American Economic Review 101 (December 2011): 2899–2922
http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.7.2899

Search and Satisficing†

By Andrew Caplin, Mark Dean, and Daniel Martin*

Many everyday decisions are made without full examination of all available options, and, as a result, 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 Herbert 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, D12, D83)

Many everyday decisions are made without full examination of all available options, and, as a result, the best available option may be missed. However, little is known about how such incomplete consideration affects choice behavior. We develop a search-theoretic choice experiment that provides new insights into how information gathering interacts with decision making.

Our central finding is that many decisions can be understood using the satisficing model of Herbert Simon (1955). Simon posited a process of item-by-item search, and the existence of a "satisficing" level of utility, attainment of which would induce the decision maker to curtail further search. Our experiments cover various settings that differ in the number of 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 an environmentally determined level of reservation utility has been realized.

One factor that has held back research on how incomplete search impacts choice is that there are no observable implications of a general model 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, we develop a new experimental

* Caplin: Center for Experimental Social Science and Department of Economics, New York University, 19 West 4th Street, New York, NY 10012 (e-mail: andrew.caplin@nyu.edu); Dean: Department of Economics, Brown University, Providence, RI 02912 (e-mail: 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 (e-mail: daniel.martin@nyu.edu). We thank Douglas Bernheim, Isabelle Brocas, Juan Carrillo, Vince Crawford, Stefano DellaVigna, David Laibson, Alessandro Lizzeri, Ulrike Malmendier, Rosemarie Nagel, Efe Ok, Antonio Rangel, Ariel Rubinstein, Yuval Salant, Andy Schotter, seminar participants at SITE, USC, Michigan, and the CEPR's European Summer Symposium in Economic Theory, and three anonymous referees for insightful suggestions.

† To view additional materials, visit the article page at
http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.7.2899.

[1] The satisficing model itself has testable implications for choice data only if it is assumed that the search order never changes. See Paola Manzini and Marco Mariotti (2007) and Yusufcan Masatlioglu and Daisuke 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 Kfir Eliaz and Ran Spiegler (2011). Ariel Rubinstein and

technique that incentivizes subjects to reveal not only their final choices, but also how their provisional choices change with contemplation time.[2] This "choice process" data provides a test bed for simple models of sequential search (see Donald Campbell 1978; and Caplin and Dean 2011).

A second barrier to research in this area is that there is no general way to define, let alone measure, the quality of decisions.[3] To overcome this conceptual problem, subjects in our experiment select among monetary prizes presented as sequences of addition and subtraction operations.[4] These calculations take time and effort to perform, making the choice problem nontrivial. As a result, we find that subjects regularly fail to find the best option when choosing from sets of such alternatives.

We use choice process data to test the satisficing model. We find that its two identifying features are supported by our data. First, 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. Second, for each of our experimental treatments, 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 characterize the satisficing model. The estimated levels of reservation utility increase with set size and with object complexity.

Choice process data provide insight into search order. We find that some subjects search from the top of the screen to the bottom, while others do not. These search modes impact choice quality: those who search down from the top do poorly if good objects are at the bottom of the screen.

Our method for eliciting choice process data impacts the incentive to search, since there is an increasing chance that later choices will not be actualized. In order to explore the impact of these incentives, we develop a simple model of optimal search with psychic costs that is rich enough to cover this case in addition to standard choice data. We find that, while a fixed reservation level is optimal in the standard case, a declining reservation level is optimal for the choice process environment. Moreover, the reservation level in a choice process environment is always below the fixed optimal level in the equivalent standard choice environment.

We test the predictions of the optimizing model by comparing behavior in the choice process experiment with that in a standard choice environment. We exploit the fact that subjects were able to, and indeed chose to, change options prior to finalizing decisions even in our standard choice experiments, creating a sequence of choices that we can interpret as choice process data. We find that standard choice data is indeed well described by the fixed reservation model. However, we find no evidence of a declining reservation level in the choice process environment. This

Yuval Salant (2006) present a model of choice from lists, in which a decision maker searches through the available options in a particular order. Efe Ok (2002) considers the case of a decision maker who is unable to compare all the available alternatives in the choice set. These models make specific assumptions about the nature of search to gain empirical traction.

[2] Compared to other novel data used to understand information search, such as those based on eye tracking or Mouselab (John Payne, James Bettman, and Eric Johnson 1993; Xavier Gabaix et al. 2006, Elena Reutskaja et al. 2011), choice process data is more closely tied to standard choice data and revealed preference methodology.

[3] See B. Douglas Bernheim and Antonio Rangel (2008); Faruk Gul and Wolfgang Pesendorfer (2008); and Botond Kőszegi and Matthew Rabin (2008) for methodological viewpoints on the classification of particular decisions as "poor" or "mistaken."

[4] Caplin and Dean (2011) characterize theoretical connections between choice process data, sequential search, and reservation stopping rules with arbitrary objects of choice.

suggests that our subjects may be satisficing for the reasons that Simon (1955) originally proposed, as a rule of thumb that performs adequately across a broad range of environments, rather than finely honing their search strategy to each choice environment they face. We find some evidence that reservation levels in choice process settings are below those in equivalent standard choice settings.

While our findings are in line with simple theories of sequential search, we consider (and reject) an alternative model 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.

The article is arranged into six sections. In Section I we introduce our experimental protocols. In Section II we describe the pattern of choice mistakes exhibited by our subjects. In Section III we test the satisficing model and show how 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. Section VI contains our estimates of the model based entirely on calculation errors rather than sequential search.

I. Experimental Design

We conducted experiments of four types. 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 imposes a time limit on subjects in an otherwise standard choice task, allowing us to understand the source of differences in behavior between experiments 1 and 2. All experiments were conducted at the Center for Experimental Social Science laboratory at New York University, using subjects recruited from the undergraduate population.

A. *Experiment 1: Standard Choice*

Our goal in this article 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 (three or seven addition and subtraction operations for each object) and the total number of available alternatives (10, 20, or 40). Figure 1 shows a ten option choice set with objects of complexity 3.[5]

[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.

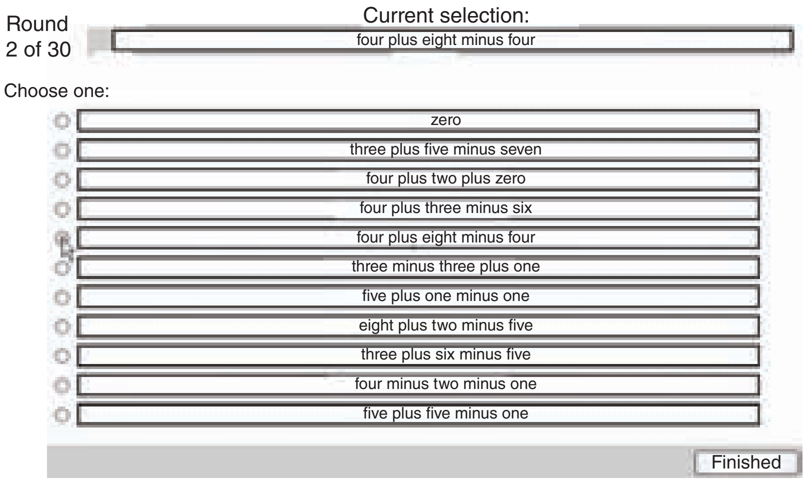

Figure 1. A Typical Choice Round

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 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 two practice rounds and between 27 and 36 regular rounds, drawn from all six treatments. At the end of the session, two regular rounds were drawn at random, and the subject received the value of the final 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.

B. *Experiment 2: Choice Process*

Choice process data tracks not only final choice, but also how subjects' provisional choices evolve with contemplation time. It is closely related to standard choice data, in that all observations represent choices, albeit indexed by time. We see these

[6] Changes that were made over the predecision 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.

data 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. 2011). While choice process data misses out on such potentially relevant clues to search behavior as eye movements, it captures the moment at which search changes a subject's assessment of the best option thus far encountered.

Our experimental design for eliciting choice process data 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. Only at the end of each round does the subject find out the time that was actualized, and what his or her selection had been at that time. This incentivizes subjects always to select the option that they perceive as best. We therefore treat their sequence of selections as recording their preferred option at each moment in time.[8]

The instructions that were given to subjects in the choice process experiment are available in the online supplemental material. They were informed that the actualized time would be drawn from a beta distribution with parameters $\alpha = 2$ and $\beta = 5$, truncated at 120 seconds.[9] The interface for selecting and switching among objects was identical to that of experiment 1. 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. Typically, a subject took part in a session consisting of two 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 three or seven operations. Moreover, exactly the same choice object values were used in the choice process and standard choice experiments. For the six treatments of experiment 2, we collected data on 1,066 choice sets from 76 subjects.

C. *Experiment 3: Varying Complexity*

Experiment 3 was designed to explore how screen position and object complexity impacts search order. All choice sets were of size 20, and the objects in each set ranged in complexity from one to nine operations. Subjects were instructed that object complexity, screen position, and object value were independent of one another. Incentives were as in experiment 2, the choice process experiment. Experiment 3 was run on 21 subjects for a total of 206 observed choice sets.

D. *Experiment 4: Time Constraint*

While the choice process experiments included time limits, the standard choice experiment did not. In order to explore whether this time limit was responsible

[8] In support of this interpretation, 58 of 76 subjects in a postexperiment 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.

for differences in behavior between the two settings, we reran 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, a total of 29 subjects chose from 407 observed choice sets.

II. Choice Performance

A. *Standard Choice Task*

Table 1 reports the results of experiment 1, the standard choice experiment. 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.

Averaging across all treatments, subjects fail to select the best option almost 38 percent of the time and leave $3.12, or 17 percent of the maximum amount, on the table in each round.[10] Both of these performance measures worsen with the size and the complexity of the choice set, reaching a failure rate of 65 percent and an average loss of $7.12 in the size 40, complexity 7 treatment. Regression analysis shows that the difference in losses between treatments is significant.[11]

B. *Choice Process Task*

Given that our analysis of the search-based determinants of choice quality is based primarily on the choice process data of experiment 2, it is important to explore how the level and pattern of final choices compares across experiments 1 and 2. To ensure that subjects in experiment 2 had indeed finalized their choices, we retain only rounds in which they explicitly press the submit button before the allotted 120 seconds. This removes 94 rounds, or 8.8 percent of our total observations. Table 1 compares failure rates and average absolute losses by treatment for choice process and standard choice tasks.

In both the choice process experiment and the standard choice experiment, subjects fail to find the best option more frequently and lose more money in larger and more complicated choice sets. However, in almost 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 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.

[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.

[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 5 percent level for size 40 compared to size 10 choice sets, and for the interaction of size 40 and complexity 7 compared to size 10 and complexity 3 choice sets.

TABLE 1—PERFORMANCE IN CHOICE PROCESS TASK (*Experiment 2*) VERSUS STANDARD CHOICE TASK (*Experiment 1*)

Failure rate (percent)		Complexity	
Set size		3	7
10	Choice process	11.38	46.53
	Standard choice	6.78	23.61
20	Choice process	26.03	58.72
	Standard choice	21.97	56.06
40	Choice process	37.95	80.86
	Standard choice	28.79	65.38

Absolute loss (dollars)		Complexity	
Set size		3	7
10	Choice process	0.42	3.69
	Standard choice	0.41	1.69
20	Choice process	1.62	4.51
	Standard choice	1.10	4.00
40	Choice process	2.26	8.30
	Standard choice	2.30	7.12

Number of observations		Complexity	
Set size		3	7
10	Choice process	123	101
	Standard choice	59	72
20	Choice process	219	172
	Standard choice	132	132
40	Choice process	195	162
	Standard choice	132	130

III. Sequential Search and Satisficing

We use the choice process data from experiment 2 to test whether a simple sequential search model with a reservation level of utility can explain the failure of people to select the best available option. We test both whether subjects appear to understand the value of each searched object in full before moving on to the next (as in the classic search models of George Stigler 1961 and John McCall 1970), and whether they appear to search until an object is found that is above a fixed reservation utility level. The power of our tests depends on observing subjects switching from one alternative to another. Fortunately, in 67 percent of rounds we observe at least one occasion on which the subject switches between options after the initial change away from $0. The mean number of such switches is 1.4.

A. *Sequential Search*

Caplin and Dean (2011) provide a method of identifying whether or not choice process data is consistent with sequential (but possibly incomplete) search. Assuming that utility is monotonically increasing in money, a necessary and sufficient condition for choice process data to be in line with sequential search is 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.[12]

In order to test whether our subjects are close to satisfying Condition 1, we use a measure of consistency proposed by Martijn Houtman and J. A. H. Maks (1985). The Houtman-Maks (HM) index is based on calculating the largest fraction of observations that are consistent with Condition 1.[13]

Figure 2 shows the distribution of HM indices for all 76 subjects. Over 40 percent of our subjects have an HM index above 0.95, while almost 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 2 also shows the distribution of HM indices for 76,000 simulated subjects with the same number of switches as our subjects but who choose at random—a measure of the power of our test (see Stephen Bronars 1987). Clearly, the two distributions are very different, as confirmed by a Kolmogorov-Smirnov 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. To identify such sequential searchers, 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. For the remainder of the paper we focus on the 68 out of 76 subjects who have an HM index above the 95th percentile of their randomly generated distribution.[14]

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 3 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. The level of mistakes as measured by the standard definition of revealed preference is far higher than by the sequential search measure. Note also that while there is strong

[12] Note that the choice process methodology identifies only 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.

[13] Specifically, we identify the smallest number of observations that need to be removed for the resulting data to be consistent with Condition 1. The HM index is the number of remaining observations, normalized by dividing through by the total number of observations.

[14] 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, a reduction only of 1 would be needed.

The correlation between these two measures is very high in our sample: the Spearman's rank correlation is 0.96. 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.

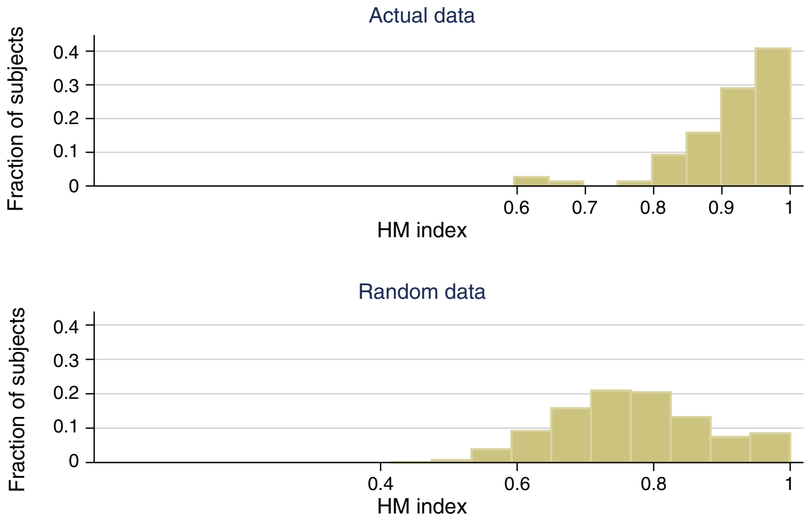

FIGURE 2. DISTRIBUTION OF HM INDICES FOR ACTUAL AND RANDOM DATA (*Experiment 2*)

evidence of increasing mistakes 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 on mistakes, and only a small effect from complexity.

B. *Satisficing and Reservation Utility*

The essential advantage that choice process data provide in testing the satisficing model is that they allow us to observe occasions in 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 4. 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.[15] 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 two switches were made, and the final segmented line for rounds in which three switches were made.

Figure 4 is strongly suggestive of satisficing behavior. First, as we would expect from the preceding 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.

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

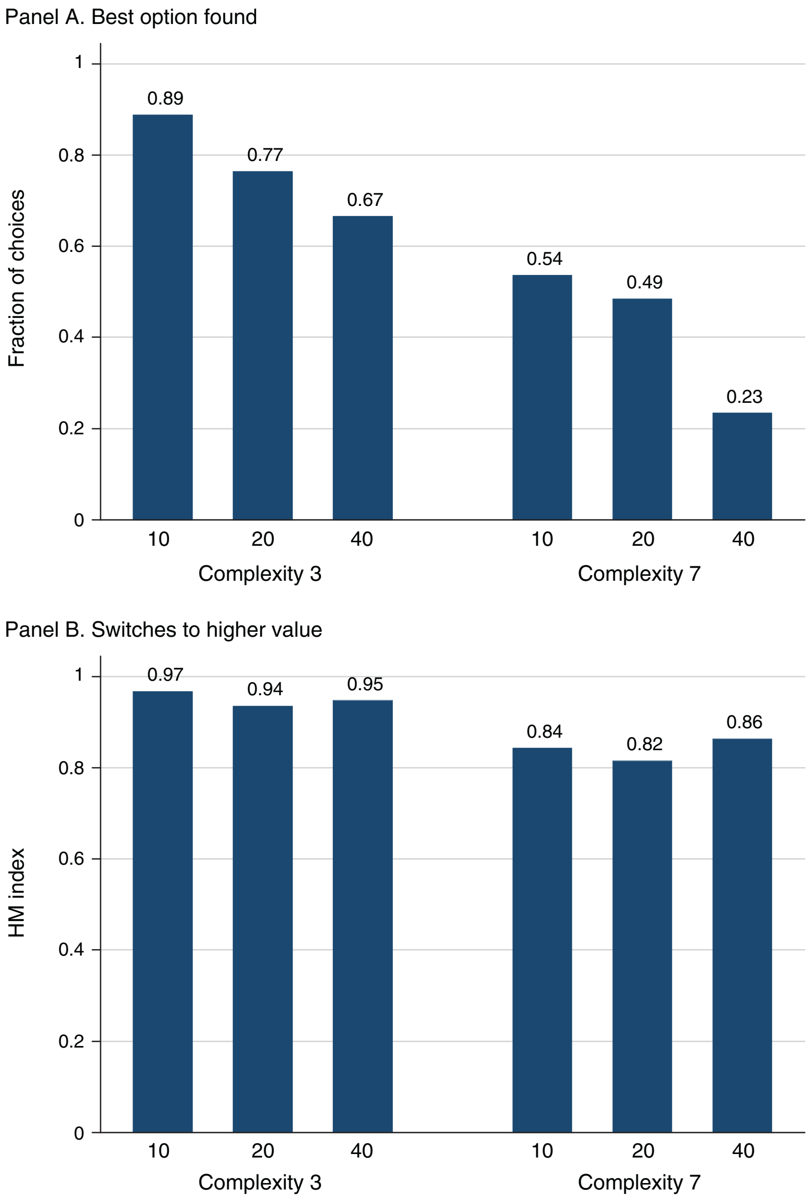

Figure 3. Proportion of Final Choices Where the Best Option was Found and Largest Proportion of Switches to Higher Value (*Experiment 2*) with HM Index above 95th Percentile

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 utilities for each treatment, we assume that all individuals in a given choice environment have the same reservation value $\overline{v}$ and experience variability ε in this value each time they decide whether or not to continue search. Further, we assume this stochasticity

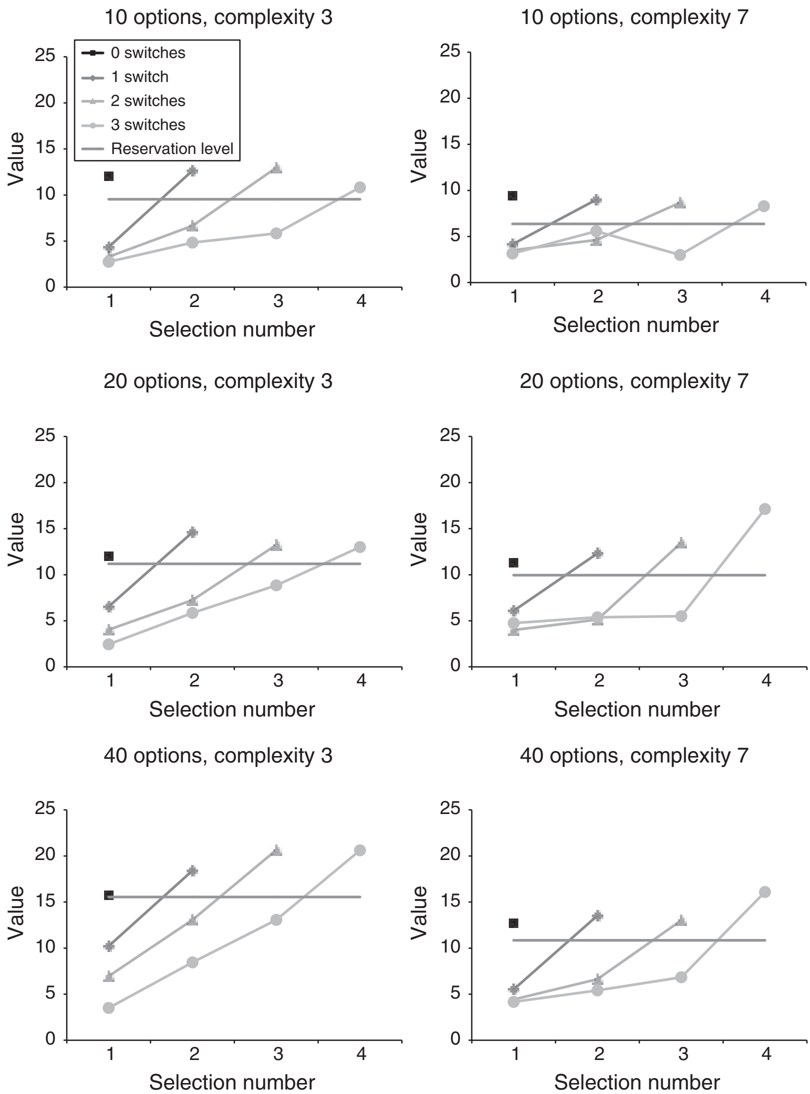

FIGURE 4. AVERAGE VALUE BY SELECTION (*Experiment 2*)

enters additively and is drawn independently and identically from the standard normal distribution.[16] Letting v be the value of the item that has just been evaluated, the decision maker (DM) stops search if and only if $v \geq \overline{v} + \varepsilon$, where $\varepsilon \sim N(0,1)$.

[16] 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 have only 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 Pavlo Blavatsky and Ganna Pogrebna (2010). Models can broadly be categorized into two types. The first are "tremble" models of the type used in David Harless and Colin 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 of this type 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 Heteroskedastic Random Errors, which assumes that the reservation value is observed with an additive, normally distributed error term. In our

To cast this as a binary choice model, let k be a decision node, v_k be the value of the object uncovered and ε_k the error. Note that the probability of stopping search is $\Phi(v_k - \overline{v})$, where Φ is the cumulative density function of the standard normal distribution, so we can estimate $\overline{v}$ using maximum likelihood.

To employ this procedure using our data, 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, he might have continued to search but run out of time before he found a better object. Second, he might have continued to search but already had selected the best option. Third, he might have stopped searching. We therefore consider a subject to have stopped searching at a decision node only if he made no further selections, pressed the submit button, and the object he had selected was not the highest value object in the choice set.

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 2 shows the estimated reservation levels for each treatment, with standard errors in parentheses.

Table 2 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 ($p < 0.001$). Second, reservation levels increase monotonically with set size (significantly different for all pairwise comparisons of set sizes for both complexity levels with $p < 0.001$).

One question that this estimation strategy does not answer is how well the reservation utility model explains our experimental data. In order to shed light on this question, we calculate the equivalent of the HM index for this model with the estimated reservation levels of Table 2. 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).

The results, aggregated across all subjects, are shown in Table 3. The estimated 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 HM indices above 75 percent, some have extremely low scores and are clearly poorly described by a reservation utility model with the given estimated reservation levels.

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.

In 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.

TABLE 2—ESTIMATED RESERVATION LEVELS (*Experiment 2*)

		Complexity			
Set size		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 parentheses.

TABLE 3—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

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 six subjects). These results are in Table 2 under the rows for "Reservation-based search types." The estimated reservation levels are similar to those for the whole sample.

C. *Reservation Utility or Reservation Time?*

A natural question is whether our data are 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 assess this possibility, we redraw in Figure 5 the graphs of Figure 4 but show the average time of each switch, rather than the average value on the vertical axis.

Figure 5 provides no support for the reservation time stopping rule. Unlike in Figure 4, 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 described in Section IIIB). 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 provide insight into the order of search, and that this information can help predict when subjects will do badly in particular choice sets.

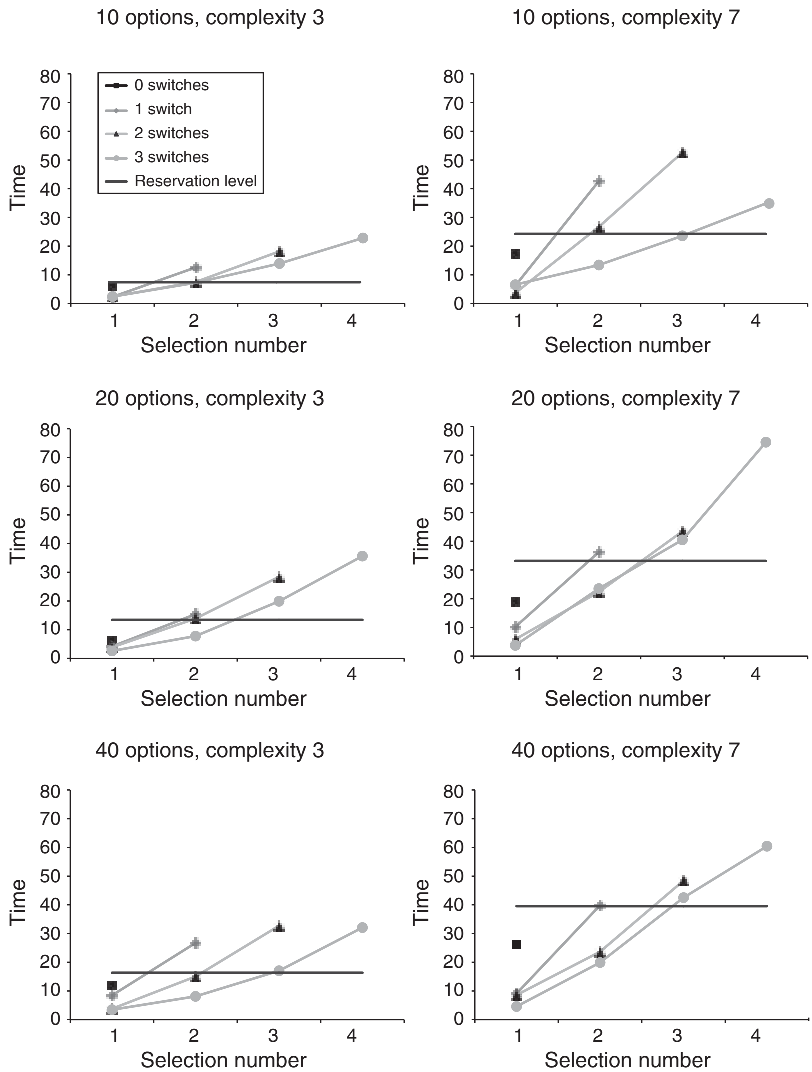

Figure 5. Average Time by Selection (*Experiment 2*)

The first finding is that subjects in experiment 2 tend to search from the top to the bottom of the screen. When we regress the order in which an object is selected on its position on screen, we find that the average screen position is significantly higher (i.e., further down the screen) for later selections.[17] This relationship is more pronounced for choice sets with simple, rather than complex, objects.[18]

[17] 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).

[18] For complexity 3 choice sets, regressing selection number on the screen position of the selection gives a coefficient of 0.036, significant at the 1 percent level, while for complexity 7 sets the coefficient is 0.018, not significant at the 10 percent level.

To assess whether subjects search from top to bottom (TB), we 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. A subject is categorized as being a TB searcher if this HM index for his search order is in the 95th percentile of a benchmark distribution constructed using random search orders. With this criterion, 53 percent of subjects in experiment 2 are well described by TB search.

While the search order HM index is determined independently of a subject's performance, we find that TB searchers do worse when the best object appears further down the screen. When we regress whether a subject found the best option onto the screen location of the best option, the coefficient is negative (-0.03) and significant at the 1 percent level for TB searchers, but is smaller in magnitude (-0.01) and insignificant at the 10 percent level for those not classified as TB searchers.

For subjects that are strict TB searchers, sequential search has particularly strong implications. Thus far, we have assumed that we know an object has been searched only if it has been chosen at some point. However, if a strict TB searcher at some point selects the object at a certain screen position, then he must have searched all objects in screen positions above it. For example, if the object in position 10 is selected, then the objects in positions 1 to 9 must have been searched through as well. In this case, the test for sequential 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.

In the low complexity choice environment, we find that subjects classified as TB searchers behave in line with this strict form of sequential search in about 92 percent of cases. They also do significantly better in this test than subjects that we do not classify as TB.[19] However, even those we categorize as TB searchers violate this condition in about 42 percent of cases for 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.

In addition to TB search, experiment 3 enables us to explore whether or not object complexity impacts search order. We find not only that subjects in general search the screen from top to bottom, but also from simple to complex objects.[20] We define a subject in this experiment to be a "Simple-Complex" (SC) searcher if they have a corresponding HM index above the 95th percentile of random search orders. Eight subjects are categorized as both TB and SC searchers, six as just TB searchers, three as just SC searchers. Only three subjects could be categorized as neither.

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 incentives have on decisions, we characterize optimal stopping strategies in a sequential search

[19] 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.

[20] Regressing selection number on the screen position and complexity of the object selected gives coefficients of 0.037 and 0.136, respectively, both significant at the 1 percent level (allowing for clustering at the subject level).

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.[21] 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.

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. These tests indicate that if anything, data from the standard choice environment are more in line with sequential search than choice process data. Indeed, there are even fewer violations of Condition 1 in experiment 1 (8 percent of rounds with a violation) than there were in experiment 2 (10 percent of rounds with a violation). Once again there was little effect of either complexity or choice set size on conformity with Condition 1.

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 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). 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 is assumed ex ante to have a utility level that is independently drawn from some distribution $F(z)$, as in our experiment which is known by the DM. There is no discounting.

To break the otherwise rigid connection between time and the number of objects searched, 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 search stops, the agent must choose one of the identified objects.[22]

To match the choice process experimental design, we allow for the possibility that search after time $t \geq 1$ will have no impact on what the DM receives. We let the nonincreasing function $J(t)$ identify the probability that the search from time t on will actually impact choice. In the standard choice environment, $J(t)$ is constant

[21] 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.

[22] This method of modeling makes the process of uncovering an option equivalent to the process of "locating" it as feasible. The strategy is more intricate if we allow unexplored options to be selected.

at 1, while in the choice process environment $J(0) = 1$, $J(t) - J(t+1) > 0$ for $1 \leq t \leq T - 1$ and $J(T+1) = 0$ (where $T = 120$ seconds).

Our characterization of the optimal search strategy is straightforward, and the proof is available in the online Appendix.

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

$$\int_{u^R(t)}^{\infty} [z - u^R(t)]\, dF(z) = \frac{\kappa}{qJ(t)}. \tag{1}$$

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

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

$$\int_{\overline{u}^R}^{\infty} \left(z - \overline{u}^R\right) dF(z) = \frac{\kappa}{q}. \tag{2}$$

This reservation level is decreasing in the cost of search κ 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 1 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 6 recreates the analysis of Figure 4 and suggests that a reservation stopping rule broadly describes the aggregate data. Table 4 shows that the estimated reservation levels for the standard choice data exhibit the same comparative statics as do those for the choice process data.[23] Table 5 shows that the estimated HM indices for these reservation levels in the standard choice data are roughly similar

[23] For the analysis of Table 4 we drop subjects who never switch in any round and who are not classified as using a reservation strategy.

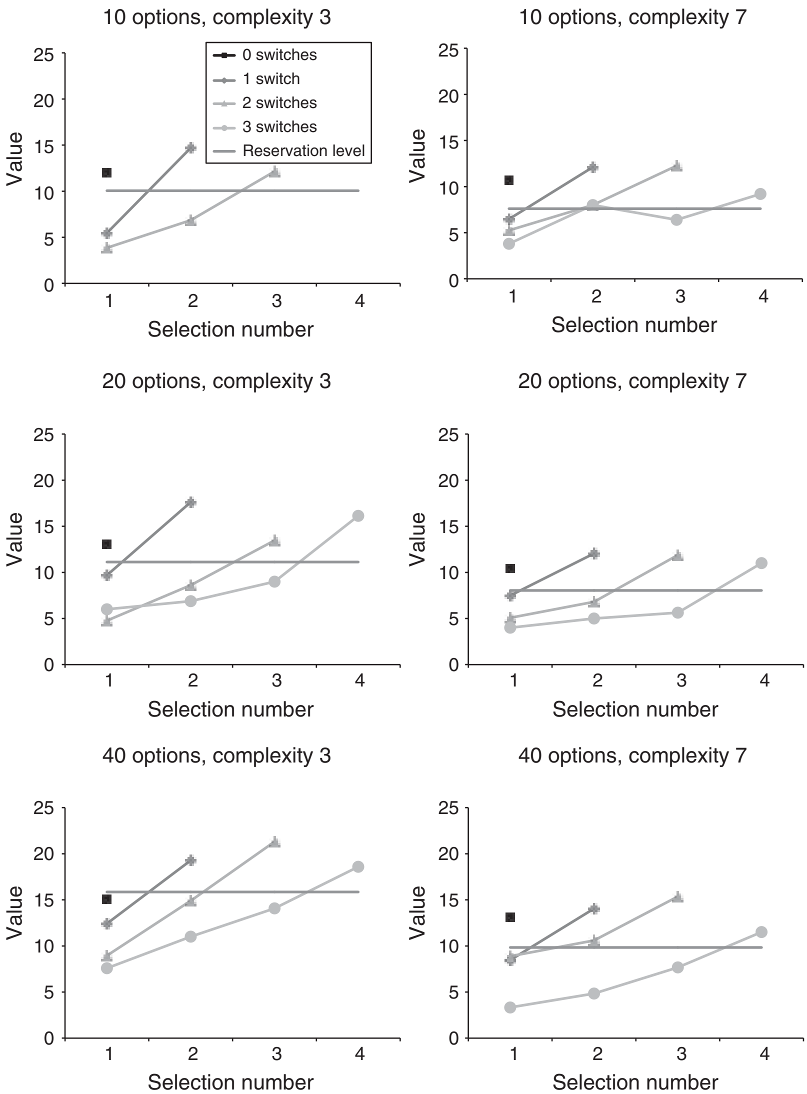

Figure 6. Average Value by Selection (*Experiment 1*)

for lower complexity and smaller for higher complexity.[24] This suggests that there is little qualitative distinction between behavior in the standard choice and choice process environments.

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. Table 4 suggests that this is broadly so for the sample pursuing reservation strategies (HM index above 0.5). As Table 4 shows, the reservation utility is lower in experiment

[24] For none of the treatments is the difference between experiments 1 and 2 in terms of compliance with the reservation utility model significant at the 5 percent level.

TABLE 4—ESTIMATED RESERVATION LEVELS
(*Experiment 1 and Experiment 2*: *reservation-based search types with a selection in every round*.)

		Complexity			
Set size		3		7	
10	Choice process	10.17	(0.22)	6.34	(0.13)
	Standard choice	10.05	(0.50)	8.41	(0.20)
20	Choice process	11.22	(0.11)	8.92	(0.09)
	Standard choice	11.73	(0.16)	8.39	(0.12)
40	Choice process	15.15	(0.10)	10.07	(0.09)
	Standard choice	16.38	(0.13)	10.39	(0.12)

Note: Standard errors in parentheses.

TABLE 5—AGGREGATE HM INDICES FOR RESERVATION-BASED SEARCH (*Experiment 1*)

	Complexity	
Set size	3	7
10	0.94	0.74
20	0.83	0.74
40	0.77	0.73

1 than in experiment 2 in four of six treatments. This difference is significant in only two cases, and in both cases experiment 1 has the lower reservation level. Lower reservation levels could also explain why subjects in the choice process experiment finished 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 two 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 33.7 percent for the standard choice experiment, 39.5 percent in the standard choice with time limit experiment, and 43.6 percent in the choice process experiment.[25] 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 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

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

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 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 y is greater than the value of x. It turns out that the declining reservation model does not offer a better description of choice process data. While the declining reservation model by definition has fewer violations in absolute terms, the *proportion* of observations that violate revealed preference is higher—24 percent for the fixed reservation model versus 32 percent 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*

Assuming that search costs are higher for more complex objects, our model of optimal search implies that reservation utility should be lower in the higher complexity environment. It implies also that optimal reservation levels are independent of the size of the choice set. The comparative statics properties of our experimentally estimated stopping rules do not align perfectly with those of the optimal stopping rule. While 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.

One possible reason for this discrepancy is that subjects may be 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 (Sheena Iyengar and Mark Lepper 2000; Maria Seuanez-Salgado 2006).[26] It is also possible that satisficing is followed as a rule of thumb, as Simon (1955) suggested. In the more everyday context with unknown object values, subjects may search more in larger sets in order to refine their understanding of what is available. They may then import this behavior into the experimental lab, despite being fully informed about the distribution of object values.

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

[26] One factor that potentially links these two findings is the concept of regret. Marcel Zeelenberg and Rik 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.

Table 6—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

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 on the error term 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.[27] We allow for scale factors to differ between treatments.

Table 6 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 size 40, complexity 3 treatment, the range between minus one and plus one standard deviation is around \$7, while the mean value of our choice objects is just \$4.

Despite these large standard deviations, the calculation error model significantly underpredicts both the frequency and the magnitude of our subjects' losses, as shown in Table 7.[28] 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.

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 7. 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.[29] Note that the

[27] 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.

[28] Alternatively, we could have estimated the scale factor to best match the failure rate and average loss found in the data, but this would ignore the actual choices that subjects made, which may contain other unpredicted patterns.

[29] 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

Table 7—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	8.35	32.47
20	Actual choices	26.03	58.72
	Simulated choices	20.13	37.81
40	Actual choices	37.95	80.86
	Simulated choices	25.26	44.39

Absolute loss (dollars)			
		Complexity	
Set size		3	7
10	Actual choices	0.42	3.69
	Simulated choices	0.19	1.86
20	Actual choices	1.62	4.51
	Simulated choices	0.62	1.78
40	Actual choices	2.26	8.30
	Simulated choices	0.75	2.48

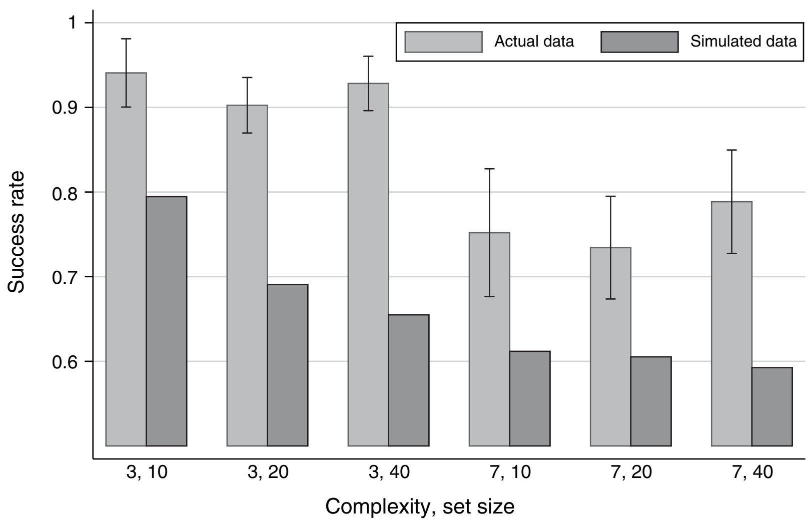

Figure 7. Comparison of the Proportion of Switches to Larger Value for Actual Data and Simulated Data from Calculation Error Model (*Experiment 2*)

predicted success rates for the calculation error model lie below the lower bounds of the 95 percent confidence interval bars for all treatments.

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.

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 various 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, suggesting the value of the search theoretic lens 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, 155–92. New York: Oxford University Press.

Blavatskyy, Pavlo R., and Ganna Pogrebna. 2010. "Models of Stochastic Choice and Decision Theories: Why Both Are Important for Analyzing Decisions." *Journal of Applied Econometrics*, 25(6): 963–86.

Bronars, Stephen G. 1987. "The Power of Nonparametric Tests of Preference Maximization." *Econometrica*, 55(3): 693–98.

Campbell, Donald E. 1978. "Realization of Choice Functions." *Econometrica*, 46(1): 171–80.

Caplin, Andrew, and Mark Dean. 2011. "Search, Choice, and Revealed Preference." *Theoretical Economics*, 6(1): 19–48.

Caplin, Andrew, Mark Dean, and Daniel Martin. 2011. "Search and Satisficing: Dataset." *American Economic Review.* http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.7.2899.

Eliaz, Kfir, and Ran Spiegler. 2011. "Consideration Sets and Competitive Marketing." *Review of Economic Studies*, 78(1): 235–62.

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): 1043–68.

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, 3–42. New York: Oxford University Press.

Harless, David W., and Colin F. Camerer. 1994. "The Predictive Utility of Generalized Expected Utility Theories." *Econometrica*, 62(6): 1251–89.

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(6): 995–1006.

Kőszegi, Botond, and Matthew Rabin. 2008. "Revealed Mistakes and Revealed Preferences." In *The Foundations of Positive and Normative Economics*, ed. Andrew Caplin and Andrew Schotter, 193–209. New York: Oxford University Press.

Manzini, Paola, and Marco Mariotti. 2007. "Sequentially Rationalizable Choice." *American Economic Review*, 97(5): 1824–39.

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

McCall, John J. 1970. "Economics of Information and Job Search." *Quarterly Journal of Economics*, 84(1): 113–26.

Ok, Efe A. 2002. "Utility Representation of an Incomplete Preference Relation." *Journal of Economic Theory*, 104(2): 429–49.

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

Reutskaja, Elena, Rosemarie Nagel, Colin F. Camerer, and Antonio Rangel. 2011. "Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study." *American Economic Review*, 101(2): 900–926.

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." Fondazione Eni Enrico Mattei Working Paper 37.

Simon, H. A. 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics*, 69: 99–118.

Stigler, G. J. 1961. "The Economics of Information." *Journal of Political Economy*, 69: 213–25.

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