# Search and Satisficing

*By* ANDREW CAPLIN, MARK DEAN, AND DANIEL MARTIN*

Draft: June 2, 2010

*Many options are available even for everyday choices. In practice, most decisions are made without full examination of all such options, so that the best available option may be missed. We develop a search-theoretic choice experiment to study the impact of incomplete consideration on the quality of choices. We find that many decisions can be understood using the satisficing model of Simon [1955]: most subjects search sequentially, stopping when a "satisficing" level of reservation utility is realized. We find that reservation utilities and search order respond systematically to changes in the decision making environment.*

*JEL: D03, D83, C91*

*Keywords: Revealed preference, search, incomplete information, bounded rationality, stochastic choice, decision time*

Many options are available even for everyday choices. In practice, most decisions are made without full examination of all such options, so that the best available option may be missed. Unfortunately, little is known concerning how such incomplete consideration impacts the quality of final decisions. One complicating factor is that incomplete search breaks the revealed preference identification of choice with preference.

In this paper we develop a search-theoretic choice experiment that provides new insights into how information gathering interacts with decision making. We find that many decisions can be understood using the satisficing model of Simon [1955]. Simon posited the existence of a satisficing level of "reservation" utility, attainment of which would induce the decision maker to curtail further search. Our experiments cover various settings that differ in the number of

* Caplin: Center for Experimental Social Science and Department of Economics, New York University, 19 West 4th Street, New York, NY 10012, andrew.caplin@nyu.edu. Dean: Department of Economics, Brown University, Providence, RI 02912, mark_dean@brown.edu. Martin: Center for Experimental Social Science and Department of Economics, New York University, 19 West 4th Street, New York, NY 10012, daniel.martin@nyu.edu. We thank Isabelle Brocas, Juan Carrillo, Vince Crawford, Stefano DellaVigna, Drew Fudenberg, Daniel Gottlieb, David Laibson, Alessandro Lizzeri, Ulrike Malmendier, Rosemarie Nagel, Efe Ok, Antonio Rangel, Ariel Rubinstein, Yuval Salant, and three anonymous referees for insightful suggestions.

options available and in the complexity of these objects, and in all cases, we find broad support for Simon's hypothesis. Most subjects search sequentially, and stop search when a "satisficing" level of reservation utility has been realized. Moreover, the satisficing levels of utility that we identify vary systematically with environmental factors: they are lower for more complicated objects and higher in larger choice sets.

One factor that has held back research on how incomplete search impacts choice is that, in principle, any pattern of behavior can be rationalized. There are no observable implications of a general theory in which the set of objects that a subject considers may be smaller than the choice set as understood by an external observer.[1] To identify such restrictions, our experiment identifies not only final choices, but also how provisional choices change with contemplation time.[2] By making visible the features of search that are not revealed in standard choice data, this "choice process" data provides a test bed for simple models of sequential search (see Campbell [1978] and Caplin and Dean [Forthcoming]).

A second barrier to research on how incomplete search impacts choice is that decision quality is generally hard to assess. Given that preferences are personal, there is no general way to define, let alone measure, the quality of decisions.[3] To overcome this conceptual problem, our experiment is designed to make the quality of all options easy to measure. Subjects in our experiment select among monetary prizes presented as sequences of addition and subtraction operations.[4] While these calculations are not especially difficult, they take time and effort to perform, making the choice problem nontrival. Given that prizes are monetary, it is simple to measure decision quality.

Our first experimental finding relates to the structure of search. We find that subjects typically switch from lower to higher value objects, in line with information being absorbed on an item-by-item basis, as in sequential search theory. The second finding relates to the rule for stopping search. For each treatment, we identify fixed reservation values such that most subjects curtail search early if, and only if, they identify an option of higher value than the reservation level. Taken together, these two findings support the satisficing model. While our basic findings relate to nonstandard choice process data, we arrive at analogous findings for standard choice data. Our

[1]See Manzini and Mariotti [2007] and Masatlioglu and Nakajima [2009] for examples of other decision theoretic models in which the decision maker's consideration set is smaller than the externally observable choice set. See also Eliaz and Spiegler [Forthcoming]. Rubinstein and Salant [2006] present a model of choice from lists, in which a decision maker searches through the available options in a particular order. Ok [2002] considers the case of a decision maker who is unable to compare all the available alternatives in the choice set.

[2]Compared to other novel data used to understand information search, such as those based on eye tracking or Mouse-lab (Payne, Bettman and Johnson [1993], Gabaix et al. [2006], Reutskaja et al. [Forthcoming]), choice process data is more closely tied to standard choice data and revealed preference methodology.

[3]See Bernheim and Rangel [2008], Gul and Pesendorfer [2008] and Koszegi and Rabin [2008] for methodological viewpoints on the classification of particular decisions as "poor" or "mistaken."

[4]Caplin and Dean [Forthcoming] characterize theoretical connections between choice process data and sequential search with arbitrary objects of choice.

findings in this regard exploit the fact that subjects were able to, and indeed choose to, change options prior to finalizing decisions even in our standard choice experiments.

In addition to characterizing the nature of search and the stopping rule, we uncover systematic interactions between search order and choice. We present subjects with choice sets in list order that contain items which vary in the number of computational operations. In this manner we identify separate effects of list order and computational complexity on search. We also identify individual differences in search order: some subjects search from the top of the screen to the bottom, while others search from simple to more complex objects. These differences impact choice: those who search down from the top do poorly if good objects are at the bottom of the screen, while those who search based on simplicity miss good objects that are complicated.

While our findings are broadly in line with simple theories of sequential search, we consider an alternative theory in which subjects search the entire choice set, but make calculation errors that lead to choice mistakes. We estimate a random utility model in which the size of the utility error depends on the size and complexity of the choice set. Fitting the model requires seemingly large perceptual errors, yet simulations based on the fitted model significantly overestimate subject performance in large and complex choice sets. Moreover, the estimated calculation errors are incompatible with the fact that subjects almost always switch from lower to higher value alternatives, in line with the principle of sequential search. For all of these reasons, the sequential search model appears to better organize our experimental findings.

The paper is arranged into six sections. In section I we introduce our experimental protocols, including not only the choice process experiments which incentivize intermediate choices, but also standard experiments with only final choices incentivized. In section II we show that failures of optimality increase in line with the number of available options and their complexity in both standard choice and choice process experiments. In section III we confirm that sequential search based on a reservation stopping rule rationalizes many of the mistakes we observe in the choice process experiments. We estimate empirically how these reservation rules vary across environments. Order effects on choice are addressed in section IV. Section V investigates the connection between standard choice experiments and choice process experiments. We organize comparisons around an optimizing model with "psychic" search costs that is rich enough to cover both settings. Section VI contains our estimates of the model based entirely on calculation errors rather than sequential search.

## I. Experimental Design

Our paper consists of four experiments. Experiment 1 measures choice quality in our experimental task in a standard choice experiment. Experiment 2 uses the choice process design to examine provisional choices within the same environment. Experiment 3 uses the choice process experiment to explore search order. Experiment 4 is a version of the standard choice experiment that incorporates a time limit. In addition to a change in incentives, experiment 2 differs from experiment 1 in that it involves a time limit. Thus, experiment 4 bridges some of the gap between the choice process experiment and the standard choice experiment.

### *A. Experiment 1: Standard Choice*

Our goal in this paper is to study whether a model of information search can explain why people sometimes fail to choose the best available option. Hence we work with objects of choice for which such failures are easy to identify: dollar amounts expressed as addition and subtraction operations. We conducted six treatments that differ in terms of complexity (3 or 7 addition and subtraction operations for each object) and the total number of available alternatives (10, 20 or 40). Figure 1 shows a 10 option choice set with objects of complexity 3.[5]

FIGURE 1. ABOUT HERE.

Each round began with the topmost option on the screen selected, which had a value of \$0 and was worse than any other option. While only the final choice was payoff relevant, subjects could select whichever option they wanted at any time by clicking on the option or on the radio button next to it.[6] The currently selected option was displayed at the top of the screen. Once subjects had finalized their selection, they could proceed by clicking on the submit button at the bottom of the screen. Subjects faced no time constraint in their choices.

The value of each alternative was drawn from an exponential distribution with $\lambda = 0.25$, truncated at \$35 (a graph of the distribution was shown in the experimental instructions – see online supplemental material).[7] The individual terms in the algebraic expression representing the

[5]Given that the subjects (New York University students) made negligible mistakes when purely numerical options were presented, we wrote out the arithmetic expressions in word form rather than in symbolic form.

[6]Changes that were made over the pre-decision period were recorded and are analyzed in section V.

[7]For each of the three choice set sizes we generated 12 sets of values, which were used to generate the choice objects for both the low and the high complexity treatments.

alternative were generated stochastically in a manner that ensured that neither the first nor the maximal term in the expression were correlated with total value.

Subjects for experiment 1 took part in a single experimental session consisting of 2 practice rounds and between 27 and 36 regular rounds, drawn from all 6 treatments. At the end of the session, two regular rounds were drawn at random, and the subject received the value of the selected object in each round, in addition to a $10 show up fee. Each session took about an hour, for which subjects earned an average of $32. In total we observed 22 undergraduate students making 657 choices in the Center for Experimental Social Science laboratory at New York University.

### *B. Experiment 2: Choice Process*

Choice data alone does not allow us to directly test for a reservation value. Experiment 2 is therefore designed to generate what we term "choice process" data. This data tracks not only final choice, but also how subjects' provisional choices evolve with contemplation time. Choice process data is explicitly designed to enable simple search theoretic models to be tested (see Caplin and Dean [Forthcoming]).

DEFINITION 1: *Given nonempty set $X$ and upper time limit $T > 0$, **choice process data** specifies, for $A \subset X$, a time indexed sequence of choices $C_A(t) \in A$ for all $1 \leq t \leq T$.*

As the definition indicates, choice process data come in the form of sequences of observed choices. We interpret $C_A(t)$ as the object that a subject chooses from set $A$ having contemplated the decision problem for time $t$.

Choice process data is closely related to standard choice data, in that all observations represent choices, albeit indexed by time. We therefore see it as complementary to other attempts to use novel data to understand information search, such as those based on eye tracking or Mouselab [Payne, Bettman and Johnson, 1993; Gabaix et al., 2006; Reutskaja et al., Forthcoming]. These approaches make aspects of the search process observable, yet do not connect these intermediate acts of search with their implications for choice. While choice process data misses out on potentially relevant cues to search behavior, it captures the moment at which search changes a subject's assessment of the best option thus far encountered.

To generate choice process data, we need to elicit for each choice set a time series of observations that record the most preferred alternative at each moment in time. Our design has two key features. First, subjects are allowed to select any alternative in the choice set at any time, changing their selected alternative whenever they wish. Second, actualized choice is recorded at a random point in time unknown to the experimental subject. At the end of each choice round,

a random time is generated, and the subject is apprised both of the time that was selected and of their choice at that time. This incentivizes subjects to always have selected their current best option in the choice set. We therefore interpret the sequence of selections as choice process data.[8]

The instructions that were given to subjects are available in the online supplemental material. As in the standard choice experiment, each round began with the topmost and worst option of $0 selected. Subjects could at any time select any of the alternatives on the screen by clicking on the alternative or the radio button next to it, with the currently selected object being displayed at the top of the screen. Unlike experiment 1, experiment 2 had a time constraint, with subjects having up to 120 seconds to complete the choice task. Subjects were instructed that at the end of the round a random time between 1 and 120 seconds would be picked according to a truncated beta distribution with parameters $\alpha = 2$ and $\beta = 5$ and that the alternative the subject had selected at this time would be recorded as their choice for that round.[9] A subject who finished in less than 120 seconds could press a submit button, which completed the round as if they had kept the same selection for the remaining time.

Experiments were again run at New York University, using subjects recruited from the undergraduate population. Typically, a subject took part in a session consisting of 2 practice rounds and 40 regular rounds. Two recorded choices were actualized for payment, which was added to a $10 show up fee.

Experiment 2 included six treatments that matched the treatments in experiment 1: choice sets contained 10, 20 or 40 alternatives, with the complexity of each alternative being either 3 or 7 operations. Moreover, exactly the same choice object values were used in the choice process and standard choice experiments. For the 6 treatments of experiment 2, we collected data on 978 choice sets from 76 subjects.

### *C. Experiment 3: Varying Complexity*

One of the features of the choice process methodology is that it can shed light on the order in which subjects search. To explore whether the complexity of an object affects the order in which it is searched, we introduced a further treatment in which object complexity varied within a choice round. Specifically, experiment 3 consisted of choice sets of size 20, and the objects in each set ranged in complexity from one to nine operations. The complexity range was selected to generate substantial differences in difficulty among choice objects, and subjects were instructed

[8] In support of this interpretation, 58 of 76 subjects in a post experiment survey responded directly that they always had their most preferred option selected, while others gave more indirect responses that suggest similar behavior (e.g. having undertaken a recalculation before selecting a seemingly superior alternative).

[9] A graph of this distribution was shown in the experimental instructions. The front-weighting in the beta distribution provides an incentive for subjects to begin recording their most preferred options at an early stage.

that object complexity and object value were independent of one another. Incentives were as in experiment 2, the choice process experiment. Experiment 3 was run on 20 subjects for a total of 206 observed choice sets.

### *D. Experiment 4: Time Constraint*

One difference between the standard choice and choice process experiments is that the latter had a time limit, whereas the former did not. In order to explore whether this time limit was responsible for differences in behavior between the two settings, we ran a further treatment. Specifically, we re-ran the standard choice experiment with a two minute time constraint, as in the choice process experiment. If subjects failed to press the submit button within 120 seconds they received $0 for that round. For this experiment, we had a total of 29 subjects and 407 observed choice sets.

## II. Choice Performance

### *A. Standard Choice Task*

Table 1 reports the results of experiment 1, the standard choice experiment. We measure the choice performance of subjects in each of the 6 treatments in two ways. The top section reports the "failure rate" – the proportion of rounds in which the subject did not choose the option with the highest dollar value. The second section reports the average absolute loss – the difference in dollar value between the chosen item and the highest value item in the choice set.

TABLE 1—ABOUT HERE.

Averaging across all treatments, subjects fail to select the best option almost 38 percent of the time. On average, subjects leave $3.12, or 17 percent of the available money on the table in each round.[10] Subjects' performance varies across treatments. All measures reported in table 1 increase both with the size and the complexity of the choice set. Failure rates vary from less than 7 percent for the size 10, low complexity (3 operations) treatment to over 65 percent for the

[10]There is no evidence for any effect of learning or fatigue on choice performance. The order in which choice rounds were presented was reversed for half the subjects, and the order of presentation did not have a significant effect on performance. In part, this may be because our experimental design is structured to remove learning effects. The decision making context, including the distribution of prizes, is known to the decision maker at the start of each experimental round.

size 40, high complexity (7 operations) treatment. Average losses range from $0.41 (3 percent of the highest valued option) in the size 10, low complexity treatment to $7.12 (33 percent of the highest valued option) in the size 40, high complexity treatment. Regression analysis shows that the difference in losses between treatments is significant.[11] There is also some evidence that the effect of complexity on choice performance is stronger in larger choice sets – the difference in loss between low and high complexity objects in size 10 choice sets is $1.28 (10.2 percentage points) and not significant at the 10 percent level. For size 40 choice sets, the difference is $4.82 (22.8 percentage points) and significant at the 1 percent level.[12]

### *B. Choice Process Task*

Our analysis of the search-based determinants of choice quality is based primarily on the choice process data of experiment 2 rather than the standard choice data of experiment 1. It is therefore important to explore how the level and pattern of final choices compares across these two environments. For purposes of comparison, we discard observations from choice process rounds in which subjects do not press the submit button before the allotted 120 seconds. In such rounds, we assume that subjects have not finalized their choice. This assumption removes 94 rounds, or 8 percent of our total observations, which we will exclude from all subsequent analyses. Table 2 compares failure rates and average absolute losses by treatment for choice process and standard choice tasks.

TABLE 2—ABOUT HERE.

The comparative statics of choice performance between treatments are very similar for the choice process experiment and the standard choice experiment. In both cases, subjects fail to find the best option more frequently and lose more money in larger and more complicated choice sets. However, in all treatments, the quality of final choice is worse in the choice process task than the standard choice task. We explore this difference in section V, where we relate it to the different

[11] Absolute loss was regressed on dummies for choice set size, complexity and interactions, with standard errors calculated controlling for clustering at the subject level. Losses were significantly higher at the 1 percent level for complexity 7 as against complexity 3 for size 20 and 40 choice sets, though not for size 10 choice sets. Losses were also significantly higher at the 1 percent level for size 40 vs. size 10 choice sets at both levels of complexity.

[12] While not the primary subject of study in the current paper, there are significant individual differences in mistakes. Estimates obtained from a regression of absolute loss on individual specific dummies, controlling for treatment effects, indicate that the 25th percentile subject does on average $1.10 better than the median subject, while the 75th percentile subject does $1.23 worse, averaging across all rounds.

incentives in the two experiments. There is less incentive to continue search in the choice process task, given that the probability of additional effort raising the payoff falls over time.

## III. Sequential Search and Satisficing

We use the choice process data from experiment 2 to test whether a simple sequential search model can explain the failure of people to select the best available option. Specifically we test the reservation-based search (RBS) model of Caplin and Dean [Forthcoming], which is a simple and general version of the satisficing model of Simon [1955]. The model is based on two features of behavior:

1) Subjects search through the available alternatives *sequentially*, understanding the value of each object before moving on to the next, as in the classic search models of Stigler [1961] and McCall [1970].

2) Search continues until an object is found that is above a fixed *reservation utility level*, at which point search stops.

In order for choice process data to be useful in testing the RBS model, it needs to provide more information than standard choice alone. The power of our tests depends on observing subjects switching from one alternative to another. Figure 2 shows histograms of the number of choice switches per round for each treatment. We define a choice switch as an occasion on which the subject selects a new alternative, excluding the initial change away from the $0 option.

FIGURE 2. ABOUT HERE.

Figure 2 demonstrates that the choice process methodology does indeed reveal more than just final choices: in 68 percent of rounds we observe at least one switch. In the remaining 32 percent of rounds, subjects switch away from the initial $0 option and then stop. This suggests the possibility that there may be changes of mind that are not recorded in the choice process data, possibly due to perceived transactions costs of recording the switch. It is important to note that none of our analyses are impacted by this possibility because our model is robust to the inclusion of a private threshold of significance that has to be crossed before a change is recorded.

## A. Sequential Search

Caplin and Dean [Forthcoming] provide a method of identifying whether or not choice process data is consistent with sequential (but possibly incomplete) search. The key to the test is understanding what type of behavior implies revealed preference in the context of sequential search. It is not the case that the final choice of $x$ over $y$ indicates that $x$ is preferred to $y$, as the decision maker may simply be unaware of $y$. However, if a subject at some point chooses $y$ and replaces it with $x$ then, under the sequential search model, they must be interpreted as preferring $x$ to $y$. The fact that $y$ has previously been chosen indicates that the subject is aware of it. However, the subject has later rejected $y$ in favor of $x$, indicating that the latter must be preferred. The test of sequential search is particularly simple in the current setting if we assume that utility is monotonically increasing in money, as it corresponds precisely to the requirement that successive recorded values in the choice process must be increasing. We refer to this as Condition 1:

**Condition 1** If option $y$ is selected at time $t$ and option $x$ is selected at time $s > t$, it must be the case that the value of $x$ is no less than the value of $y$.[13]

Condition 1 is necessary and sufficient for a subject to be characterized by sequential search with a utility function that increases in money.

In order to test whether our subjects are close to satisfying Condition 1, we use a measure of consistency proposed by Houtman and Maks [1985]. The Houtman-Maks (HM) index is based on calculating the largest fraction of observations that are consistent with a condition, which can be determined by finding the minimum number of observations that have to be removed before the condition is satisfied. The underlying idea is that a data set which requires fewer such removals is "closer" to satisfying Condition 1 than one that requires more removals. In this case, we specifically ask how many selections have to be removed from a subject's data set before Condition 1 is satisfied. The resulting HM index is normalized by dividing through by the total number of observations, so that the HM index takes a value between 0 and 1, which can be interpreted as the largest fraction of a data set that satisfies condition 1.[14]

[13]Note that the choice process methodology only identifies a subset of searched objects: anything that is chosen at some point we assume must have been searched, but there may also be objects that are searched but never chosen, which we cannot identify. Combining our technology with a method of identifying what a subject has searched (for example Mouselab or eye tracking) would therefore be of interest.

[14]To give a concrete example, consider that for one subject we observe they initially select an option worth 7, then one worth 6, then 8 and 9. Such a subject would not be consistent with condition 1, as their initial switch would be to a lower value. However, if we removed their second selection, the choice process data would show them switching from value 7 to value 8 to value 9 – in line with condition 1. Thus this subject would have an HM index of 0.75, as 1 of their 4 observations would have to be removed to make their data consistent with condition 1. We consider this subject closer to satisfying condition 1 than one who switched from value 7 to value 6 to value 9 to value 8. We would have to remove two observations from this subject's data to make them consistent with condition 1, giving them an HM index of 0.50.

FIGURE 3. ABOUT HERE.

Figure 3 shows the distribution of HM indices for all 76 subjects. Over 40 percent of our subjects have an HM index above 0.95, while 70 percent have an HM index above 0.9 – meaning that over 90 percent of their switches are consistent with Condition 1, and therefore consistent with sequential search. Figure 3 also shows the simulated distribution of HM indices for 76,000 subjects who chose at random – a measure of the power of our test (see Bronars [1987]). Clearly, the two distributions are very different, as confirmed by a Kolmogorov-Smirnoff test ($p < 0.0001$).

This analysis suggests that, for the population as a whole, sequential search does a good job of describing search behavior. We can also ask whether the behavior of a particular subject is well described by the sequential search model – if so, we describe this subject as a sequential search type. To identify such types, we compare each subject's HM index with the HM indices of 1,000 simulations of random data with exactly the same number of observations in each round as that subject. We find that only 1 subject has an HM index below the median HM index of the corresponding random choice process data, and only 8 subjects have an HM index lower than the 95th percentile. For the remainder of the paper we focus on the 68 out of 76 subjects who have an HM index above this benchmark.[15]

FIGURE 4. ABOUT HERE.

One feature of the sequential search model is that it revives the concept of revealed preference in a world of incomplete information. Panel A of figure 4 shows how close our subjects are to satisfying the standard rationality assumption in each of our treatments, by showing the proportion of rounds in which the best alternative is chosen. Panel B shows how close our subjects are to satisfying rationality for sequential search in each treatment by calculating the HM index with respect to Condition 1. Two key facts stand out in this figure. First, the level of irrationality

[15] An alternative measure of the failure of condition 1 would be to calculate the minimum total change in payoff needed in order to adjust the data to satisfy condition 1. For example, if an object worth 12 was selected first and then one worth 4, we would have to make a reduction of 8 to bring the data in line with condition 1. On the other hand, if a subject selected 5 and then 4, only a reduction of 1 would be needed.

The correlation between these two measures is very high in our sample: the Spearman's rank correlation is 0.86. However, our subjects perform worse relative to the random benchmark according to this measure than according to the standard HM index. Using the new measure, 62 out of 76 subjects can be categorized as sequential search types using the 95th percentile of random choice simulations. This suggests that, when our subjects mistakenly switch to worse objects, they sometimes make large errors in terms of dollar value.

as measured by the standard definition of revealed preference is far higher than the sequential search measure. Second, while there is strong evidence of increasing irrationality in larger and more complex choice sets according to the standard measure, such effects are minimal according to the sequential search measure. Using the latter, there is no effect of set size, and only a small effect of complexity on mistakes.

### *B. Satisficing and Reservation Utility*

In his pioneering model of bounded rationality, Simon [1955] suggested that decision makers do not optimize, but rather search through a decision set until they achieve a satisfactory level of utility. One factor that has held back research on satisficing behavior is that the model has typically been interpreted in terms of its implications for final choices alone. The problem in this regard is that the simplest form of satisficing cannot be separated from utility maximization on the basis of choice alone: both are characterized by final choices that obey the weak axiom of revealed preference.[16]

In this section we use choice process data to shed new light on satisficing behavior. The essential advantage that choice process data provides is that it opens up to observation occasions on which subjects continue to search having uncovered unsatisfactory objects. This allows us to directly test the reservation stopping rule and estimate reservation values for our different treatments.

The first indication that our subjects exhibit satisficing behavior is shown in figure 5. This shows how the value of the selected object changes with order of selection for each of our six treatments. Each graph has one isolated point and three segmented lines. The isolated point shows the average object value for those who stop at the first object chosen.[17] The first segmented line shows the average value of each selection from rounds in which one switch was made. The next segmented line shows the average value of each selection in rounds where 2 switches were made, and the final segmented line for rounds in which 3 switches were made.

FIGURE 5. ABOUT HERE.

Figure 5 is strongly suggestive of satisficing behavior. First, as we would expect from the

[16]This is true in the version of the satisficing model in which decision makers always search through choice objects in the same order and the set of satisficing objects is fixed. If the order of search can change over time, then the satisficing model has no implications for final choice.

[17]Following the initial switch away from the zero value option.

proceeding section, aggregate behavior is in line with sequential search: in all but one case, the average value of selections is increasing. Second, we can find reservation values for each treatment such that aggregate behavior is in line with satisficing according to these reservations. The horizontal lines drawn on each graph show candidate reservation levels, estimated using a technique we describe below. In every case, the aggregate data show search continuing for values below the reservation level and stopping for values above the reservation level, just as Simon's theory predicts.

## ESTIMATING RESERVATION LEVELS

In order to estimate reservation utility for each treatment, we consider a stochastic generalization of the reservation strategy. We assume that all individuals in a given choice environment have the same constant reservation value $\overline{v}$ and experience variability $\varepsilon$ in this value each time they decide whether or not to continue search. Further, we assume this stochasticity enters additively and is drawn independently and identically from the standard normal distribution.[18] Let $v$ be the value of the item that has just been evaluated, and so the decision maker (DM) uses the following strategy to determine whether to continue searching through the choice set:

$$\begin{aligned} \text{search stops if } v &\geq \overline{v} + \varepsilon, \\ \text{search continues if } v &< \overline{v} + \varepsilon, \end{aligned}$$

where $\varepsilon \sim N(0, 1)$.

We can recast this procedure as a binary choice model. Let $k$ be a decision node, $v_k$ be the value of the object uncovered and $x_k$ be the choice made at that decision node, with $x_k = 1$ if search stops and $x_k = 0$ if search continues. Then

$$x_k = 1(v_k - \overline{v} - \varepsilon_k \geq 0), \tag{1}$$

[18] There are at least two ways to interpret the additive error term in this model. The first is that subjects calculate each option perfectly but only have a rough idea of their reservation value. The second is that subjects have a clear idea of their reservation value, but see the value of each option with some error.

The existing literature regarding stochastic choice models is summarized in Blavatsky and Pogrebna [2007]. Models can broadly be categorized into two types. The first are "tremble" models of the type used in Harless and Camerer [1994]. For any given decision, there is a constant probability that the subject will make a mistake. All types of mistake are then equally probable. The second type assumes that the value of each option is observed with some stochastic error. Different models assume different error structures, but all assume that small errors are more likely than large ones.

Our estimation technique uses a model from the second category - the Fechner Model of Heteroscedastric Random Errors which assumes that the reservation value is observed with an additive, normally distributed error term. In our setting, we find the tremble class of models implausible - neither intuition nor the data supports the idea that small errors are as likely as large ones.

It terms of the precise distribution of the error term, we tested other common alternatives – logistic and extreme value errors. The results under these alternative assumptions were essentially the same.

where $1(\cdot)$ is the indicator function.

An individual will stop searching if $\varepsilon_k \leq v_k - \overline{v}$, so the probability of stopping search is $\Phi(v_k - \overline{v})$, where $\Phi$ is the cumulative density function of the standard normal distribution. Similarly, search will continue if $\varepsilon_k > v_k - \overline{v}$, so the probability of search continuing is given by $1 - \Phi(v_k - \overline{v}) = \Phi(\overline{v} - v_k)$.

Thus, to estimate the parameter $\overline{v}$ with maximum likelihood, we use the log likelihood function

$$\ln \mathcal{L} = \sum_{k=1}^{K} \left[ x_k \ln\left(\Phi\left(v_k - \overline{v}\right)\right) + (1 - x_k) \ln\left(\Phi\left(\overline{v} - v_k\right)\right)\right] \tag{2}$$

and find the value of $\overline{v}$ that maximizes $\ln \mathcal{L}$.

To employ this procedure using our data, we consider each selection made by a subject as a decision node. We need to identify when search has stopped, and when it has continued. The latter is simple: search continues if a subject switches to another alternative after the current selection. Identifying stopped search is slightly more complicated. If we observe that a subject does not make any more selections after the current one, then there are three possibilities. First, they might have continued to search, but run out of time before they found a better object. Second, they might have continued to search, but already have selected the best option. Third, they might have stopped searching. We therefore consider a subject to have stopped searching at a decision node only if they made no further selections, pressed the submit button, and the object they had selected was not the highest value object in the choice set.

Choice process data is clearly useful for directly estimating reservation values. If we ignore data on the choice process and instead consider only standard choice data, we cannot use the same estimation strategy because it requires observations of subjects continuing to search as well as observations in which they stop searching. Choice data is composed entirely of the latter, so it only indicates when search has stopped, not when it continues.

## RESULTS: ESTIMATED RESERVATION LEVELS

Because we assume that all individuals have the same distribution of reservation values in a given environment, we pool together all selections within each treatment for the 68 participants whose choice data is best modeled with sequential search. Table 3 shows the estimated reservation levels for each treatment, with standard errors in parentheses.

Table 3 reveals two robust patterns in the estimated reservation levels. First, reservation levels decrease with complexity: using a likelihood-ratio test, estimated reservation levels are significantly lower for high complexity treatments than for low complexity treatments at all set sizes

TABLE 3—ABOUT HERE.

($p < 0.001$). Second, reservation levels increase monotonically with set size (significantly different across set sizes for both complexity levels with $p < 0.001$).

One question that this estimation strategy does not answer is how well the RBS model explains our experimental data. In order to shed light on this question, we calculate the equivalent of the HM index for the RBS model with the estimated reservation levels of table 3. For each treatment, we calculate the fraction of observations which obey the reservation strategy (i.e. subjects continue to search when they hold values below the reservation level and stop when they have values above the reservation level).

TABLE 4—ABOUT HERE.

The results, aggregated across all subjects, are shown in table 4. The estimated RBS model describes about 86 percent of observations for treatments with simple objects and about 78 percent for complicated objects. Both of these percentages are significantly higher than the random benchmark of 50 percent (where people arbitrarily stop or continue at each decision node) at the 1 percent level.

There is significant heterogeneity across individuals with respect to how well they follow a fixed reservation stopping rule. While the majority of subjects have an HM index above 75 percent, some have extremely low scores and are clearly poorly described by an RBS model with the given estimated reservation levels. In order to ensure these individuals are not affecting our estimates in table 2, we repeat the estimation of reservation strategies without those subjects who have an HM index below 50 percent (an additional 6 subjects). These results are in table 3 under the rows for "Reservation-based search types." The estimated reservation levels are very similar to those for the whole sample.

### *C. Reservation Utility or Reservation Time?*

A natural question is whether our data is consistent with other stopping rules. One obvious candidate is a stopping rule based on a reservation time, in which subjects search for a fixed time and select the best option found subject to this time constraint. In order to test this possibility, we redraw the graphs of figure 5, but show the average time of each switch, rather than the

average value. If subjects are using a fixed stopping time strategy then we expect the graphs to look like those in figure 5 – for each treatment we should be able to find a stopping time such that, on average, subjects continue searching until the stopping time has been breached, stopping immediately thereafter.

FIGURE 6. ABOUT HERE.

The results of the above analysis are shown in figure 6. The graphs provide no support for the reservation time stopping rule. Unlike in figure 5, there is generally no "reservation time" such that subjects continue to search for times below this level and stop for times above that level (the horizontal lines on each graph show a reservation stopping time estimated using the procedure describes in section B). Instead, those who identified a high value object with their first selection stopped quickly, while those who made the most switches took significantly longer. This is precisely as the reservation utility model would suggest, and runs counter to the predictions of the reservation time model.

## IV. Search Order and Choice

In this section we show that choice process data provides insight into the order of search, and that this information can help predict when subjects will do badly in particular choice sets.

### *A. Aggregate Search Order*

#### TOP-BOTTOM SEARCH

We begin by examining whether the position of an object on the screen helps to predict when it will be searched. In experiment 2, people tend to search from the top of the screen to the bottom of the screen: the average screen position tends to increase (i.e. be further down the screen) with later selections. While the relation is not completely monotonic, regression analysis confirms that it is significant.[19] This relationship is more pronounced for choice sets with simple, rather than complex objects.[20]

[19] Regressing selection number on the screen position of the selection gives a coefficient of 0.028, significant at the 1 percent level (allowing for clustering at the subject level).

[20] For complexity 3 choice sets, regressing selection number on the screen position of the selection gives a coefficient of 0.035, significant at the 1 percent level, while for complexity 7 sets the coefficient is 0.018, not significant at the 10 percent level.

### Simple-Complex Search

As explained in the experimental design section, experiment 3 presented subjects with choice rounds in which objects varied in complexity within the same round, enabling us to explore the impact of complexity on search order. On average, subjects tend to search simple objects before complex ones. Complexity is uncorrelated with value, but presumably is correlated with search cost, so it is reasonable for subjects to adopt such a search order.

In experiment 3, we find that subjects in general search the screen from top to bottom, and from simple to complex objects. While neither relationship is completely monotonic, regression analysis confirms that both are significant.[21]

## *B. Individual Search Order*

We can augment our analysis of aggregate search behavior by looking at the search patterns of individual subjects. For the subjects who participated in experiments 2 and 3, we looked for subjects whose behavior is consistent with "Top-Bottom" (TB) search, and those whose behavior is consistent with "Simple-Complex" (SC) search. The former are subjects whose search order takes them from the top to the bottom of the screen, while the latter are subjects whose search takes them from simple to complex objects.

We categorize subjects by calculating the HM indices assuming each of these two search orders. We first assume that a subject is a TB searcher and calculate the fraction of observations that are consistent with this search order – in other words, the fraction of observations for which objects selected later appear further down the screen. We then repeat the procedure assuming that the subject is an SC searcher. A subject is categorized as being a TB or SC searcher if their HM index for that search order is in the 95th percentile of a benchmark distribution constructed using random search orders. With this criterion, the majority of subjects in experiment 2 are well described by TB search. In experiment 3, eight subjects are categorized as both TB and SC searchers, six as just TB searchers, three as just SC searchers, and only three subjects are categorized as neither.

The assumption that subjects search in a particular order suggests another, stronger test of sequential search. Thus far, we have assumed that we only know an object has been searched if it has been chosen at some point. However, if we believe that a subject is searching (for example) from top to bottom, and we see them select the object in screen position 10, then we could further assume that they have searched all objects in positions 1 to 9. In this case, the test for sequential

[21] Regressing selection number on the screen position and complexity of the object selected gives coefficients of 0.034 and 0.132 respectively, both significant at the 1 percent level (allowing for clustering at the subject level).

search is whether or not, at any given time, the value of the currently chosen object is higher than all the objects that fall earlier in the assumed search order.

TABLE 5—ABOUT HERE.

Table 5 documents the results of this more stringent test of sequential search for experiments 2 and 3, assuming that people do search from top to bottom. For each treatment it shows the proportion of observations in which the chosen object has a higher value than any alternative that appears above it in the list. We report the results of this test for all subjects, and only those subjects who we classify as TB searchers.

In the low complexity choice environment, subjects who are classified as top down searchers perform relatively well on this test. Across all treatments, they behave in line with top to bottom sequential search in about 75 percent of cases. They also do significantly better in this test than subjects that we do not classify as TB.[22] However, even those we categorize as TB searchers often violate this condition in more complicated choice sets. This suggests that, in more complicated choice sets, even subjects who generally search from top to bottom may not fully examine all of the objects along the way.

### *C. Search Order and Choice*

We provide two examples from experiment 3 that illustrate how knowledge of a subject's search order helps to identify choice sets in which they will fail to find the best option. Example 1 is from a round in which the highest valued item has low complexity but occurs at the end of the list (see panel A of figure 7 – the best option is highlighted in green). We would expect this to be a choice set in which TB searchers would do worse and SC searchers would do better. This turns out to be the case. Pure TB searchers find the best option less often (80 percent of the time) than pure SC searchers (100 percent of the time).[23] Unfortunately, due to the small sample size, these numbers are not significant at the standard levels of significance.

Example 2 is from a round in which the highest valued item has high complexity but occurs very early in the list (panel B of figure 7). In this case, we would expect TB searchers to do better

[22]Controlling for selection number and position on screen, the coefficient on being a Top-Bottom searcher is negative and significant ($p = 0.005$) in a regression where success or failure of top down sequential search is the dependent variable.

[23]In order to avoid potential circularity in our argument, we reestimate subjects' search types excluding these two example rounds. The results are unchanged.

FIGURE 7. ABOUT HERE.

and SC searchers to do worse. Indeed, pure top-bottom searchers find the best option more often (80 percent of the time) and pure simple-complex searchers find it less often (66 percent).

## V. Choice Process and Standard Choice Data

The choice process experiment has incentives that are different from those operating in a standard choice environment. To understand the impact that these nonstandard incentives have on decisions, we characterize optimal stopping strategies in a sequential search model that covers both the standard experiment and the choice process experiment. We also explore behavioral differences between experiments. In this respect we take advantage of the fact that, in experiment 1, subjects were able to, and indeed did, select options prior to hitting the submit button and finalizing their choices.[24] We can use these intermediate clicks to test our search models in the standard choice environment of experiment 1, just as we did in experiment 2. Figure 8 illustrates the distribution of recorded choice switches in experiment 1 and 2.

FIGURE 8. ABOUT HERE.

### *A. Condition 1 in Experiment 1*

We use the intermediate choice data from experiment 1 to explore evidence for Condition 1, the sequential search condition, in the standard choice environment. Figures 9 and 10 suggest that subjects in experiment 1 do indeed search sequentially. Figure 9 compares the distribution of HM indices for the choice process and standard choice experiments, which shows that, if anything, data from the standard choice environment are more in line with sequential search than choice process data. Figure 10 repeats the analysis of figure 4 for standard choice data, showing by treatment the proportion of observations that do not violate standard revealed preference and do not violate the sequential search model. Again, we see little effect of treatment on violations of the sequential search based measure. Indeed, there are even fewer violations of Condition 1 for the complex objects than there were in experiment 2.

[24]While there was no direct financial incentive for changing the selection in experiment 1, there may be a psychological incentive if object selection aids memory.

FIGURE 9. ABOUT HERE.

FIGURE 10. ABOUT HERE.

### B. A Model of Optimal Search

Given that Condition 1 applies generally in both experiments 1 and 2, we develop an optimizing model of sequential search that covers both experimental designs. The search cost is specified in utility terms, as in Gabaix et al. [2006]. The DM is an expected utility (EU) maximizer with a utility function $u : X \to \mathbb{R}$ on the choice set $X$. We endow the searcher with information on one available option of utility $\bar{u}_0$ at time $t = 0$, a period in which no choice is to be made. We normalize $u : X \to \mathbb{R}$ so that the endowed prize has an EU of zero. At each subsequent time $1 \leq t \leq T$, the DM faces the option of selecting one of the options already searched, or examining an extra option and paying a psychological search cost $\kappa > 0$ (in EU units). To break the otherwise rigid connection between time and the number of objects searched, we introduce stochasticity into the search time. We introduce parameter $q \in (0, 1)$ as the probability that searching an object in hand for one period will result in its identity being known. If this does not happen, the same geometric probability applies in the following periods. Once identified, the new object is immediately available to replace the object inherited from the prior period. Once search stops, the agent must choose one of the uncovered objects.[25]

The agent's search strategy from any nonempty finite subset $A \subset X$ is based only on the size $M$ of the set of available objects in $A$, not the identities of these objects. Each available prize i

The optimal search strategy depends on the surplus function $S : \mathbb{R} \to \mathbb{R}_+$, defined by

$$S(x) = \int_x^\infty [z - x]\, dF(z). \tag{3}$$

This is the expected value of searching one more period, assuming that additional search affects the outcome. Note that $S(x)$ is continuous and strictly diminishing in $x$ to the upper bound of the EU distribution, where it hits zero. Note also that $S(0) = \int_0^\infty z dF(z) \equiv E(z)$.

It is useful in solving the model to define the strictly increasing function $\rho : \mathbb{R} \to \mathbb{R}_+$, and to define $\bar{x}$ as the point at which this function takes value 1,

$$\rho(x) \equiv \frac{\kappa}{qS(x)}, \tag{4}$$

$$\rho(\bar{x}) = 1. \tag{5}$$

By way of interpretation, $\rho(x)$ measures the ratio of the costs of one more search to the expected benefits in terms of surplus generation. Note that $\bar{x} > 0$ (or $\frac{\kappa}{q} < E(z)$) is necessary for the problem to be nontrivial, since otherwise it is better to stop immediately and accept the object of value 0.

THEOREM 2: *For any time $t$, $1 \le t \le T$, define the reservation utility level $u^R(t)$ as the unique solution to the equation,*

$$\rho(u^R(t)) = J(t). \tag{6}$$

*It is uniquely optimal to stop search and select the best prior object searched of utility $\bar{u}_{t-1}$ if $\bar{u}_{t-1} > u^R(t)$, to continue search if $\bar{u}_{t-1} < u^R(t)$, with both strategies optimal if $\bar{u}_{t-1} = u^R(t)$.*

In the standard choice environment, $J(t) = 1$ for all $t$. Theorem 2 implies that the optimal strategy is a fixed reservation level $\bar{u}^R$ defined as the solution to the following equation:

$$\int_{\bar{u}^R}^\infty (z - \bar{u}^R) dF(z) = \frac{\kappa}{q}. \tag{7}$$

This reservation level is decreasing in the cost of search $\kappa$, but is invariant to both the size of the choice set and the number of options that remain unsearched.

In the choice process environment, $J(t)$ is decreasing. Theorem 2 therefore implies that the optimal strategy is defined by a declining reservation level that depends only on $J(t)$, not the size of the choice set or the number of remaining alternatives. For any time $t > 0$, the reservation level in the choice process environment will be below the level in the equivalent standard choice environment. This result is intuitive: for any $t > 0$, the probability of further search affecting the outcome is higher in the standard choice environment than the choice process environment.

### *C. Stopping Rules in Experiments 1 and 2*

The theoretical model suggests that, if anything, standard choice data should be better explained by the satisficing model than the choice process data. We begin by repeating the analysis of section III to determine whether this is the case. We find that the standard choice experiments are indeed well explained by a fixed reservation rule. Figure 11 recreates the analysis of figure 5, and suggests that a reservation stopping rule broadly describes the aggregate data. Table 6 shows that the estimated reservation levels for the standard choice data exhibit the same comparative statics as do those for the choice process data.[26] Table 7 shows that the estimated HM indices for these reservation levels in the standard choice data are roughly similar for lower complexity and smaller for higher complexity. This suggests that there is little qualitative distinction between behavior in the standard choice and choice process environments.

FIGURE 11. ABOUT HERE.

TABLE 6—ABOUT HERE.

TABLE 7—ABOUT HERE.

[26]For the analysis of table 6 we drop subjects who never switch in any round.

The optimal stopping model suggests that there should be two differences between the standard choice data and the choice process data. First, reservation levels should be lower in the choice process environment than in the standard choice environment. The analysis of table 6 offers some support for this. Estimated reservation levels are lower in 4 of the 6 treatments. Lower reservation levels would also explain why subjects in the choice process environment finish searching more quickly than those in the standard choice environment.

While differing incentives could explain why final choice performance is worse in the choice process environment than in the standard choice environment, another possibility is more mundane – experiment 2 had a time limit while experiment 1 did not. Experiment 4 allows us to determine which of these is the case, as it replicates the pure choice environment of experiment 1, but with a 2 minute time limit. The results suggest that the time limit is responsible for some, but not all of the difference. The average failure rate across all treatments is 31.2 percent for the standard choice experiment, 36.5 percent in the standard choice with time limit experiment, and 40.8 percent in the choice process experiment.[27] The difference in incentives does appear to impact performance in experiment 2 relative to that in experiment 1, over and above the effect of the time limit.

The theoretical model shows that, while a fixed reservation strategy is optimal in the standard choice data case, a declining reservation strategy is optimal in the choice process environment. We have already seen that a fixed reservation model does a good job of explaining both data sets, but it could be that a generalization that allows for a declining reservation level does even better.

We use a revealed preference approach to test for the possibility of a declining reservation level. The revealed preference implication of a declining reservation level is straightforward. If a subject stops searching and chooses an object $x$ at time $t$, but continues searching having found object $y$ at time $s > t$, it must be the case that $x$ is preferred to $y$. This is because the value of $x$ must be above the reservation value at time $t$, which is in turn above the reservation level at time $s$. Moreover, the value of $y$ must be below the reservation level at time $s$ as search is continuing. Thus $x$ must be preferred to $y$. In contrast, the revealed preference implication of a fixed reservation level is that $x$ is preferred to $y$ if search stops with $x$ at some time $t$ but continues with $y$ at some time $s$, *regardless of the relationship between $t$ and $s$*. Note that the fixed reservation model is a special case of the declining reservation model.

Armed with these observations, we can ask whether the declining reservation model helps to explain more of the choice process data than the fixed reservation model, by asking how many

[27]To calculate the average across all treatments, we calculate the average loss for each treatment and average across these.

times the relevant revealed preference condition is violated. We classify data as violating a particular revealed preference condition if option $x$ is revealed preferred to option $y$, but the value of $x$ is greater than the value of $y$. It turns out that the declining reservation model does not offer a better description of choice process data. While the declining reservation model has fewer violations in absolute terms (as has to be the case – any violation of the declining reservation model is also a violation of the fixed reservation model), the *proportion* of observations that violate revealed preference is higher – 24% for the fixed reservation model versus 32% for the declining reservation. Thus, our revealed preference approach finds little evidence that our subjects are responding to the choice process environment by implementing a declining reservation strategy.

### *D. Comparing Behavior across Treatments*

Our model of optimal search makes predictions concerning how reservation levels should vary between our experimental treatments. First, the optimal reservation level falls as the per unit search cost rises. Thus, assuming that search costs are higher for more complex objects, optimal reservation levels are lower in the higher complexity environment. Second, optimal reservation levels are independent of the size of the choice set: there is no increase in the optimal reservation level as the size of the choice set increases.

The comparative statics properties of our experimentally estimated stopping rules do not align perfectly with those of the optimal stopping rule. While we do find that subjects reduce their reservation level in response to higher search costs, they also tend to *increase* their reservation level as the size of the choice set increases.

There are two possible reasons for this discrepancy. First, subjects may be behaving optimally with respect to a different maximization problem. For example, our optimal model assumes that no learning takes place with respect the distribution of values of the objects in the choice set. While our subjects are explicitly told the distribution from which values are drawn, it may be that they in fact try to learn this distribution for every new choice set. In such a case, estimated reservation levels would tend to be greater in larger choice sets.

A second possibility is that subjects are acting suboptimally by increasing their reservation levels in larger choice sets: they are searching “too much” in larger choice sets relative to smaller ones. This may relate to findings from the psychology and experimental economics literature that show that people may prefer smaller choice sets (Iyengar and Lepper [2000], Seuanez-Salgado [2006]). One factor that potentially links these two findings is the concept of regret. Zeelenberg and Pieters [2007] show that decision makers experience more regret in larger choice sets and suggest that this can lead them to search for more information.

## VI. A Pure Random Error Model

Our explanation for subjects' failure to pick the objectively best option is based on incomplete sequential search. However, another possibility is that these failures result from calculation errors – subjects search the entire choice set but make errors when evaluating each option. In order to test this alternative explanation, we consider a simple model of complete search with calculation errors. We put a simple structure on the error process – subjects are modeled as if they see the true value of each object with an error that is drawn independently from an extreme value distribution. The mode of this distribution is 0, and the scale factor is allowed to vary with complexity level and set size. With these assumptions, we can estimate the scale factor for each treatment using logistic regression. Specifically, we find the scale factor that best predicts the actual choice in each choice set.[28] We allow for scale factors to differ between treatments.

Table 8 shows the estimated standard deviations from the calculation error model. This provides the first piece of evidence to suggest that the calculation error model is implausible. In large and complicated choice sets, the standard deviation needed to fit the data becomes very large – for example, in the complexity 3, size 40 treatment, the range between minus one and plus one standard deviation is \$7, while the mean value of our choice objects is just \$4.

TABLE 8—ABOUT HERE.

TABLE 9—ABOUT HERE.

Despite these large standard deviations, the calculation error model significantly underpredicts both the frequency and magnitude of our subjects' losses, as shown in table 9.[29] The prediction of subject performance under the estimated calculation error model was based on 1,000 simulations of each observed choice set, in which a draw from the estimated distribution was added to the value of each option and the object of highest total value was identified as being chosen.

[28]For example, if a value of 10 was chosen by a subject from {7, 10, 12}, then our estimation strategy would find the scale factor that gives the highest probability to choosing 10, given that all options are seen with their own error. With this approach, enough error must be applied so that the noisy signal of 10 appears larger than the noisy signal of 12, but not so much error that the noisy signal of 7 appears larger than the noisy signal of 10.

[29]Alternatively, we could have estimated the scale factor to best match the number of mistakes or magnitude of mistakes found in the data, but this would ignore the actual choices that subjects made, which may contain other unpredicted patterns.

FIGURE 12. ABOUT HERE.

A final problem with the calculation error model is that it should lead to far more violations of sequential search than we in fact observe. Were subjects to be making calculation errors of the magnitude required to explain final choices, we would expect to see them switch to worse objects more often than they do. We demonstrate this in figure 12. For this figure, the prediction of subject performance under the estimated calculation error model is based on simulations of choice process data assuming that values are observed with treatment-specific error.[30] Note that the predicted success rates for the calculation error model lie below the lower bounds of the 95 percent confidence interval bars for all treatments.

## VII. Concluding Remarks

We introduce a choice-based experiment that bridges the gap between revealed preference theory and the theory of search. We use it to classify search behaviors in different decision making contexts. Our central finding concerns the prevalence of satisficing behavior. Models of sequential search based on achievement of context dependent reservation utility closely describe our experimental data. These results suggest that the search theoretic lens may be of significant value in systematizing our understanding of boundedly rational behavior.

## REFERENCES

Bernheim, B. Douglas, and Antonio Rangel. 2008. "Choice-Theoretic Foundations for Behavioral Welfare Economics." In *The Foundations of Positive and Normative Economics*, ed. Andrew Caplin and Andrew Schotter. New York: Oxford University Press.

Blavatsky, Pavlo and Ganna Pogrebna. 2007. "Models of Stochastic Choice and Decision Theories: Why Both are Important for Analyzing Decisions." IEW - Working Papers iewwp319, Institute for Empirical Research in Economics.

[30] Simulated data was generated as follows. For each sequence of choice process data observed in experiment 2, we simulated 1,000 sequences of the same length. For each sequence, a draw from the value distribution (rounded to the nearest integer) was treated as the initial selection. The sum of this value and a draw from the treatment-specific error distribution was then compared to the sum of a second draw from the value distribution and a draw from the treatment-specific error distribution. If the latter sum was higher than the initial sum, then we assumed a switch occurred, and the value of the second draw from the value distribution was carried forward as the current selection. Otherwise we assumed that no switch occurred, and so the initial selection remained the current selection. Another draw from the value and error distributions was then made, and compared to the current selection plus error. This process was then repeated until the number of simulated switches was equal to the length of actual switches in sequence taken from experiment 2. We then calculated the ratio of correct switches (where the true value of the new selection was higher than the true value of the current selection) to the total number of switches.

Bronars, Stephen. 1987. “The Power of Nonparametric Tests of Preference Maximization.” *Econometrica*, 55(3): 693-698.

Campbell, Donald. 1978. “Realization of Choice Functions.” *Econometrica*, 46(1): 171-180.

Caplin, Andrew, and Mark Dean. Forthcoming. “Search, Choice, and Revealed Preference.” *Theoretical Economics*.

Eliaz, Kfir, and Ran Spiegler. Forthcoming. “Consideration Sets and Competitive Marketing.” *Review of Economic Studies*.

Gabaix, Xavier, David Laibson, Guillermo Moloche, and Stephen Weinberg. 2006. “Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model.” *American Economic Review*, 96(4): 1043-1068.

Gul, Faruk, and Wolfgang Pesendorfer. 2008. “The Case for Mindless Economics.” In *The Foundations of Positive and Normative Economics*, ed. Andrew Caplin and Andrew Schotter. New York: Oxford University Press.

Harless, David, and Colin Camerer. 1994. “The Predictive Utility of Generalized Expected Utilities Theories.” *Econometrica*, 62(6): 1251-1289.

Houtman, Martijn, and J. A. H. Maks. 1985. “Determining all Maximal Data Subsets Consistent with Revealed Preference.” *Kwantitatieve Methoden*, 19: 89-104.

Iyengar, Sheena, and Mark Lepper. 2000. “When Choice is Demotivating: Can One Desire Too Much of a Good Thing?” *Journal of Personality and Social Psychology*, 79: 995-1006.

Koszegi, Botond, and Matthew Rabin. 2008. “Revealed Mistakes and Revealed Preferences.” In *The Foundations of Positive and Normative Economics*, ed. Andrew Caplin and Andrew Schotter. New York: Oxford University Press.

Manzini, Paola, and Marco Mariotti. 2007. “Sequentially Rationalizable Choice.” *American Economic Review*, 97(5): 1824-1840.

Masatlioglu, Yusufcan, and Daisuke Nakajima. 2009. “Choice by Iterative Search.” Unpublished.

McCall, John. 1970. “Economics of Information and Job Search.” *Quarterly Journal of Economics*, 84(1): 113-126.

Ok, Efe. 2002. “Utility Representation of an Incomplete Preference Relation.” *Journal of Economic Theory*, 104: 429-449.

Payne, John, James Bettman, and Eric Johnson. 1993. *The Adaptive Decision Maker*. Cambridge, MA: Cambridge University Press.

Reutskaja, Elena, Johannes Pulst-Korenberg, Rosemarie Nagel, Colin Camerer, and Antonio

Rangel. Forthcoming. “Economic Decision Making under Conditions of Extreme Time Pressure and Option Overload: An Eyetracking Study.” *American Economic Review*.

Rubinstein, Ariel, and Yuval Salant. 2006. “A Model of Choice from Lists.” *Theoretical Economics*, 1(1): 3-17.

Seuanez-Salgado, Maria. 2006. “Choosing to have Less Choice.” FEEM Working Paper No. 37. http://ssrn.com/sequential searchtract=889311.

Simon, Herbert. 1955. “A Behavioral Model of Rational Choice.” *Quarterly Journal of Economics*, 69(1): 99-118.

Stigler, George. 1961. “The Economics of Information.” *Journal of Political Economy*, 69(3): 213-225.

VanLehn, Kurt. 1982. “Bugs are Not Enough: An Analysis of Systematic Subtraction Errors.” *Journal of Mathematical Behavior*, 3(2): 3-71.

Zeelenberg, Marcel, and Rik Pieters. 2007. “A Theory of Regret Regulation.” *Journal of Consumer Psychology*, 17(1): 3-18.

## Appendix – Proof of Theorem 1

PROOF:

We establish the result first for the final period, and then induct backward. Suppose that search continues until period $T$ and that there is at least one unsearched item left. Let $\bar{u}^T = (\bar{u}_0, ..., \bar{u}_{T-1}) \in \mathbb{R}^T$ denote the vector of highest utility objects encountered in prior periods, and let $H(T, \bar{u}^T)$ capture the expected utility in hand at time $T$ based on the possibility that the search clock had stopped strictly prior to period $T$,

$$H(T, \bar{u}^T) = \sum_{s=0}^{T-1} [J(s) - J(s+1)]\bar{u}_s.$$

If no search is conducted in period $T$, the payoff from stopping is $\pi^S(T, \bar{u}^T)$,

$$\pi^S(T, \bar{u}^T) = H(T, \bar{u}^T) + J(T)\bar{u}_{T-1}.$$

If search continues for one last period, then the payoff for the final period is still $\bar{u}_{T-1}$ unless a new object is identified (probability $q$), that object has utility higher than $\bar{u}_{T-1}$, and the random

choice time continues to period $T$,

$$\pi^C(T, \bar{u}^0) = H(T, \bar{u}^T) - \kappa + J(T)\left[\bar{u}_{T-1} + qS(\bar{u}_{T-1})\right].$$

Hence continued search is an optimal strategy if and only if,

$$qJ(T)S(\bar{u}_{T-1}) \geq \kappa,$$

or,

$$\rho(\bar{u}_{T-1}) = \frac{\kappa}{qS(\bar{u}_{T-1})} \leq J(T) = \rho(u^R(T)).$$

Since $\rho$ is a strictly increasing function, this implies that continued search is optimal if and only if $\bar{u}_{T-1} \leq u^R(T)$, stopping search is optimal if and only if $\bar{u}_{T-1} > u^R(T)$, establishing the result for period $T$.

Now assume that the identified strategy is optimal if search continues in period $t+1 \geq 2$, and consider the optimal strategy in period $t$ with prior maxima $\bar{u}^t = (\bar{u}_0, ..., \bar{u}_{t-1})$ and with $H(t, \bar{u}^t)$ the fixed expected utility should the search clock have stopped prior to period $t$. Continued search for one and only one period costs $\kappa$, yielding the expected surplus above $\bar{u}_{\tau-1}$ if the new search is effective. Hence it is worthwhile if and only if,

$$J(t)\left[\bar{u}_{t-1} + qS(\bar{u}_{t-1})\right] \geq \kappa,$$

or,

$$\rho(\bar{u}_{t-1}) \leq J(t) = \rho(u^R(t)).$$

Given that $\rho$ is strictly increasing, one and only one additional period of search dominates stopping if $\bar{u}_{t-1} < u^R(t)$, stopping immediately is strictly superior if $\bar{u}_{t-1} > u^R(t)$, while they are indifferent if $\bar{u}_{t-1} = u^R(t)$. To establish the induction hypothesis requires only that an individual for whom it is optimal to stop when considering one period continuation will not continue on the basis of expected gains in later periods. This can be ruled out, since if $\bar{u}_{t-1} \geq u^R(t)$, then since $u^R(t) > u^R(t+1)$ the induction hypothesis implies that search will certainly not continue beyond period $t+1$, making the single period argument in favor of stopping definitive.

Figure 1: A typical choice round

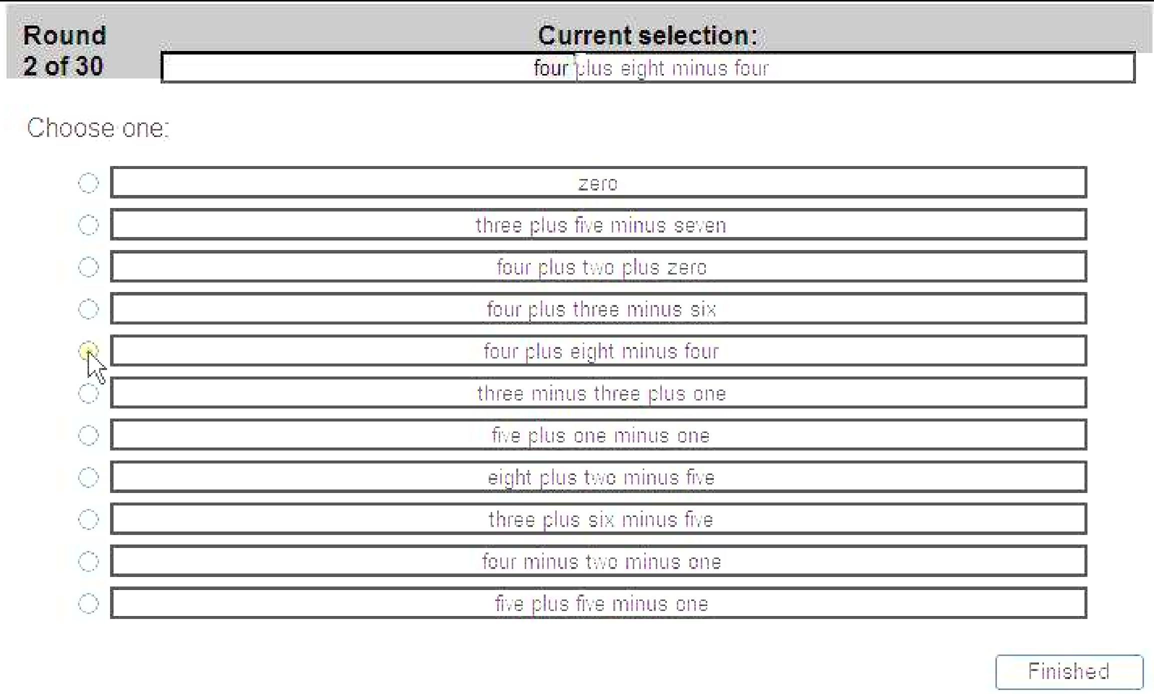


Figure 2: Number of switches per choice round (experiment 2)

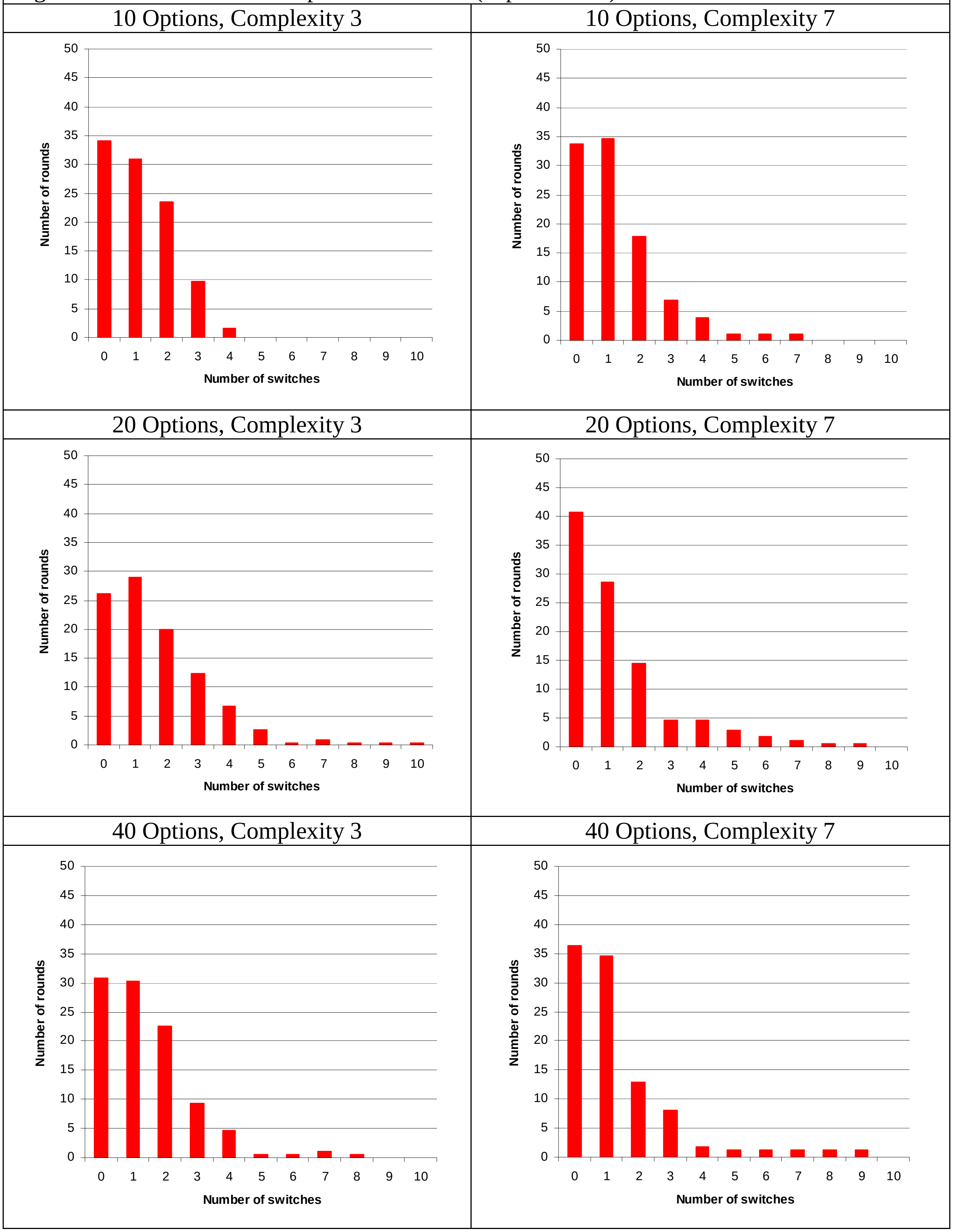


Figure 3: Distribution of HM indices for actual and random data (experiment 2)

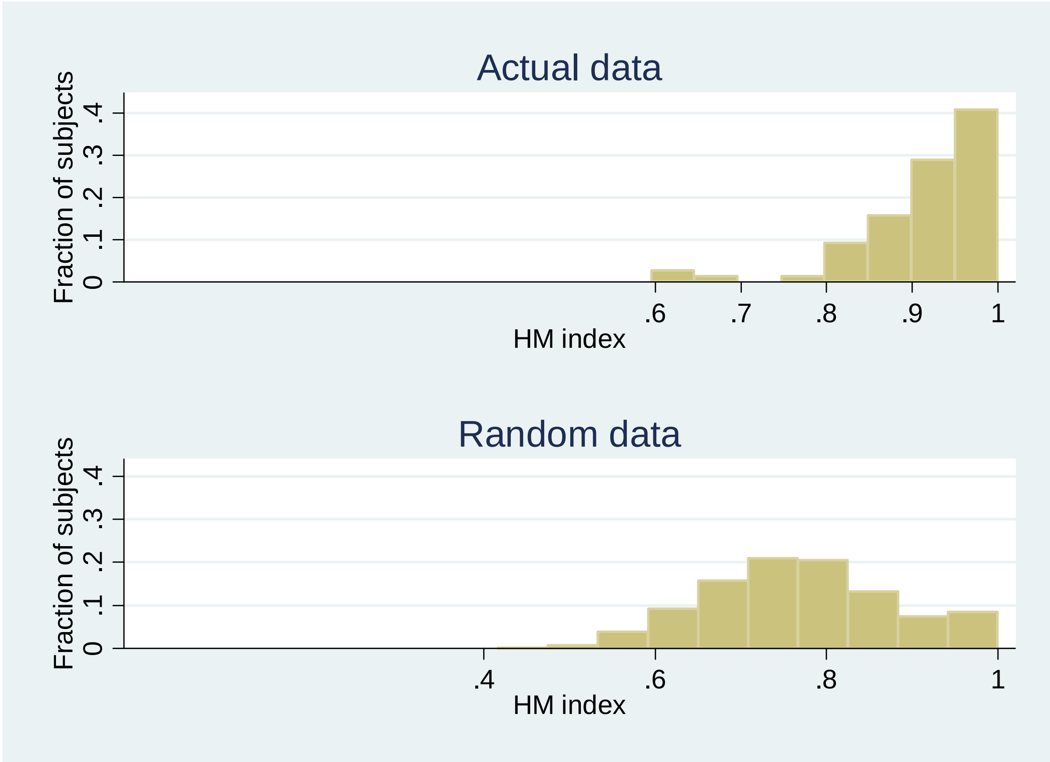
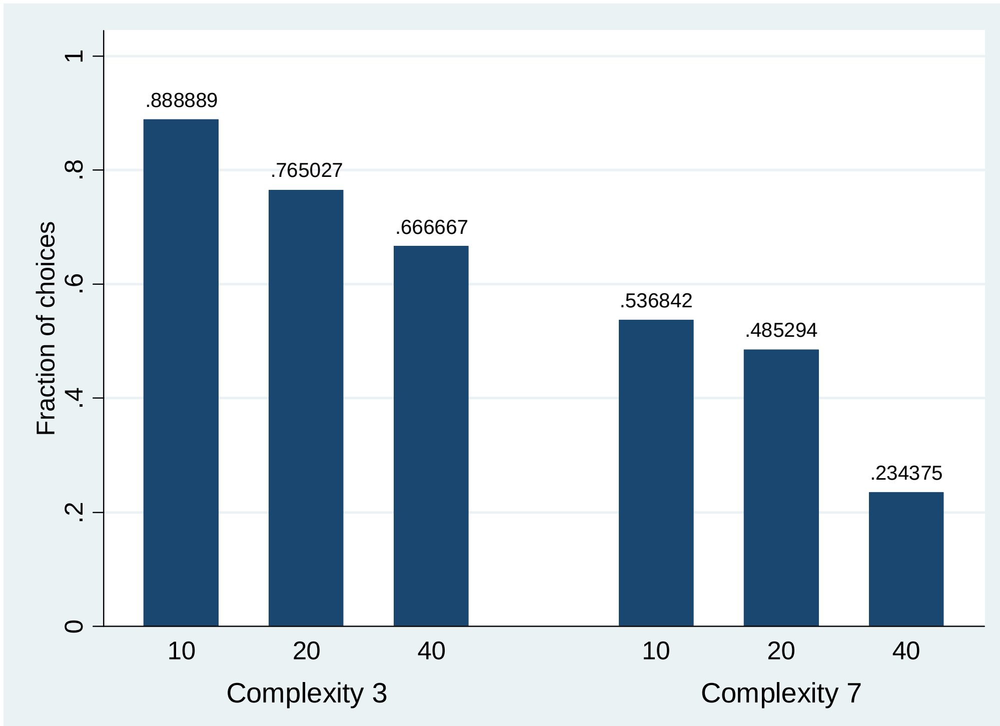


Figure 4: Proportion of choices consistent with rationality according to the standard model and the sequential search model (experiment 2)

Panel A: Standard model




Panel B: Sequential search model

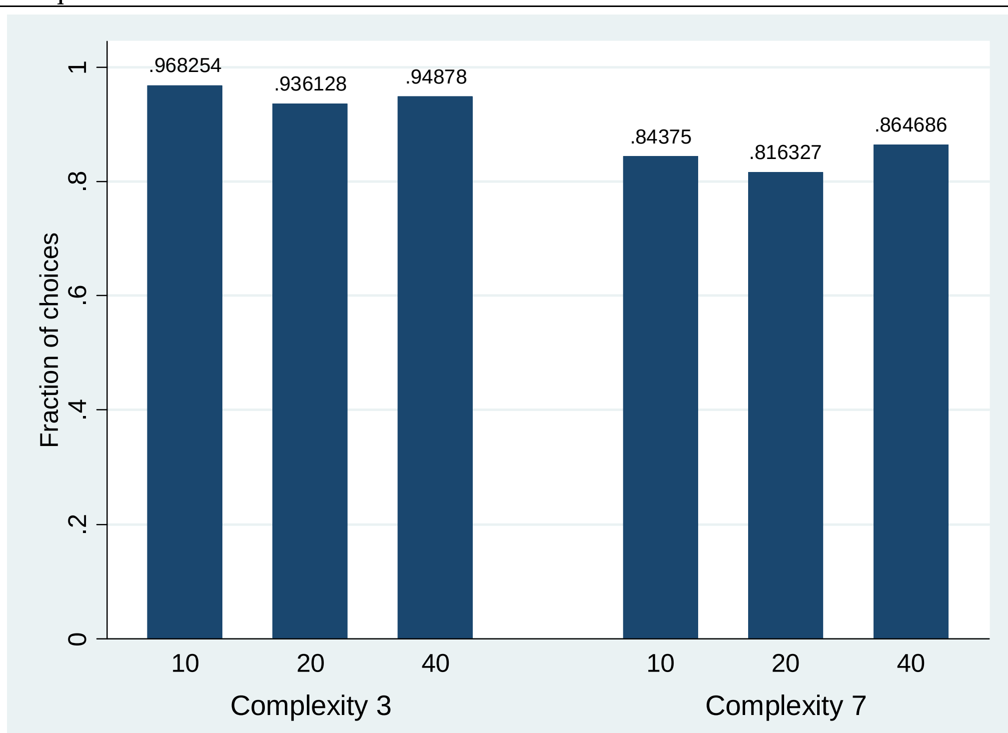
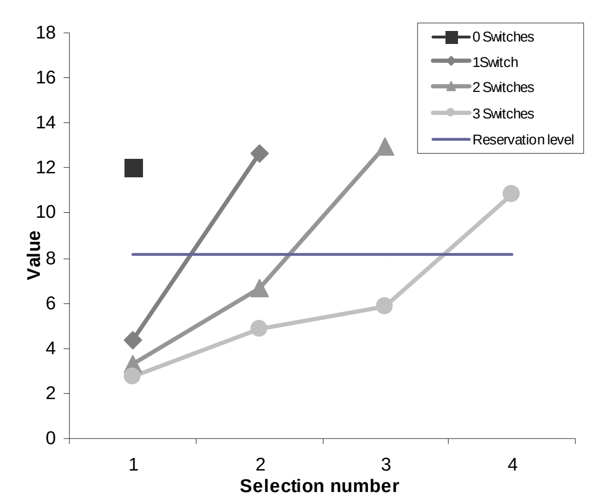
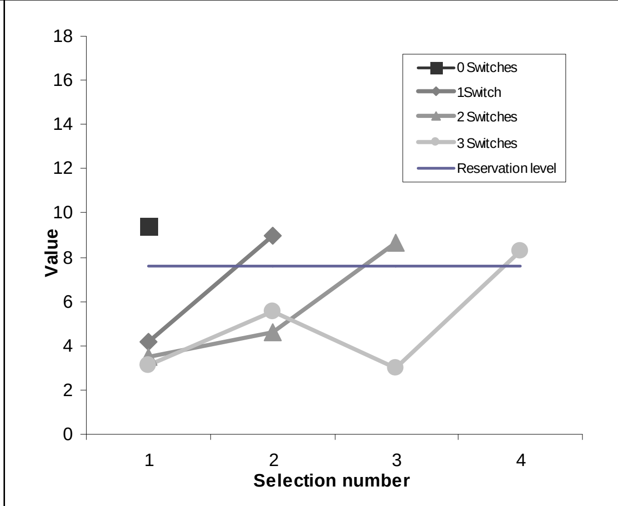


Figure 5: Average value by selection (experiment 2)

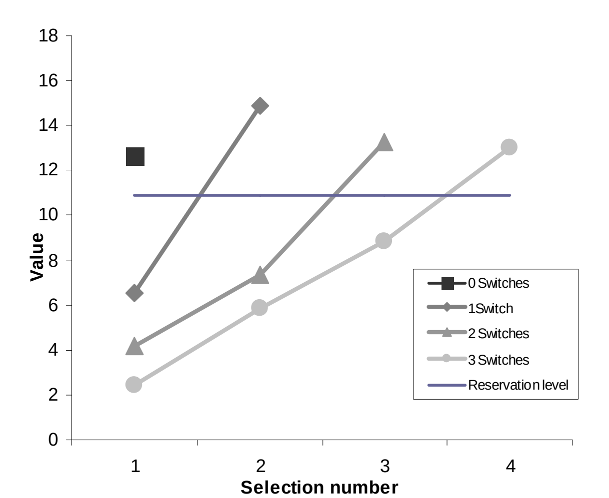
10 Options, Complexity 3



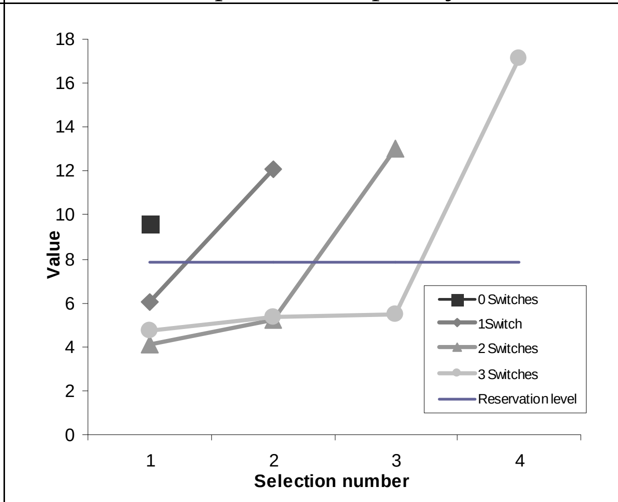

10 Options, Complexity 7



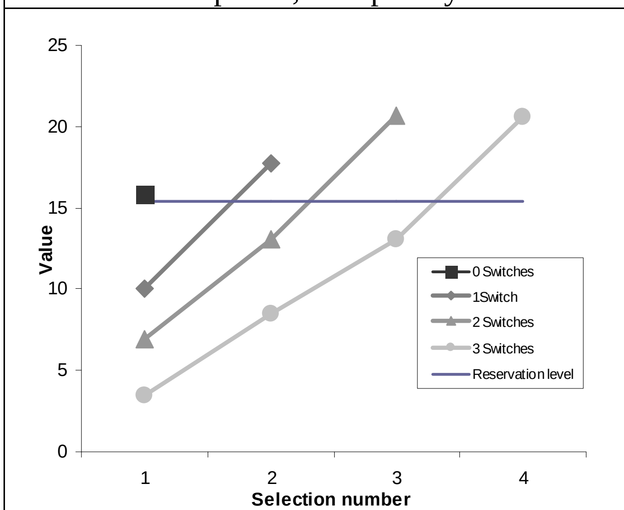

20 Options, Complexity 3




20 Options, Complexity 7




40 Options, Complexity 3




40 Options, Complexity 7

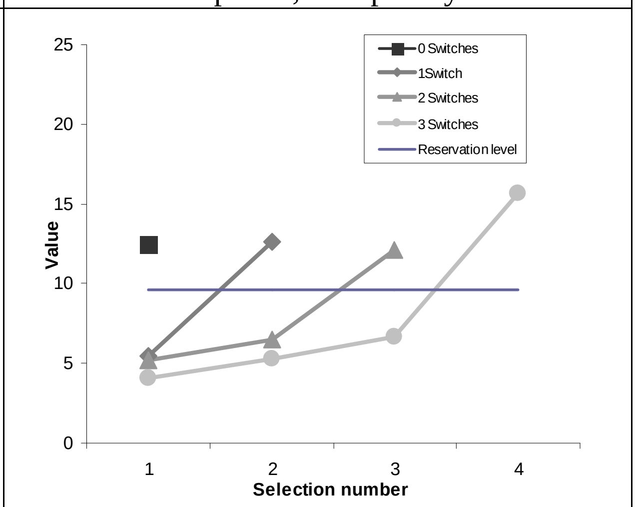
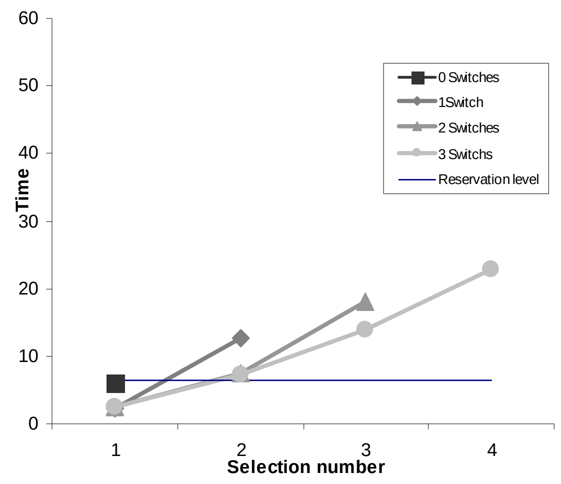
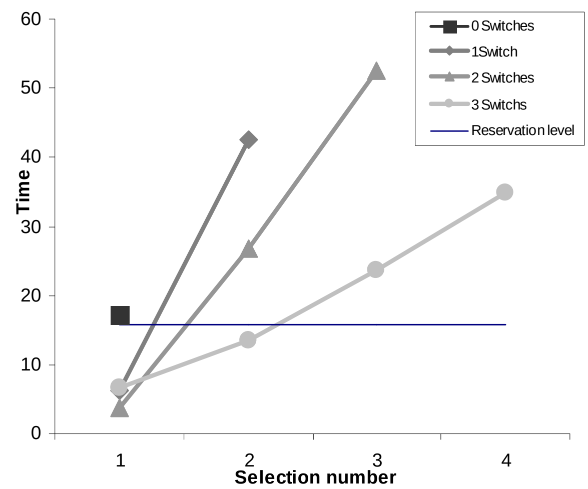


Figure 6: Average time by switch (experiment 2)

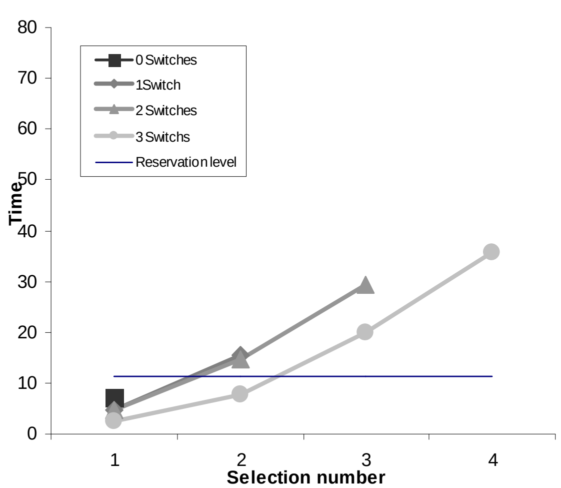
10 Options, Complexity 3



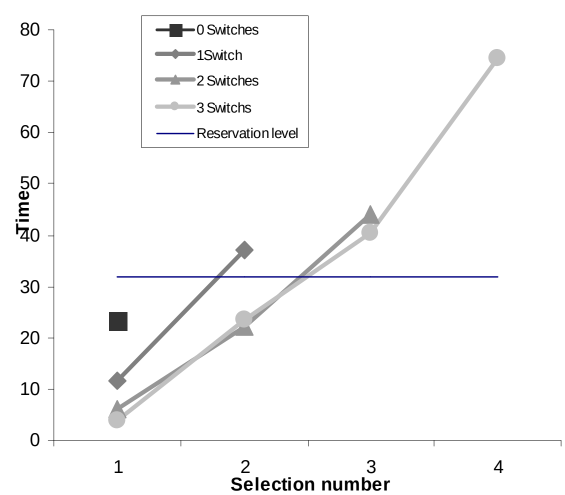

10 Options, Complexity 7



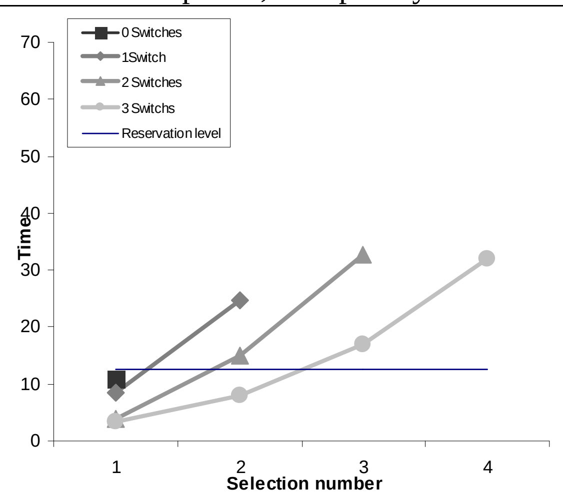

20 Options, Complexity 3




20 Options, Complexity 7




40 Options, Complexity 3




40 Options, Complexity 7

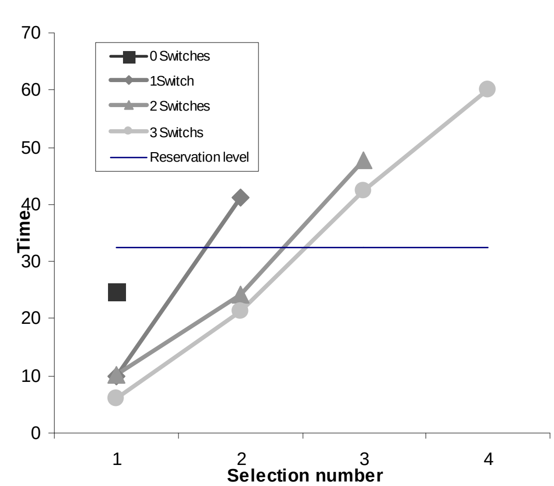
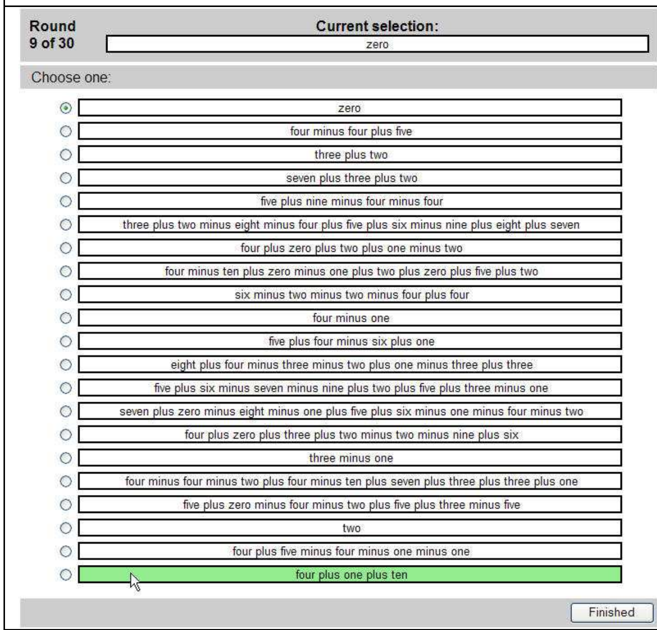


Figure 7: Choice sets with the highest valued object at the bottom of the screen and simple (panel A) and at the top of the screen and complex (panel B)

Panel A




Panel B

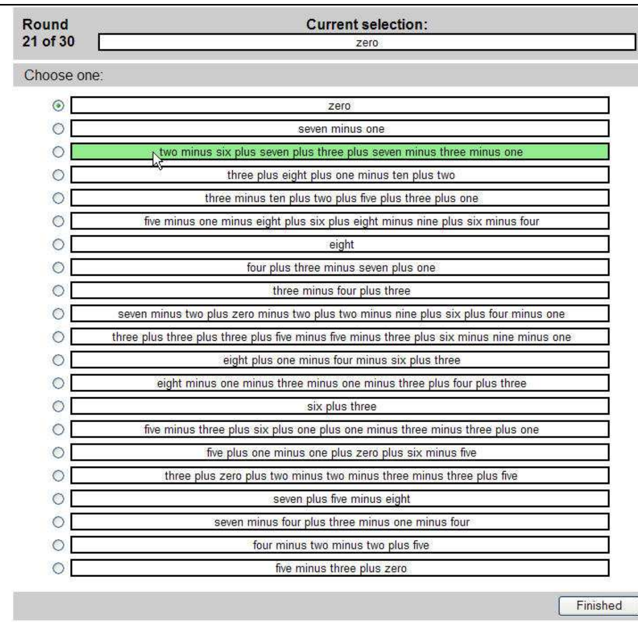


Figure 8: Number of switches for the choice process task (experiment 2) and the standard choice task (experiment 1)

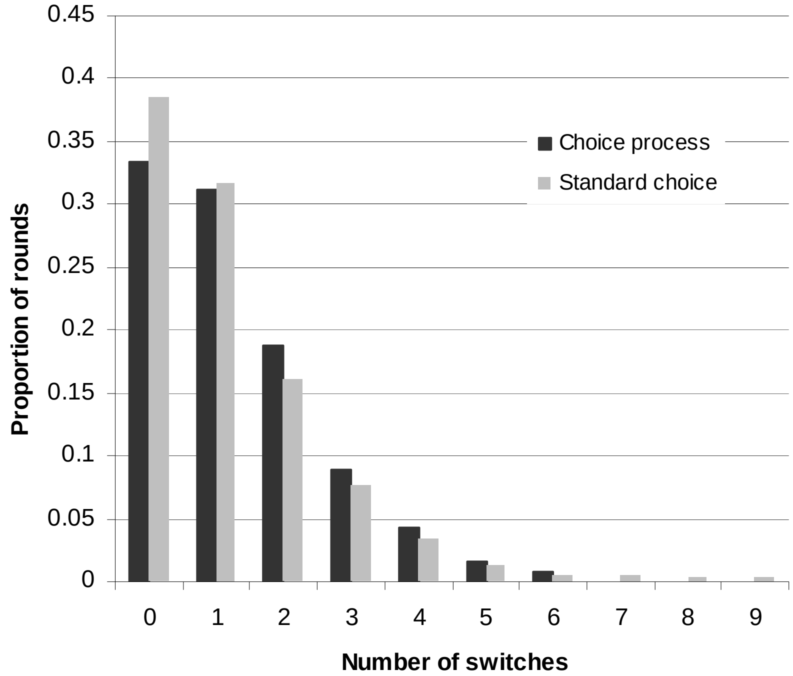
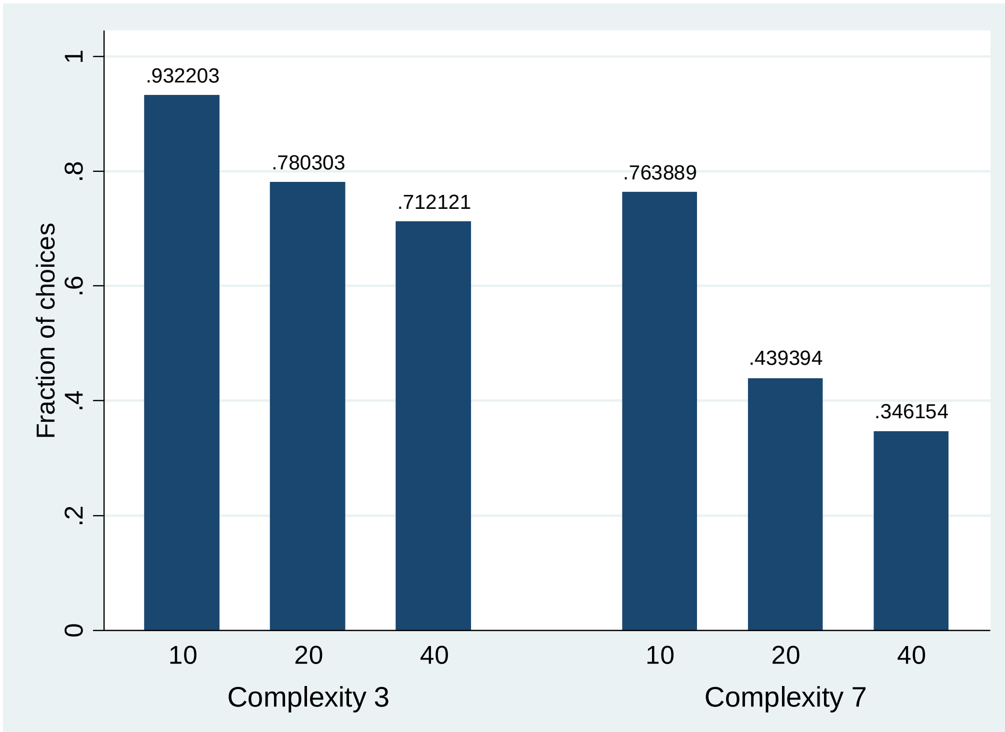



Figure 9: HM indices for the choice process task (experiment 2) and the standard choice task (experiment 1)

Figure 10: Proportion of choices consistent with rationality according to the standard model and the sequential search model (experiment 1)

Panel A: Standard model




Panel B: Sequential search model

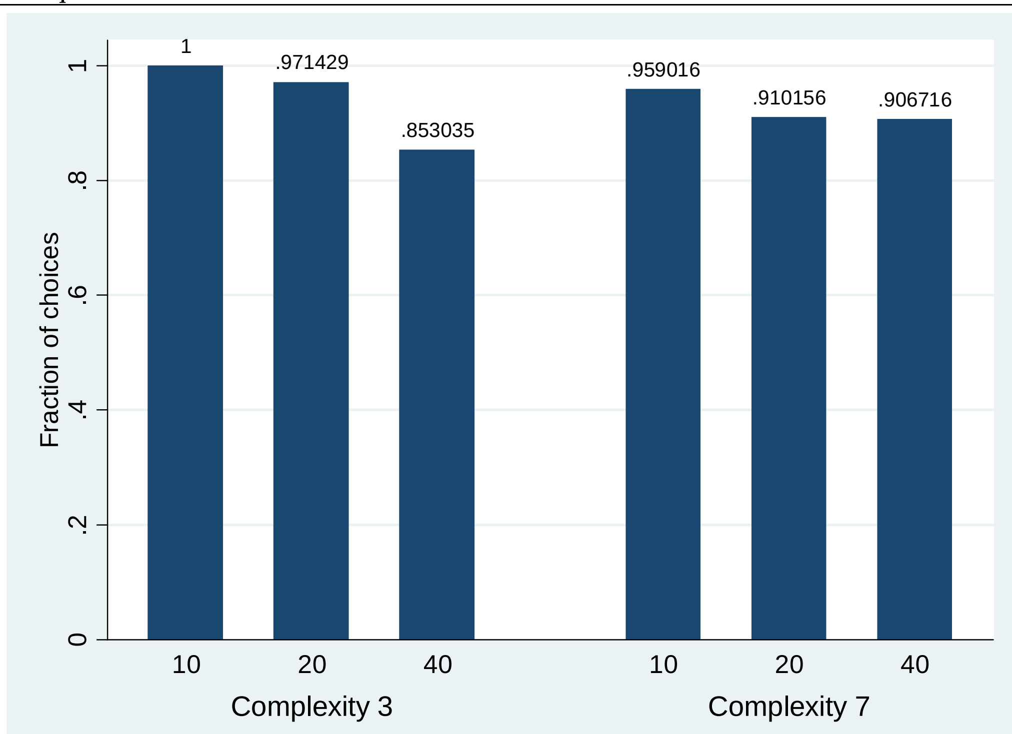
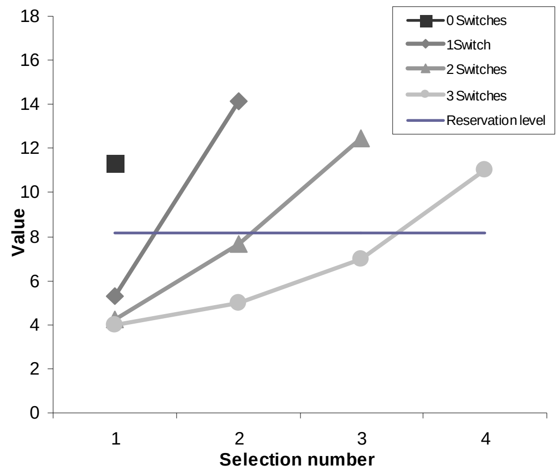
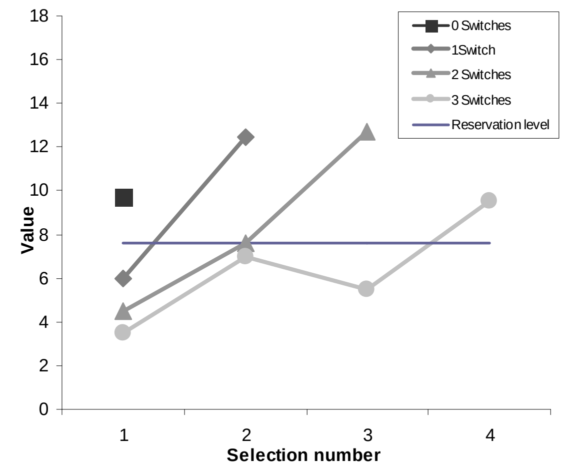


Figure 11: Average value by switch (experiment 1)

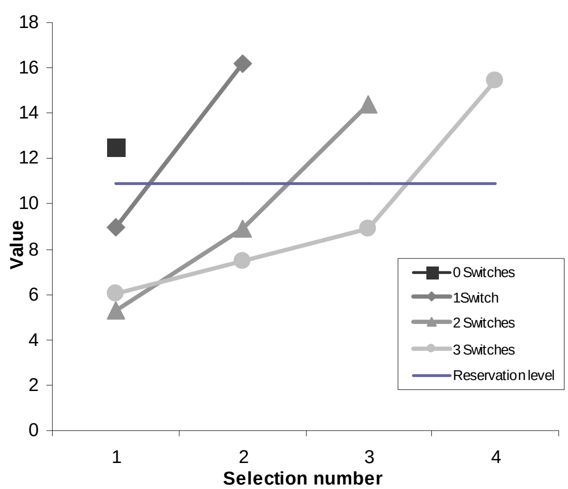
10 Options, Complexity 3



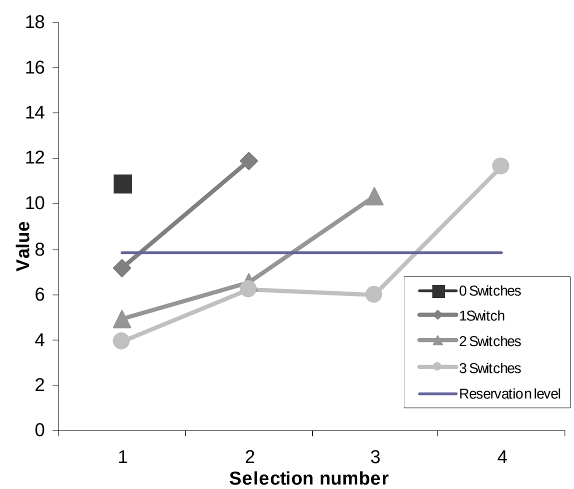

10 Options, Complexity 7



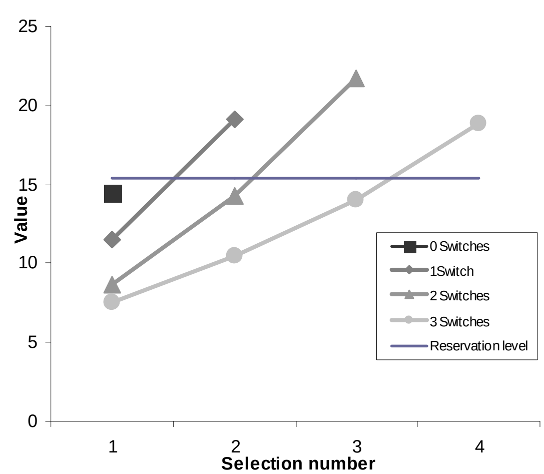

20 Options, Complexity 3




20 Options, Complexity 7




40 Options, Complexity 3




40 Options, Complexity 7

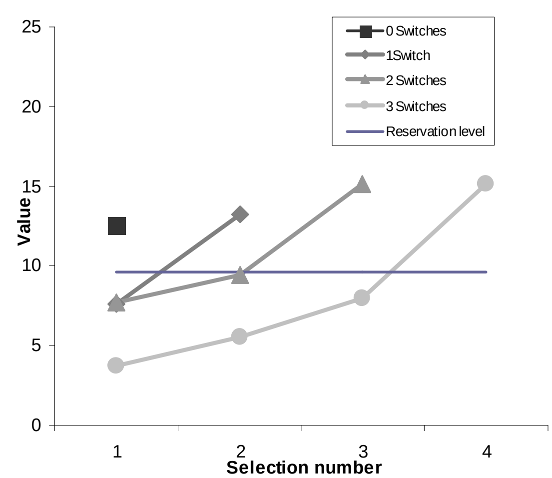
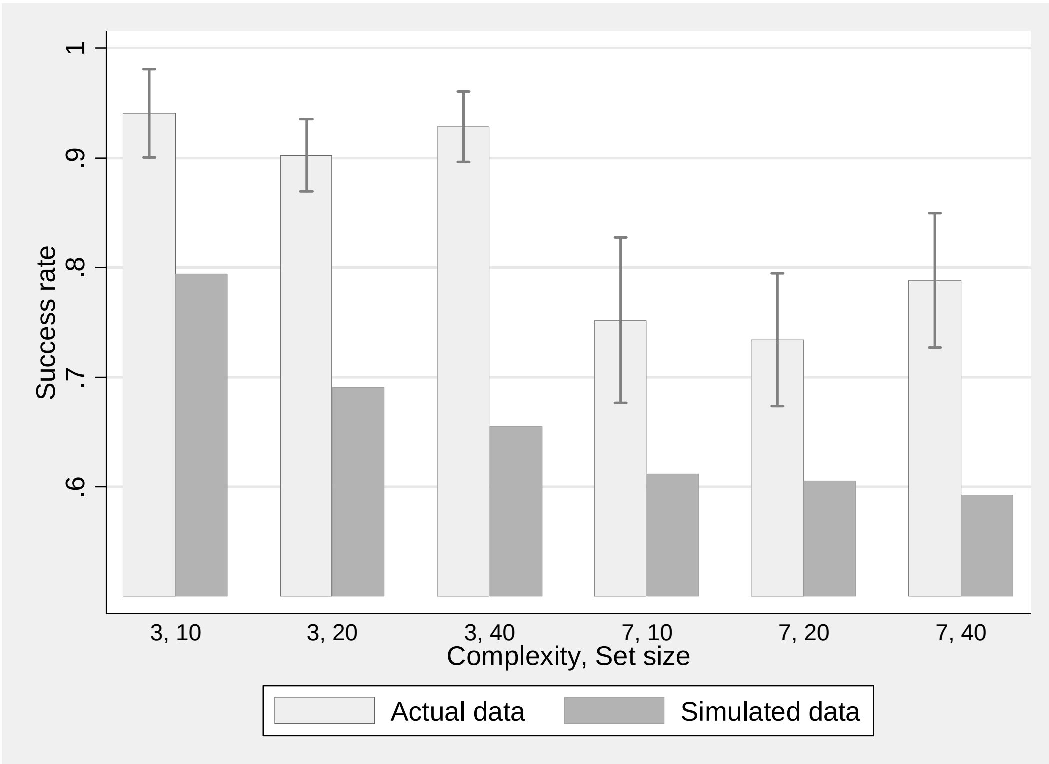


Figure 12: Comparison of the proportion of switches to larger value for actual data and simulated data from calculation error model (experiment 2)




Note: Interval bars represent 95 percent confidence intervals

Table 1: Performance in standard choice task (experiment 1)

| | Complexity | | Total |
|---|---|---|---|
| **Failure rate (percent)** | | | |
| **Set size** | 3 | 7 | |
| 10 | 6.78 | 23.61 | 16.03 |
| 20 | 21.97 | 56.06 | 39.02 |
| 40 | 28.79 | 65.38 | 46.95 |
| **Total** | 21.98 | 52.69 | 37.60 |
| **Absolute loss (dollars)** | | | |
| | **Complexity** | | **Total** |
| **Set size** | 3 | 7 | |
| 10 | 0.41 | 1.69 | 1.11 |
| 20 | 1.10 | 4.00 | 2.55 |
| 40 | 2.30 | 7.12 | 4.69 |
| **Total** | 1.46 | 4.72 | 3.12 |
| **Number of observations** | | | |
| | **Complexity** | | **Total** |
| **Set size** | 3 | 7 | |
| 10 | 59 | 72 | 131 |
| 20 | 132 | 132 | 264 |
| 40 | 132 | 130 | 262 |
| **Total** | 323 | 334 | 657 |

Table 2: Performance in choice process task (experiment 2) vs. standard choice task (experiment 1)

**Failure rate (percent)**

| **Set size** | | **Complexity** 3 | 7 | **Total** |
|---|---|---|---|---|
| 10 | Choice process | 11.38 | 46.53 | 27.23 |
| | *Standard choice* | *6.78* | *23.61* | *16.03* |
| 20 | Choice process | 26.03 | 58.72 | 40.41 |
| | *Standard choice* | *21.97* | *56.06* | *39.02* |
| 40 | Choice process | 37.95 | 80.86 | 57.42 |
| | *Standard choice* | *28.79* | *65.38* | *46.95* |
| **Total** | Choice process | 27.00 | 64.14 | 43.62 |
| | *Standard choice* | *21.98* | *52.69* | *37.60* |

**Absolute loss (dollars)**

| **Set size** | | **Complexity** 3 | 7 | **Total** |
|---|---|---|---|---|
| 10 | Choice process | 0.42 | 3.69 | 1.90 |
| | *Standard choice* | *0.41* | *1.69* | *1.11* |
| 20 | Choice process | 1.62 | 4.51 | 2.89 |
| | *Standard choice* | *1.10* | *4.00* | *2.55* |
| 40 | Choice process | 2.26 | 8.30 | 5.00 |
| | *Standard choice* | *2.30* | *7.12* | *4.69* |
| **Total** | Choice process | 1.58 | 5.73 | 3.44 |
| | *Standard choice* | *1.46* | *4.72* | *3.12* |

**Number of observations**

| **Set size** | | **Complexity** 3 | 7 | **Total** |
|---|---|---|---|---|
| 10 | Choice process | 123 | 101 | 224 |
| 20 | Choice process | 219 | 172 | 391 |
| 40 | Choice process | 195 | 162 | 357 |
| **Total** | Choice process | 537 | 435 | 972 |

Table 3: Estimated reservation levels (experiment 2)

| **Set size** | | **Complexity** 3 | 7 |
|---|---|---|---|
| 10 | Sequential search types | 9.54 (0.20) | 6.36 (0.13) |
| | *Reservation-based search types* | *10.31 (0.23)* | *6.39 (0.13)* |
| 20 | Sequential search types | 11.18 (0.12) | 9.95 (0.10) |
| | *Reservation-based search types* | *11.59 (0.13)* | *10.15 (0.10)* |
| 40 | Sequential search types | 15.54 (0.11) | 10.84 (0.10) |
| | *Reservation-based search types* | *15.86 (0.12)* | *11.07 (0.10)* |

Note: Standard errors in parenthesis

| Table 4: Aggregate HM indices for reservation-based search (experiment 2) | | |
|---|---|---|
| | **Complexity** | |
| **Set size** | 3 | 7 |
| 10 | 0.90 | 0.81 |
| 20 | 0.87 | 0.78 |
| 40 | 0.82 | 0.78 |

Table 5: Proportion of selections with a larger value than any alternative above them on the screen (experiments 2 and 3)

| **Set size** | | **Complexity** 3 | 7 | Mixed |
|---|---|---|---|---|
| 10 | All subjects | 0.56 | 0.40 | |
| | *Top-bottom searchers* | *0.78* | *0.51* | |
| 20 | All subjects | 0.41 | 0.38 | 0.41 |
| | *Top-bottom searchers* | *0.75* | *0.43* | *0.44* |
| 40 | All subjects | 0.52 | 0.27 | |
| | *Top-bottom searchers* | *0.71* | *0.37* | |

| Table 6: Estimated reservation levels (experiment 1 and experiment 2) | | | |
|---|---|---|---|
| | | **Complexity** | |
| **Set size** | | 3 | 7 |
| 10 | Choice process | 9.73 (0.22) | 5.79 (0.13) |
| | *Standard choice* | *10.05 (0.50)* | *8.04 (0.19)* |
| 20 | Choice process | 11.18 (0.12) | 9.82 (0.10) |
| | *Standard choice* | *11.31 (0.15)* | *8.28 (0.12)* |
| 40 | Choice process | 15.54 (0.11) | 10.84 (0.10) |
| | *Standard choice* | *15.98 (0.13)* | *8.68 (0.11)* |

Table 7: Aggregate HM indices for reservation-based search (experiment 1)

| | **Complexity** | |
|---|---|---|
| **Set size** | 3 | 7 |
| 10 | 0.94 | 0.78 |
| 20 | 0.83 | 0.73 |
| 40 | 0.74 | 0.71 |

| Table 8: Estimated standard deviations (in dollars) for the calculation error model (experiment 1 and experiment 2) | | | |
|---|---|---|---|
| | | **Complexity** | |
| **Set size** | | 3 | 7 |
| 10 | Choice process | 1.91 | 5.32 |
| | *Standard choice* | *1.90* | *3.34* |
| 20 | Choice process | 2.85 | 5.23 |
| | *Standard choice* | *2.48* | *4.75* |
| 40 | Choice process | 3.54 | 7.25 |
| | *Standard choice* | *3.57* | *6.50* |

Table 9: Performance of actual choices and simulated choices using the calculation error model (experiment 2)

| Failure rate (percent) | | | |
|---|---|---|---|
| | | Complexity | |
| **Set size** | | 3 | 7 |
| 10 | Actual choices | 11.38 | 46.53 |
| | *Simulated choices* | *8.72* | *31.75* |
| 20 | Actual choices | 26.03 | 58.72 |
| | *Simulated choices* | *19.94* | *38.58* |
| 40 | Actual choices | 37.95 | 80.86 |
| | *Simulated choices* | *26.05* | *44.87* |

| Absolute loss (dollars) | | | |
|---|---|---|---|
| | | Complexity | |
| **Set size** | | 3 | 7 |
| 10 | Actual choices | 0.42 | 3.69 |
| | *Simulated choices* | *0.19* | *1.80* |
| 20 | Actual choices | 1.62 | 4.51 |
| | *Simulated choices* | *0.61* | *1.85* |
| 40 | Actual choices | 2.26 | 8.30 |
| | *Simulated choices* | *0.79* | *2.55* |