and the preferences of the individuals using the sorting) and the location of the partitioning in the resulting ranking. The lower the quality of the sorting and the smaller the partitioned set, other things being equal, the more likely it is that the optimal

choice for any particular individual lies outside of the partitioned set. Thus, from a normative expected utility perspective, SPSs seem almost always to be welfare reducing, increasingly as the quality of the sorting declines.

Behaviorally, cognitive limitations provide an important countervailing consideration. As the complexity and number of the alternatives in a set grows, individuals increasingly lack the cognitive capacity to engage in the expected utility search strategy assumed in the normative model. At some point along that continuum, individuals may experience choice overload (Chernev, Böckenholt, and Goodman 2015; Iyengar and Lepper 2000). By narrowing the number of choices within the immediate perception of the individual, SPSs focus individuals' attention on the top alternatives and promote a "compensatory" decision strategy among those choices that is closer to the expected utility search strategy of the normative model than the "screening" decision strategies that individuals employ when they have to decide among too many options (Payne, Bettman and Johnson 1993). The empirical process results in Study 2 strongly support this prediction because we find that consumers indeed focus much more strongly on the best alternatives in the case of partitioning combined with high sorting quality. In addition, the results of all three studies show that consumers are more likely to select the best alternative from the set because they apply a more compensatory decision strategy. The mediation analysis in Study 2 shows that SPSs facilitate more compensatory consideration of attributes of the alternatives in the partitioned set, even allowing individuals to partly compensate for the less than perfect match between the sorted list and their preferences (also see the positive results of the medium quality sorting condition in Study 1).

### *Managerial implications*

Managerially, there are likely to be different practical paths for improving sorting quality and for improving partitioning. To improve sorting quality, data and algorithms can be fine-tuned over time and enhanced to improve user model fit. User models can be highly individual and contextualized to fit well with each consumer's needs. This is a relatively continuous process and higher sorting quality will generally improve the quality of consumer choice outcomes in each step. Among the many factors that

can affect sorting quality in practice are: (1) Heterogeneity of preferences among the individuals using the list, (2) The ability of the choice architect to observe those preferences and reliably match them to subpopulations of choosers, (3) The capacity of the algorithm used by the sorting system to take that heterogeneity into account by providing different sortings for different subpopulations, and (4) The ability of the choice architect to accurately observe the relevant attributes of the goods or services being ranked (e.g., not all may wish to share all relevant product features). Because of difficulties involved in addressing these factors, all real-world sorting systems will be imperfect. However, as we found in Study 1, even a medium quality sorting can be quite helpful to consumers.

To improve partitioning, a different, step-wise approach is likely to be more effective. As implied by our theorizing, there are two opposing drivers of optimal partition size: first, the degree of error in the recommendation algorithm's quality and, second, the degree of error in the consumers' own decision making process. All other things being equal, greater error in the recommendation algorithm counsels a larger partitioned set, while greater error in the consumers' decision making counsels a smaller partitioned set. Accordingly, small partitioned sets would be optimal when recommendation error is small and consumer decision error is large. In this situation, the best course of action is to give consumers' less opportunity to make a mistake. By contrast, larger partitioned sets would be optimal in the opposite situation, when decision error is small and recommendation error is large. In this situation, consumers are able to choose sufficiently well among the alternatives, and efforts to increase the impact of the (low quality) sorting are harmful. We believe that markets such as health insurance may provide good examples of the former situation: the extent of error in consumer decision making is relatively high due to the complexity of the product (Johnson et al. 2013), while firms may be able to classify individuals quite well in terms of their insurance needs. In those conditions, a relatively small SPS would be appropriate. On the other hand, in markets such as apartment rentals, consumers themselves may be more informed about their preferences than the recommendation website, and also more capable of determining what they like best. In that case a relatively larger SPS would be better, encouraging broader search by consumers.

Thus, there is not a “one-size fits all” optimal partitioning. Rather, there are different optimal sizes for different markets. A practical way for a firm to begin to find the optimal partitioning size for the products that it offers is to run experimental trials with different set sizes to measure which recommended set size generates the highest quality consumer choice for (one or more of) the markets that it operates in. Depending on the observed variations in consumer choice quality, the firm could then select the best performing recommended set size.

In this context it is useful to distinguish between markets in which an objectively identifiable decision standard exists (or a close proxy thereof – such as in the case of determining the best priced insurance product in Study 2), and markets in which variations in personal taste make it difficult to identify such a standard. In the former case, establishing the best performing recommended set size is straightforward: Multiple recommended set sizes can be tested and the best performing set size is selected. In the latter case, we propose that a useful proxy to identify consumer decision quality can be drawn from the literature on consumer choice consistency in discrete choice modelling. This literature has shown that the complexity of a choice task systematically affects the level of random error in consumer choice (Dellaert, Donkers and Van Soest 2012; DeShazo and Fermo 2002; Louviere 2001; Swait and Adamowicz 2001). This empirical result and the corresponding modelling framework can be used to estimate error scale differences between choice models estimated on the same total product set, but with different partitioning (recommended set) sizes. The partitioning size that generates the lowest choice model error scale for a model estimated on the total product set, reflects the condition under which consumer choices are most consistent across the recommended set and the total set (note that this applies even when there is no objective benchmark of an optimal choice). This partitioning can then be selected by the firm to implement as the most beneficial partitioning size to consumers, because consumers have been shown to make the most consistent overall choices in that condition.

Given that recommendation systems are already quite helpful and user models are improving, we propose that choice architecture ensembles that promote focus on the best alternatives and a compensatory decision-making strategy among a set of options selected on the basis of a well-informed

sorting should lead to better matches between individuals and options than choice architecture that simply presents individuals with all of the options ordered by rank. If we look at the level of heterogeneity that would be best to implement, individual level partitioning might theoretically be optimal, but we expect that this would be practically challenging to implement. Also, consumers may have concerns when they observe variations in recommended set sizes between individuals or variations over their own subsequent interactions with the firm. Therefore establishing stable segment-level recommended set sizes after an initial period of experimentation may be preferable. In addition, while introducing SPSs seems likely to reduce the possibility that individuals whose optimal plan lies outside the partitioned set will find and choose that optimal plan, the research on website search provides substantial reason to doubt that they would find that plan even in the absence of partitioning (Johnson et al. 2004). Thus, provided that the quality sorting is “good enough,” partitioning should represent a welfare improvement.

#### *Limitations and Future research*

Our research provided a balanced mix of hypothetical experiments and a field study that examined whether SPSs improve consumer choice outcomes. The results strongly support the hypothesized positive effect, as well as a negative effect of partitioning when sorting quality is low. However, it is likely that there are other applications or contingencies under which SPSs have more or less impact. Accordingly, it would be valuable to have subsequent research that addresses the impact of, for example, variations in the ‘softness’ of the partitioning in the sorted list, which could be varied by making it easier or harder to see alternatives further down the recommended list. Future research could also address more extensively the economic and consumer welfare implications of introducing SPSs in various markets, depending on consumer search costs, consumer decision complexity, the possibilities for providing high quality sorted recommendations, and other considerations.

Future research could perhaps develop algorithms and models to simultaneously capture the combined impact of (heterogeneous) consumer preferences, consumer decision errors, search costs, and the firm’s modeling errors in an automated SPS generation process. Conceptually, an automated SPS

generating process could perhaps resemble a web-morphing process as proposed by Hauser et al. (2009), but it would still differ in many ways, for example in that it would need to be designed towards generating SPSs (rather than webpages) at the individual level. An alternative, but equally interesting approach to advance the SPS generation process could be to design a more interactive process by developing a dialogue between consumer and firm centered on how consumers would most like to be advised in making their decision. In this process SPSs could also be supplemented by other choice architecture tools to encourage high quality decision making. For example, by customizing information, using more accessible representations of numbers and probabilities, or expressing different attributes on similar scales (Johnson et al. 2012). Future research could investigate the impact of such alternative choice architecture tools.

Finally, it would also be interesting to apply SPS that balance freedom of choice and decision support in other types of decisions and in other decision contexts. For example, in the medical domain, shared decision making between patient and physician could potentially benefit from recommendations that would balance preference models of the two different perspectives of patient and physician. Similarly, in financial decision making there may be opportunities to assist (consumer) investors in making better decisions by focusing their choices on a small set of good investment strategies, while still allowing consumers to indicate their preferred strategy within the set.

## REFERENCES

- Adam, Klaus (2001), "Learning While Searching for the Best Alternative," *Journal of Economic Theory*, 101(1), 252-281.
- Ansari, Asim, Skander Essegaier, and Rajeev Kohli (2000), "Internet Recommendation Systems," *Journal of Marketing Research*, 37(3), 363-375.
- Bhargava, Saurabh, George Loewenstein, and Justin Sydnor (2017) "Choose to Lose: Health Plan Choices from a Menu with Dominated Options," *Quarterly Journal of Economics*, 132(3), 1319–1372.
- Besedeš, Tibor, Cary Deck, Sudipta Sarangi, and Mikhael Shor (2015), "Reducing Choice Overload without Reducing Choices." *Review of Economics and Statistics*, 97(4), 793-802.
- Chakravarti, Amitav, Chris Janiszewski, and Gülden Ülkümen (2006), "The Neglect of Prescreening Information," *Journal of Marketing Research*, 43(4), 642-653.
- Chernev, Alexander, Ulf Böckenholt, and Joseph Goodman (2015), "Choice Overload: A Conceptual Review and Meta-Analysis," *Journal of Consumer Psychology*, 25(2), 333–358.
- Chung, Tuck Siong, Roland T. Rust, and Michel Wedel (2009), "My Mobile Music: An Adaptive Personalization System for Digital Audio Players," *Marketing Science*, 28(1), 52-68.
- Dellaert, Benedict G.C., Bas Donkers, and Arthur van Soest (2012), "Complexity Effects in Choice Experiment-Based Models," *Journal of Marketing Research*, 49(3), 424-434.
- Dellaert, Benedict G.C., and Gerald Häubl (2012), "Searching in Choice Mode: Consumer Decision Processes in Product Search with Recommendations," *Journal of Marketing Research*, 49(2), 277-288.
- DeShazo, J.R. and German Fermo (2002), "Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency," *Journal of Environmental Economics and Management*, 44(1), 123-143.

- Diehl, Kristin (2005), "When Two Rights Make a Wrong: Searching too Much in Ordered Environments," *Journal of Marketing Research*, 42(3), 313-322.
- Ericson, Keith Marzilli, and Amanda Starc (2012), "Heuristics and Heterogeneity in Health Insurance Exchanges: Evidence from the Massachusetts Connector," *American Economic Review*, 102(3), 493-497.
- Ge, Xin, Gerald Häubl, and Terry Elrod (2011), "What to Say When: Influencing Consumer Choice by Delaying the Presentation of Favorable Information," *Journal of Consumer Research*, 38(6), 1004-1021.
- Hanoch, Yaniv, Stacy Wood, Andrew Barnes, Pi-Ju Liu, and Thomas Rice (2011), "Choosing the Right Medicare Prescription Drug Plan: The Effect of Age, Strategy Selection, and Choice Set Size," *Health Psychology*, 30(6), 719-727.
- Häubl, Gerald, and Valerie Trifts (2000), "Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids," *Marketing Science*, 19(1), 4-21.
- Häubl, Gerald, Benedict G.C. Dellaert, and Bas Donkers (2010), "Tunnel Vision: Local Behavioral Influences on Consumer Decisions in Product Search," *Marketing Science*, 29(3), 438-455.
- Hauser, John R., Glen L. Urban, Guilherme Liberali, and Michael Braun (2009), "Website Morphing," *Marketing Science*, 28(2), 202-223.
- Hauser, John R., and Birger Wernerfelt (1990), "An Evaluation Cost Model of Consideration Sets," *Journal of Consumer Research*, 16(4), 393-408.
- Hayes, Andrew F. (2013), *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. Guilford Press.
- Hogarth, Robin M., and Hillel J. Einhorn (1992), "Order Effects in Belief Updating: The Belief-Adjustment model," *Cognitive Psychology*, 24(1), 1-55.
- Iyengar, Sheena S., and Mark R. Lepper (2000), "When Choice is Demotivating: Can One Desire Too Much of a Good Thing?" *Journal of Personality and Social Psychology*, 79(6), 995-1006.



- Johnson, Eric J., Ran Hassin, Tom Baker, Allison T. Bajger, and Galen Treuer (2013), "Can Consumers Make Affordable Care Affordable? The Value of Choice Architecture," *PLoS ONE*, 8(12), e81521, doi:10.2139/ssrn.2291598.
- Johnson, Eric J., and Robert J. Meyer (1984), "Compensatory Choice Models of Noncompensatory Processes: The Effect of Varying Context," *Journal of Consumer Research*, 11(1), 528–541.
- Johnson, Eric J., Wendy W. Moe, Peter S. Fader, Steven Bellman, and Gerald L. Lohse (2004), "On the Depth and Dynamics of Online Search Behavior," *Management Science*, 50(3), 299-308.
- Johnson, Eric J., Suzanne B. Shu, Benedict G.C. Dellaert, Craig Fox, Daniel G. Goldstein, Gerald Häubl, Richard P. Larrick, John W. Payne, Ellen Peters, David Schkade, Brian Wansink, and Elke U. Weber (2012), "Beyond Nudges: Tools of a Choice Architecture," *Marketing Letters*, 23(2), 487-504.
- Lohse, Gerald L., and Eric J. Johnson (1996), "A Comparison of Two Process Tracing Methods for Choice Tasks," *Organizational Behavior and Human Decision Processes*, 68(1), 28–43.
- Louviere, Jordan J. (2001), "What If Consumer Experiments Impact Variances as well as Means?" *Journal of Consumer Research*, 28(3), 506-511.
- Lu, Shasha, Li Xiao, and Min Ding (2016), "A Video-Based Automated Recommender (VAR) System for Garments," *Marketing Science*, 35(3), 484-510.
- Lynch Jr, John G., and Dan Ariely (2000), "Wine Online: Search Cost Affect Competition on Price, Quality, and Distribution," *Marketing Science*, 19(1), 83-103.
- Moorthy, Sridhar, Brian T. Ratchford, and Debabrata Talukdar (1997), "Consumer Information Search Revisited: Theory and Empirical Analysis," *Journal of Consumer Research*, 23(4), 263-277.
- Payne, John W. (1976), "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," *Organizational Behavior and Human Performance*, 16(2), 366-387.
- Payne, John W., James R. Bettman, and Eric J. Johnson (1993), *The Adaptive Decision Maker*, Cambridge, England: Cambridge University Press.

- Reisen, Nils, Ulrich Hoffrage, and Fred W. Mast (2008) "Identifying Decision Strategies in a Consumer Choice Situation," *Judgment and Decision Making*, 3(8), 641.
- Roberts, John H., and James M. Lattin (1991), "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research*, 28(4),429-440.
- Rosenfield, Donald B., Roy D. Shapiro, and David A. Butler (1983), "Optimal Strategies for Selling an Asset," *Management Science*, 29(9), 1051-1061.
- Scheibehenne, Benjamin, Rainer Greifeneder, and Peter M. Todd (2010), "Can There Ever be Too Many Options? A Meta-Analytic Review of Choice Overload," *Journal of Consumer Research*, 37(3), 409-425.
- Swait, Joffre, and Wiktor Adamowicz (2001), "Choice Environment, Market Complexity, and Consumer Behavior: A Theoretical and Empirical Approach for Incorporating Decision Complexity into Models of Consumer Choice," *Organizational Behavior and Human Decision Processes*, 86(2), 141-167.
- Thaler, Richard H., and Cass R. Sunstein (2008), *Nudge*. New Haven, US: Yale University Press.
- Ubel, Peter A., David A. Comerford, and Eric Johnson (2015), "Healthcare.gov 3.0— Behavioral Economics and Insurance Exchanges," *New England Journal of Medicine*, 372(8), 695-698.
- Weitzman, Martin (1979), "Optimal Search for the Best Alternative," *Econometrica*, 47(3), 641-65.
- Willemsen, Martijn C., Ulf Böckenholt, and Eric J. Johnson (2011), "Choice by Value Encoding and Value Construction: Processes of Loss Aversion," *Journal of Experimental Psychology: General*, 140(3), 303-324.
- Willemsen, Martijn C., and Eric J. Johnson (2011), "Visiting the Decision Factory: Observing Cognition with MouselabWEB and other Information Acquisition Methods," In Schulte-Mecklenbeck, M., Kühberger, A. & Ranyard, R. (Eds.). *A Handbook of Process Tracing Methods for Decision Making*, (pp. 21-42), New York: Taylor & Francis.

- Wilson, Timothy D., Thalia Wheatley, Jonathan M. Meyers, Daniel T. Gilbert, and Danny Axsom (2000), "Focalism: A Source of Durability Bias in Affective Forecasting," *Journal of Personality and Social Psychology*, 78(5), 821-836.
- Wong, Charlene A., Arthur T. Jones, Janet Weiner, Robert J. Town, and Tom Baker (2016), "For Third Enrollment Period, Marketplaces Expand Decision Support Tools to Assist Consumers," *Health Affairs*, 35(4), 680-687.
- Xiao, Bo, and Izak Benbasat (2007), "Consumer Decision Support Systems for E-Commerce: Design and Adoption of Product Recommendation Agents," *MIS Quarterly*, 31(1), 137-209.
- Zhao, Xinshu, John G. Lynch, and Qimei Chen (2010), "Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis," *Journal of Consumer Research*, 37(2), 197-206.

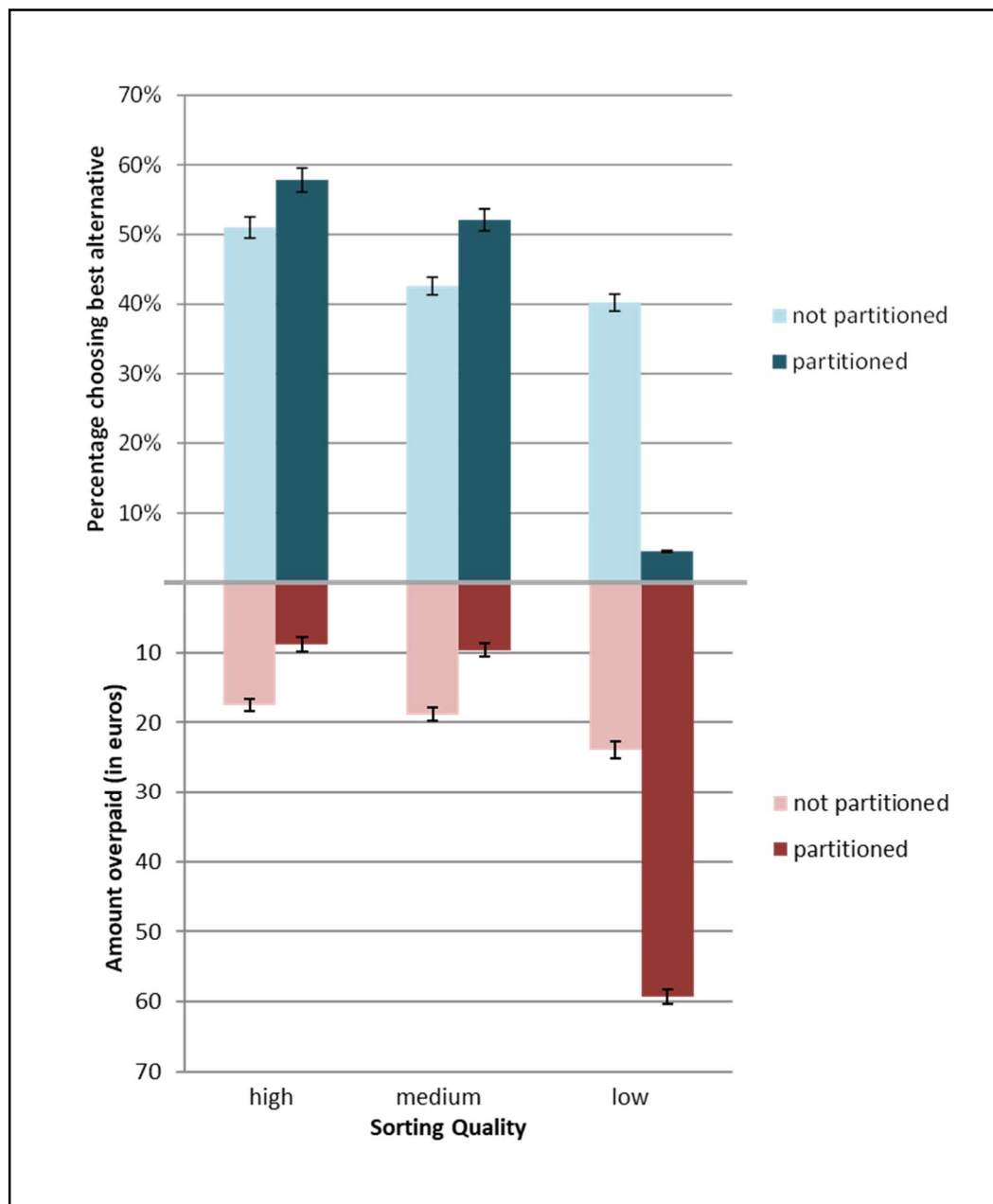
*TABLE 1*  
*STUDY 2: MEDIATION ANALYSIS FOR BEST CHOICE*<sup>§</sup>

<i>Unstandardized regression coefficients of different model pathways</i>	<i>Parameter Estimate</i>	<i>t Value</i>
Partitioning on Focus	-1.41**	-24.94
Sorting Quality on Focus	1.82**	32.22
Interaction of Partitioning and Sorting Quality on Focus	1.66**	29.24
Partitioning on Compensatory Decision Making (CDM)	0.04**	4.44
Sorting Quality on CDM	-0.02**	-2.13
Interaction of Partitioning and Sorting Quality on CDM	0.01	0.58
Focus on Choice of Best Alternative	0.69**	7.70
CDM on Choice of Best Alternative	0.82**	4.96
Partitioning on Choice of Best Alternative	0.02	0.37
Sorting Quality on Choice of Best Alternative	-0.14	-1.86
Interaction of Partitioning and Sorting on Choice of Best Alternative	-0.07	-1.11
<i>Bootstrap tests of indirect (mediated moderation) effects</i>	<i>Effect</i>	<i>S.E.</i>
Indirect effect of Focus (Sorting quality is low)	-2.10*	0.20
5000 bootstraps C.I. at 95%	LL=-2.54	UL=-1.78
Indirect effect of Focus (Sorting quality is high)	0.17*	0.02
5000 bootstraps C.I. at 95%	LL=0.13	UL=0.21
Indirect effect of CDM (Sorting Quality is low)	0.03*	0.01
5000 bootstraps C.I. at 95%	LL=0.01	UL=0.06
Indirect effect of CDM (Sorting Quality is high)	0.04*	0.01
5000 bootstraps C.I. at 95%	LL=0.02	UL=0.07
Moderating effect of Sorting Quality on indirect effect of Focus	2.27*	0.21
5000 bootstraps C.I. at 95%	LL=1.92	UL=2.73
Moderating effect of Sorting Quality on indirect effect of CDM	0.01	0.02
5000 bootstraps C.I. at 95%	LL=-0.02	UL=0.04

<sup>§</sup> We also included a choice task order variable in the model that captures the average difference between the first and the second set of choices per respondent. For expositional clarity this variable is not reported in the table. While the PROCESS module we applied for estimation does not allow for random effects, we also ran the analysis for only the first choice for each participant and obtained similar findings. n=828 (with two choice tasks observed for most, but not all respondents).

\*\* Parameter estimate significant at  $p < .01$ ; \* Parameter estimate significant at  $p < .05$ .

FIGURE 1  
 PERCENTAGE ACCURATE CONSUMER CHOICES AND YEARLY AMOUNT OVERPAID DUE  
 TO NOT CHOOSING THE BEST PRODUCT (STUDY 1)



**FIGURE 2**  
*PERCENTAGE ACCURATE CONSUMER CHOICES AND AMOUNT OVERPAID DUE TO NOT CHOOSING THE BEST PRODUCT (STUDY 2)*

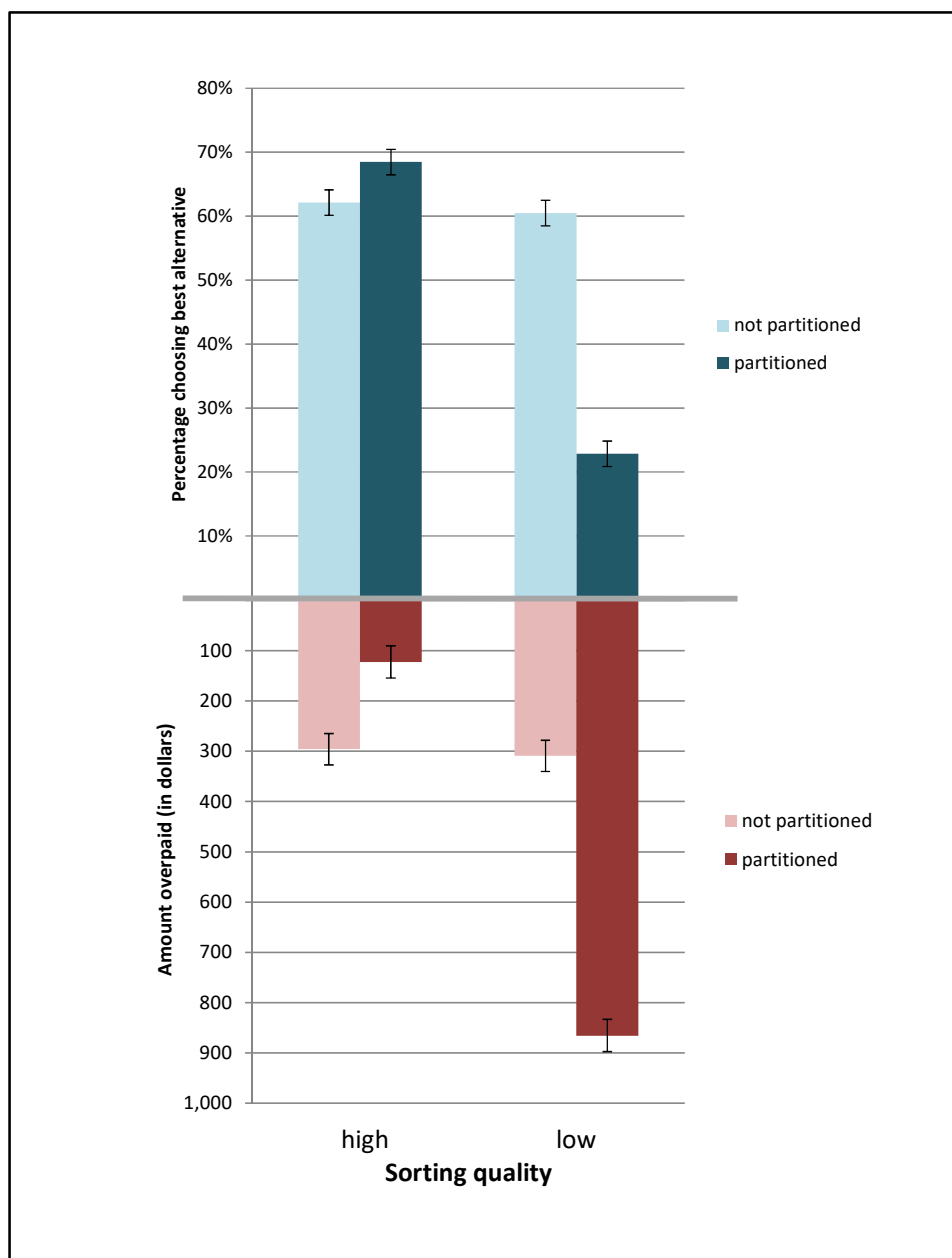
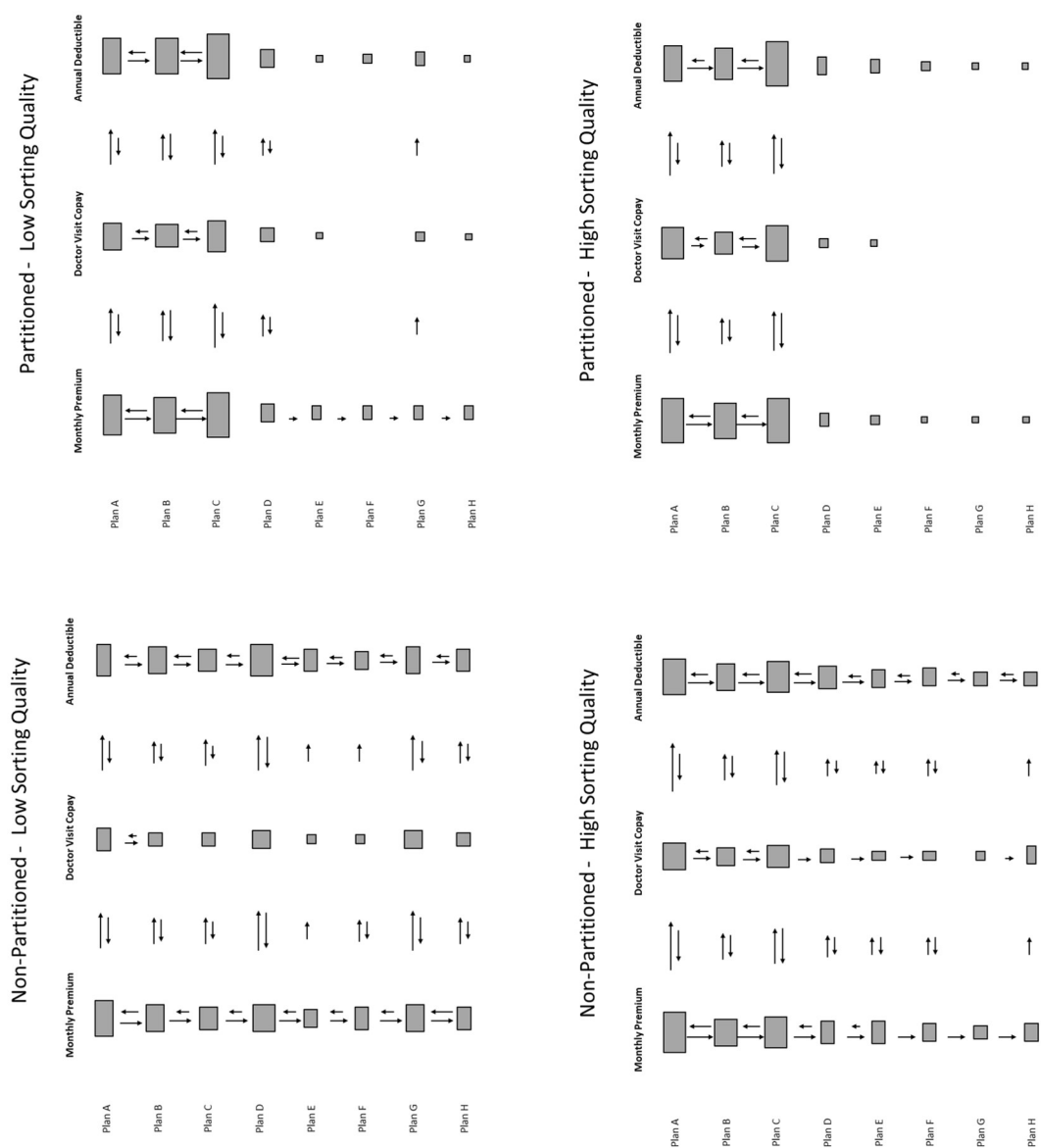
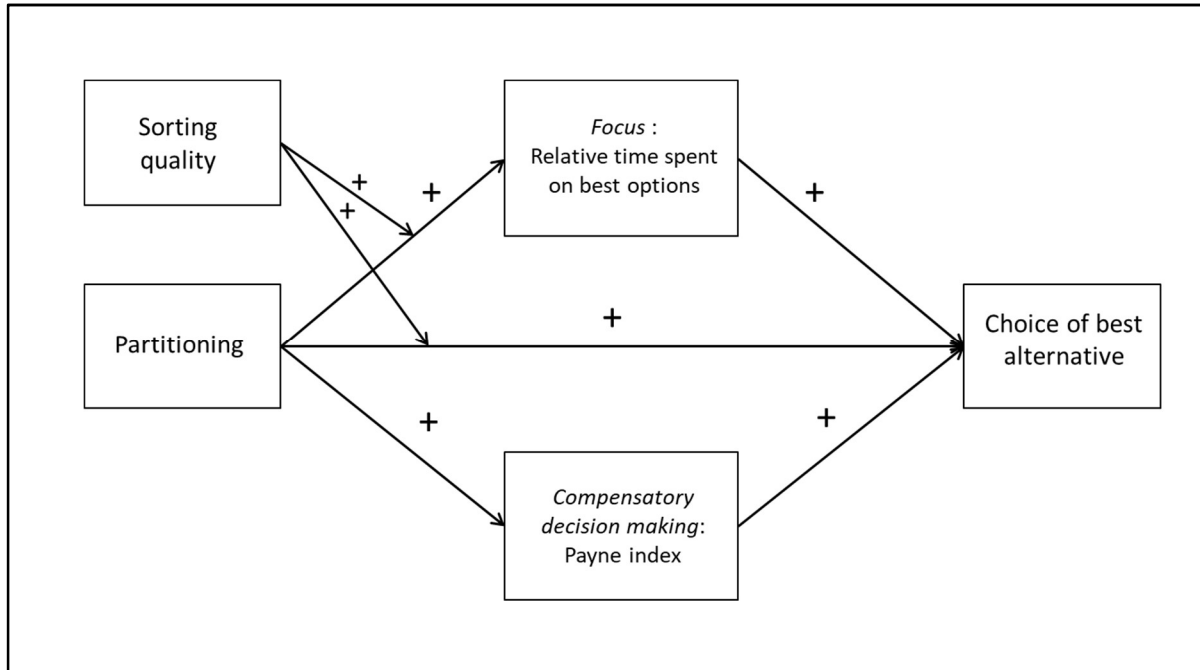


FIGURE 3  
STUDY 3: PROCESS DATA MOUSELABWEB<sup>£</sup>



£ Boxes in the figure correspond to the information acquisition for the attribute cells that participants viewed. The horizontal length of each box represents average inspection time and vertical height average number of times inspected. The length of the arrows between the boxes represents the number of transitions participants made. Results are averaged across the two tasks.

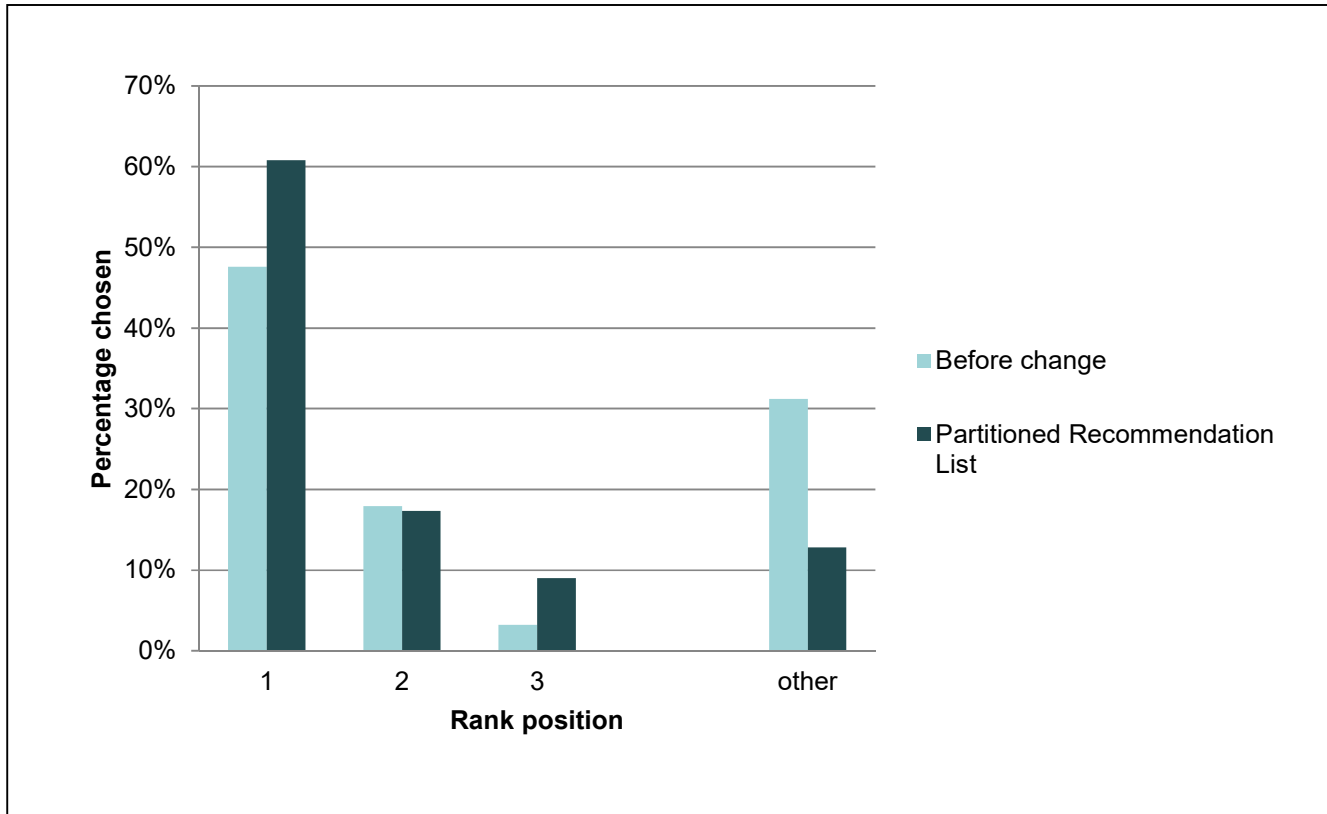
FIGURE 4  
PROPOSED MEDIATED MODERATION STRUCTURE FOR PROCESS DATA<sup>£</sup>



£ Note: The moderation effects of sorting quality are indicated by paths that point directly to other paths. High sorting quality increases the positive effect of partitioning on focus and choice of the best alternative (H1). However, low sorting quality, consistent with H2, reverses the effect of partitioning and leads to lower focus and decision quality.



*FIGURE 5*  
*FIELD STUDY: CONSUMER CHOICES BEFORE AND AFTER THE INTRODUCTION OF SORTED*  
*PARTIONED SETS*





*STUDY 2–NON-PARTITIONED*

Monthly Premium	Doctor Visit Copay	Annual Deductible	I choose...
Premium A	Copay A	Deductible A	<input type="button" value="Plan A"/>
Premium B	Copay B	Deductible B	<input type="button" value="Plan B"/>
Premium C	Copay C	Deductible C	<input type="button" value="Plan C"/>
Premium D	Copay D	Deductible D	<input type="button" value="Plan D"/>
Premium E	Copay E	Deductible E	<input type="button" value="Plan E"/>
Premium F	Copay F	Deductible F	<input type="button" value="Plan F"/>
Premium G	Copay G	Deductible G	<input type="button" value="Plan G"/>
Premium H	Copay H	Deductible H	<input type="button" value="Plan H"/>

There are no further options available. Please make a choice from the eight options shown above and then click "Next >>".

## STUDY 2-PARTITIONED

Monthly Premium	Doctor Visit Copay	Annual Deductible	I choose...
Premium A	Copay A	Deductible A	<input type="button" value="Plan A"/>
Premium B	Copay B	Deductible B	<input type="button" value="Plan B"/>
Premium C	Copay C	Deductible C	<input type="button" value="Plan C"/>

If you would like to look for more options, click "See More Options"

Have you made your choice? If you have selected an option click "Next >>"

## STUDY 3: FIELD STUDY

BEFORE: SORTING BASED ON PRICE, NO PARTITIONING

51 autoverzekeringen voor je BMW 3-Serie

Kenteken: 99-HGT-8 Schadevrije jaren: 2 Km per jaar: 20.001 – 25.000 km

Wijzig je gegevens

Hulp nodig?  
035 - 626 55 44  
Vielia Verbeek  
Teamleider klantadvies

Independer.nl Top 3 **Alle verzekeringen**  Af te sluiten bij Independer.nl

Toon prijzen met: **WA+ Volledig Casco (All Risk)** **Standaard eigen risico** **Maandpremie** Sorteer op: **prijs**

Vergelijk	Prijs-kwaliteit	Polis-voorwaarden	Klantcijfer	Eigen risico	Independer.nl Keurmerk	Maandpremie	
<input type="checkbox"/> PREMIO	99	7,0	7,7 > 497 reviews	€ 135	<input checked="" type="checkbox"/>	77,55	<a href="#">Verder</a> > <a href="#">Toon Details</a>
<input type="checkbox"/> Klaverblad Verzekeringen	87	8,2	7,9 > 1027 reviews	€ 135	<input checked="" type="checkbox"/>	78,58	<a href="#">Verder</a> > <a href="#">Toon Details</a>
<input type="checkbox"/> AllSecur	100	8,1	8,1 > 2154 reviews	€ 0	<input checked="" type="checkbox"/>	80,54	<a href="#">Verder</a> > <a href="#">Toon Details</a>
<input type="checkbox"/> FBTO	91	6,5	8,0 > 57 reviews	€ 150	<input checked="" type="checkbox"/>	83,11	<a href="#">Verder</a> > <a href="#">Toon Details</a>
<input type="checkbox"/> delta lloyd	91	7,4	8,0 > 710 reviews	€ 135	<input checked="" type="checkbox"/>	<del>106,27</del> 85,32	<a href="#">Verder</a> > <a href="#">Toon Details</a>
<input type="checkbox"/> delta lloyd	89	7,4	8,0 > 710 reviews	€ 135	<input checked="" type="checkbox"/>	<del>106,27</del> 85,32	<a href="#">Verder</a> > <a href="#">Toon Details</a>
<input type="checkbox"/> London	85	6,2	7,3 > 12 reviews	€ 250	<input checked="" type="checkbox"/>	85,51	<a href="#">Verder</a> > <a href="#">Toon Details</a>
<input type="checkbox"/> FBTO	85	5,5	8,0	€ 150	<input checked="" type="checkbox"/>	86,31	<a href="#">Verder</a>

AFTER: SORTING BASED ON PRICE-QUALITY, PARTITIONING BASED ON TOP 3

Jouw eigen risico:  
 € 375 eigen risico (standaard)

Type verzekering:  
 Alleen restitutie

Prijzen per:  
 Maand

[Alle 26 verzekeringen](#)

### Top 3 beste prijs-kwaliteit

van verzekeringen die direct via Independer af te sluiten zijn

▶ Hoe werkt de top 3?

Verzekeraar

---

Prijs-kwaliteit

---

Klantscore

---

Vrije zorgkeuze

▶ Zoek je alleen restitutie?

Laagste prijs

1

Ditzo

■■■■■■■ 100

7,7 3155 reviews

✓

▶ Toon meer

---

bij verzekeraar €99,95

via Independer  
89,95

▶ Toon details

Verder ▶

Vergelijk

2

Klamer

■■■■■■■ 98

8,0 734 reviews

✓

▶ Toon meer

---

bij verzekeraar €90,95

via Independer  
90,95

▶ Toon details

Verder ▶

Vergelijk

3

OHRA Direct geregeld

■■■■■■■ 97

7,9 10543 reviews

✓

▶ Toon meer

---

€100,74

via Independer  
91,87

Je voordeel:  
€8,87 per maand

▶ Toon details

Verder ▶

Vergelijk

Vergelijk met:

- Maak je keuze -

Vergelijk hier met één van de verzekeraars uit de markt

*WEB APPENDIX B*  
*ADDITIONAL MODEL ESTIMATES STUDY 2*

*TABLE WB.1*  
*STUDY 2: BEST PRODUCT - CHOICE MODEL<sup>a</sup>*

	<i>Parameter Estimate</i>	<i>t Value</i>
Intercept	0.22	1.57
Std. Deviation of Intercept	3.14**	10.99
Partitioning <sup>§</sup> (yes = 1, no = -1)	-0.78**	-4.97
Sorting Quality <sup>§</sup> (high = 1, low = -1)	1.15**	7.04
Partitioning x Sorting Quality	1.09**	6.69
Choice task order <sup>§</sup> (first = 1, second = 0)	-0.25**	-3.11

<sup>a</sup> Random effects binary logit model; McFadden  $R^2 = .20$ ;  $n = 828$  (with two choice tasks observed for most but not all respondents). Best alternative chosen coded 1, other alternatives chosen coded 0.

\*\* Parameter estimate significant at  $p < .01$ ; \* Parameter estimate significant at  $p < .05$ .

<sup>§</sup> Variable dummy coding in parentheses. The choice task order variable captures the average difference between the first and the second set of choices per respondent. We also tested for possible interactions between order and the experimental variables, but these did not affect our conclusions and are therefore not reported in the table.

*TABLE WB.2*  
*STUDY 2: AMOUNT OVERPAID - REGRESSION MODEL<sup>b</sup>*

	<i>Parameter Estimate</i>	<i>t Value</i>
Intercept	413.97**	15.73
Partitioning <sup>§</sup> (yes = 1, no = -1)	94.00**	6.01
Sorting Quality <sup>§</sup> (high = 1, low = -1)	-189.14**	-12.11
Partitioning x Sorting Quality	-182.94**	-11.70
Choice task order <sup>§</sup> (first = 1, second = 0)	25.77	1.65

<sup>b</sup> Random effects regression model;  $R^2 = .16$ ;  $n = 828$  (with two choice tasks observed for most but not all respondents).

\*\* Parameter estimate significant at  $p < .01$ ; \* Parameter estimate significant at  $p < .05$ .

<sup>§</sup> Variable dummy coding in parentheses. The choice task order variable captures the average difference between the first and the second set of choices per respondent. We also tested for possible interactions between order and the experimental variables, but these did not affect our conclusions and are therefore not reported in the table.

*WEB APPENDIX C*  
*A CONTROLLED REPLICATION OF THE FIELD STUDY IN THE LAB*

In this study we conducted an experiment in a controlled hypothetical choice environment that mimicked our field study to investigate the hypothesized beneficial effect of SPSs on consumer choice. We simulated, in a simplified and more controlled setting, the consumer decision task that is presented by health insurance exchange websites, which allow consumers to compare and purchase health insurance plans. The website we used as the basis for constructing the experiments is based in The Netherlands (where our field data was collected). We manipulated the two key components of SPSs independently from one another: (1) The sorting criterion on which products were ranked in the list (a more restricted sorting based on price only vs. a higher quality sorting based on both and price and quality) and (2) the presence of partitioning in the list or not. Participants were from a nationally representative on-line panel.

*Method*

Participants were randomly assigned to one of four treatment conditions in a between-subjects design. In all conditions participants were presented with a recommendation list of 10 health insurance plans. Plans were sorted on one of two measures of attractiveness: a price-based sorting (sorted from most to least attractive price), and a price-quality based sorting (sorted from most to least attractive price-quality score). Insights from discussions with experts in the field and with representatives of regulatory authorities suggested that both price and quality are important for making high quality health insurance decisions, therefore the price-quality sorting represents a better simple user model.

We also varied whether or not the set was partitioned: The partitioned condition showed participants the three most highly recommended product with the option to click through to the full list (ranked on the same sorting criterion). The full recommendation list condition showed the sorted list of all recommended products directly to participants without any partitioning.



Products were described in terms of key health insurance product features as used on the health insurance comparison website from which we obtained our field data. Participants could click on a help link to see a short explanation of each feature. While the presence of partitioning and the order of presentation varied, participants had access to the same set of products and explanations in all conditions, thus any changes in observed choices must be due to the manipulated factors.

Participants were asked to make a choice from the health insurance plans presented as if they had to select a health insurance plan for themselves. The plans on the recommendation list closely reflected a realistic product selection in the participants' health insurance market (The Netherlands) for adults. The actual plans and the price-quality sorting were taken from the website used to conduct the field experiment described later. Brand names were changed into fictitious brands to avoid unobserved brand associations that participants might have. The only minor modifications that were made were to avoid the presence of clearly dominating alternatives in the list.

For this study, a total of 858 participants of over 18 years old were recruited through a large scale online panel run by a marketing research company. The average age of the participants was 46.7 years old, 50.9% of the participants were female and 36.9% had a bachelor or higher degree, 38.7% of participants had bought health insurance for themselves or for someone else in their household within the past two years.

### *Results*

Table WC.1 provides an overview of the results from the study. For each condition we present the rank position in the sorted list that the participants selected. Rank 1 represents the most highly recommended product, rank 2 the second highest recommended product, etc.. To determine the significance of the observed differences in the ranks chosen between conditions we used an ordered probit model to analyze the effects of partitioning, the price only based sorting vs. price-quality based sorting, and the interaction between these two factors on the observed ranks of participants' selected alternatives between conditions. The ordered probit is most appropriate for this analysis as it takes into account the

fact that, although a higher rank position is superior to a lower rank position, we cannot observe the metric distance in attractiveness between ranks.

– INSERT TABLE WC.1 ABOUT HERE –

The results of the analysis again strongly support H1. We find that there are significant effects of sorting quality ( $\beta=.80$ ;  $p<0.001$ ) and partitioning ( $\beta=1.83$ ;  $p<0.001$ ) and in the expected direction. We find no significant interaction of partitioning and sorting criterion. As a robustness check, we also estimated the ordered probit model while controlling for participants' experience with buying health insurance in the past two years (yes/no). The effect of this control variable was not significant, and adding it in the model did not affect the significance of the other results.

We also explored two process measure proxies to obtain an understanding if the underlying consumer decision making was shifting in the expected direction. As a first proxy for how thorough participants' evaluation of the alternatives was, we divided the total time spent in the experiment by the number of alternatives that was visible to participant (taking into account whether or not the participant in the partitioned condition had clicked through to see the full recommendation list). We find that, as expected, in the partitioned condition participants spent significantly more time per alternative ( $M_{\text{PARTITIONED}}=43.5$ ;  $M_{\text{NOT-PARTITIONED}}=19.0$ ;  $t=-17.7$ ;  $p<0.001$ ). As a second proxy we looked at how many participants clicked through to look at products beyond the recommended set in the partitioned condition. We find that 33.3% of the participants made use of this possibility, and that there is no significant difference between high and low quality sorting ( $M_{\text{HIGH}}=32.8\%$ ;  $M_{\text{LOW}}=33.8\%$ ;  $p>.84$ ), suggesting a strong focus on the top three alternatives when partitioning is present in both sorting quality conditions.

*Discussion*

The results of this study provide strong support for the hypothesized positive impact of SPSs on consumer choice (H1) that we observed in the Field study, and in particular for the positive effect of partitioning and high quality sorting on promoting higher ranked consumer decision outcomes.

Note that in Studies 1 and 2 in the paper we conducted principle-agent based experiments in which participants preferences were determined exogenously by the researcher and which allowed for objectively determining the outcome quality of the participants' choice task. The results of these studies provided further support for this positive effect of SPSs (H1), as well as for the negative effect of partitioning when combined with low sorting quality (H2).

*TABLE WC.1*  
*CONSUMER CHOICES OF RANK POSITIONS ACROSS CONDITIONS*

Rank position chosen	<i>Price-quality based sorting</i>				<i>Price based sorting</i>			
	<i>Without partitioning</i>		<i>With partitioning</i>		<i>Without partitioning</i>		<i>With partitioning</i>	
	<i>Number times selected</i>	<i>Percentage</i>	<i>Number times selected</i>	<i>Percentage</i>	<i>Number times selected</i>	<i>Percentage</i>	<i>Number times selected</i>	<i>Percentage</i>
First ranked alternative	29	13.5%	50	22.8%	17	8.6%	49	21.7%
Second ranked alternative	32	14.9%	121	55.3%	0	0.0%	9	4.0%
Third ranked alternative	2	0.9%	7	3.2%	31	15.7%	116	51.3%
...								
Other alternatives	152	70.7%	41	18.7%	150	75.7%	52	23.0%
Total	215	100%	219	100%	198	100%	226	100%