



# Market User Interface Design

## Citation

Seuken, Sven, David C. Parkes, Eric Horvitz, Kamal Jain, Mary Czerwinski, and Desney Tan. 2012. "Market user interface design." In Proceedings of the 13th ACM Conference on Electronic Commerce (EC '12), June 04 - 08, 2012, Valencia, Spain, 898-915. New York, NY: ACM Press.

## Published Version

doi:10.1145/2229012.2229080

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:33009590>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

# Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# Market User Interface Design

SVEN SEUKEN, University of Zurich  
DAVID C. PARKES, Harvard University  
ERIC HORVITZ, Microsoft Research  
KAMAL JAIN, eBay Research Labs  
MARY CZERWINSKI, Microsoft Research  
DESNEY TAN, Microsoft Research

Despite the pervasiveness of markets in our lives, little is known about the role of user interfaces (UIs) in promoting good decisions in market domains. How does the way we display market information to end users, and the set of choices we offer, influence users' decisions? In this paper, we introduce a new research agenda on "market user interface design." Our goal is to find the optimal market UI, taking into account that users incur cognitive costs and are boundedly rational. Via lab experiments we systematically explore the market UI design space, and we study the automatic optimization of market UIs given a behavioral (quantal response) model of user behavior. Surprisingly, we find that the behaviorally-optimized UI performs worse than the standard UI, suggesting that the quantal response model did not predict user behavior well. Subsequently, we identify important behavioral factors that are missing from the user model, including loss aversion and position effects, which motivates follow-up studies. Furthermore, we find significant differences between individual users in terms of rationality. This suggests future research on personalized UI designs, with interfaces that are tailored towards each individual user's needs, capabilities, and preferences.

Categories and Subject Descriptors: J.4 [**Computer Applications**]: Social and Behavioral Sciences—*Economics*

Additional Key Words and Phrases: Market Design, UI Design, Behavioral Economics, Experiment.

## 1. INTRODUCTION

Electronic markets are becoming more pervasive but a remaining research challenge is to develop user interfaces (UIs) to promote effective outcomes for users. This is important because markets often present users with a very large number of choices, making it difficult for users to find the optimal choice. For example, the markets for digital content which we can access via Amazon or iTunes are growing exponentially in size. Soon, we will have to deal with many complex markets in unfamiliar domains, and react to more frequent price changes. The smart grid domain is a prime example for such a domain. As we are asked to make market decisions more and more frequently, deliberation gets costly and we cannot spend too much time on individual decisions. This is where Herb Simon's 40-year old quote still says it best:

*"...a wealth of information creates a poverty of attention..."*

Herbert A. Simon (1971), pp. 40–41.

---

Author's addresses: S. Seuken, Department of Informatics, University of Zurich; D. C. Parkes, School of Engineering & Applied Sciences, Harvard University; E. Horvitz and M. Czerwinski and D. Tan, Microsoft Research; K. Jain, eBay Research Labs.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or [permissions@acm.org](mailto:permissions@acm.org).

EC'12, June 4–8, 2012, Valencia, Spain.

Copyright 2012 ACM 978-1-4503-1415-2/12/06...\$10.00.

Because humans incur cognitive costs when processing information [Miller 1956], a wealth of information, or a wealth of choices in market environments makes attention a scarce resource. Yet, traditional economic models assume agents to be perfectly rational, with unlimited time and unbounded computational resources for deliberation. We address this discrepancy by explicitly taking behavioral considerations into account when designing market UIs. The same way that color-coded planes make the job of an air-traffic controller easier, it is our goal to design market UIs that make economic decision making easier, thereby improving social welfare. In this sense, our approach is in line with the “choice architecture” idea put forward by Thaler et al. [2010].

A market UI can best be defined via two questions: first, what information is displayed to the user? Second, how many and which choices are offered to the user? Our goal is to develop a computational method that finds the optimal market UI, given a behavioral user model. Using behavioral models may lead to different market UIs for multiple reasons. For example, taking into account that users make mistakes, it may be best not to offer some choices that can lead to particularly bad outcomes (e.g., spending too much of your budget in one step). So far, the market design literature has largely ignored the intersection of market design and UI design. We argue that this intersection is particularly important because the complexity of the UI defines the cognitive load imposed on users. Furthermore, the UI defines how, and how well, users can express their preferences. Thus, when designing an electronic market, the design of the market’s UI may be as important as the market’s economic design.

### 1.1. Overview of Results

This paper introduces a new research agenda on “market user interface design.” We first present a systematic, empirical exploration of the effect that different UI designs have on users’ performance in economic decision making. Then we study the automatic optimization of market UIs based on a behavioral quantal response model.<sup>1</sup> We situate our study in a hypothetical market for 3G bandwidth where users can select the desired speed level, given different prices and values. While there is a possibly infinite set of choices (possible speed levels), the market UI only exposes some finite number. As the market UI designer, we get to decide how many and which choices to offer.

The participants of our experiments play a series of single-user games, facing a sequential decision-making problem with inter-temporal budget constraints. We vary a) the number of choices offered to the users (3, 4, 5, or 6), b) whether prices are fixed or dynamic, and c) whether choice sets are fixed or adaptive. Additionally, we also learn a quantal response model based on parts of the experimental data, and use computation to *automatically* optimize the market UI given the behavioral model. We then compare the behaviorally optimized UI with a standard UI. Because the market UI has a finite number of choices, the optimization algorithm must make a trade-off between having some choices at the lower end of the speed levels (which may be the best choice when values are low and prices are high) and some choices at the upper end (which may be the best choice when values are high and prices are low). Our main results are:

- (1) Users’ realized value increases as we increase the number of choices from 3 to 4 to 5, with no statistically significant difference between 5 and 6 choices.
- (2) The realized value is higher with adaptive choice sets compared to fixed choice sets.
- (3) The total realized value is *lower* when using the UI that is optimized for behavioral play, compared to the UI that is optimized for perfectly-rational play.

---

<sup>1</sup>Due to space constraints, some aspects of our study are omitted in this version of the paper. For more details, please see the appendix of the expanded version of this paper, available at: [www.ifi.uzh.ch/ce/publications/MarketUserInterfaceDesign.pdf](http://www.ifi.uzh.ch/ce/publications/MarketUserInterfaceDesign.pdf).

The third result was particularly surprising and prompted a more detailed analysis of users' decisions. We find that the quantal response model was too simplistic with significant negative consequences for market UI design. Our analysis suggests that we omitted important behavioral factors like *loss aversion* and *position effects*. Furthermore, we identify large differences between individual users' *level of rationality*. We find that for the "less rational" users there was no statistically significant difference in realized value using the UI optimized for rational play or optimized for behavioral play. However, the more rational users suffered, because the UI optimization took away too many of the valuable choices, making the decision problem easier, but reducing the total realized value. Thus, this result points towards the need for personalized market UIs that take into account each user's individual level of rationality.

## 1.2. Related Work

Prior research has identified a series of behavioral effects in users' decision making. Buscher et al. [2010] show that the amount of *visual attention* users spend on different parts of a web page depends on the task type and the quality of the information provided. Dumais et al. [2010] show that these "gaze patterns" differ significantly from user to user, suggesting that different UIs may be optimal for different groups of users. In a study of the cognitive costs associated with decision making, Chabris et al. [2009] show that users allocate time for a decision-making task according to cost-benefit principles. Because time is costly, more complex UIs put additional costs on users.

In addition to UI complexity, emotional factors are also important in decision making. Consider the "jam experiment" by Iyengar and Lepper [2000], which shows that customers are happier with the choices they make when offered 6 different flavors of jam compared to 24 different flavors. Schwartz [2005] identifies multiple reasons why more choices can lead to decreased satisfaction, including *regret*, *missed opportunities*, *the curse of high expectations*, and *self blame*. While emotional factors are important in many domains, in this paper we do not aim to study them directly. Instead we focus on users' cognitive limitations and corresponding bounded rationality.

Some research on UIs for recommender systems addresses aspects related to our work. Knijnenburg et al. [2012] study which factors explain the user experience of recommender systems. Chen and Pu [2010] propose methods for dynamically changing a recommender system UI, based on user feedback, to help users find suitable products in very large domains. Hauser et al. [2009] present a completely automated approach for dynamically adapting user interfaces for virtual advisory websites. They are able to infer users' cognitive styles based on click-stream data and then adjust the look and feel of a website accordingly. However, in contrast to our work, their goal is to increase users' purchase intentions, while our goal is to help users make better decisions.

Horvitz and Barry [1995] present a framework for the design of human-computer interfaces for time-critical applications in non-market-based domains. Their methodology trades off the costs of cognitive burden with the benefits of added information. Johnson et al. [1988] show that the way information is displayed (e.g., fractional vs. decimal probability values) has an impact on user decision making. The authors briefly discuss the implications of their findings for the design of information displays.

The work most closely related to ours is SUPPLE, introduced by Gajos et al. [2010], who present a system that can *automatically* generate user interfaces that are adapted to a person's devices, tasks, preferences, and abilities. They formulate the UI generation as an optimization problem and find that automatically-generated UIs can lead to significantly better performance compared to manufacturers' defaults. While their approach is in line with our goal of "automatic UI optimization," they do not consider a market context. They build a model of users' pointing and dragging performance and optimize their UIs for accuracy, speed of use, and users' subjective preferences for UI

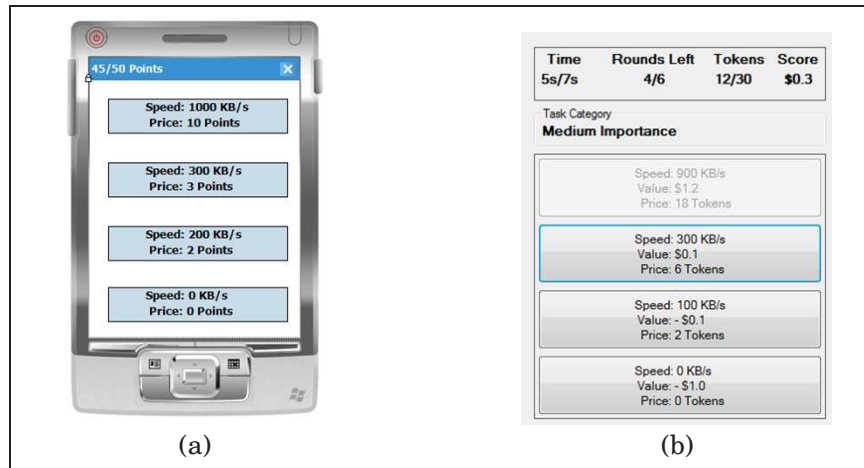


Fig. 1. (a) Mockup of the bandwidth market UI. (b) Screenshot of the market game used in the experiments.

layouts. In contrast, we build a *behavioral user model* and optimize for *decision quality* in market environments where users are dealing with values, prices, and budgets.

In our own prior work [Seuken et al. 2010c], we have introduced the goal of designing “hidden markets” with *simple* and *easy-to-use* interfaces. In related work [Seuken et al. 2010a,b], we have presented a UI for a P2P backup market, demonstrating that it is possible to hide many of the market’s complexities, while maintaining a market’s efficiency. Similarly, Teschner and Weinhardt [2011] show that users of a prediction market make better trades when using a simplified market interface, compared to one that provides the maximum amount of information and trading options. This paper is in the same vein as this prior work, but presents the first *systematic* exploration of the market UI design space, thereby opening up a new field of empirical research.

## 2. THE BANDWIDTH ALLOCATION GAME

The experiments described in this paper were conducted as part of a larger user study on people’s experiences and preferences regarding smartphone usage.<sup>2</sup> The international smartphone market is a Billion-dollar market with more than 100 Million users worldwide. With an ever growing set of bandwidth-hungry applications on these phones, the efficient allocation of 3G or 4G bandwidth is an important problem. According to Rysavy Research [2010], the demand for 3G bandwidth will continue to grow exponentially over the next few years and it will be infeasible for the network operators to update their infrastructure fast enough to satisfy future demand. The common approach for addressing the problem of bandwidth demand temporarily exceeding supply is to slow down every user in the network and to impose fixed data usage constraints. Obviously, this introduces large economic inefficiencies because different users have different values for high speed vs. low speed Internet access at different points in time.

Now imagine a hypothetical market-based solution to the 3G bandwidth problem. The main premise is that users sometimes do tasks of high importance (e.g., send an email attachment to their boss) and sometimes of low importance (e.g., random browsing). If we assume that users are willing to accept low performance now for high performance later, then we can optimize the allocation of bandwidth by shifting excess

<sup>2</sup>While future experiments will explore how our results translate to other domains, it is important to note that the four design levers we study constitute *within-experiment* variations. Thus, any *change* in behavior can be attributed to changes in the UI and are likely not specific to this domain.

demand to times of excess supply. Figure 1 (a) shows a mock-up application for such a bandwidth market. Assume that at the beginning of the month each user gets 50 points, or tokens. As long as there is more supply than demand, a user doesn't need to spend his tokens. However, when there is excess demand and the user wants to go online, then a screen pops up (as shown in Figure 1 (a)), requiring the user to make a choice. Each speed level has a different price (in tokens). For simplicity, we assume that when a user runs out of tokens, he gets no access or some very slow connection.<sup>3</sup> This domain is particularly suitable to studying market UIs because we can easily change many parameters of the UI, including the number of choices, whether prices stay fixed or keep changing, and the particular composition of the choice set.

## 2.1. Game Design

Figure 1 (b) shows a screenshot of the market game we designed for our experiments, mirroring the mockup of the market application, except that now the value for each choice is no longer private to each user, but determined by the game. Note that this is a single-user game on top of a *simulated market domain*. Each game has 6 rounds. At the beginning of a game, a user has 30 tokens available to spend over the course of the 6 rounds. In each round, the user has to select one of the choices. Each choice (i.e., a button in Figure 1 (b)) has three lines: the first line shows the *speed* of that choice in KB/s. The second line shows the *value* of that choice in dollars. The value represents the dollar amount that is added to a user's *score* when that choice is selected. The third line shows the *price* of that choice in tokens. When the user selects a particular choice, the corresponding number of tokens is subtracted from his budget and the corresponding value is added to his *score* which is displayed in the top right corner of the window. The score after the 6th round is the final score for the game.

Next to the score is a label displaying the user's current budget, which always starts at 30 in round 1 and then goes down as the user spends tokens. With the user's budget decreasing during a game, choices that have a price higher than the user's current budget become unavailable and are greyed out (as is the case for the top choice in Figure 1 (b)). To the left of the user's budget, the game shows the number of rounds that are left until the game is over. Finally, at the very left of the window, we show the user how much time he has left to make a decision in this particular round (e.g., in Figure 1 (b) the user still has 5 seconds left to make a decision in the current round).<sup>4</sup>

In every round, the user is in one of three task categories (high importance, medium importance, and low importance), which is displayed in the *task category label*. Every round, one of these three categories is chosen randomly with probability 1/3. Note that this corresponds to the original premise that users are doing tasks of different importance at different points in time. The task category determines the values of all choices. Effectively, the user has three *concave value functions* that map bandwidth levels to values. Table I shows an overview of the values the user can expect in the three categories for a game with 4 choices.<sup>5</sup> As one would expect, selecting the higher speed choices in the "high importance" category gives the user very high value, while

<sup>3</sup>In this paper, we do not concern ourselves with different business models or market designs for this domain. In particular, we do not address the question whether users should be allowed to pay money to buy more tokens. We do not suggest that this particular business model of using a fixed number of tokens per month should be adopted. Instead, we merely use this hypothetical market application as a motivating domain for our experiments into market UI design.

<sup>4</sup>We put users under time pressure to induce a certain error rate which allows for a meaningful comparison of different market UIs.

<sup>5</sup>The values shown in Table I are only the averages of the values in each category. In every round, the actual value for each choice is perturbed upwards or downwards with probability 1/3, to introduce additional stochasticity in the game. This avoids that users can memorize a fixed set of values for each task category.

Table I. The Values in the 3 different Task Categories

	High Imp.	Medium Imp.	Low Imp
900 KB/s	\$1.7	\$1.1	\$0.4
300 KB/s	\$0.5	\$0.2	- \$0.2
100 KB/s	-\$0.3	- \$0.3	- \$0.5
0 KB/s	-\$1	- \$0.9	- \$0.8

choosing low speeds in the high importance category leads to a severe penalty. In contrast, in the “low importance” category the user can earn less value for selecting high speeds, but is also penalized less for selecting the lowest speed.

The user’s problem when playing the game is to allocate the budget of 30 tokens optimally over 6 rounds, not knowing which categories and values he will face in the future. In some of our experiments, we randomly vary the prices charged for each of the choices from round to round. Thus, the user may also have uncertainty about which price level (out of 3 possible price levels) he will be facing next. This problem constitutes a sequential decision making problem under uncertainty.<sup>6</sup>

## 2.2. MDP Formulation and Q-Values

Each game can formally be described as a finite-horizon Markov Decision Problem (MDP) without discounting:

- **State Space:**  $CurrentRound \times CurrentBudget \times CurrentCategory \times CurrentValueVariation \times CurrentPriceLevel$ .
- **Actions:** Each choice affordable in the current round given current budget.
- **Reward Function:** The value of each choice.
- **State Transition:** The variables  $CurrentRound$ ,  $CurrentBudget$ , and  $CurrentScore$  transition deterministically given the selected choice. The other variables  $CurrentCategory$ ,  $CurrentValueVariation$  and  $CurrentPriceLevel$  transition stochastically.

The largest games we consider have approximately 7 million state-action pairs. Using dynamic programming, we can solve games of this size quickly (in less than 20 seconds). Thus, we can compute the optimal MDP policy, and we always know exactly which choice is best for each possible situation (game state) that can arise. Note that this policy is, of course, computed assuming that the future states are not known; only the model and the transition probabilities as described above are known.

Solving the MDP involves the computation of the  $Q$ -values for each state-action pair. For every state  $s$  and action  $a$ , the  $Q$ -value  $Q(s, a)$  denotes the expected value for taking action  $a$  in state  $s$ , and following the optimal MDP policy for every subsequent round. Thus, the optimal action in each state is the action with the highest  $Q$ -value, and by comparing the differences between the  $Q$ -values of two actions, we have a measure of how much “worse in expectation” an action is, compared to the optimal action.

## 2.3. The Quantal-Response Model

A well-known theory from behavioral economics asserts that agents are more likely to make errors the smaller the cost for making that error. This can be modeled formally with the *quantal response model* ([McKelvey and Palfrey 1995]) which predicts the likelihood that a user chooses action  $a_i$  in state  $s$  to be:

$$P(a_i | s) = \frac{e^{\lambda \cdot Q(s, a_i)}}{\sum_{j=0}^{n-1} e^{\lambda \cdot Q(s, a_j)}}$$

<sup>6</sup>Note that to play the game optimally, the user only needs to know the values and the prices of each choice, but not the *speeds*. However, we also include the *speed* information to label the buttons such that it is easier for users to recognize what has changed in the current round (e.g., values and/or prices).

where  $n$  denotes the total number of actions,  $Q(s, a_i)$  denotes the Q-value of action  $a_i$  in state  $s$ , and  $\lambda \geq 0$  is a precision parameter indicating how sensitive users are to differences between Q-values.  $\lambda = 0$  corresponds to random action selection, and  $\lambda = \infty$  corresponds to perfectly-rational action selection, i.e., always choosing the optimal action. Based on experimental results, one can compute a maximum-likelihood estimate for  $\lambda$ , i.e., maximizing the likelihood of the observed data. Equipped with a particular  $\lambda$ , this constitutes a user model which we use to optimize the UI for behavioral play (see [Wright and Leyton-Brown 2010] for a comparison of behavioral models).

### 3. EXPERIMENT DESIGN

Before we discuss the experiment design, let’s briefly pause to understand what exactly is within the control of the market UI designer, and what is not. Remember that in theory, there is an infinite set of choices (possible speed levels), but we assume that any market UI can only expose a fixed number of choices to the user. The UI designer decides 1) how many choices and 2) which exact choices to offer. For example, as in Figure 1 (b), we can provide 4 choices, i.e., 0 KB/s, 100 KB/s, 300 KB/s, and 900 KB/s. Alternatively, we can provide 3 choices, for example 0 KB/s, 500 KB/s, and 1000 KB/s. Note that by picking the choices, we only choose the market interface; the user’s *value function* which maps speed levels to values doesn’t change. Of course, higher speed levels have higher value for the user, but they also have a higher price.

In addition to the constraint of having a fixed number of choices, we also require the choice set to be fixed ex-ante and stay fixed throughout a game. In particular, the choices cannot depend on the state of the game (round, budget, category, value variation, price level).<sup>7</sup> In fact, the UI remains fixed for the 10 to 15 games that users play per treatment. For example, in the treatment with 5 choices, the user gets the same 5 choices in every round. Of course, in each of the possibly millions of different game states, a different choice is optimal. If the user could choose his speed freely, perhaps the optimal speed in some state would be 378 KB/s. But our UIs only offer a fixed number of choices. Of course, despite this constraint, for every state in the game, one of the available choices is still the best, and by solving the MDP we know which one it is. But in the real world, a UI designer would also only get to pick *one* UI (possibly knowing a distribution over situations a user will be in). We as the experimenters adopt the same viewpoint: we select one fixed UI, knowing the distribution of game states that a user will encounter, but we cannot change the UI during a game.

#### 3.1. Design Levers

We study the following four *market UI design levers*:

- (1) **Number of Choices:** This design lever describes how many choices (i.e., the number of buttons) were available to the users (3, 4, 5, or 6).
- (2) **Fixed vs. Dynamic Prices:** In the *fixed price* treatment, each choice always costs a fixed number of tokens (2 tokens per 100KB/s). With *dynamic prices*, one of 3 price levels is chosen randomly with probability 1/3 in each round, where the price per 100 KB/s is either 1, 2, or 3 tokens (thus, 500KB/s cost either 5, 10, or 15 tokens).<sup>8</sup>
- (3) **Fixed vs. Adaptive Choice Sets:** In the *fixed choice set* treatment, the users always have the same set of choices available to them in every round (e.g., always 0 KB/s, 100 KB/s, 300 KB/s, and 900KB/s). In the *adaptive choice set* treatment, the

<sup>7</sup>With the exception of the *Adaptive Choice Set* treatment, where we specify not one but three different UIs, one for each category (i.e., high, medium, low).

<sup>8</sup>The motivation for testing this design lever is that in some domains, balancing supply and demand may be possible with other means than dynamic prices. However, a detailed discussion of this idea is beyond the scope of this paper. We also don’t present results regarding this particular design lever in this paper.



decision within the UI design as to which choices to offer is allowed to vary with the category (e.g., in the high category, more high speed choices are available; in the low category, more low speed choices are available).

- (4) **Rational vs. Behavioral UI Optimization:** This describes which method is used to determine the composition of the choice sets (i.e., fixing the number of choices, which particular speed levels are available to users). In the *Rational-Optimization* treatment, the choice sets are optimized based on the MDP model assuming perfectly rational play. In the *Behavioral-Optimization* treatment, the choice sets are optimized assuming behavioral play according to the quantal response model.

### 3.2. Methodology and Experimental Set-up

We recruited 53 participants (27 men, 26 women) from the Seattle area with non-technical jobs. All participants had at least a Bachelors degree and we excluded participants who majored in computer science, economics, statistics, math, or physics. They were fluent English speakers, had normal (20/20) or corrected-to-normal vision, and were all right-handed. All of them used a computer for at least 5 hours per week. Their median age was 39, ranging from 22 to 54. None of the participants worked for the same company, but all of them had some familiarity with smartphones. We ran one participant at a time with each session lasting about 1.5 hours. The users first filled out a pre-study questionnaire (5 minutes). Then they went through a training session where the researcher first explained all the details of the game and gave them the opportunity to play 12 training games (20 minutes). Then they participated in the experiment (55 minutes) and finally completed a post-study survey (10 minutes). The participants were compensated in two ways. First, they received a software gratuity that was independent of their performance (users could choose one item from a list of Microsoft software products). Second, they received an Amazon gift card via email with an amount equal to the total score they had achieved over the course of all games they had played. The expected score for a random game, assuming perfect play, was around \$1. With random action selection, the expected score was highly negative. After each game, we showed the user his score from the last game and his accumulated score over all games played so far.<sup>9</sup> The final gift card amounts of the 53 users varied between \$4.60 and \$43.70, with a median amount of \$24.90.

### 3.3. Time Limits

To study the effect of the UI design on a user's ability to make economic decisions we need a reasonably complex decision problem, such that it is neither too easy nor too difficult for users to find the optimal decision. We achieve this by making decision time a scarce resource, as prior research has shown that users make worse decisions when under time pressure [Gabaix et al. 2006]. We impose an **exogenous** time limit of 12 (7) seconds per round. If a user doesn't make a choice within this time limit, the lowest choice (with 0KB/s for 0 tokens and a highly negative value) is chosen, and the game transitions to the next round. The time resets in every round. To warn the user, the game starts beeping three seconds before the end of a round.

In addition to the games with a fixed time limit, the users also played a series of games with an **endogenous** time limit. They had 240 seconds to play many games repeatedly; once a user finished one game, there was a 15 second break, and then the next game started. Thus, the cost for spending more time on a decision was internal-

---

<sup>9</sup>Originally, we had 56 participants in our study, but we had to exclude 3 participants from the first experiment (2 males, 1 female) because they did not understand the game well enough and achieved a negative overall score. However, we performed all regression analyses with and without those 3 users, and obtained qualitatively similar results.

Table II. Design of Experiment 1. Each participant played between 40 and 50 games. The design lever *Number of Choices* was a within-subject factor, the design lever *Fixed vs. Dynamic Prices* was a between-subjects factor.

Number Of Choices	12-second game	7-second game	240-second game
3	4 ×	4 ×	1 ×
4	4 ×	4 ×	1 ×
5	4 ×	4 ×	1 ×
6	4 ×	4 ×	1 ×

Table III. Design of Experiment 2. Every participant played between 40 and 50 games. Both design levers *Fixed vs. Adaptive Choice Sets* and *UI Optimization* were within-subject factors.

Treatment Variation	12-second game	7-second game	240-second game
Fixed Choice Sets & Rational Optimization	4 ×	4 ×	1 ×
Adaptive Choice Sets & Rational Optimization	4 ×	4 ×	1 ×
Fixed Choice Sets & Behavioral Optimization	4 ×	4 ×	1 ×
Adaptive Choice Sets & Rational Optimization	4 ×	4 ×	1 ×

ized by the participants. We used this time treatment to study the effect of fixed vs. dynamic prices on decision time. However, we do not discuss this aspect in this paper.

### 3.4. Treatment Variations

The study was split into two separate experiments. Experiment 1 involved 35 participants, and we tested the design levers *Number of Choices* (within-subject factor) and *Fixed vs. Dynamic Prices* (between-subject factor). Table II depicts the experiment design for each individual user. We randomized the order in which the users played the games with 3, 4, 5, or 6 choices. For each of those treatments, every user started with the four 12-second games, then played the four 7-second games, and then the 240-second endogenous time game. In Experiment 2 we had 18 participants and we tested the design levers *Fixed vs. Adaptive Choice Sets* and *Rational vs. Behavioral UI Optimization* (both within-subject factors). Here, all games had four choices and dynamic prices. See Table III for a depiction of the experiment design for each individual participant. Again, we randomized the order of the treatments.

### 3.5. Computational UI Optimization

To allow for a fair comparison of different market UIs (e.g., one with 4 choices vs. one with 5 choices), we chose each of these UIs *optimally*, given the constraints imposed by the treatment. The only choice that was always included was the 0 KB/s choice (for 0 tokens). Here, “optimally” means that we selected the one fixed UI with the highest *ExpectedOptimalValue* given the underlying market model (i.e., distribution of game states). To make this optimization computationally feasible, we discretized the search space, with 100KB/s being the smallest unit. Our search algorithm took as input the design parameters (e.g., 3 choices and optimized for rational play), iterated through all possible combinations of choices (i.e., all possible combinations of speed levels), solved the resulting MDP for each combination, and output the UI with the highest *ExpectedOptimalValue*. The optimization algorithm makes a trade-off between having some choices at the lower end of the speed levels (e.g., 200 KB/s which may be the best choice when values are low and prices are high) and some choices at the upper end of the speed levels (e.g., 900 KB/s which may be the best choice when values are high and prices are low). Note that this means that for a particular game state, the “optimal” choice for that state will not always be among the set of offered choices. Using this UI optimization approach, we guarantee that for every particular set of design criteria, we present the user with the best possible UI given the constraints.

### 3.6. Hypotheses

The larger the number of choices, the higher the expected value of the game assuming optimal play. Yet, Malhotra [1982] has shown that information overload leads to poorer decisions. We hypothesized that at first the benefit from having more choices outweighs the additional cognitive load (H1), but that as the number of choices gets large, the added cognitive costs become the dominant factor (H2). Similarly, using *Adaptive-ChoiceSets* tailors the available choices to the particular task category, which should make the decision easier for the user. On the other hand, the fact that the choices may change from round to round might also make it harder for users to find the optimal one. We hypothesized that the overall effect is positive (H3). Finally, the behavioral optimization eliminates some choices that may be useful in some game states because the behavioral model deems them as too risky. Thus, a user might suffer without those choices, or he might benefit, because the risky choices are eliminated. We hypothesized that the overall effect is positive (H4). To summarize, our four hypotheses are:

*H1: The realized value increases as we increase the number of choices.*

*H2: The realized value first increases as we increase the number of choices, but ultimately decreases.*

*H3: The realized value is higher when using adaptive choice sets, compared to using fixed choice sets.*

*H4: The realized value is higher when using behavioral optimization, compared to using rational optimization.*

## 4. EXPERIMENTAL RESULTS

In this section we describe the results regarding our hypotheses. As the regression technique we use *Generalized Estimating Equations (GEE)*, an extension of generalized linear models that allows for the analysis of correlated observations [Nelder and Wedderburn 1972]. This gives us consistent coefficient estimates with robust standard errors despite using repeated measures from individual users.

### 4.1. Number of Choices

The first design lever we analyze is *NumberOfChoices*. We measure the effect of this design lever by analyzing the dependent variable *RealizedValue*, which is a randomness-adjusted version of the user's total score. Consider first the graph in Figure 2. While the top line, representing *ExpectedOptimalValue*, monotonically increases as the number of choices is increased, the bottom line, representing *RealizedValue*, only increases as we go from 3 to 4 to 5 choices, but then slightly decreases as we go from 5 to 6 choices (with largely overlapping error bars). One possible explanation is that the disadvantage from adding cognitive load as we go from 5 to 6 choices outweighs the theoretical benefits of having one more choice available. For more insights, we now turn to the statistical data analysis.

Consider column (1) of the table in Figure 3 where we present the results of the regression analysis with indicator variables for the different treatments. The coefficients for *NumChoices* are with respect to *NumChoices=6*. We see that the effects of *NumChoices=3* and *NumChoices=4* are statistically significant at  $p < 0.001$ . Furthermore, the coefficient for *NumChoices=5* is positive, but not statistically significant. Thus, the data does not provide enough evidence that there is also a statistically significant decrease in *RealizedValue* as we go from 5 to 6 choices. In column (2) of the table in Figure 3, we add more covariates to the analysis to test the robustness of the results. We see that *dynamicPrices* has a statistically significant effect on the realized value,

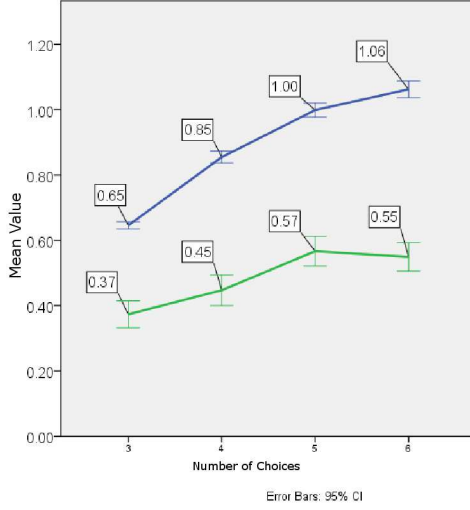


Fig. 2. Mean values for 3, 4, 5, and 6 choices. The blue line (on top) corresponds to *ExpectedOptimalValue*. The green line (on the bottom) corresponds to *RealizedValue*.

while the covariates *7-secondGame* (controlling for whether it was a 7-second or a 12-second game) and *GameCounter* (controlling for possible learning effects) do not have a statistically significant effect. While adding those covariates, the results regarding *NumChoices* remain qualitatively unchanged. Thus, we obtain the following results:

**RESULT 1 (NUMBER OF CHOICES).** *We reject the null hypothesis in favor of H1, i.e., the realized value per game significantly increases as we increase the number of available choices from 3 to 4 to 5. Regarding H2, we cannot reject the null hypothesis, i.e., we don't have enough evidence to conclude whether the realized value per game ultimately plateaus or decreases as we increase the number of available choices further.*

#### 4.2. Fixed vs. Adaptive Choice Sets

We now move on to the analysis of the data from Experiment 2 where we studied the two design levers *Fixed vs. Adaptive Choice Sets*, and *UIOptimization*. For this experiment, we fixed the number of available choices to four and only considered dynamic prices. Due to space constraints we cannot discuss the details of the design lever *Fixed vs. Adaptive Choice Sets*. Here, we only state the final result which we obtain from the regression analysis presented in Table IV:

**RESULT 2 (FIXED VS. ADAPTIVE CHOICE SETS).** *We reject the null hypothesis in favor of H3, i.e., the realized value is significantly higher with adaptive choice sets, compared to fixed choice sets.*

#### 4.3. UI Optimization for Rational vs. Behavioral Play

For the design lever *UIOptimization* we compare two different UIs, one optimized for perfectly rational play, and one optimized for behavioral play. For the behavioral optimization, we first built a behavioral model based on the data from experiment 1. We computed different likelihood-maximizing  $\lambda$ -parameters for the quantal response model depending on 1) the total number of choices in the particular game, 2) the number of choices left in a particular round, and 3) whether prices were fixed or dynamic.

Fig. 3. GEE for the dependent variable *Realized Value*, studying the effect of *NumChoices*. Standard errors are given in parentheses under the coefficients. The individual coefficient is statistically significant at the \*10% level, the \*\*5% level, the \*\*\*1% level, and at the \*\*\*\*0.1% level.

Factors/Covariates	(1)	(2)
Intercept	0.549**** (0.0374)	0.610**** (0.0605)
NumChoices=3	-0.176**** (0.0451)	-0.177**** (0.0451)
NumChoices=4	-0.102**** (0.0275)	-0.108**** (0.0283)
NumChoices=5	0.018 (0.0286)	0.021 (0.0309)
NumChoices=6	0	0
DynamicPrices=0		-0.169**** (0.0396)
7-SecondGame		-0.012 (0.0236)
GameCounter		0.001 (0.0010)
Model Fit (QICC)	149.743	147.719



Fig. 4. (a) Market UI Optimization Method. (b) A sample UI optimized assuming perfectly rational play. (c) A sample UI optimized assuming behavioral play.

Then we solved the resulting MDP, where now Q-values are computed assuming that the user will follow the “behavioral strategy” when playing the game. Finally, we selected the UI with the highest expected value according to this “behavioral MDP.” Figure 4 (a) shows a diagram illustrating our market UI optimization methodology.

To get some intuition for what happens under behavioral optimization, consider Figures 4 (b) and (c) where we display two sample UIs, one optimized for perfectly rational play, and one optimized for behavioral play. Note that both UIs are the result of a computational search algorithm. The only difference between the two UIs is the top choice: the UI that was optimized for perfectly rational play gives the user the 900KB/s choice, while the UI that was optimized for behavioral play gives the user the 400KB/s choice. This result is understandable in light of how the UI optimization algorithm works and the behavioral vs. optimal user model. The quantal response model assigns each action a certain likelihood of being chosen, corresponding to the Q-values of those actions. Now, consider the top choice in Figure 4 (b), which has a high value, but which can also cost between 9 and 27 tokens (this is a game with dynamic prices). Thus, in the worst case, the user spends 27 out of his 30 tokens with one click, and then has only 3 tokens left for the remaining 5 rounds. Even if it is very unlikely that the user selects this action, the negative effect of an occasional mistake would be very large. Consequently, the UI optimized for behavioral play shown in Figure 4 (c) does not have such high-value high-cost choices, reducing the negative effect of mistakes.

Now, consider Table IV for the effect of *BehavioralOptimization* on *RealizedValue*. We see that the coefficient for *BehavioralOptimization* is *negative* and statistically significant ( $p < 0.001$ ).<sup>10</sup> Thus, the UI optimization assuming behavioral play had a negative effect on *RealizedValue* and we obtain the following result:

**RESULT 3 (UI OPTIMIZATION).** *We cannot reject the null hypothesis in favor of H4. Instead, we find that using the behaviorally optimized UI leads to a realized value which is not higher but in fact significantly lower, compared to using the UI optimized for rational play.*

<sup>10</sup>We also performed this analysis separately for i) fixed choices sets, and ii) adaptive choice sets. The effect of the behavioral optimization is negative in both cases, however, it is only statistically significant when using adaptive choice sets.

Table IV. GEE for dependent variable *RealizedValue*, studying *AdaptiveChoiceSets* and *BehavioralOptimization*.

Factors/Covariates	(1)
Intercept	0.462**** (0.0501)
AdaptiveChoiceSets?	0.077** (0.0376)
BehavioralOptimization?	-0.111**** (0.0334)
Model Fit (QICC)	106.927

Table V. GEE for dependent variable *OptChoice*, studying *Lambda* and *QvalueDiff*. Standard errors are given in parentheses under the coefficients. The individual coefficient is statistically significant at the \*10% level, the \*\*5% level, the \*\*\*1% level, and at the \*\*\*\*0.1% level.

Factors	(1)		(2)	
	B	Exp(B)	B	Exp(B)
Intercept	-0.816**** (0.1408)	0.442****	-1.529**** (0.1593)	0.217****
Lambda	0.150**** (0.0180)	1.162****	0.161**** (0.0197)	1.175****
QvalueDiff			5.868**** (0.4353)	353.713****
Fit (QICC)	(3771.953)		(3589.063)	

This result is very surprising, in particular because the behavioral UI optimization had a *negative* instead of a positive effect on *RealizedValue*. Upon finding this result, we hypothesized that the quantal response model was too simple for a UI optimization, ignoring some important behavioral factors. Given prior behavioral research, possible candidate factors were *loss aversion* and *position effects*. The goal of the analysis in the next section is to find empirical support for our hypothesis that behavioral factors which we omitted in our UI optimization had a significant impact on users’ decisions.

## 5. BEHAVIORAL DECISION ANALYSIS

In this section, we analyze individual *rounds* of a game to understand which factors influence users’ decision making. Here, we only consider the data from 7-second games from experiment 1. We primarily analyze the dependent variable *OptChoice*, which is 1 if the user clicked on the optimal choice in a particular round, and 0 otherwise.

### 5.1. Degree of Rationality

We first test whether individual users exhibit significant differences in their play according to the quantal response model. We compute a separate maximum-likelihood parameter  $\lambda_i$  for each user  $i$ . This parameter can be seen as measuring how “rational” a user’s play was. In fact, the users exhibited large differences, with a minimum  $\lambda$  of 3.9, a maximum of 9.0, and a median of 6.8. Table V presents the regression results for *OptChoice*. In column (1), we see that the parameter *Lambda* has a statistically significant effect ( $p < 0.001$ ). Looking at the odds ratio (Exp(B)), we see that the odds of choosing the optimal choice are 16% higher for a user with  $\lambda = x$  compared to a user with  $\lambda = x - 1$ .<sup>11</sup> Thus, for the analysis of *OptChoice* it is important to control for  $\lambda$ .

### 5.2. Q-Value Differences

We now analyze the factor *QvalueDiff* which denotes the difference between the Q-values of the best and second-best action. In column (2) in Table V we see that *QvalueDiff* is statistically significant ( $p < 0.001$ ) with an odds ratio of 354. Note that this is the odds ratio for a *one unit change* in the Q-value difference. Yet, in our data, the mean of the Q-value difference is 0.11. The odds ratio for a change of 0.1 is 1.8. Thus, holding *Lambda* constant, if the Q-value difference between the best and second-best choice increases by 0.1, the odds for choosing the optimal choice increase by 80%.

<sup>11</sup>We also analyzed two other user-specific factors: *Age* and *Gender*. There was no statistically significant effect of *Age* on either *OptChoice* or *RealizedValue*. For *Gender*, there was no effect with respect to *RealizedValue*, but there was a small statistically significant effect ( $p < 0.1$ ) on *Optchoice*: female users were slightly more likely to miss the optimal choice, but male users made bigger mistakes when they missed the optimal choice. However, the factor *Lambda* already captures user-specific cognitive differences, and thus we do not need to also control for *Gender* in the regression analyses.

Table VI. GEE for dependent variable *OptChoice*, studying UI complexity, position effects, and loss aversion. Standard errors are given in parentheses under the coefficients. The individual coefficient is statistically significant at the \*10% level, the \*\*5% level, the \*\*\*1% level, and at the \*\*\*\*0.1% level.

Factors/Covariates	(1)		(2)		(3)	
	B	Exp(B)	B	Exp(B)	B	Exp(B)
Intercept	0.283 (0.2182)	1.327	-0.495* (0.2489)	0.610*	-0.616*** (0.2339)	0.540***
Lambda	0.167**** (0.0206)	1.181****	0.162**** (0.0223)	1.176****	0.158**** (0.0238)	1.171****
QvalueDiff	5.062**** (0.4297)	157.888****	4.421**** (0.5061)	83.196****	4.595**** (0.4989)	98.962****
NumChoices	-0.391**** (0.0446)	0.677****	-0.087* (0.0487)	0.916*	-0.065 (0.0583)	0.937
OptRelativeRank=5			-3.853**** (0.9808)	0.021****	-4.046**** (1.0396)	0.017****
OptRelativeRank=4			-1.893**** (0.4499)	0.151****	-1.853**** (0.4959)	0.157****
OptRelativeRank=3			-1.201**** (0.2706)	0.301****	-1.188**** (0.3382)	0.305****
OptRelativeRank=2			-0.614** (0.2807)	0.541**	-0.522 (0.3351)	0.593
OptRelativeRank=1			-0.160 (0.2283)	0.852	-0.170 (0.2512)	0.844
OptRelativeRank=0			0	1	0	1
OptimalChoiceNegative=1					-1.299**** (0.2247)	0.273****
CurrentCategory=2					1.532**** (0.2059)	4.626****
CurrentCategory=1					0.033 (0.1295)	1.034
CurrentCategory=0					0	1
Goodness of Fit (QICC)	3476.044		3345.116		3288.243	

### 5.3. UI Design: Number of Choices

We now study how the UI design affects users' ability to make optimal choices. Consider Table VI column (1), where we add *NumChoices* to the regression. This factor denotes the number of choices in the game (i.e., 3, 4, 5, or 6 choices). We see that the factor has a large and highly statistically negative effect ( $p < 0.001$ ) on *OptChoice*. Holding all other factors constant, increasing the number of choices by 1 reduces the odds for selecting the optimal choice by 32%. Naturally, a more complex UI (i.e., more choices) makes it harder for users to find the optimal choice.<sup>12</sup>

### 5.4. Incomplete Search: Position Effects

By design, the game exhibits a strong ordering effect: the values of the choices decrease monotonically from top to bottom, as do the prices. Thus, it is conceivable that users scan the choices in a linear way, either from top to bottom or from bottom to top. Given that they are under time pressure, incomplete search effects may be expected, and prior research has shown that this can lead to significant position effects [Dumais et al. 2010; Buscher et al. 2010]. We can control for positional effects by adding information about the position of the optimal choice to the regression. Consider column (2) in Table VI where we added six indicator variables to the regression. *OptRelativeRank* denotes

<sup>12</sup>We also analyzed *NumChoicesLeft* which denotes the number of choices that were still affordable during a game situation, given the prices of the current choices and the user's budget. However, when controlling for *NumChoices* we found that *NumChoicesLeft* does not have a statistically significant effect on *OptChoice*.

Table VII. UI optimization: Effects on optimal and realized value.

<i>Behavioral Optimization?</i>	<i>Optimal Value</i>	<i>Realized Value</i>
no	1.02	0.50
yes	0.78	0.39

the “relative rank” or “relative position” of the optimal choice, taking into account the currently unavailable choices. Consider a game with 6 choices as an example. If there are currently 4 choices left and the optimal choice is the third from the top, then the absolute position of that choice would be 2, but the relative rank is 0.<sup>13</sup>

In column (2) of Table VI we see that *OptRelativeRank* has a very strong, and highly statistically significant negative effect on *OptChoice*. Note that all coefficient estimates are relative to *OptRelativeRank*=0. The lower the rank of the optimal choice, the less likely the users were to choose the optimal action. As we go from *OptRelativeRank*=0 to *OptRelativeRank*=5, the coefficients decrease monotonically, and except for *OptRelativeRank*=1, all of the effects are statistically significant. Compared to the case when the optimal choice has rank 0, holding everything else constant, if *OptRelativeRank*=4 the odds of choosing the optimal action decrease by 84%, and if *OptRelativeRank*=5, the odds decrease by 98%. Thus, in particular for the very low ranks, the position effect is indeed very strong and highly statistically significant ( $p < 0.001$ ), and because our user model did not take it into account, this presents a possible explanation for why the UI optimization failed.

### 5.5. Loss Aversion

Loss-aversion, i.e., people’s tendency to avoid losses more than they appreciate gains, is a well-known effect in behavioral economics [Tversky and Kahneman 1991]. Thus we hypothesized to find it in our data as well. Consider column (3) of Table VI where we added *OptimalChoiceNegative* to the regression, an indicator variable that is 1 when the value of the optimal choice is negative, and 0 otherwise. Additionally, we also added the factor *CurrentCategory* to the regression, controlling for the different value distributions in different game situations. *OptimalChoiceNegative* has a large negative coefficient, and is statistically significant ( $p < 0.001$ ). Thus, whether the optimal choice has a positive or negative value makes a large difference on users’ behavior, providing strong evidence for the loss aversion hypothesis.<sup>14</sup>

## 6. TOWARDS PERSONALIZED MARKET USER INTERFACES

We have seen that behavioral factors such as position effects and loss aversion play a significant role in users’ decision making, offering potential answers to the question why the UI optimization failed. We now come back to the design lever *UI Optimization* that we studied in experiment 2. In further analyses of *OptChoice* (not presented here), we find that the behavioral UI optimization indeed made the decision problem easier for the users: they were 17% more likely to select the optimal choice when using the UI optimized for behavioral play. Given that the users made better choices but their *RealizedValue* still decreased, this suggests that the UI optimization eliminated too many valuable choices. In some sense, it was “too aggressive”.

<sup>13</sup> We have also performed the same analyses using absolute rank and obtained qualitatively similar results.

<sup>14</sup> Note that this loss aversion effect cannot be explained with classical risk aversion, which is based on diminishing marginal utility of wealth [Koeszegi and Rabin 2007]. In our games, users repeatedly face many small-scale risks with almost no effect on their overall wealth. Thus, risk aversion is not a convincing explanation for the observed behavior. Another possible explanation for the observed effect could be myopia. It is conceivable that users have a limited look-ahead horizon when making decisions during the game, and thus do not fully account for the effect of running out of budget towards the end of the game. However, in further statistical analyses we could not find evidence for this hypothesis.



Table VIII. GEE for dependent variable *RealizedValue* for *SmallLambda*=0, studying the effect of *BehavioralOptimization*.

Factors/Covariates	(1)
Intercept	0.360**** (0.0331)
Lambda	0.048**** (0.0068)
AdaptiveChoices=1	0.061 (0.0424)
BehavioralOptimization=1	-0.172**** (0.0371)
Model Fit (QICC)	30.050

Table IX. GEE for dependent variable *RealizedValue* for *SmallLambda*=1 studying the effect of *BehavioralOptimization*.

Factors/Covariates	(1)
Intercept	-0.308**** (0.0377)
Lambda	0.186**** (0.0094)
AdaptiveChoices=1	0.099* (0.0572)
BehavioralOptimization=1	-0.068 (0.0495)
Model Fit (QICC)	74.808

Now consider Table VII, which shows what happened to *OptimalValue* and *RealizedValue* under the behavioral optimization. By using the behavioral optimization, we decreased the optimal value (achievable for a perfectly rational player) from \$1.02 to \$0.78. Thus, we “took away” approximately \$0.24 per game. Note that we never expected the users to come even close to the optimal values, but instead we expected them to do better using the behaviorally optimized UI such that the *Realized Value* would actually increase. However, as we can see in the last column of Table VII, the *Realized Value* also dropped from \$0.50 to \$0.39. Relative to the optimal value, the users did better in the re-optimized game – but in absolute terms they did worse.

A potential explanation is that by coincidence, the users in experiment 2 acted “more rationally” than the users in experiment 1. However, the best fitting  $\lambda$ -parameters for experiments 1 and 2 are very similar, and thus, the data does not support this hypothesis. Yet, we found another unexpected result regarding users’ level of rationality in experiment 2. As before, we computed a  $\lambda_i$ -parameter for each user, as well as one  $\lambda$  corresponding to the best fit across all users. Next, we computed a binary variable *SmallLambda* for each user which denotes whether that user’s  $\lambda_i$  is smaller or larger than the average  $\lambda$ . Thus, *SmallLambda* denotes whether a user belongs to the *more rational* or to the *less rational* group of users.

Now consider Tables VIII and IX, where we study the effect of *BehavioralOptimization* on *RealizedValue*, separating users into the *more rational users* (on the left) and the *less rational users* (on the right). For *SmallLambda*=0 (the more rational users), the effect of *BehavioralOptimization* is particularly negative: for those users we made the game a lot worse by doing the re-optimization. However, for *SmallLambda*=1 (the less rational users) the effect of *BehavioralOptimization* is close to zero, and not statistically significant. Thus, for less rational users, the behaviorally-optimized UI was easier to use, but the resulting *RealizedValue* was practically the same.

This finding suggests a new research direction on *personalized market user interfaces*, with the goal to tailor the UI to the capabilities, needs, and preferences of individual users. To achieve this, we must access user-specific, behavioral and non-behavioral data. This is available in many domains, in particular in the smartphone domain. Once we have an estimate of a user’s “degree of rationality,” we can provide each user with a market UI that is specifically optimized for that particular user.

## 7. CONCLUSION

In this paper, we have introduced a new research agenda on *market user interface design*. Our long-term goal is to understand how UI design choices for markets affect users’ abilities to make good economic decisions, and how we can develop automated methods to optimize market UIs. In studying this question, it is crucial to take a be-

havioral approach, deviating from a perfectly rational agent model. Thus, our research explores a design space in which human limited cognition meets computing.

We ran a behavioral economics lab experiment, testing the effect of different market UI design levers. In regard to the number of choices, we found that the realized value increases as we go from 3 to 4 to 5 choices, with no significant effect going from 5 to 6 choices. In future experiments we will also test 7 and 8 choices, to see if the realized value ultimately decreases again. For the design lever *Fixed vs. Adaptive Choice Sets*, we found that the realized value is significantly higher with adaptive choice sets.

Finally, the most interesting design lever was the behavioral UI optimization. An unexpected result is that the realized value was *lower* when using the behavioral UI optimization. This suggests that the quantal response model was too simplistic to accurately predict user behavior. In future research, we will consider other behavioral models that are better supported by neuroeconomic experiments like the drift-diffusion model [Fehr and Rangel 2011]. In a subsequent decision analysis, we found that our model ignored important behavioral factors like loss aversion and position effects. Yet, the most intriguing result concerns how *less rational* and *more rational* users differed regarding the effect of the UI optimization. While there was no significant difference regarding the realized value for the less rational users, the more rational users lost a lot of value in the UI optimization due to precluded opportunities. This result points towards the need to estimate each individual user's level of rationality based on behavioral data obtained over time, to generate personalized market UIs. Taking this idea a step further, we can also take a user's value for time into account. Thus, there are still many opportunities for research at the intersection of market design, intelligent agents, UI design, and behavioral economics, ranging from better behavioral models, to algorithms for learning user preferences and automated UI optimization.

## ACKNOWLEDGMENTS

We would like to thank Al Roth, Alain Cohn, Haoqi Zhang, Yiling Chen, five anonymous reviewers, and seminar/conference participants at Microsoft Research, Yahoo! Research, Harvard University, Brown University, the NBER Market Design Working Group Meeting, the 2nd Conference on Auctions, Market Mechanisms and Their Applications, the Toulouse Workshop on the Psychology and Economics of Scarce Attention, and the Aarhus Workshop on New Trends in Mechanism Design for helpful comments. Part of this work was done while Seuken was an intern at Microsoft Research. Seuken gratefully acknowledges the support of a Microsoft Research PhD Fellowship during that time.

## References

- BUSCHER, G., DUMAIS, S., AND CUTRELL, E. 2010. The Good, the Bad, and the Random: An Eye-Tracking Study of Ad Quality in Web Search. In *Proceedings of the 33rd Annual International ACM SIGIR Conference*. Geneva, Switzerland.
- CHABRIS, C. F., LAIBSON, D. I., MORRIS, C. L., SCHULDT, J. P., AND TAUBINSKY, D. 2009. The Allocation of Time in Decision-Making. *Journal of the European Economic Association* 7, 628–637.
- CHEN, L. AND PU, P. 2010. Experiments on the Preference-based Organization Interface in Recommender Systems. *ACM Transactions on Computer-Human Interaction (TOCHI)* 17(1), 1–33.
- DUMAIS, S., BUSCHER, G., AND CUTRELL, E. 2010. Individual Differences in Gaze Patterns for Web Search. In *Proceedings of the Information Interaction in Conext Symposium (IIX)*. New Brunswick, NJ.
- FEHR, E. AND RANGEL, A. 2011. Neuroeconomic Foundations of Economic Choice—Recent Advances. *Journal of Economic Perspectives* 25(4), 3–30.
- GABAIX, X., LAIBSON, D., MOLOCHE, G., AND WEINBERG, S. 2006. Costly Informa-

- tion Acquisition: Experimental Analysis of a Boundedly Rational Model. *American Economic Review* 96 (4), 1043–1068.
- GAJOS, K. Z., WELD, D. S., AND WOBROCK, J. O. 2010. Automatically Generating Personalized User Interfaces with Supple. *Artificial Intelligence* 174, 910–950.
- HAUSER, J. R., URBAN, G. L., LIBERALI, G., AND BRAUN, M. 2009. Website Morphing. *Marketing Science* 28(2), 202–223.
- HORVITZ, E. AND BARRY, M. 1995. Display of Information for Time-Critical Decision Making. In *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence (UAI)*. Montreal, Canada.
- IYENGAR, S. S. AND LEPPER, M. R. 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, E. J., PAYNE, J. W., AND BETTMAN, J. R. 1988. Information Displays and Preference Reversals. *Organizational Behavior and Human Decision Processes* 42(1), 1–21.
- KNIJNENBURG, B. P., WILLEMSSEN, M. C., GANTNER, Z., SONCU, H., AND NEWELL, C. 2012. Explaining the User Experience of Recommender Systems. *User Modeling and User-Adapted Interaction (UMUAI)* 22, 1–64.
- KOESZEGI, B. AND RABIN, M. 2007. Reference-Dependent Risk Attitudes. *American Economic Association* 97(4), 1047–1073.
- MALHOTRA, N. K. 1982. Information Load and Consumer Decision Making. *Journal of Consumer Research: An Interdisciplinary Quarterly* 8(4), 419–430.
- MCKELVEY, R. AND PALFREY, T. 1995. Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior* 10, 6–38.
- MILLER, G. A. 1956. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review* 63, 81–97.
- NELDER, J. AND WEDDERBURN, R. 1972. Generalized Linear Models. *Journal of the Royal Statistical Society* 135 (3), 370–384.
- RYSAVY RESEARCH. 2010. Mobile Broadband Capacity Constraints and the Need for Optimization. Research report available at <http://www.rysav.com/papers.html>.
- SCHWARTZ, B. 2005. Can There Ever be too Many Flowers Blooming? *Culture Choice* 3, 1–26.
- SEUKEN, S., CHARLES, D., CHICKERING, M., AND PURI, S. 2010a. Market Design and Analysis for a P2P Backup System. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC)*. Cambridge, MA.
- SEUKEN, S., JAIN, K., TAN, D., AND CZERWINSKI, M. 2010b. Hidden Markets: UI Design for a P2P Backup Application. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*. Atlanta, GA.
- SEUKEN, S., PARKES, D. C., AND JAIN, K. 2010c. Hidden Market Design. In *Proceedings of the 24th Conference on Artificial Intelligence (AAAI)*. Atlanta, GA.
- SIMON, H. A. 1971. Designing Organizations for an Information-Rich World. In *Computers, Communication, and the Public Interest*, M. Greenberger, Ed. Johns Hopkins Press.
- TESCHNER, F. AND WEINHARDT, C. 2011. Evaluating hidden market design. In *Proceedings of the 2nd Conf. on Auctions, Market, and Market Applications*.
- THALER, R. H., SUNSTEIN, C. R., AND BALZ, J. P. 2010. Choice Architecture. Unpublished. Available at SSRN: <http://ssrn.com/abstract=1583509>.
- TVERSKY, A. AND KAHNEMAN, D. 1991. Loss Aversion in Riskless Choice: A Reference Dependent Model. *Quarterly Journal of Economics* 106(4), 1039–1061.
- WRIGHT, J. R. AND LEYTON-BROWN, K. 2010. Beyond Equilibrium: Predicting Human Behavior in Normal-Form Games. In *Proceedings of the 24th Conference on Artificial Intelligence (AAAI)*.