Consumer Choice between Recommendation Algorithms - Experimental Evidence

Felix Schleef*

February 28, 2026

[Link to Latest Version]

Abstract

Regulators are increasingly concerned with the power of large online platforms to bias consumer recommendations. In light of these concerns, I study experimentally whether giving consumers the choice between several recommendation algorithms can improve consumer welfare. I develop a novel experimental setting in which I first elicit the subjects' preferences and then expose them to choice tasks intermediated by personalized recommendation algorithms. When faced with a costly choice between two algorithms, subjects have a positive willingness-to-pay for better recommendations. However, subjects underestimate the potential gain from the better algorithm: a $1 increase in estimated gain raises willingness-to-pay by only $0.15 on average. These findings suggest that giving consumers power over recommendation algorithms to curtail potential abuse is not straightforward. However, it may be a viable business for platforms to offer improved recommendation algorithms to consumers for a fee.

*HEC Paris. E-mail: felix.schleef@gmail.com. I would like to thank Roxana Fernández Machado, Thibaud Vergé, Michelagelo Rossi, Louis Pape, Guillaume Hollard and Yves le Yaouanq. This research was supported by funding from the ANR-21-CE26-0018-01.

1 Introduction

Choices on online platforms are often too numerous for consumers to compare all relevant products (Natan, 2024). When consumers search products, platforms therefore employ recommendation systems. Recommendation systems shape which products a consumer sees and in which order they are searched (Ursu, 2018). With the design of its recommendation system, a platforms affect both consumer choice as well as the competition among firms (Armstrong, Vickers and Zhou, 2009). While these algorithms can be used to help consumers, platforms can also exploit the consumers' reliance on recommendations (Peitz, 2023, Lam, 2021).

Regulators have in the recent past have been particularly concerned with self-preferencing, i.e. the power of hybrid platforms to inflate the ratings of items that the platform sells itself. So far, the answer from European regulators has been to ban self-preferencing for large platforms through the Digital Markets Act (*Digital Markets Act*, 2022). However, this regulation maybe difficult to enforce, difficult to comply with[1] and leaves other consumer-welfare related concerns, such as price discrimination through recommendations, unaddressed.

In light of these concerns, I study an alternative proposal for regulation, which would enable the consumer to choose between different recommendation systems. In recent times most prominently discussed by Fukuyama and coauthors (Fukuyama, Richman, Goel, Katz, Melamed and Schaake, 2020) as "middleware", the concept of enabling consumers to have more control over their interaction with online services is not new[2]. In the context of recommendation systems, Resnick and Varian (1997) envisioned a future, where e-commerce stores and recommendation systems would be two separate entities, owned and operated by two different firms. The e-commerce stores would be competing intensely with each other,

[1] Given that the past purchase data on large platforms was generated in a setting with potentially biased (or self-preferencing) algorithms, it may be difficult to generate truly unbiased rankings, even when using an unbiased algorithm. How to de-bias data for recommendation applications is an active area of research (Chen, Dong, Qiu, He, Xin, Chen, Lin and Yang, 2021, Lin, Liu, Pan and Ming, 2021).

[2] This paper also relates to recent contributions of Bergemann, Cremer, Dinielli, Groh, Heidhues, Schafer, Schnitzer, Morton, Seim and Sullivan (2023), Esber, Kominers, Kornfeld, Belsky, Chitnis, Chmielinski, Deighton, Eliot, Rossini, Searls, Shah and Williams (2024) and Posner and Weyl (2018) that consider ways for consumers to better govern and make use of their personal data.

and the consumer could use the recommendation system that they perceive to best fit their needs. Nowadays, e-commerce platforms such as Amazon or content platforms like Spotify or Netflix come bundled with their own recommendation system, and the consumer does not have an alternative recommendation system to choose from, unless they leave the platform. This leads to the question whether it would be a promising policy idea to request large platforms to let consumers choose from multiple recommendation systems. Apart from how the incentives for platforms might change, for such a policy to be effective it is necessary that consumers are able to judge different recommendation systems well.

I use a laboratory experiment to study how subjects use and assess different recommendation algorithms. I first elicit the subjects risk aversion, which I then use to create and order lists of three-outcome lotteries with two different algorithms: one that selects and orders lotteries based on the subjects expected utility (expected utility algorithm), and one that weighs the expected utility of the subject and the objective of the platform (platform algorithm). I then let subjects make choices from lists of lotteries that have been selected and ordered by these two algorithms. After 10 rounds of the choice task with each algorithm, I elicit their willingness to pay for either algorithm - the subject gets to change the probability with which they are going to face either the expected utility or the platform algorithm for a last set of 10 choice tasks.

I document three key findings. First, I find that in both algorithmic treatments, subjects are more likely to choose items that are towards the top of the list. This is consistent with our empirical knowledge of consumer behavior on online platforms, where defaults and particularly salient options are more likely to be chosen. Second, the platform algorithm leads subjects to choose lotteries with lower expected values, and lotteries that are further down the list of ordered lotteries. The experiment therefore successfully creates a setting where the platform algorithm induces worse outcomes for the subjects, and where subjects adapt their behavior to the different algorithms. Third, I find that subjects have on average a positive willingness-to-pay for the expected utility algorithm, but that there is substantial heterogeneity. Subjects that have a higher expected gain from using the expected utility algorithm have a higher willingness-to-pay, however the willingness to pay increases only by \$0.15 for an additional \$1 of estimated gain. This indicates that subjects underestimate the potential benefit of having the expected utility algorithm.

The novel contribution of this project is twofold. First, I develop a parsimonious experimental design which allows researchers to study the effects of *personalized* recommendation algorithms. Existing field experiments on platforms cannot observe algorithm objectives, while existing lab experiments typically study non-personalized rankings. My design bridges this gap: by eliciting risk preferences and using them to construct personalized rankings in real time, I create a setting where the researcher controls both the algorithm's objective and has knowledge of the consumer's preferences. This enables clean measurement of how algorithm alignment affects consumer outcomes. The framework can accommodate varying degrees of congruence or conflict between platform and consumer interests de Cornière and Taylor (2019), and could be extended to study information disclosure, advertising, framing, or other interventions.

As a second contribution, I then apply this framework to the question: can consumers choose effectively between algorithms that differ in alignment with their interests? I find that while consumers do have positive willingness-to-pay for a better algorithm, they substantially undervalue the gains—paying only \$0.15 for each \$1 of expected benefit. This suggests that simply giving consumers the choice between algorithms may not be sufficient to protect consumer welfare. However, the fact that willingness-to-pay on average is positive suggests that better recommendations are an important value proposition for platform companies, which could be monetized independently (through premium algorithms) or used as a competitive differentiator.

I contribute to a growing empirical literature on rankings and platform recommendation systems (Donelly, Kanodia and Morozov, 2023; Zhang, Ferreira, Matos and Belo, 2021; Farronato, Fradkin and MacKay, 2023; Reimers and Waldfogel, 2023; Lee and Musolff, 2021; Greminger, 2022; Greminger, 2023; Kaye, 2024; Ghose, Ipeirotis and Li, 2014; De los Santos and Koulayev, 2017; Derakhshan, Golrezaei, Manshadi and Mirrokni, 2022). Ursu (2018) studies the effect of rankings on consumer choices using experimental data from the online travel platform Expedia. A subset of consumers looking for hotels were shown a random ranking, while the others were shown a ranking by "relevance" according to Expedia. The subjects were not aware that their rankings were created differently. She finds that rankings affect consumer search, but do not affect purchase decisions after controlling for search. Donelly et al. (2023) investigate the effects of recommendation systems using an experiment on the furniture shopping platform Wayfair. The experiment randomly assigned shoppers to see either personalized or non-personalized rankings of products. They find that personalized rankings induce more active consumer search and increase purchase probabilities. Importantly however, the experiment was conducted unbeknownst to the user - the users may have expected their product search results to be personalized based on their experience on similar E-commerce websites, or they might have changed their search behavior if they had known that their rankings are not personalized. Zhang et al. (2021) run an experiment on a Video-on-demand system. They randomize both the slot (i.e. the ranking of the movie, how saliently it is presented) and the price of the movie. They find that movies that are prominently displayed have less price elastic demand. Similarly to the previous studies, subjects are not aware that the algorithm has changed for them as compared to when they used the same service before the researchers intervention. Farronato et al. (2023) document that products of Amazon brands are more prominently displayed than other products. They also show that steering can be potentially very effective, since in 72.1% of searches consumers stay on the first page of search results.

This paper also relates to the emergent experimental literature on platform recommendation systems. Boldrini and Clavorà Braulin (2025) conducts a framed experiment related to the Amazon "Buybox" to study the effect of information provision about the algorithm on consumer choices. Düll, Karle, Martin and Schumacher (2025) study the performance of search cost estimation methods when subjects have priors about the returns to search (which could be induced by recommendation algorithms). Most closely related to my work is the paper by Fong, Natan and Pantle (2024), where they conduct an experiment in order to disentangle position specific search costs from beliefs about the recommendation (or ranking) algorithm. They find that both mechanisms are important, and that not accounting for beliefs leads to overestimating position-specific search costs. Lastly, they also show that consumers learn about algorithms in experimental settings where the recommendation algorithm changes, and they are neither informed about the change nor the algorithms. Their findings are important for the simulation of counterfactual recommendation procedures, as overestimated search costs lead to overestimating gains from potentially better algorithms. Instead, I elicit preferences for different algorithms directly.

Regarding the literature on recommendation algorithms beyond their use as ranking tools on platforms, I relate to Chen, Wu and Zhong (2023), from whom I borrow the three outcome lottery design. Chen et al. (2023) study how subjects make binary choices between three outcome lotteries when they can use recommendation systems. They find, that subjects tend to follow the recommendations, and that they make better and faster decisions when they have access to recommendations. They also document that subjects are willing to pay a fee to have access to the recommendations. To generate recommendations, they use algorithms that are employed in real-life recommendation systems, such as collaborative filtering. In contrast to their research, I do not consider technical intricacies of different recommendation algorithms. I start from the view that platforms use accurate utility estimates for their recommendations, and then investigate how consumers choose from ordered lists, rather than how they make a binary choice. Their finding that consumers are willing to pay for recommendations rather than not having recommendations is an important result, and and I will be able to contrast this with my findings on whether subjects are willing to pay for a recommendation system that is aligned with their preferences rather than one that is not aligned. Caplin, Dean and Martin (2011) study how complexity affects choices from lists. They show that many choices from lists are consistent with sequential search, which is corroborates the steering power of rankings. While I will not vary the complexity of choices in my experiment, the experimental setting could be easily amended to vary complexity by increasing the number of lottery choices. Kaye (2024) studies outcomes of firms and consumers with recommendation algorithms where firms price endogenously. He finds that personalization improves the matches of consumer to products, but it implies that firms are matched with consumers that have more inelastic demand for their offering. Facing this new demand, firms find it optimal to increase their price, negating the welfare benefits that consumers derive from the improved matches. Importantly for my experiment, he emphasizes the role of consumer expectations about recommendation algorithms. In experiments on platforms like Expedia, consumers may expect the items to be ordered by a recommendation system, and may infer something about unobserved product attributes from the items position in a list, even if the recommendation system has been deactivated for the user as part of an experiment. I therefore emphasize to subjects in the experiment, that the items on the lists that they are able to choose from are selected and ordered by different algorithms.

I also relate to an active theoretical literature on biased intermediation on platforms (Hagiu and Jullien, 2011; Hagiu and Jullien, 2014; Armstrong et al., 2009; Armstrong and Zhou, 2011; Reimers and Waldfogel, 2023; de Cornière and Taylor (2019); Bergemann and Bonatti (2023)). Theoretical contributions highlight consumer heterogeneity and the role of information in platform intermediation. de Cornière and Taylor (2019) outline a model where recommendations by the intermediary can either harm or benefit consumers, depending on whether the intermediaries' and the consumers' interests are conflicting or congruent. Heidhues, Köster and Kőszegi (2023) focus on how recommendations can affect consumers when they make mistakes. Mistakes in this model fall into two categories - either the consumer purchases something that they should not have purchased, or they do not purchase something that they should have purchased. However, this model abstracts from consumer search. Bourreau and Gaudin (2022) study biased intermediation on streaming platforms. In this setting, the user does not pay per song or movie they stream, but the platform bears a marginal cost for each stream. The platform chooses which items to recommend, and the subscription price. They show that the platform optimally steers consumers to items with lower marginal costs. In the model of Bergemann and Bonatti (2023), the platform has an informational advantage over the user - the platform knows the user-item match specific utility exactly, while the user is uncertain. In this setting, the platform matches products and users with targeted ads, and monetizes by charging sellers for the advertising. Reimers and Waldfogel (2023) develop a discrete choice framework that sheds light on how biased intermediation can be studied without data from inside the platform. They trace out the a welfare frontier, where total welfare is divided by between buyers and sellers on the platform, and label departures from this welfare frontier as biased intermediation. Using simulations, they illustrate that regressing rank on observable product characteristics as well as a dummy that equals to 1 if a product is sold by the platform (or "platform-preferred") does not produce reliable evidence for or against biased intermediation.

While I do not study biased intermediation theoretically, my experiment relies on these contributions in important ways. Firstly, as in de Cornière and Taylor (2019), I study a setting where alignment of interest between the user and the platform vary from user to user. In the experiment, users that are risk neutral or close to risk neutral have completely conflicting interests with the platform, while the interests of either very risk averse or risk seeking consumers are more aligned. I will be able to study if consumers are more or less likely to "make mistakes" or choose options the lead to a lower expected utility based on certain characteristics, as in Heidhues et al. (2023), and even though I abstract from "mistake-based" steering, the experimental setting could be extended to include such alternative steering algorithms. As in Bourreau and Gaudin (2022) I study a setting where user choose items at no cost. The choice between different recommendation systems, where the recommended items are free for the user is similar to platform competition between streaming services. For the design of the algorithms in the experiment, I rely extensively on Reimers and Waldfogel (2023). In my conceptual framework, I use their simple logit formulation of rank-dependent expected utility, and I rely on their results on consumer- and seller optimal algorithms for the design of my experiment. Lastly, mirroring Bergemann and Bonatti (2023), I assume that the platform knows the users preferences, and can therefore compute the (rank-independent) utility of each choice for each user.

The paper proceeds as follows. In Section 2, I describe a conceptual framework motivating the different recommendation algorithms used in the experiment. Section 3 describes the design of the experiment. In section 4, I show the results. Section 5 concludes.

2 Conceptual Framework

Just as consumers on e-commerce platforms, subjects in my experiment will be faced with lists of items. While in the platform case, these items are likely to be products (or songs or movies), in the case of my experiment, they are three-outcome lotteries. The outcomes of the lotteries are $x_1 = \$0$, $x_2 = \$5$ and $x_3 = \$10$, and I denote the corresponding probabilities as p_1, p_2 and p_3. I follow Reimers and Waldfogel (2023) logit approach in modeling the subject's utility. The subject chooses among L lotteries[1] based on the lotteries characteristics, and its ranking in the list r_j. I model the subject's utility of choosing a lottery j ranked at r_j as:

$$u_{ij} = \delta_{ij} + \gamma_i r_j + \xi_j + \epsilon_{ij} \tag{1}$$

where γ_i denotes the user-specific effect of rank on utility, ξ_j is are unobserved lottery attributes (for example, lotteries where the probability of one outcome is equal to 0 might be more easy to understand for subjects, leading to them being chosen more frequently), and ϵ_{ij} is an extreme-value error. Lastly, δ_{ij} denotes the rank-independent component of utility, which I choose to model using a constant relative risk aversion (CRRA) utility function:

$$\delta_{ij} = \begin{cases} \sum_{k=1}^{3} p(c_k) \frac{c_k^{1-\omega_i}}{1-\omega_i} & \text{if} \quad \omega \neq 1 \\ \sum_{k=1}^{3} p(c_k) ln(c_k) & \text{if} \quad \omega = 1 \end{cases}$$

The variable k indexes the three different outcomes (c_k) of lottery j, and ω_i is the individual-specific risk aversion. Given that there is no outside option in the experimental setting, the choice probability for a lottery j ranked at r_j can be computed as:

$$P_{ij} = \frac{e^{v_{ij}}}{\sum_{j \in J} e^{vij}}$$

with $v_{ij} = \delta_{ij} + \gamma_i r_j + \xi_j$. The ranking of a lottery affects the subjects utility of choosing a lottery, and therefore also the probability that the lottery will be chosen.

[1]Contrary to most demand modeling approaches, there is no outside option in the experiment.

2.1 Ranking algorithms

In the experiment, subjects face two different algorithms. The first algorithm, which I will call the expected utility algorithm, is fully aligned with the interests of the subjects and orders lotteries. The other algorithm, which I will call the platform algorithm, weighs consumer utility and the expected cost of providing a lottery to a subject.

Expected Utility Algorithm:

The expected utility algorithm ranks lotteries by the following index:

$$I_{ij} = \underbrace{\sum_{k=1}^{3} p(c_k) \frac{c_k^{1-\omega_i}}{1-\omega_i}}_{\delta_{ij}} \tag{2}$$

Reimers and Waldfogel (2023) show that this ordering by *rank-independent* expected utility is optimal for consumers.

Platform Algorithm:

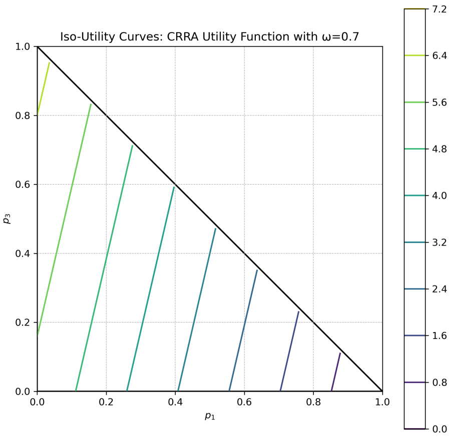
The platform algorithm ranks lotteries by the following index:

$$I'_{ij} = \kappa_1 S_u(\delta_{ij}) - \kappa_2 S_\pi \left(\sum_{k=1}^{3} p(c_k) c_k \right) \tag{3}$$

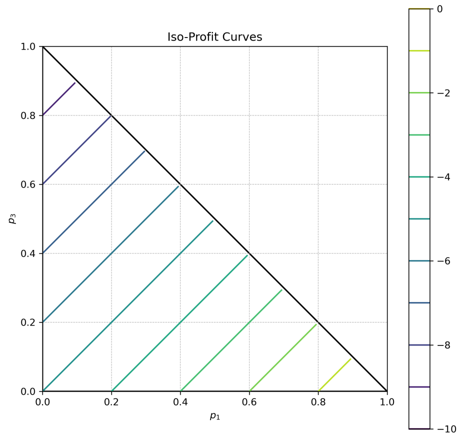
where κ_1, κ_2 are the weights on the consumer and the "platform" objective respectively, where $S_\pi(.)$ and $S_u(.)$ are min-max scalers: $S_u(x) = x \frac{1-\omega_i}{10^{1-\omega_i}}$ and $S_\pi(x) = \frac{x}{10}$. The minimums and maximums correspond to the profits and utilities in the case where the subject receives the degenerate lottery $(p_1 = 1, p_2 = 0, p_3 = 0)$ and $(p_1 = 0, p_1 = 0, p_3 = 1)$ respectively. Intuitively, these scaling functions are required, because otherwise the value that is associated with the utility of the subject could be much higher or lower than the number associated with the platforms profit, since it is scaled by their risk aversion. However, there is no reason why a platform should disadvantage users with a given set of preferences more than others, therefore the values are appropriately rescaled.

To give some intuition for how the two algorithms work, I plot iso-utility, iso-profit and iso-index curves in the Marschak-Machina triangle (Marschak, 1974, Machina, 1982). The case of a risk averse individual is displayed in figure 1. For a risk averse individual with a risk preference parameter of $\omega = 0.7$, the utility is increasing in the direction of the outcome 10$ with probability 100%. The iso-utility curves are steeply increasing, illustrating that the individual is willing to accept a lottery with a much lower chance of the highest outcome of 10$ in order to decrease the probability of the worst outcome 0$. The iso-profit curves that capture the objective of the platform are increasing in the direction of the outcome 0$ with probability 100%. In this setting, there is no way for the "platform" to gain anything, the best it can do is to minimize its losses. The iso-profit curves are 45 degree lines, illustrating that the platform is risk-neutral - all lotteries with equal expected value are equally costly to the platform in expectation. The iso-index curves plot equal levels of the index I'. The index reaches its highest level at the point where the subject receives the outcome 5$ with 100% probability. This is because the subject benefits from lower risk - ensuring that the probability of the worst outcome is close to 0 - which is costless for the platform to provide. The figure 2 illustrates further that the interests of the platform and the subject are not completely opposed. Consider a lottery at the intersection of the iso-index and iso-utility curves in the Marschak-Machina triangle. The area shaded in green shows the lotteries that are simultaneously preferred by the subject, but that are also going to be ranked higher than the lottery at the intersection of the iso-index and iso-utility curve in the graph. The red-shaded area marks the lotteries that would be ranked above the lottery at the intersection of the iso-index and iso-utility curve, but for which the lottery at the intersection would be preferred by the subject.

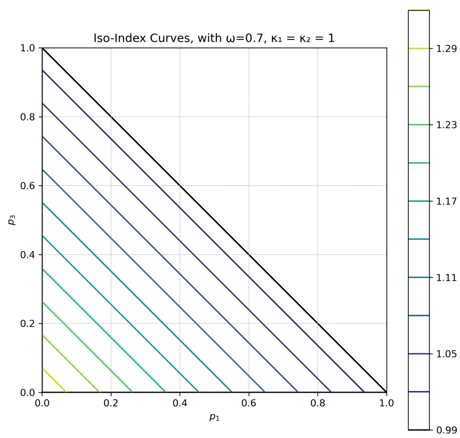
In a similar vein, figure 3 shows the iso-utility and iso-index curves for a risk seeking individual with a risk preference of $\omega_i = -2$. While the utility is still increasing in the direction of the lottery that pays the outcome 10$ with 100% probability, the slopes of the iso-utility curves are now flatter than 45 degree lines, indicating that the risk-seeking subject is willing to accept a higher probability of receiving the worst outcome of 0$ in



(a) CRRA iso-utility curves with $\omega = 0.7$



(b) Iso-profit curves



(c) Iso-index curves, $\omega = 0.7, \kappa_1 = \kappa_2 = 1$

Figure 1: Iso-utility, Iso-profit, and Iso-index curves with $\omega = 0.7$ (Risk-averse)

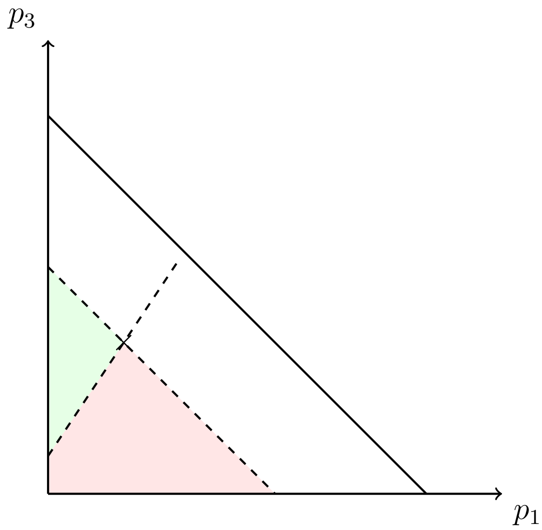
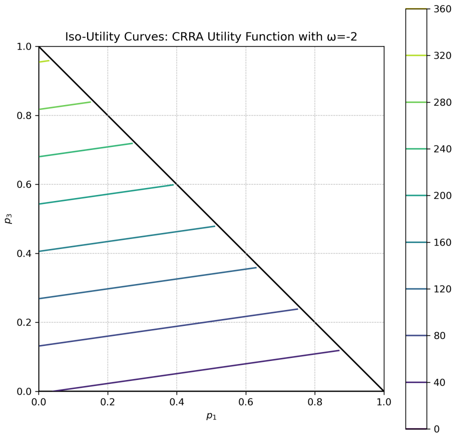
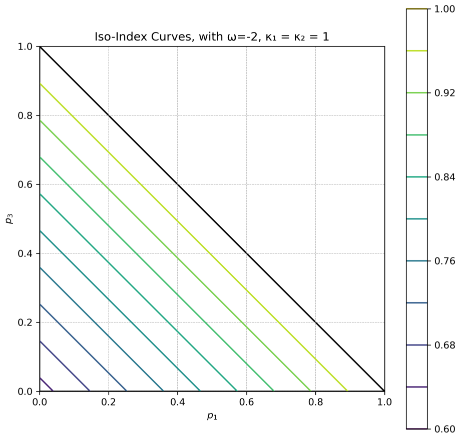


Figure 2: Iso-utility and iso-index curves for the lottery (0.4,0.4,0.2) with $\omega = 0.7$ (Risk-averse)

order to slightly increase the probability of the best outcome, 10$. Combining this with the the iso-profit curves that represent the objective of the platform, the index I' is now increasing when the lotteries are further from the lottery that pays 5$ with probability 100%, exactly opposite to the risk aversion case. In order to create a symmetric choice situation for risk seeking and risk averse individuals in the experiment, I constrain the set of lotteries to lotteries where $p_1 \leq 50\%$ and $p_3 \leq 50\%$. Figure 4 mirrors figure 2 in illustrating the lotteries that are preferred/not preferred by the subject and the index I' criterion respectively.



(a) CRRA iso-utility curves with $\omega = -2$



(b) Iso-index curves, $\omega = -2, \kappa_1 = \kappa_2 = 1$

Figure 3: Iso-utility, and Iso-index curves with $\omega = -2$ (Risk-seeking)

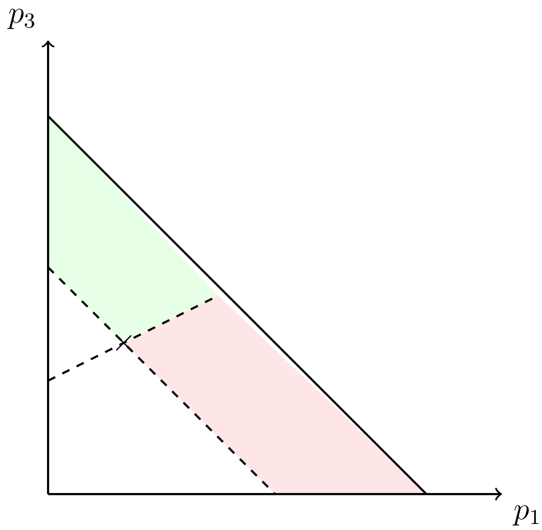


Figure 4: Iso-utiltiy and iso-index curves for the lottery (0.4, 0.4, 0.2) with $\omega = -2$ (Risk-seeking)

3 Experimental Design

The experiment features three kinds of tasks: a risk preference elicitation task, a choice task, and a willingness-to-pay elicitation task.

The risk preference elicitation task is in the style of Johnson, Baillon, Bleichrodt, Li, van Dolder and Wakker (2021). As shown in figure A9, subjects are presented with the prospect of gaining either the outcome of a lottery that pays \$10 or \$0 with equal probability or receiving a sure but unknown monetary amount X, randomly drawn with uniform probability between \$0 and \$10. Subjects are asked to declare a threshold such that they would prefer receiving X if it is above the threshold, and they would prefer receiving the lottery if X is below the threshold. The threshold of a risk neutral is \$5, whereas the thresholds of risk averse (risk seeking) individuals are lower (higher) than \$5.

The central task of this experiment is the choice task, which subjects complete 30 times. A screenshot of the choice task is shown in figure 5. It is designed to resemble choices from ordered lists on online platforms. For each round of this task, subjects are asked to choose from lists of three-outcome lotteries. The lotteries have been ordered by either the platform or the expected utility algorithm as described in section 2[1]. The algorithms also *lotteries* in the sense that only the top 10 lotteries according to the algorithms ranking criterion are visible to subject. These algorithms take into account the subjects risk aversion that has been calculated using their choice in the first task and assuming constant relative risk aversion[2]. The outcomes of all lotteries are 10\$, 5\$, and 0\$, and the respective probabilities are drawn to ensure that the probability of 10\$ and 0\$ does not exceed 50%. Subjects do not know the properties of the algorithms, but they are aware of the different algorithms that select and order the lotteries. Throughout the choice task, subjects see either "Algorithm 1" or "Algorithm 2" at the top of the page, colored green or blue (color is randomized across individuals). They can thus learn through there repeated interactions whether algorithm 1

[1]The experiment uses 30 sets of 20 lotteries each with probabilities ($p_1 \leq 0.5, \quad p_2 \leq 1, \quad p_3 \leq 0.5$)

[2]Since platform algorithm does not produce an ordering in case the subject is risk neutral, risk neutral subjects are randomly treated as either very slightly risk averse ($\omega = 0.01$) or slightly risk seeking ($\omega = -0.01$).

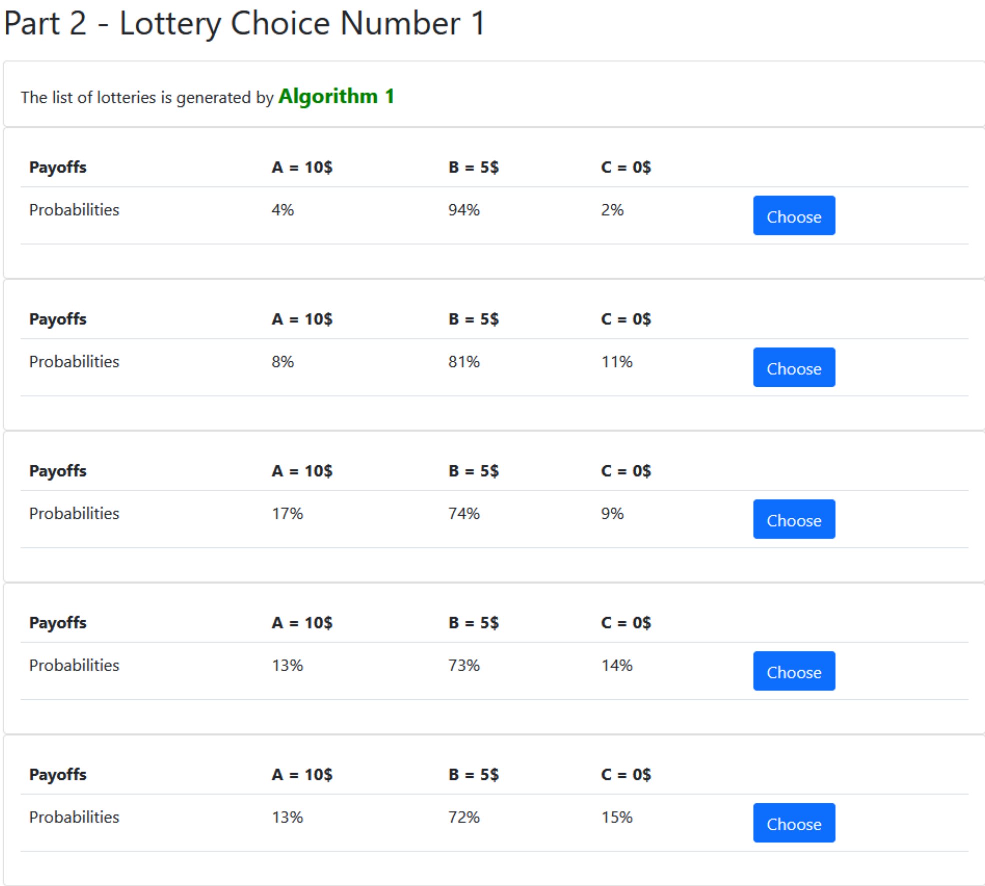


Figure 5: A screenshot of round 1 of the choice task for a risk-averse subject. The lotteries are selected and ordered by the platform algorithm which ranks safe but low expected value lotteries high. The subject sees that the lotteries in this round of the choice task are selected and ordered by algorithm 1.

or algorithm 2 creates better recommendations for them.

Finally, I elicit subjects willingness-to-pay for one algorithm over the other. For this step, I give subjects an endowment of \$1. Using a slider, subjects can change the probability that they will face either algorithm 1 or algorithm 2 in the next set of 10 choice rounds. Choosing equal probability between the two algorithms (50% - 50%) is costless. Changing the probabilities by 1%, i.e. to 49% - 51% costs \$0.02. The cost increases linearly, so that the cost of setting the probability to 1 to have either algorithm for the last 10 choice rounds is equal to the full endowment - \$1.

Timing	Description
t=1	Instructions • Consent form and general Instructions about the experiment - figure A1, A2 • Instructions on the risk elicitation task - figure A3, A4 • Instructions on the lottery choice task - figure A5 • Instructions on the Willingness-to-pay (WTP) elicitation mechanism - figure A6, A7
	Start of the main experiment - figure A8
t=2	Risk preference elicitation task - figure A9
t=3	10 rounds of choice task - figure A11 Choice task with either platform or expected utility algorithm (in case of platform algorithm, expected utility algorithm in the next set of 10 choice tasks and vice-versa; order randomized).
t=4	Buffer Screen informing subjects of a change in the algorithm - figure A12
t=5	10 rounds of choice task - figure A13
t=6	Willingness to pay elicitation - figure A14, A15 • Subjects receive an endowment of 1\$ • Subjects can choose the probabilities $(x,\ 1-x)$ that they face the expected utility algorithm or the platform algorithm respectively for the next set of 10 choice tasks. • The cost of changing the probability is equal to $p = \$(x-0.5)\times 2$
t=7	10 rounds of choice task With either expected utility of platform algorithm based on a random draw with the probabilities (x, 1-x).
t = 8	Survey - figure A16, A17
t = 9	Disclosure of the realization of the bonus payment, end of the experiment - figure A18

Table 1: Timing of the experiment

3.1 Incentives

When multiple lotteries are paid out, the subjects might realize that risk is diversified and exhibit higher risk appetite than in other situations (Azrieli, Chambers and Healy, 2018). I minimize this problem by randomly selecting one part to reward. As suggested by Plott and Zeiler (2005), subjects will be exposed to training and practice rounds with the elicitation mechanism in an anonymous fashion. The practice rounds will however not be incentivized, as this is at odds with the recommendation of Azrieli et al. (2018) to select one round to reward at random.

I divide the experiment into three parts, of which one is randomly selected for the bonus payment with equal probability. The three parts are risk preference elicitation ($t = 2$), the first 20 rounds of the choice task ($t = 3$, $t = 4$), and the willingness-to-pay elicitation task and the final set of choice rounds ($t = 5$, $t = 6$), as outlined in table 1.

In case the randomly selected part is part 1, subjects get a bonus according to their choice. In case the randomly drawn number X was above their threshold, they receive the X as the bonus, otherwise they receive the outcome of the lottery that pays 10$ with probability 50% (and 0$ with probability 50%).

If the randomly selected part is part 2, I select 1 of the 20 rounds at random and the subject gets the outcome of the lottery they chose in that round. For example, a subject may have chosen the lottery [P(10$)=0.03, P(5$)=0.9, P(0$)=0.07] in the randomly selected round. The outcome drawn based on these probabilities is the subjects bonus payoff.

Lastly, if the randomly selected part is part 3, I add the unspent endowment from the willingness-to-pay elicitation to their bonus payoff. Furthermore, I draw one of the last 10 choice rounds at random, and the outcome drawn based on the probabilities of the lottery that they chose in that round is added to the bonus payoff as well.

3.2 Hypotheses

The primitive data set consists of each participant's choice in the risk-preference elicitation task. Their choices in the 30 rounds of the choice task, as well as their elicited willingness-to-pay. For the choice task, the dataset contains the full list of lotteries that the subject was presented with in each round and their respective rank. I further compute the following values: using the participant's choice in the risk elicitation task, I compute their risk aversion assuming a CRRA utility function. Using the data from the choice task, I compute the average expected value of the chosen lotteries for each section of the experiment ($t = 3$,

$t = 4$, $t = 6$) and the corresponding average expected utility using the computed coefficient of risk aversion and assuming a CRRA utility function.

The first set of hypotheses concern consumer behavior and outcomes when they are subject to the two different algorithms. Since the goal of the variation between the two algorithms is that one is more aligned with the interests of the consumer than the other, the general hypothesis is that consumer outcomes are better when they are subject to the expected utility algorithm. The second set of hypotheses concern willingness-to-pay. I assume that willingness-to-pay for the expected utility algorithm is positive. I further assume that subjects who have close to risk neutral risk preferences have a higher willingness-to-pay for the expected utility algorithm that subjects that are very risk averse or very risk-seeking. For the full set of hypotheses, as pre-registered using the AEA RCT Registry with the ID AEARCTR-0014017, please refer to appendix A.2.

4 Results

The sample contains 298 individuals that have completed three sets of choice rounds each. For the analysis, the data therefore contains of 298 individuals, 894 sets of choice rounds and 8940 choice rounds. Given that subjects have 10 lotteries available to them in each choice rounds, the full dataset contains 89400 records, 1 chosen lottery for each choice round and 9 non-chosen lotteries. Furthermore, subjects completed the experiment in 21 minutes on average, and received an average bonus payment of $5.79 in addition to the completion fee of $4 for a total average reward of $9.79 ($27.97 per hour).

Of the 298 individuals, 113 have made a choice in the risk elicitation task that is consistent with risk-aversion, 66 have made a risk-neutral choice, and 119 have made a choice that is consistent with risk-seeking preferences.

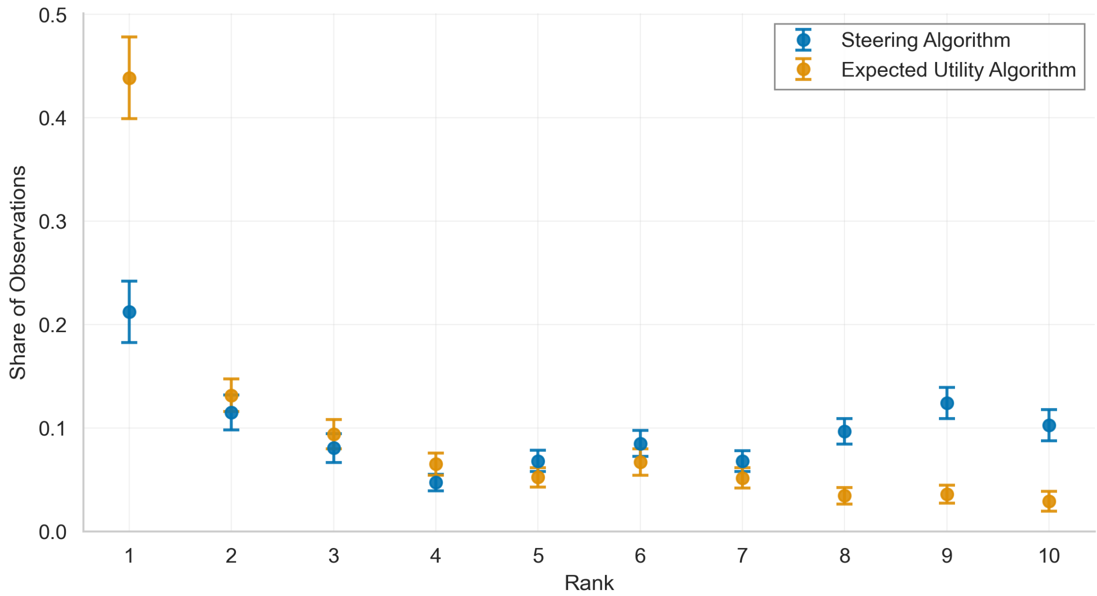


Figure 6: Share of choices with each rank conditional on algorithm, 1-10

4.1 Algorithms and consumer behavior

A comparison of means of the average expected value, average rank, and average variance of chosen lotteries are displayed in figures 7, 8, and 9. Overall, the findings are consistent with the hypothesis that subjects have better outcomes when they are faced with the expected utility algorithm.

Subjects choose on average lotteries with a lower expected payoff when they are faced with the platform algorithm than when they are faced with the expected utility algorithm. The average difference is equivalent to \$0.72, a difference of around 13%. Using a two-sided t-test, the average expected value with the platform algorithm and the expected utility algorithm are significantly different at the 1% level.

Secondly, subjects choose on average lotteries with a higher rank when they are faced with the platform algorithm than when they are faced with the expected utility algorithm. Using a two-sided t-test, the average expected value with the platform algorithm and the expected utility algorithm are significantly different at the 1% level. The difference in the average rank of chosen lotteries between the platform algorithm and the expected utility algorithm is

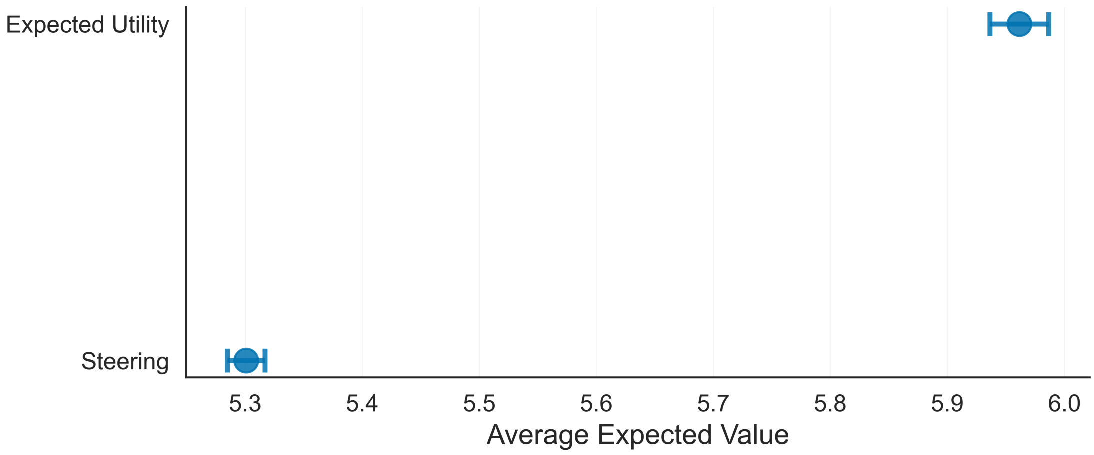


Figure 7: Difference in average expected value between chosen lotteries in the platform algorithm condition and chosen lotteries in the expected utility algorithm condition.

1.86. The average rank of lotteries chosen under the expected utility algorithm is somewhat surprising at 3.08. If subjects' risk aversion were correctly elicited in the first task, they should always prefer the top-ranked choice when faced with the expected utility algorithm. There are several potential explanations for this. Firstly, subjects might misunderstand the risk elicitation task. Secondly, the risk preferences of subjects might not be stable throughout the experiment. Third, subjects may make mistakes when comparing lotteries. Fourth, subjects may not search the list of lotteries from top to bottom. Fifth, the assumed CRRA utility function might not reflect the actual preferences of subjects.

Lastly, subjects choose on average lotteries with more extreme variances when they are faced with the steering algorithm, than when they are faced with the expected utility algorithm. Risk averse subjects choose lower-variance lotteries than risk-seeking subjects in both treatments. However, the gap between risk-seeking and risk-averse subjects substantially widens in the steering treatment. This is consistent with the algorithm offering to meet the subjects risk preference while at the same time reducing the expected value of the recommended lotteries.

All in all, these results confirm the hypothesis that subjects have better outcomes when they are faced with the expected utility algorithm. This means that the experiment successfully induced the desired variation that makes the abstract lottery choice comparable to consumer choices on online marketplaces. Firstly, subjects choose lotteries with a lower average expected value when they are faced with the platform algorithm. This means, that in the marketplace context, it is potentially profitable for the platform to use such an algorithm. Secondly, subjects chose lotteries with a higher ranked when faced with the platform algorithm. This suggests, that subjects understood some of the important differences between algorithms, namely that the best choice is more likely to be on top of the list when subjects face the expected utility algorithm.

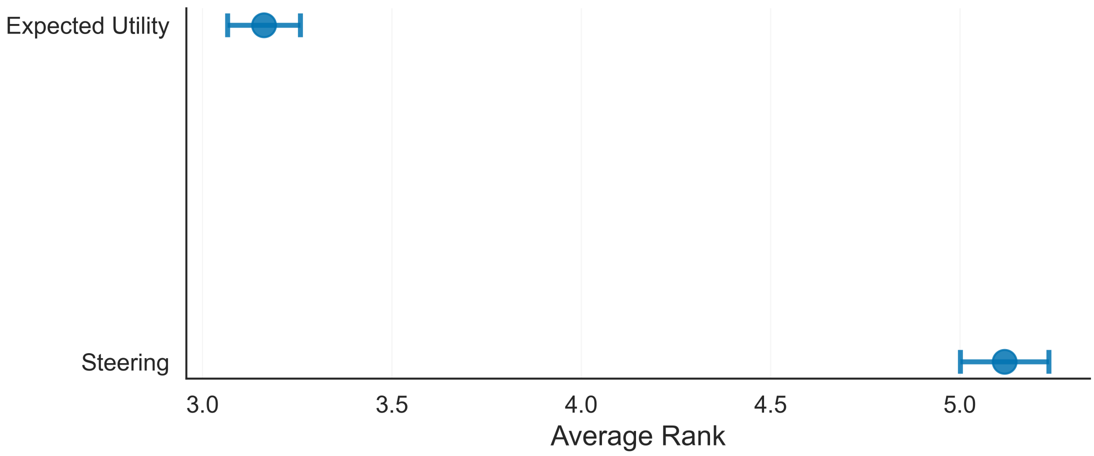


Figure 8: Difference in average rank between chosen lotteries in the platform algorithm condition and chosen lotteries in the expected utility algorithm condition.

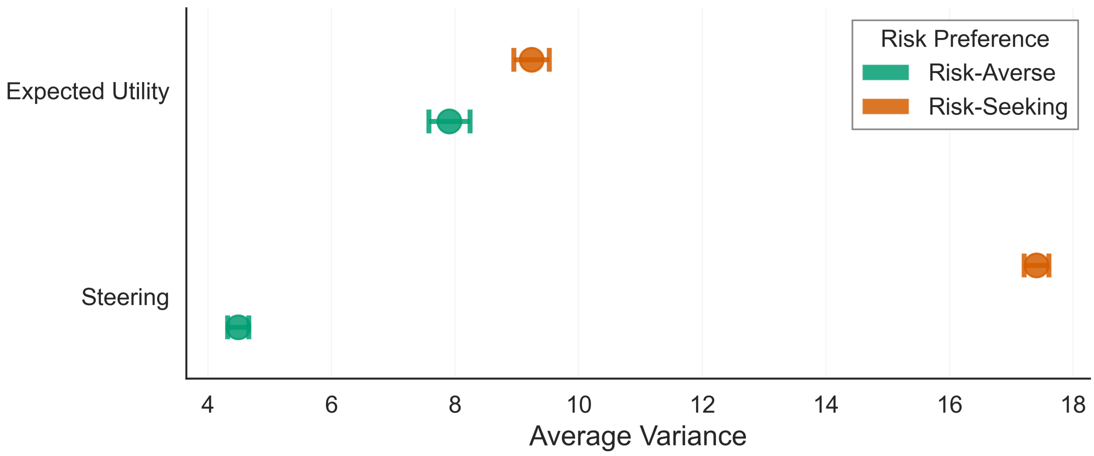


Figure 9: Difference in average variance between chosen lotteries in the platform algorithm condition and chosen lotteries in the expected utility algorithm condition.

To test the hypothesis that subjects that are close to risk neutral have worse outcomes than subjects that are either very risk averse or risk neutral, I estimate the following regression, where ω denotes the elicited coefficient of risk aversion:

$$\bar{y}_{is} = \alpha + \beta_1\omega + \beta_2\omega^2 + \mathbb{1}(A^s = \text{Platform})[\theta_0 + \theta_1\omega + \theta_2\omega^2]$$

	(1)	(2)
ω	-0.098***	-0.209***
	(0.023)	(0.022)
ω^2	-0.008**	-0.004
	(0.003)	(0.008)
Platform algorithm	-0.677***	-0.681***
	(0.028)	(0.029)
Platform $\times\omega$	0.087***	0.187***
	(0.021)	(0.019)
Platform $\times\omega^2$	0.006**	0.005
	(0.003)	(0.007)
Constant	5.982***	5.983***
	(0.032)	(0.033)
CE Restricted	No	Yes
Observations	5960	5700
R-squared	0.295	0.342

Standard errors (in parentheses) are clustered at the participant level.

CE Restricted: Sample limited to CE $\in$ [0.5, 9.5].

*** p<0.01, ** p<0.05, * p<0.1

Table 2: Regression results: Average Expected Value

The regressions show that across outcomes, risk neutral subjects fare worse when they are faced with the platform algorithm than their risk-averse or risk-seeing counterparts. The interaction term of the platform algorithm dummy and the squared risk aversion coefficient is positive, indicating that risk-averse or risk-seeking subjects have higher payoffs than risk-neutral subjects. Consequentially, risk-averse and risk-seeking subjects also choose lotteries with a lower rank when faced with the platform algorithm, indicating that the platform algorithm is indeed more aligned with their preferences than the ones of risk-neutral subjects.

	(1)	(2)
ω	-0.074	0.072
	(0.052)	(0.091)
ω^2	0.011*	0.028
	(0.007)	(0.036)
Platform algorithm	2.089***	2.151***
	(0.211)	(0.233)
Platform $\times\omega$	-0.298***	-0.755***
	(0.105)	(0.138)
Platform $\times\omega^2$	-0.036**	-0.054
	(0.015)	(0.061)
Constant	3.079***	3.040***
	(0.110)	(0.122)
CE Restricted	No	Yes
Observations	5960	5700
R-squared	0.137	0.153

Standard errors (in parentheses) are clustered at the participant level.

CE Restricted: Sample limited to CE $\in$ [0.5, 9.5].

*** p<0.01, ** p<0.05, * p<0.1

Table 3: Regression results: Average Rank

4.2 Willingness-to-pay

In the following section, I study the elicited willingness-to-pay for the expected utility algorithm. I elicit willingness-to-pay by letting subjects choose the probability with which they face either the expected utility or the platform algorithm for the last set of 10 choice rounds. Subjects receive an endowment of $1. Choosing equal probability between the two algorithms is costless, meaning that if this section of the experiment is selected for their final payoff, the subjects get $1 in addition to the outcome of a randomly selected lottery that they chose in the final set of 10 choice rounds. Choosing 100% probability for either the expected utility or the platform algorithm costs the full endowment of $1. For intermediate probabilities, the cost increases linearly by $0.02 for each percentage point.

The distribution of elicited willingness to pay is plotted in figure 11. The willingness-to-pay is coded to lie between -1 and 1, where -1 is a willingness to pay of $1 for the platform algorithm, and 1 is a willingness to pay of $1 for the expected utility algorithm. The

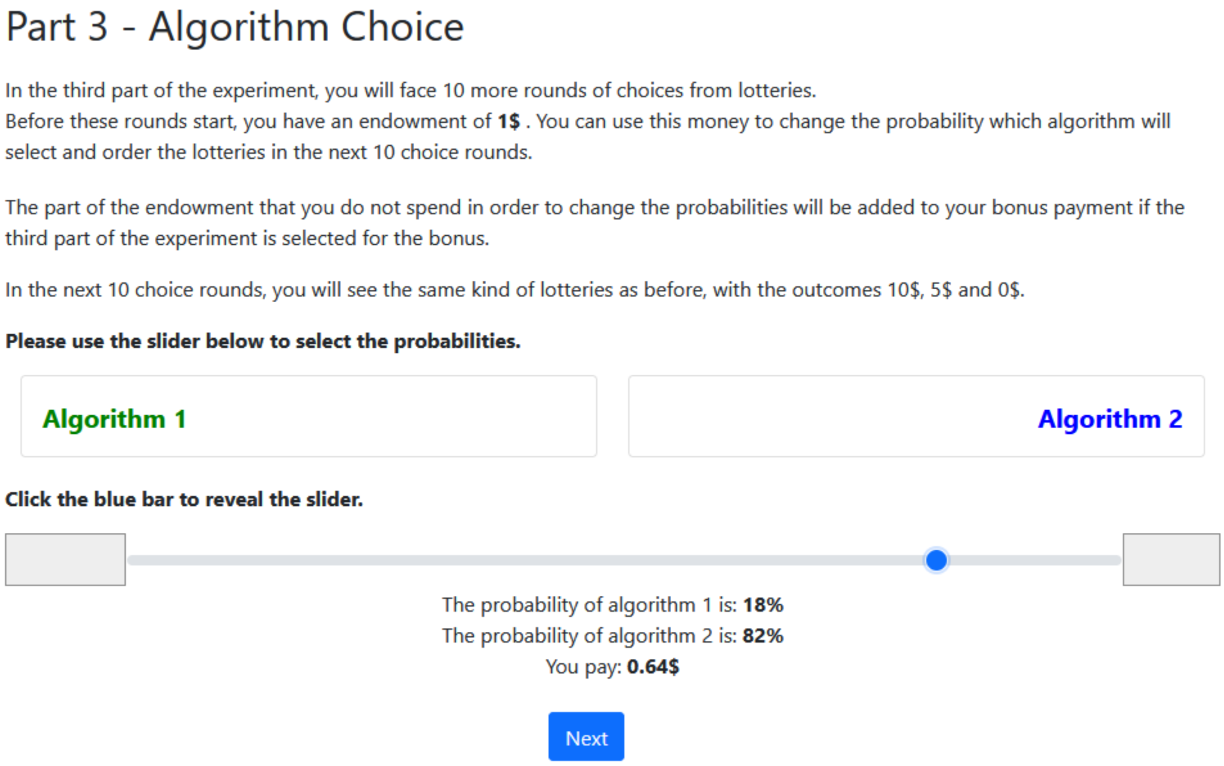


Figure 10: Screenshot - Choice of probability of algorithm 1 or algorithm 2 for the last 10 choice rounds

graph shows that the distribution has a large mass point at 0. 60 subjects selected equal probability in the willingness to pay elicitation, corresponding to a willingness to pay of 0. 150 subjects selected a positive willingness to pay, and 23 subjects selected a willingness to pay of \$1 for the expected utility algorithm. 88 subjects selected a negative willingness to pay, meaning they are willing to give up part of their endowment in order to increase the probability that they face the platform algorithm in the next set of choice rounds. The average willingness-to-pay is \$0.13.

In order to study how the willingness-to-pay is affected by risk preference and the individual benefit from having the expected utility algorithm, I create a measure of the average benefit that a subject derives from having the expected utility algorithm compared to the platform algorithm. I divide the benefit into two parts: a monetary part that is due to choosing better lotteries when faced with the expected utility algorithm, and a part that captures

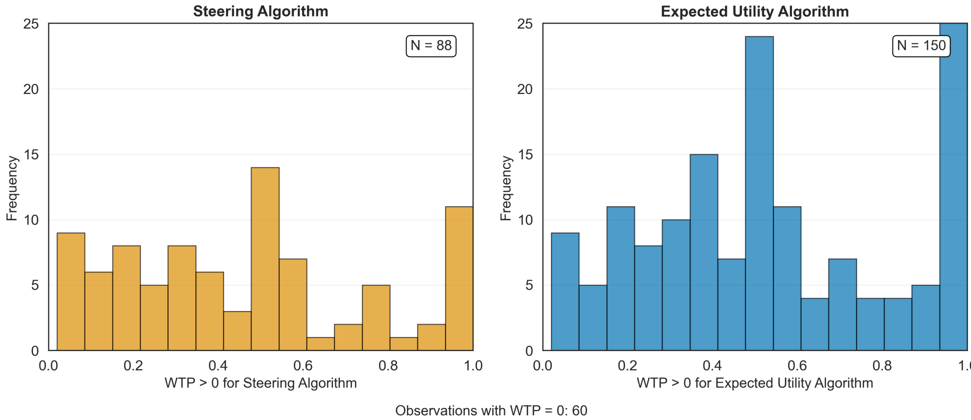


Figure 11: Distribution of elicited willingness-to-pay

the smaller cognitive costs of being faced with the expected utility algorithm rather than with the platform algorithm.

To calculate the individual expected utility algorithm benefit, I take the individual difference in average value of lotteries chosen when this subject is faced with the expected utility algorithm compared to when they are faced with the platform algorithm $\Delta\text{Expected Value}_i = \bar{\text{Expected Value}}_i^{EU} - \bar{\text{Expected Value}}_i^{P}$, where EU and P denote the expected utility and the steering algorithm respectively. I proxy the cognitive cost difference by the individual average difference in time spent on the page, $\Delta s_i = \bar{s}_i^{EU} - \bar{s}_i^{P}$, where $\bar{s}_i^{P}$, $\bar{s}_i^{EU}$ are the average number of seconds spent on a choice page when faced with the platform algorithm or the expected utility algorithm respectively. I scale this time difference by the 10 rounds that comprise the last set of choice rounds and assume that subjects have an opportunity cost of 20$ (the average payoff of this experiment is equivalent to 27.97$). The total individual benefit of having the expected utility algorithm rather than the platform algorithm therefore is:

$$\text{Time Value (\$)} = \frac{20 \times 10 \times \Delta s_i}{60 \times 60}$$

I then estimate the regression:

$$WTP_i = \alpha + \beta_1 \text{Time Value (\$)} + \beta_2 \Delta\text{Expected Value} + \beta_3 \omega_i + \beta_4 \omega_i^2 + \epsilon$$

Variable	Coefficient	Std. Error
Constant	0.0352	(0.0559)
ω	0.0120	(0.0225)
ω^2	0.0001	(0.0030)
Time Value ($)	0.0040	(0.0029)
Δ Expected Value	0.1516**	(0.0700)
Observations	298	
R^2	0.0215	
Adjusted R^2	0.0082	

Robust standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01

Table 4: Willingness to Pay for Algorithm

The results of the regression of willingness to pay on ΔExpected Value, Time Value as well as ω and ω^2 are displayed in table 4. The results suggest that a higher individual expected gain ΔExpected Value from having the expected utility algorithm positively correlates with a higher willingness to pay. However, the increase in willingness to pay is modest compared to the expected gain - an increase in the expected gain of 1$ increases willingness to pay only by $\sim$ 0.15$. The coefficient for Time Value is not significantly different from zero, suggesting that subjects did not take time savings into account when determining their willingness to pay for the better algorithm. The coefficient for ω^2 is a close to zero, suggesting that subjects with extreme risk preferences are not more or less willing to pay in order to have the expected utility algorithm, even though compared to risk-neutral subjects, their preferences are more aligned with the platform algorithm. In this regression, the coefficient for the risk preference parameter ω is also negative and significant at the 1% level. This may be an artifact of the willingness to pay elicitation method, where subjects are asked to give up part of an endowment in order to change the probability to have the expected utility algorithm for the next set of choice rounds. Risk averse individuals may simply be less willing to give up the sure endowment, and therefore have a lower willingness to pay. Extreme values of risk aversion seem to have at most a very small effect on willingness to pay. The regression reveals that subjects willingness-to-pay increases with the individual expected gain, but by a comparatively small amount.

5 Conclusion

In this paper, I develop an online experiment in order to study the effect of recommendation algorithms on consumer choices. The experiment consists of two parts: a first, standard risk elicitation task, and a choice task where the subjects are asked to choose from lists of lotteries, where the lists have been selected and ordered by the different recommendation algorithms. This way of designing the experiment has the important advantage, that no prior information about the characteristics of the subjects, or any previous choice data is necessary in order to inform the recommendation algorithms. The recommendation algorithms can simply use the elicited risk preference and with some assumption use this parameter to estimate the expected utility for all lotteries. The expected utility can than be combined with other information in order to create an index according to which lotteries are selected and ranked. I then use this experimental framework to study the question of whether subjects have a positive willingness to pay for a better recommendation algorithm. It is important to point out, that this experimental framework can in principle be applied to study many other questions related to recommendation systems, including the role of complexity, what happens when information about the algorithm is disclosed, and how subjects search when they are faced with recommendation systems.

Regarding the willingness to pay for the better (expected utility) algorithm, I find that it is positive on average. Individuals that derive a higher benefit from having the expected utility algorithm have a higher willingness to pay for it - however, subjects are only willing to pay \$0.15 for an additional \$1 of expected gain from having the better algorithm. Overall, the results from this experiment suggest, that from a regulatory perspective, giving consumers the option between different recommendation systems may not be ideal, since subjects seem to underestimate the benefit from having the better recommendation algorithm. However, it could be viable for platform businesses to offer better recommendations for a fee, or to offer the choice of different recommendation systems in order to differentiate themselves from competitors[1]. Future studies could focus on the role of information that subjects have when choosing between algorithms. In this experiment, the only information that helps subjects in their choice comes from their experience with the two algorithms. It would be interesting how choices vary with additional information about the algorithm behavior, and whether subjects are willing to incur costs to learn this information.

[1]The platform Bluesky is essentially offering different recommendation systems by allowing users to subscribe to different "feeds".

References

Armstrong, Mark and Jidong Zhou, "Paying for Prominence," *The Economic Journal*, November 2011, *121* (556), F368–F395.

_ , John Vickers, and Jidong Zhou, "Prominence and consumer search," *The RAND Journal of Economics*, 2009, *40* (2), 209–233.

Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy, "Incentives in Experiments: A Theoretical Analysis," *Journal of Political Economy*, August 2018, *126* (4), 1472–1503.

Bergemann, Dirk and Alessandro Bonatti, "Data, Competition, and Digital Platforms," April 2023.

_ , Jacques Cremer, David Dinielli, Carl-Christian Groh, Paul Heidhues, Maximilian Schafer, Monika Schnitzer, Fiona M. Scott Morton, Katja Seim, and Michael Sullivan, "Market Design for Personal Data," *Yale Journal on Regulation*, 2023, *40*, 1056.

Boldrini, Michela and Francesco Clavorà Braulin, "Taming Tech Giants's Algorithms: An analysis of consumer's choices on Amazon.pdf," November 2025.

Bourreau, Marc and Germain Gaudin, "Streaming platform and strategic recommendation bias," *Journal of Economics & Management Strategy*, 2022, *31* (1), 25–47.

Caplin, Andrew, Mark Dean, and Daniel Martin, "Search and Satisficing," *American Economic Review*, December 2011, *101* (7), 2899–2922.

Chen, Jiawei, Hande Dong, Yang Qiu, Xiangnan He, Xin Xin, Liang Chen, Guli Lin, and Keping Yang, "AutoDebias: Learning to Debias for Recommendation," in "Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval" SIGIR '21 Association for Computing Machinery New York, NY, USA July 2021, pp. 21–30.

Chen, Yiting, Ziye Wu, and Songfa Zhong, "Risk-Taking Behavior under AI-Based Recommendation," April 2023.

de Cornière, Alexandre and Greg Taylor, "A model of biased intermediation," *The RAND Journal of Economics*, 2019, *50* (4), 854–882.

Derakhshan, Mahsa, Negin Golrezaei, Vahideh Manshadi, and Vahab Mirrokni, "Product Ranking on Online Platforms," *Management Science*, June 2022, *68* (6), 4024–4041.
Digital Markets Act

***Digital Markets Act**, September 2022.*

Donelly, Robert, Ayush Kanodia, and Ilya Morozov, *"Welfare Effects of Personalized Rankings," mimeo, March 2023.*

Düll, Adrian, Heiko Karle, Simon Martin, and Heiner Schumacher, *"Evaluating Search Cost Models: Estimation and Prediction," 2025.*

Esber, Jad, Scott Duke Kominers, Leora Kornfeld, Scott Belsky, Apurva Chitnis, K. Chmielinski, John Deighton, David Eliot, Carolina Almeida Antunes Rossini, Doc Searls, Kinjal Shah, and Reid Williams, *"The Next Phase of the Data Economy: Economic & Technological Perspectives," November 2024.*

Farronato, Chiara, Andrey Fradkin, and Alexander MacKay, *"Self-Preferencing at Amazon: Evidence from Search Rankings," AEA Papers and Proceedings, May 2023, 113, 239–243.*

Fong, Jessica, Olivia R. Natan, and Ranmit Pantle, *"Consumer Inferences from Product Rankings: The Role of Beliefs in Search Behavior," September 2024.*

Fukuyama, Francis, Barak Richman, Ashish Goel, Roberta R Katz, A Douglas Melamed, and Marietje Schaake, *"Middleware for Dominant Digital Platforms: A Technological Solution to a Threat to Democracy," 2020.*

Ghose, Anindya, Panagiotis G. Ipeirotis, and Beibei Li, *"Examining the Impact of Ranking on Consumer Behavior and Search Engine Revenue," Management Science, July 2014, 60 (7), 1632–1654.*

Greminger, Rafael P., *"Optimal Search and Discovery," Management Science, May 2022, 68 (5), 3904–3924.*

_ , *"Heterogeneous Position Effects and the Power of Rankings," December 2023.*

Hagiu, Andrei and Bruno Jullien, *"Why do intermediaries divert search?," The RAND Journal of Economics, 2011, 42 (2), 337–362.*

_ **and** _ , *"Search diversion and platform competition," International Journal of Industrial Organization, March 2014, 33, 48–60.*

Heidhues, Paul, Mats Köster, and Botond Kőszegi, *"Steering Fallible Consumers," The Economic Journal, May 2023, 133 (652), 1430–1465.*

Johnson, Cathleen, Aurélien Baillon, Han Bleichrodt, Zhihua Li, Dennie van Dolder, and Peter P. Wakker, *"Prince: An improved method for measuring incentivized preferences," Journal of Risk and Uncertainty, February 2021, 62 (1), 1–28.*

Kaye, Aaron, *"The Personalization Paradox: Welfare Effects of Personalized Recommendations in Two-Sided Digital Markets," January 2024.*

Lam, Tai H., *"Platform search design and market power," Job Market Paper, 2021, Northwestern University.*

Lee, Kwok Hao and Leon Musolff, *"Entry Into Two-Sided Markets Shaped By Platform-Guided Search," September 2021.*

Lin, Zinan, Dugang Liu, Weike Pan, and Zhong Ming, *"Transfer Learning in Collaborative Recommendation for Bias Reduction," in "Proceedings of the 15th ACM Conference on Recommender Systems" RecSys '21 Association for Computing Machinery New York, NY, USA September 2021, pp. 736–740.*

los Santos, Babur De and Sergei Koulayev, *"Optimizing Click-Through in Online Rankings with Endogenous Search Refinement," Marketing Science, July 2017, 36 (4), 542–564.*

Machina, Mark J., *""Expected Utility" Analysis without the Independence Axiom," Econometrica, 1982, 50 (2), 277–323.*

Marschak, Jacob, *"Rational Behavior, Uncertain Prospects, and Measurable Utility (1950)," in Jacob Marschak, ed., Economic Information, Decision, and Prediction: Selected Essays: Volume I Part I Economics of Decision, Dordrecht: Springer Netherlands, 1974, pp. 5–39.*

Natan, Olivia R., *"Choice Frictions in Large Assortments," October 2024.*

Peitz, Martin, *"How to Apply the Self-Preferencing Prohibition in the DMA," Journal of European Competition Law & Practice, July 2023, 14 (5), 310–315.*

Plott, Charles R. and Kathryn Zeiler, *"The Willingness to Pay-Willingness to Accept Gap, the "Endowment Effect," Subject Misconceptions, and Experimental Procedures for Eliciting Valuations," American Economic Review, June 2005, 95 (3), 530–545.*

Posner, Eric A. and E. Glen Weyl, *Radical Markets: Uprooting Capitalism and Democracy for a Just Society, Princeton University Press, May 2018.*

Reimers, Imke and Joel Waldfogel, *"A Framework for Detection, Measurement, and Welfare Analysis of Platform Bias," October 2023. Series: Working Paper Series.*

Resnick, Paul and Hal R Varian, *"Recommender systems," COMMUNICATIONS OF THE ACM, 1997, 40 (3).*

Ursu, Raluca M., *"The Power of Rankings: Quantifying the Effect of Rankings on Online Consumer Search and Purchase Decisions," Marketing Science, August 2018, 37 (4), 530–552.*

Zhang, Xiaochen, Pedro Ferreira, Miguel Godinho de Matos, and Rodrigo Belo, *"Welfare Properties of Profit Maximizing Recommender Systems: Theory and Results from a Randomized Experiment," Management Information Systems Quarterly, March 2021, 45 (1), 1–45.*

A.1 Experiment Design

Consent Form

Welcome to the ARS experiment.

We are inviting you to participate in an online economics experiment about individual decision-making. This study is being led by Felix Schleef, Department of Economics, ENSAE-CREST-IP Paris, and has been approved by the Institutional Review Board of the Institut Louis Bachelier (https://www.institutlouisbachelier.org/en/louis-bachelier-institutional-review-board-2/).

To successfully participate in the experiment, you must complete the experimental session during which you will be asked to make decisions. The experiment will start on the next page. To complete the session, you must answer all questions and perform all mandatory tasks. We expect today's session to last approximately 20 minutes.

The experiment is taking place on this website (https://ars-experiment-9e8858efe794.herokuapp.com/room/ars_prolific). If you complete the full experimental session you will receive a completion fee of 4$, as well as a bonus reward that will depend on your decisions during the experiment (and potentially on luck). The total reward including this bonus will be approximately equal to 10$ on average, and can be as high as 15$. The payout will be made to you in the next two to three days via the Prolific system. If you get screened out before the end of the experiment, you will receive a participation fee of 1$.

You are free to refuse to participate in this experiment, or to remove your consent and drop out of the experiment at any point in time. However, if you do not complete the experiment, you will not receive any payout. Therefore, please start this experiment only if you can commit to finishing the session.

The data collected during the experiment will be stored in a database. No identifying information will be collected apart from your Prolific ID, which we need to reward you with the bonus payout. Your Prolific ID will be irrevocably deleted from the database as soon as the data collection is over. The resulting anonymized data set (that is, without your Prolific ID or any other identifying information) might be made public in the future if a scientific article based on it is published.

We anticipate that your participation in this experiment presents no greater risk than everyday use of the Internet.

If you have questions about this study, you may contact the lead researcher at felix.schleef@ensae.fr. If you have any questions, concerns, or complaints regarding your rights as a subject in this experiment, you can contact the Institutional Review Board (IRB) at irb@institutlouisbachelior.org.

By checking the box below, you confirm that you have read and understood the above instructions, and that you consent in participating in the experiment.

I understand the instructions and consent in participating in the experiment:

☐

Next

Figure A1: Screenshot - Consent Form

Tutorial - Introduction

Welcome to the experiment! Before we dive into the main tasks, we'll guide you through a brief tutorial. This is a crucial step to familiarize yourself with the choice mechanisms you'll encounter and understand how these choices can influence your bonus rewards.

In the next few screens, you'll encounter simplified scenarios resembling the decisions you'll make during the experiment.
Please note: The choices you make in this tutorial section will not affect your bonus payout. They are designed solely to help you understand the different mechanisms at play.

Let's ensure you feel comfortable with the process and fully prepared to maximize your bonus potential. Ready to begin?

Next

Figure A2: Screenshot - Tutorial Introduction

Tutorial - Lottery vs Safe Option

This is an example of one of the choices that you will be asked to make in the experiment. The choice you make on this screen is not relevant for your bonus payoff, because it is part of the tutorial.

In this part of the experiment, you receive either a **lottery** or a fixed amount of money **X$**.
The **lottery** pays 10$ with 50% probability and 0$ with 50% probability.
The value of **X$** is a random number between 0$ and 10$ that was selected before the start of the experiment.
Please give us instructions by setting a **threshold** such that:

- "If the money amount **X$** is equal to or above the **threshold**, then give me **X$**."
- "If the money amount **X$** is below the **threshold** then give me the **lottery**."

Lottery	10$ with 50% probability, 0$ with 50% probability
Fixed Amount	X$ with 100% probability

On this page, you are asked to set a threshold so that your preferred choice gets implemented. Consider different values that X could take. If X was 6$, would you prefer receiving X, or would you prefer receiving the lottery? If you would prefer having 6$, you should set your threshold below 6$. Similarly, consider if X was 2$ - if you would prefer to receive the lottery, you should set your threshold above 2$. Click on the blue bar below to reveal the slider, and drag it around to select your threshold.

Click the blue bar to reveal the slider and set your threshold value.

Next

Figure A3: Screenshot - Tutorial - Risk Preference Elicitation

Tutorial - Lottery vs. Safe Option - Results

This is how the result for this choice task could look like:

The value X was 8.9$. Your selected threshold was 3.5$. X is above your threshold, therefore your bonus payoff from this part would have been 8.9$.

Next

Figure A4: Screenshot - Tutorial - Risk Preference Elicitation Result

Tutorial - Lottery Choice Number

On this page, you have the choice between different lotteries. Clicking "Choose" will select this lottery and take you to the next screen.
Consider the first lottery: you will receive 10$ with a probability of 50%; 5$ with a probability of 30% and 0$ with a probability of 20%. These probabilities mean that if 10 people play, on average 7 will receive 10$; 3 will receive 5$ and 2 will receive 0$.
The choice you make on this screen is not relevant for your bonus payoff, because it is part of the tutorial.

Payoffs	**A = 10$**	**B = 5$**	**C = 0$**	
Probabilities	50%	30%	20%	Choose

Payoffs	**A = 10$**	**B = 5$**	**C = 0$**	
Probabilities	48%	20%	32%	Choose

Payoffs	**A = 10$**	**B = 5$**	**C = 0$**	
Probabilities	72%	10%	18%	Choose

Payoffs	**A = 10$**	**B = 5$**	**C = 0$**

Figure A5: Screenshot - Tutorial - List Choice

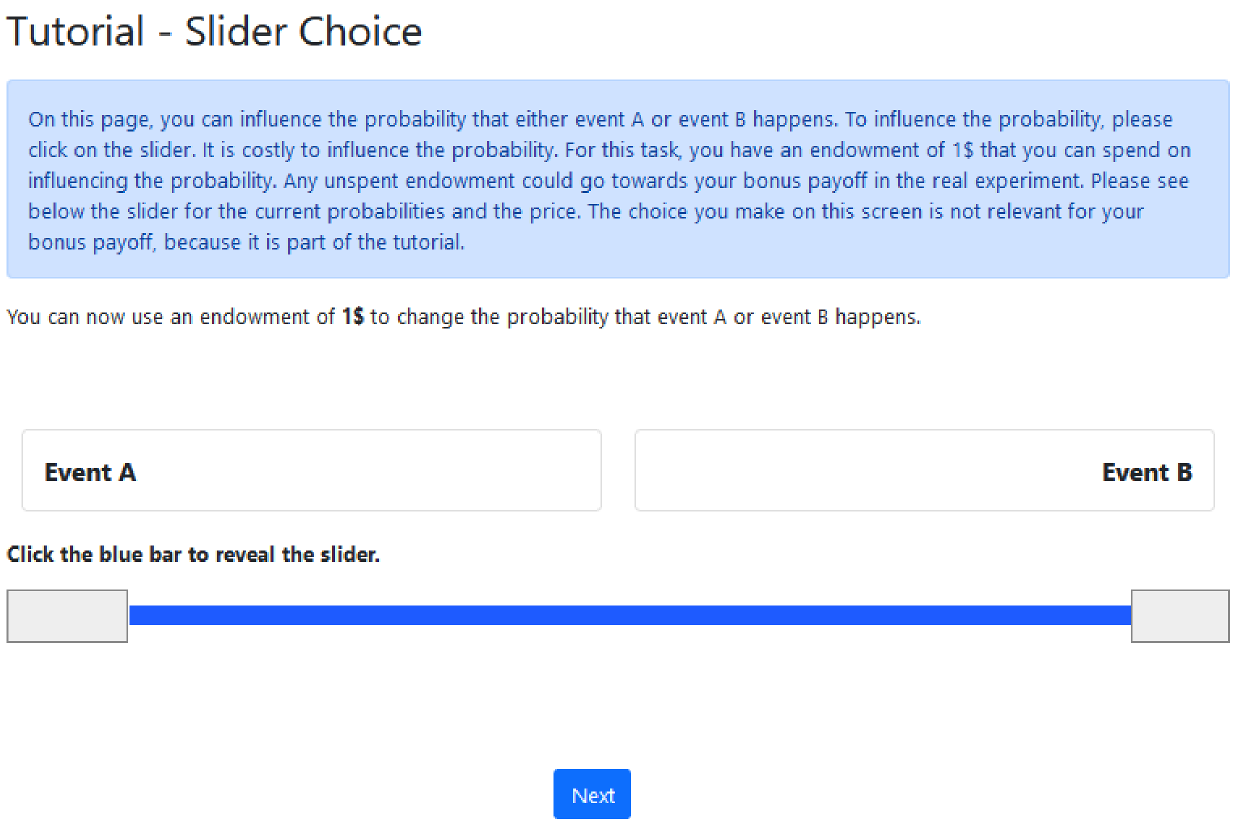


Figure A6: Screenshot - Tutorial - Slider Choice

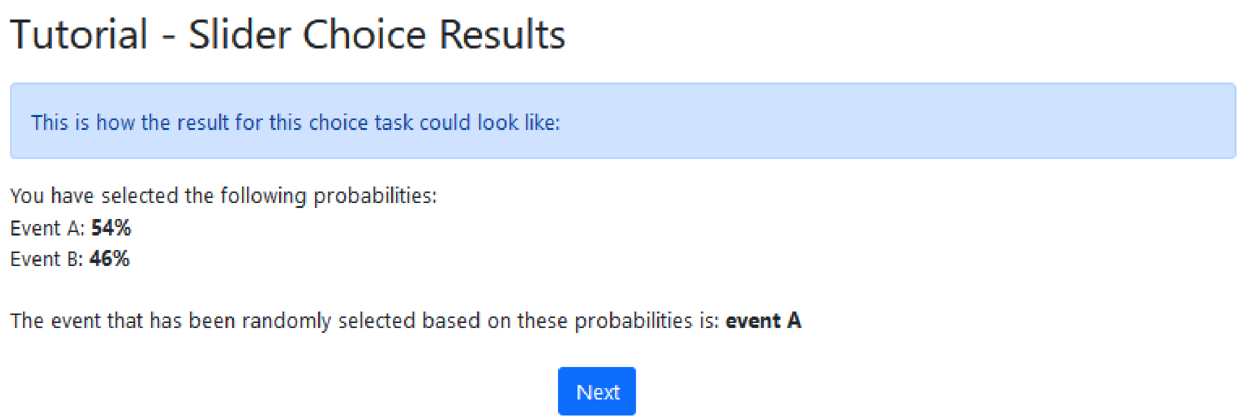


Figure A7: Screenshot - Tutorial - Slider Choice Result

Introduction

This experiment consists of four parts. In the three first parts, you will be asked to make different choices.
All choices that you make in the first three parts are relevant for your bonus payoff.
The last part is a survey, which is not relevant for your payoff.
You will only be paid if you complete the full experiment. For this reason, please only start the experiment if you have at least 20 minutes to complete it.

Next

Figure A8: Screenshot - Experiment Introduction

Part 1 - Lottery vs Safe Option

In this part of the experiment, you receive either a **lottery** or a fixed amount of money **X$**.
The **lottery** pays 10$ with 50% probability and 0$ with 50% probability.
The value of **X$** is a random number between 0$ and 10$ that was selected before the start of the experiment.
Please give us instructions by setting a **threshold** such that:

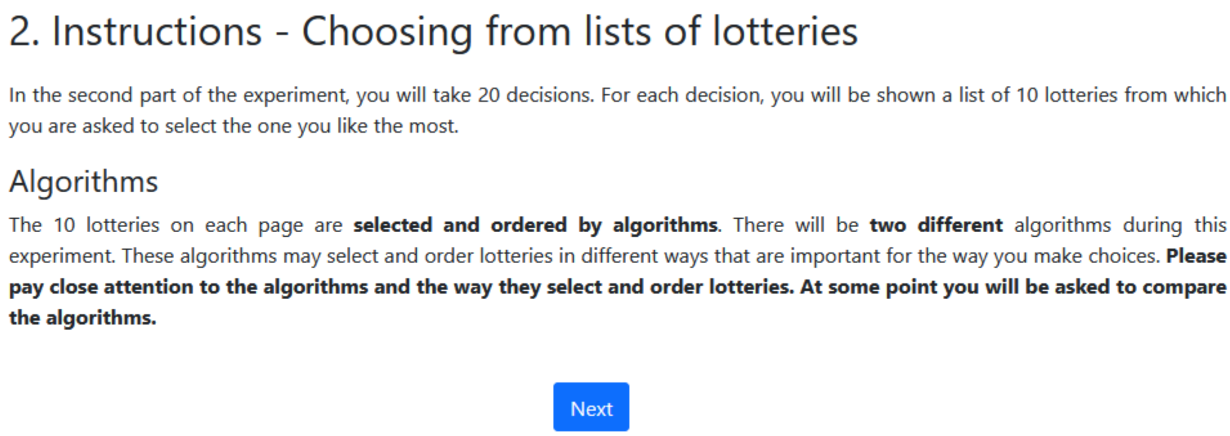
- "If the money amount **X$** is equal to or above the **threshold**, then give me **X$**."
- "If the money amount **X$** is below the **threshold** then give me the **lottery**."

Lottery	10$ with 50% probability, 0$ with 50% probability
Fixed Amount	X$ with 100% probability

Click the blue bar to reveal the slider and set your threshold value.

Next

Figure A9: Screenshot - Risk preference elicitation stage



2. Instructions - Choosing from lists of lotteries

In the second part of the experiment, you will take 20 decisions. For each decision, you will be shown a list of 10 lotteries from which you are asked to select the one you like the most.

Algorithms

The 10 lotteries on each page are **selected and ordered by algorithms**. There will be **two different** algorithms during this experiment. These algorithms may select and order lotteries in different ways that are important for the way you make choices. **Please pay close attention to the algorithms and the way they select and order lotteries. At some point you will be asked to compare the algorithms.**

Next

Figure A10: Screenshot - Introduction choice experiment

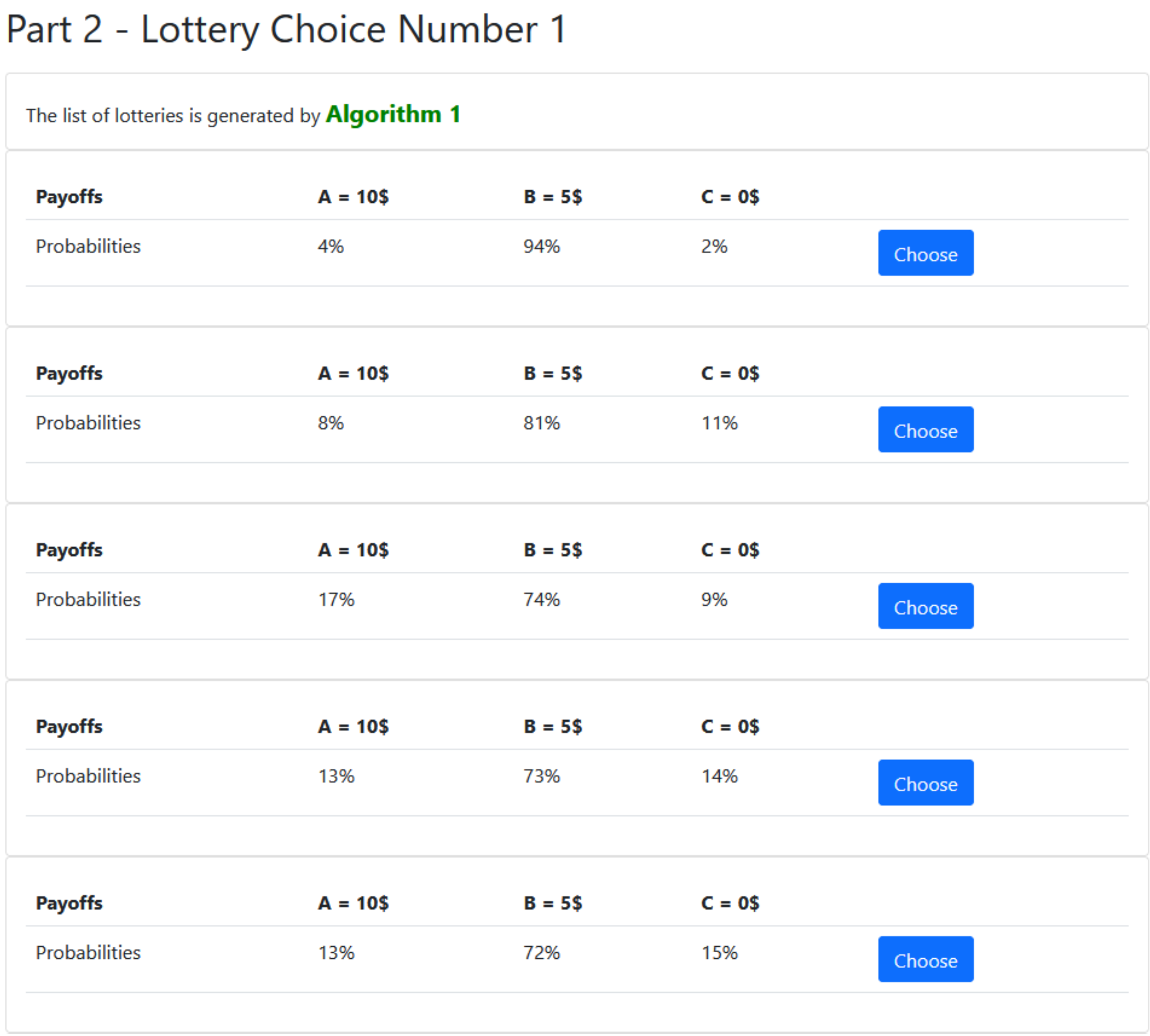


Figure A11: Screenshot - Choice screen where lotteries are selected and ordered by algorithm 1 (Platform algorithm)

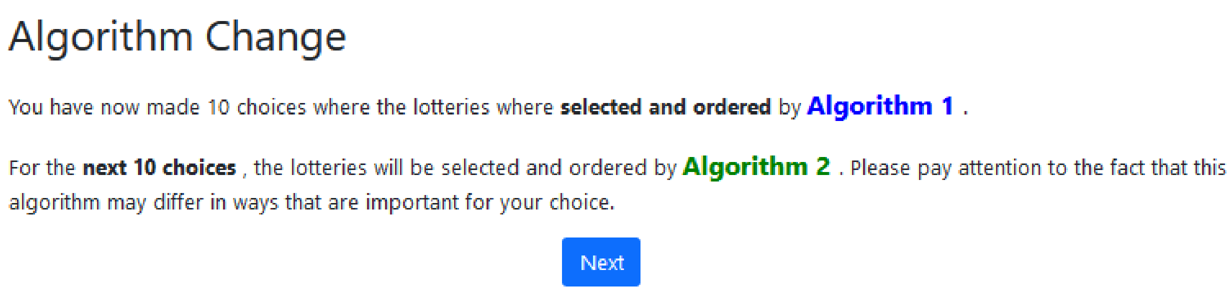


Figure A12: Screenshot - Buffer Screen between rounds 10 and 11

Part 2 - Lottery Choice Number 11

The list of lotteries is generated by **Algorithm 2**

Payoffs	**A = 10$**	**B = 5$**	**C = 0$**	
Probabilities	7%	92%	1%	Choose

Payoffs	**A = 10$**	**B = 5$**	**C = 0$**	
Probabilities	21%	68%	11%	Choose

Payoffs	**A = 10$**	**B = 5$**	**C = 0$**	
Probabilities	35%	44%	21%	Choose

Payoffs	**A = 10$**	**B = 5$**	**C = 0$**	
Probabilities	1%	99%	0%	Choose

Payoffs	**A = 10$**	**B = 5$**	**C = 0$**	
Probabilities	46%	26%	28%	Choose

Figure A13: Screenshot - Choice screen where lotteries are selected and ordered by algorithm 2 (Expected utility algorithm)

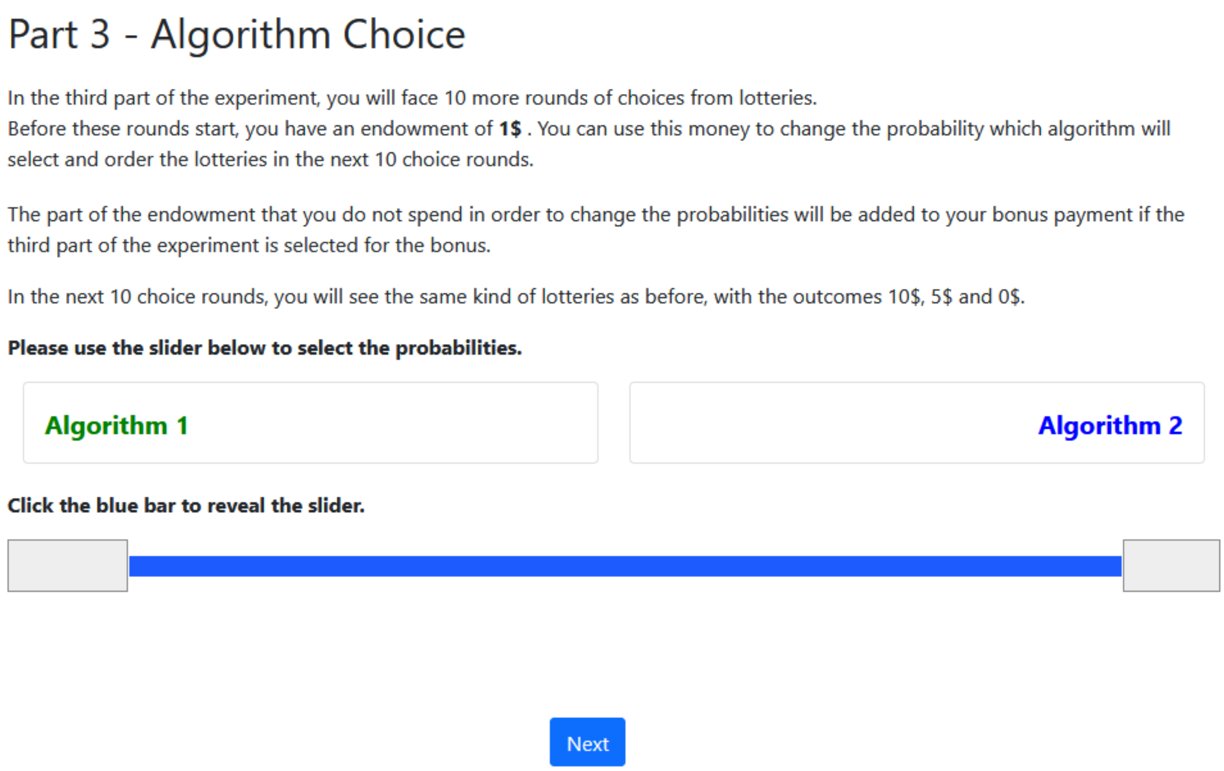


Figure A14: Screenshot - Choice between algorithms

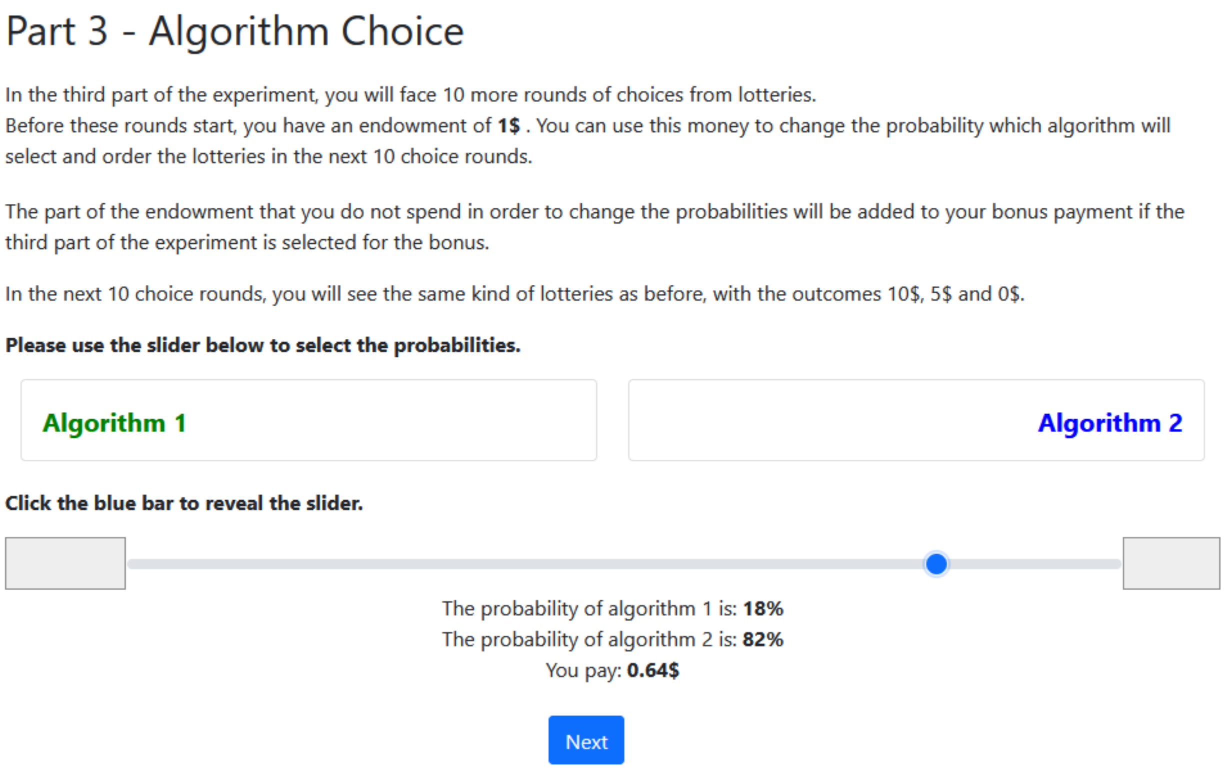
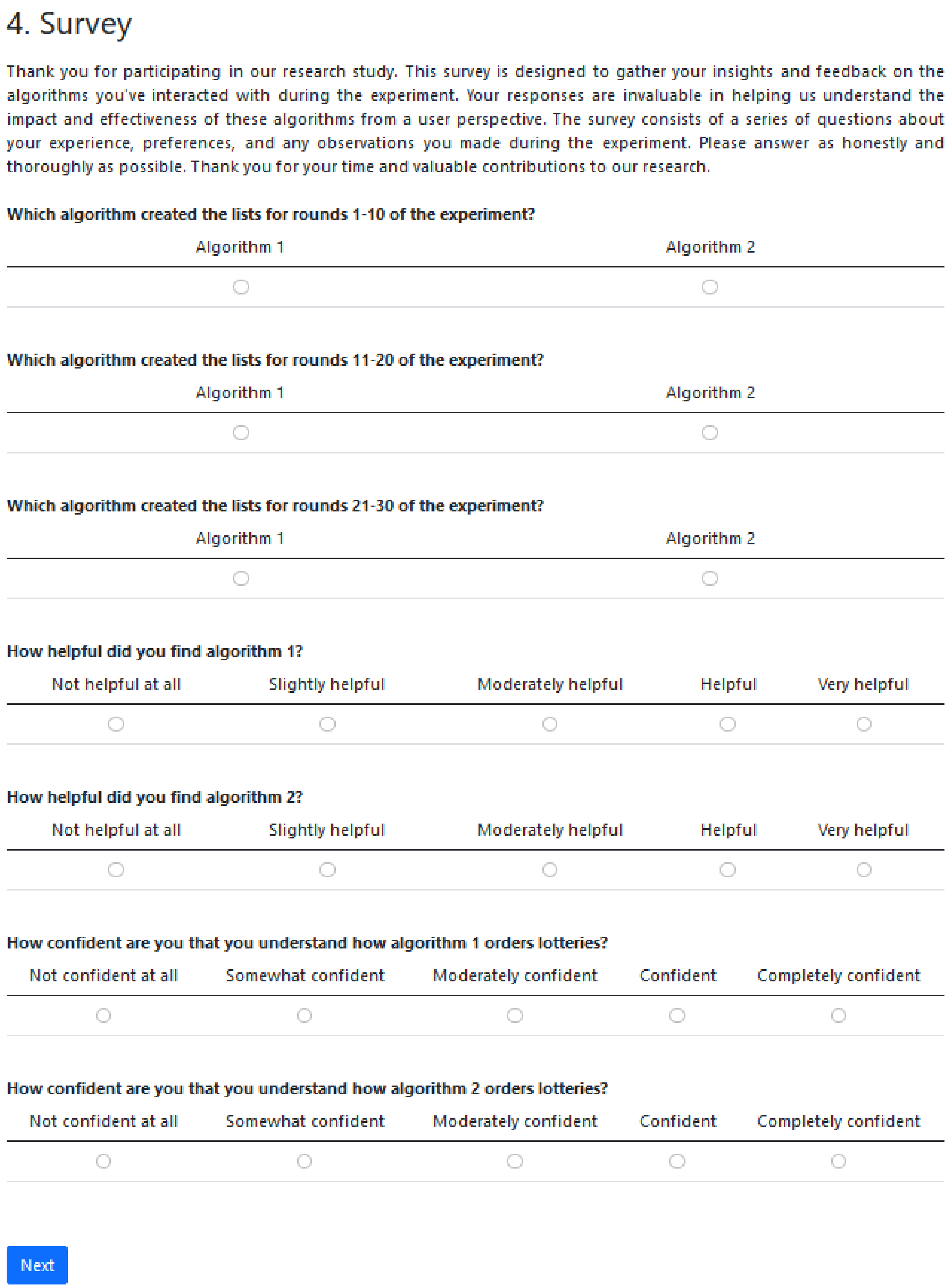


Figure A15: Screenshot - Choice of probability of algorithm 1 or algorithm 2 for the last 10 choice rounds



4. Survey

Thank you for participating in our research study. This survey is designed to gather your insights and feedback on the algorithms you've interacted with during the experiment. Your responses are invaluable in helping us understand the impact and effectiveness of these algorithms from a user perspective. The survey consists of a series of questions about your experience, preferences, and any observations you made during the experiment. Please answer as honestly and thoroughly as possible. Thank you for your time and valuable contributions to our research.

Which algorithm created the lists for rounds 1-10 of the experiment?

Algorithm 1	Algorithm 2
○	○

Which algorithm created the lists for rounds 11-20 of the experiment?

Algorithm 1	Algorithm 2
○	○

Which algorithm created the lists for rounds 21-30 of the experiment?

Algorithm 1	Algorithm 2
○	○

How helpful did you find algorithm 1?

Not helpful at all	Slightly helpful	Moderately helpful	Helpful	Very helpful
○	○	○	○	○

How helpful did you find algorithm 2?

Not helpful at all	Slightly helpful	Moderately helpful	Helpful	Very helpful
○	○	○	○	○

How confident are you that you understand how algorithm 1 orders lotteries?

Not confident at all	Somewhat confident	Moderately confident	Confident	Completely confident
○	○	○	○	○

How confident are you that you understand how algorithm 2 orders lotteries?

Not confident at all	Somewhat confident	Moderately confident	Confident	Completely confident
○	○	○	○	○

Next

Figure A16: Screenshot - Survey

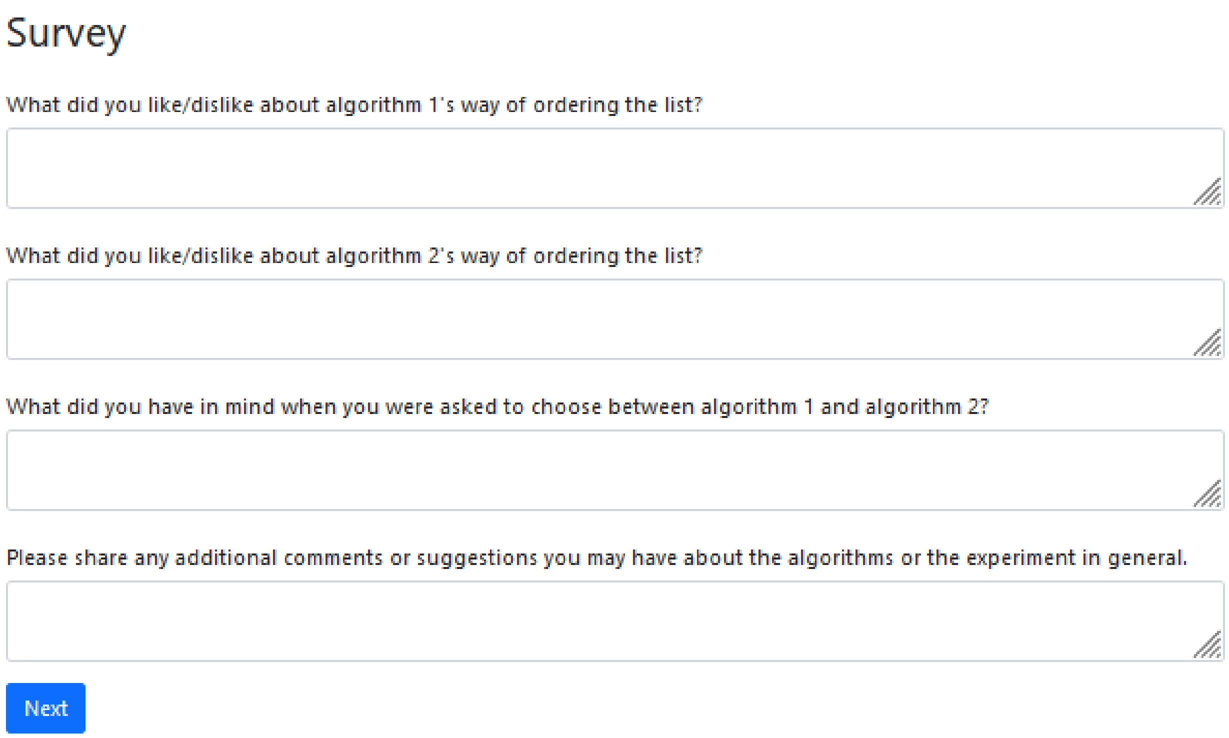


Figure A17: Screenshot - Survey open questions

End of the Experiment

Payoff Procedure

This experiment consisted of three parts. For your bonus pay, one of these parts gets randomly selected. Please see below for a description of the procedure that is used to calculate your payoff in each case.

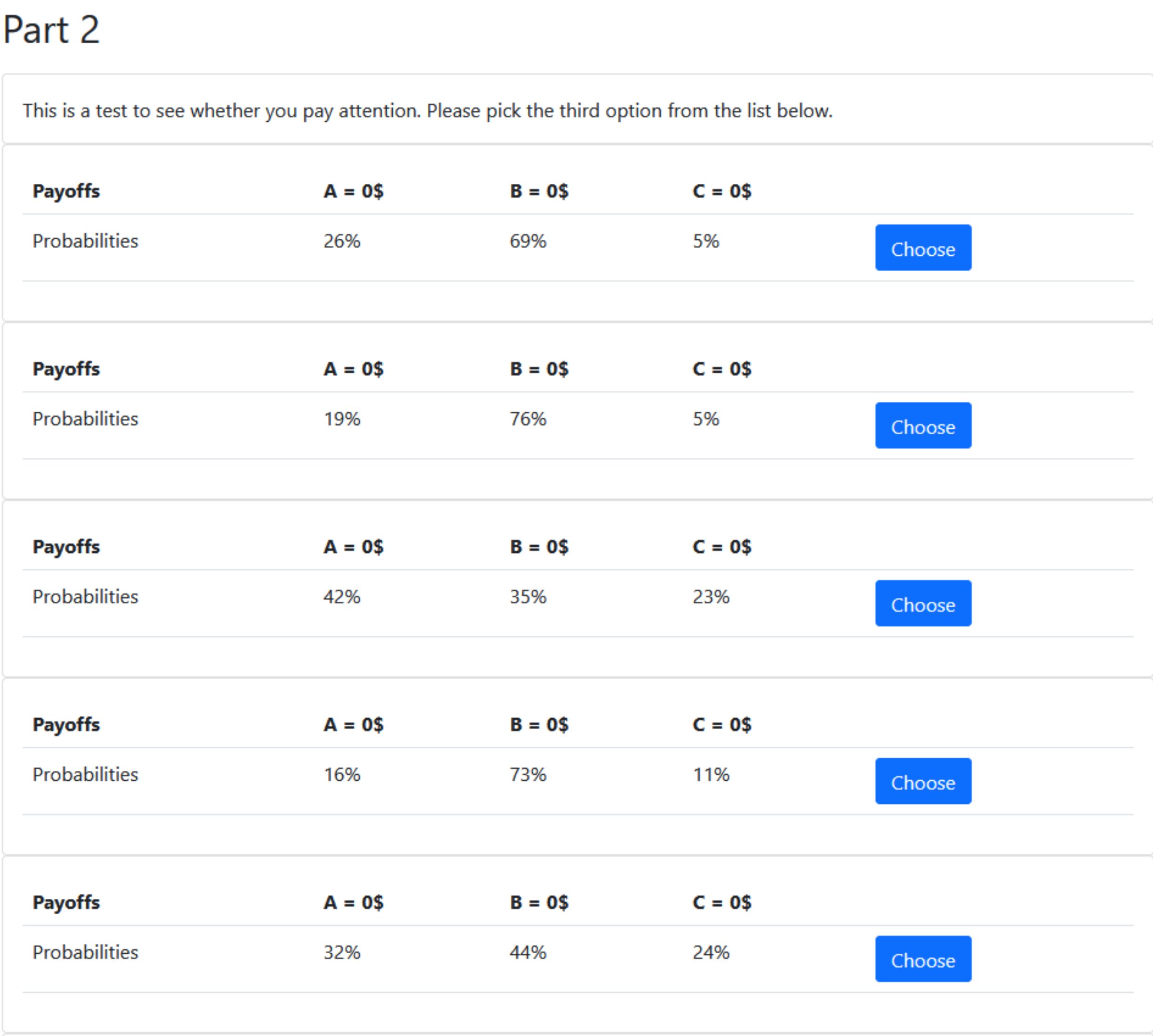
- **Part 1: Choosing between a lottery and a safe option**
  In this part, you were asked to declare a threshold above which you would want to switch from receiving a lottery to receiving a fixed payment. To calculate your payment, a value between 0 and 10 is drawn randomly. If the threshold you selected is below the randomly drawn value, you receive the randomly drawn value. Otherwise, the lottery that pays 10$ with probability 50% and 0$ with probability 50% is drawn and the outcome is added to your payoff.
- **Part 2: Choosing from lists of lotteries - round 1-20**
  In this part, you were asked to make choices from lists of lotteries, in rounds 1-10 the lotteries were selected and ordered by algorithm 1, in rounds 11-20 the lotteries where selected and ordered by algorithm 2. To calculate your payoff, one round is randomly selected. The outcome of the lottery is drawn and added to your payoff.
- **Part 3: Choosing between algorithms and choosing from lists of lotteries - round 21-30**
  In this part, you were given the opportunity to change the probability with which either algorithm 1 or algorithm 2 would create the lists of lotteries for the rounds 21-30. To calculate your payoff, one round is randomly selected. The outcome of the lottery is drawn and added to your payoff. In addition, you receive the endowment that you did not spend on changing the probability of the lottery.

Your Payoff

The part that determines your bonus payment is part 2.
The randomly selected round is round number 18. You chose the lottery **[P(10) = 49%, P(5) = 28%, P(0) = 23%]** in this round.
The outcome of the lottery was drawn to be 10$. **Your bonus payoff from the experiment is therefore $10.00.**
By clicking on the "Finish Experiment" button below, you return to the Prolific website to receive your payment.

Finish Experiment

Figure A18: Screenshot - End of the experiment and payment screen



Part 2

This is a test to see whether you pay attention. Please pick the third option from the list below.

Payoffs	**A = 0$**	**B = 0$**	**C = 0$**	
Probabilities	26%	69%	5%	Choose

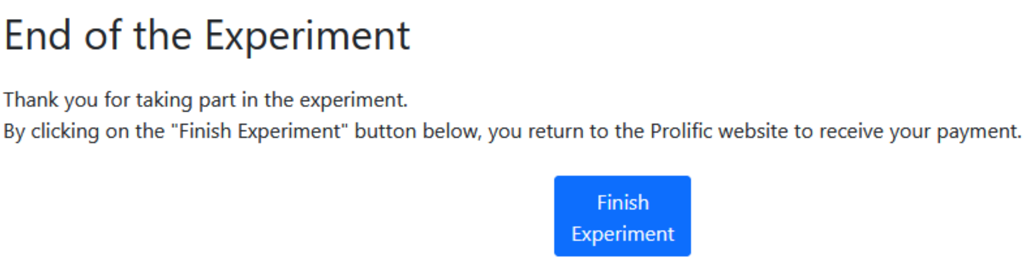
Payoffs	**A = 0$**	**B = 0$**	**C = 0$**	
Probabilities	19%	76%	5%	Choose

Payoffs	**A = 0$**	**B = 0$**	**C = 0$**	
Probabilities	42%	35%	23%	Choose

Payoffs	**A = 0$**	**B = 0$**	**C = 0$**	
Probabilities	16%	73%	11%	Choose

Payoffs	**A = 0$**	**B = 0$**	**C = 0$**	
Probabilities	32%	44%	24%	Choose

Figure A19: Screenshot - Attention check



End of the Experiment

Thank you for taking part in the experiment.
By clicking on the "Finish Experiment" button below, you return to the Prolific website to receive your payment.

Finish Experiment

Figure A20: Screenshot - Attention check fail

A.2 Hypotheses

Hypothesis 1: *Subjects choose lotteries with lower expected utility on average when facing the platform algorithm compared to when they face the expected utility algorithm.*

Hypothesis 2: *Subjects choose lotteries with lower expected value on average when facing the platform algorithm compared to when they face the expected utility algorithm.*

Hypothesis 3: *Subjects choose lotteries that are further down the list on average when facing the platform algorithm compared to when they face the expected utility algorithm.*

Hypothesis 4: *Subjects with high search costs are more negatively effected by the platform algorithm than subjects with low search costs.*

Hypothesis 5: *Subjects that are risk neutral or close to risk-neutral are more negatively effected by the platform algorithm than subjects with high risk aversion or high preference for risk.*

Hypothesis 6: *Subjects have a positive willingness-to-pay for the expected utility algorithm.*

Hypothesis 7: *Subjects with high search costs have a higher willingness-to-pay for the expected utility algorithm than subjects with low search costs.*

Hypothesis 8: *Subjects that are risk neutral or close to risk neutral have a higher willingness-to-pay for the expected utility algorithm than subjects with high risk aversion or high preference for risk.*