Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model

By Xavier Gabaix, David Laibson, Guillermo Moloche, and Stephen Weinberg*

The directed cognition model assumes that agents use partially myopic option-value calculations to select their next cognitive operation. The current paper tests this model by studying information acquisition in two experiments. In the first experiment, information acquisition has an explicit financial cost. In the second experiment, information acquisition is costly because time is scarce. The directed cognition model successfully predicts aggregate information acquisition patterns in these experiments. When the directed cognition model and the fully rational model make demonstrably different predictions, the directed cognition model better matches the laboratory evidence. (JEL D83)

Decision-making requires cognitive operations, including information acquisition and information processing. Economists assume that agents act as if they were choosing these (costly) operations optimally. But models of optimal cognition pose significant conceptual challenges. Such models are generally intractable. Only very simple settings admit *analytic* solutions. Moreover, even *computational* (i.e., numerical) tractability fails as the complexity of the problem increases. In addition, models of

optimal cognition suffer from the infinite regress problem: if cognition is costly, then optimizing cognition is also costly, leading one to optimize the optimization, and so on ad infinitum (John Conlisk, 1996; Barton L. Lipman, 1991; Herbert Simon, 1955).

Instead of trying to model optimal cognition, we study a partially myopic and tractable alternative. The directed cognition model uses approximate option-value calculations to direct cognition to mental activities with high shadow values (Gabaix and Laibson, 2005). The current paper applies the directed cognition model to a problem of information acquisition, or search. In this context, the model assumes the following iterative search structure: *At each decision point, agents act **as if** their next set of search operations were their last opportunity for search.*

Such decision-making, although partially myopic, nevertheless helps agents focus on information that is likely to be useful and ignore information that is likely to be redundant. The directed cognition model is also tractable. The model *can* be computationally solved in highly complex settings. The model does not suffer from the curse of dimensionality or the infinite regress problem.

The current paper experimentally evaluates the directed cognition model. We find that laboratory behavior matches the predictions of the directed cognition algorithm. We begin with a relatively simple choice problem for which it is possible to compute optimal choices, and show that the directed cognition model outperforms rationality. We then turn to a complex (and

more realistic) choice problem for which it is not possible to compute optimal choices, and demonstrate that the directed cognition model predicts aggregate subject behavior.

Section I describes the setup and the results for the first ("simple") experiment. Section II describes the setup for the second ("complex") experiment. Section III describes the implementation of the directed cognition algorithm in the complex experiment. Section IV summarizes the results of the second experiment and compares those results to the predictions of the model. Section V concludes.

I. First Experiment: Choice among Three Simple Goods

Consider the following decision problem, which is a special case of the class of problems investigated by John C. Gittins (1979) and Martin Weitzman (1979). An agent chooses among three uncorrelated projects. The three projects have respective (stochastic) payoffs X_1, X_2, and X_3. If project i is a "winner" then $X_i = V_i$; otherwise, $X_i = 0$. Project i is a winner with probability p_i.

The agent can sequentially investigate projects that are not known winners, thereby revealing their state. Such information acquisition costs c_i per project. The agent may stop acquiring information at any time and choose one project among the *known* winners, which we will refer to as "taking" a project.

For example, in one of our experimental games we adopt the parameters reported in Table 1 (with information acquisition cost $c_i = 1$ for all projects).

Before we derive the optimal strategy, imagine how a typical subject would start to play this game. Would the subject begin by *taking* project 1 (the only known winner at the moment), by paying a dollar to *investigate* project 2, or by paying a dollar to *investigate* project 3?

The optimal sequence of information acquisition can be derived using a Gittins-Weitzman (GW) index (Gittins, 1979; and Weitzman, 1979). Assuming risk neutrality,[1] the GW index

TABLE 1—INVESTMENT GAME A

	Payoff in winning state	Probability of winning state
Project 1	$V_1 = \$1$	$p_1 = 1.00$
Project 2	$V_2 = \$21$	$p_2 = 0.09$
Project 3	$V_3 = \$10$	$p_3 = 0.76$

Notes: A subject chooses among three uncorrelated projects. If project i is a "winner" (which happens with probability p_i), its payoff is V_i; otherwise, its payoff is 0. The subject can sequentially investigate projects that are not known winners, thereby revealing their state. Such information acquisition costs \$1 per project. The subject may stop acquiring information at any time and choose one project among the *known* winners. Would a subject begin by taking project 1 (the only known winner at the moment), by paying \$1 to investigate project 2, or by paying \$1 to investigate project 3?

value Z_i is project i's reservation value, such that:

$$E[(X_i - Z_i)^+] = c_i.$$

Intuitively, the GW index is the value of a *fictitious* outside option that makes the agent just willing to pay cost c_i to reveal the true value of X_i instead of immediately taking the outside option Z_i.[2] For an uncertain project in our setting, the GW index value is $p_i(V_i - Z_i) = c_i$, i.e.,

$$Z_i = \frac{p_i V_i - c_i}{p_i}.$$

After a project has been investigated, $c_i = 0$ and p_i is either zero or one. The GW value will be $Z_i = V_i$ for a known winner and $Z_i = 0$ for a known loser.

Gittins (1979) and Weitzman (1979) show that the optimal sequence of actions tracks the project with the highest value of Z_i. If the highest value of Z_i corresponds to a project with an unknown payoff, then the agent should acquire information about that project.

---

[1] We have also implemented the analysis for risk-averse and loss-averse preferences. Using standard calibrations for risk aversion, loss aversion, and narrow framing (e.g., Amos Tversky and Daniel Kahneman, 1992) does not change our theoretical predictions and our experimental findings. In other words, in our experiment, the predictions of the rational model

and the DC model do not change once the assumption of risk neutrality is replaced with calibrated levels of curvature in the utility function.

[2] If the agent does not explore the value of X_i, her payoff is the fictitious outside option Z_i. The expected benefit from exploration is

$$E[\max(X_i, Z_i)] - Z_i = E[(X_i - Z_i)^+].$$

Equating this expected benefit to the cost of search, c_i, yields the Gittins-Weitzman index.

TABLE 2—ALGORITHMS FOR INVESTMENT GAME A

	Gittins-Weitzman	Directed cognition
t	Investigate project 2	Investigate project 3
$t + 1$	Take project 2 if winner	Take project 3 if winner
$t + 2$	Else investigate project 3	Else investigate project 2
$t + 3$	Take project 3 if winner	Take project 2 if winner
$t + 4$	Else take project 1	Else take project 1

If the highest value of Z_i corresponds to a winning project, then the agent should take that investment project, thereby ending that game.

For the game above, the agent should acquire information about project 2 (the long shot), taking that project if it turns out to be a winner. Otherwise, the agent should acquire information about project 3, taking that project if it turns out to be a winner and taking project 1 as a last resort. (See Table 2.)

Studying this class of games enables us to run an empirical horse race between the optimal search model (i.e., the GW algorithm) and the directed cognition model (Gabaix and Laibson, 2005).

A. *Applying the Directed Cognition Algorithm*

The directed cognition (DC) model, "solves" problems by searching *as if* each search operation were the last search operation. To apply directed cognition, we calculate the expected benefit and cost of each available search operation as if this operation were the last one executed before a final investment project is taken.

Let S_t be the value of the best known winning project at time t. The (myopic) expected benefit from investigating project i is

$$E[\max(X_i, S_t)] - S_t = E[(X_i - S_t)^+].$$

The cost of this search operation is $c_i = 1$. DC selects the search operator with the highest gain—the difference between benefits and cost:

$$G_i \equiv E[(X_i - S_t)^+] - c_i$$
$$= p_i(V_i - S_t) - c_i.$$

This formula also describes the myopic gain from taking a known winner: $c_i = 0$, and $p_i =$

1. The directed cognition algorithm investigates or takes the investment with the highest G value. The algorithm iterates if an investment has not yet been taken.

For the game above, the directed cognition algorithm predicts that the agent will acquire information about project 3, taking that project if it turns out to be a winner. Otherwise, the agent will acquire information about project 2, taking that project if it turns out to be a winner and taking project 1 as a last resort. (See Table 2.)

The DC algorithm recognizes the option value of being able to reject the next investigated project if that project does not turn out to be profitable. But the algorithm does not recognize the option value of being able to investigate sequences of uncertain projects. An optimal search calculation needs to evaluate such sequences: "Project 3 has a higher expected value than project 2, but project 2 has a small chance of a high payoff. If I investigate project 2 first, and it is not a winner, then I can proceed to investigate project 3." Because directed cognition is myopic it cannot recognize such chains of reasoning. Hence, games like those in Table 1 are well suited to distinguish between optimal search and directed cognition.

For any game, let the low-probability project have probability p_i of being a winner and value V_i when it is a winner. Let the high-probability project have probability $p_j > p_i$ of being a winner and value V_j when it is a winner. Consider a parameterization in which

$$(1) \qquad \frac{p_i V_i - c}{p_i} > \frac{p_j V_j - c}{p_j} > 1$$

and

$$(2) \qquad p_i(V_i - 1) < p_j(V_j - 1).$$

TABLE 3—FRACTION OF SUBJECTS WHO PLAY THE FIRST MOVE ACCORDING TO
GITTINS-WEITZMAN OR ACCORDING TO DIRECTED COGNITION

	Percentage GW	Percentage DC	Difference	t-test	p-value
Game A	33	65	32	3.80	0.0001
Game B	35	63	28	3.30	0.0010
Game C	42	57	16	1.80	0.0726
Game D	28	70	42	5.30	0.0000
Game E	34	62	28	3.32	0.0009
A–E average	34	63	29	4.55	0.0000

Then the GW algorithm begins with an investigation of the *low-probability project* and the DC algorithm begins with an investigation of the *high-probability project.*

We study five games (A–E) with such noncongruent predictions and five other games with congruent predictions (F–J).[3] From a heuristic perspective, the noncongruent games are ones in which the low-probability investment has high variance and a low expected value.

B. *Experimental Results*

One hundred twenty-nine subjects[4] received a mean total payoff of $26.51, with a standard deviation of $13.40. Payoffs ranged from $10 to $63. Subjects played ten randomly ordered investment games like the one in Table 1. Each of the games includes a low-probability project, a high-probability project, and a sure thing. The experimental protocol and a Web-based simulation of the actual experiment are available at http://www.e-aer.org/data/sept06/20030922_data.zip.

Five of the ten investment games (games A–E) have a GW strategy that differs from the directed cognition strategy, and this difference appears in the first move. The first panel in Table 3 reports the proportions of first moves in each of these games that match the GW strategy

TABLE 4—FRACTION OF FIRST MOVES MATCHING
GITTINS-WEITZMAN AND DIRECTED COGNITION

	Percentage GW and DC
Game F	73
Game G	67
Game H	73
Game I	74
Game J	83
F–J average	74

and the proportions of first moves that match the directed cognition strategy. For all of the five games, the GW proportions are below the DC proportions. We reject the null hypothesis that the GW and DC proportions are equal for four of the five games.

Averaging over all five games, 34 percent of the moves follow the GW strategy, and 63 percent of the moves follow the directed cognition strategy. We reject the null hypothesis that these means are equivalent with a p value lower than 10^{-4}. When the predictions of the two models differ, directed cognition predicts subjects' choices better than optimal search.

The probabilities do not sum to one because neither GW nor DC predicts selection of the sure-thing investments on the first move. A small number of subjects made this choice.

Five of the ten investment games (games F–J) have a GW strategy that matches the DC strategy. Table 4 reports the proportion of first moves in each of these games that matches the GW and DC strategy. On average, 74 percent of the moves follow the GW and DC strategy.[5]

---

[3] The payoffs and probabilities of the risky projects in games A–J are: game A: ($21, $p = 0.09$) and ($10, $p = 0.76$); game B: ($19, $p = 0.11$) and ($10, $p = 0.79$); game C: ($23, $p = 0.09$) and ($13, $p = 0.72$); game D: ($18, $p = 0.12$) and ($10, $p = 0.81$); game E: ($20, $p = 0.12$) and ($12, $p = 0.85$); game F: ($22, $p = 0.48$) and ($11, $p = 0.74$); game G: ($24, $p = 0.34$) and ($9, $p = 0.70$); game H: ($18, $p = 0.52$) and ($11, $p = 0.74$); game I: ($25, $p = 0.39$) and ($9, $p = 0.70$); game J: ($10, $p = 0.09$) and ($8, $p = 0.85$). In all games, the sure thing is $1.

[4] Subjects are Harvard undergraduates and Harvard summer school students. Of the total, 61 percent report having taken at least one statistics course.

[5] Our gambles may be special. For instance, suppose that the risky projects were ($10, $p = 0.8$) and ($70, $p = 0.1$), an example suggested to us by a referee. For this game, subjects might first explore the low-probability prize, contradicting directed cognition. Such a result could, however, be due to probability-reweighting. Prospect theory

TABLE 5—FRACTION OF ALL MOVES MATCHING GITTINS-WEITZMAN OR DIRECTED COGNITION

	Percentage GW	Percentage DC	Difference	*t*-test	*p*-value
Game A	66	79	13	3.67	0.0002
Game B	68	80	11	3.21	0.0013
Game C	70	76	06	1.76	0.0785
Game D	65	83	17	5.00	0.0000
Game E	66	78	11	3.23	0.0012
A–E average	67	79	12	4.39	0.0000

We now turn to an analysis of all moves (not just the first move). Table 5 reports the fraction of *all* moves that are correctly predicted by the GW and DC strategies for games A–E. We recalculate strategy predictions after every move, conditional on the player's other selections. While the strategies predict different first moves, they often predict the same play for subsequent moves. Because of this overlap, the strategies are not mutually exclusive, and probabilities do not sum to one. This analysis does not distinguish the strategies as cleanly as the first-move analysis presented above. We present the results here as a robustness check.

For all five games, the GW proportions are below the DC proportions. We reject the null hypothesis that both models' proportions are equal for four of the five games. Averaging over all five games, 67 percent of the moves follow the GW strategy, and 79 percent of the moves follow the DC strategy. We reject the null hypothesis that these means are equivalent with a *p* value less than 10^{-4}.

Table 6 reports the fraction of moves (82 percent) that follow the GW and DC strategy for games F–J.

We now calculate the number of games in which a subject followed *all* of the moves prescribed by one of the two algorithms (i.e., GW or DC). Specifically, we calculate the number of games in which each subject followed a particular algorithm perfectly from start to finish, including the final choice.[6] We

TABLE 6—FRACTION OF ALL MOVES MATCHING GITTINS-WEITZMAN AND DIRECTED COGNITION

	Match GW/DC
Game F	83%
Game G	78%
Game H	83%
Game I	83%
Game J	86%
F–J average	82%

break our games into two subgroups: the noncongruent games (A–E), for which the two algorithms' predictions diverge, and the congruent games (F–J).

With respect to the noncongruent games (A–E), 47 percent of subjects did not follow the GW algorithm in *any* of the five games, while only 16 percent of subjects failed to follow the DC algorithm in *any* of the five games. Moreover, only 11 percent of subjects follow the GW algorithm in *all* five games, while 22 percent of subjects follow the DC algorithm in *all* five games. On average, subjects play 1.6 noncongruent games (A–E) exactly following the GW algorithm and 2.6 noncongruent games exactly following the DC algorithm. The entire frequency distribution is reported in Table 7 (for both the congruent and noncongruent games). Using these data, the DC algorithm outperforms GW ($p = 0.001$).

We have also analyzed the ability of the models to predict the final outcomes of the experiment. Using such outcomes as the focus of the study, the GW and DC models are statistically indistinguishable. The lack of resolution is partially due to the noisiness of the

predicts that small probability events are overweighted in decisions.

[6] If a subject follows the search strategy of a particular algorithm, he would make a bizarre mistake if he didn't also follow the final choice predictions. Final choices are made among "sure-thing" payoffs. Accordingly, mistakes on final

choices are very rare: they happen in 1.9 percent of the games that are played.

TABLE 7—EMPIRICAL FREQUENCIES OF THE NUMBER OF GAMES IN WHICH A MODEL COMPLETELY PREDICTS THE SUBJECT'S STRATEGY AND THE SUBJECT'S FINAL CHOICE

	Games A–E			Games F–J
	Frequency			Frequency
Number of games	GW strategy	DC strategy	Number of games	GW/DC strategy
0	47%	16%	0	5%
1	12%	19%	1	9%
2	9%	16%	2	12%
3	11%	12%	3	19%
4	11%	16%	4	22%
5	11%	22%	5	33%
Mean	1.6	2.59	Mean	3.45
Median	1	3	Median	4
Standard deviation	1.83	1.81	Standard deviation	1.51

Notes: The table displays the number of games in which the full strategy of the player, including search pattern and final choices, coincided with the predictions of the GW versus DC model. Those predictions are different in games A–E, and the same in games F–J. DC makes a successful prediction on a greater number of games than GW ($p = 0.001$).

lotteries. This analysis is reported in the AER Web Appendix.

II. Second Set of Experiments: Choice among *N* Complex Goods

The search experiment above studies a problem that is *simple* enough to admit an analytic optimal solution. Real world problems, however, tend to be far more complex. Indeed, many real world problems do not have an analytic optimal solution or even an optimal solution that can be practically calculated numerically. We would like to have models that can successfully predict behavior in such complex environments.

With these goals in mind, we analyze a second experiment that captures some of the complex factors that arise in real world problems. In this "complex" experiment, subjects choose one good from a set of *N* goods, each of which has numerous attributes.

In this complex experiment, decision time is a scarce resource, and information acquisition is measured continuously. We make time scarce in two different ways. First, we give subjects an *exogenous* amount of time to choose one good from a set of goods—a choice problem with an exogenous time budget. Here we measure how subjects allocate time as they acquire information about each good's attributes before making a final selection.

Then we give the subjects an open-ended sequence of choice problems like the one above. In this treatment, the subjects keep facing different choice problems until a total budget of time runs out. The amount of time a subject allocates to each choice problem is now an *endogenous* variable. Because payoffs are cumulative and each choice problem has a positive expected value, subjects have an incentive to move through the choice problems quickly. But moving too quickly reduces the quality of their decisions.

Following other economists (Colin F. Camerer et al., 1993; Miguel A. Costa-Gomes et al., 2001; Eric J. Johnson et al., 2002; and Costa-Gomes and Vincent P. Crawford, forthcoming), we use the "Mouselab" programming language to measure subjects' information acquisition.[7] Information is hidden "behind" boxes on a computer screen. Subjects use the computer mouse to open the boxes. Mouselab records the order and duration of information acquisition. Since we allow only one screen box to be open at any point in time, the Mouselab software enables us to pinpoint what information the subject is acquiring on a

---

[7] John W. Payne et al. (1993) developed the Mouselab language in the 1970s. Mouselab is one of many "process tracing" methods. For example, Payne et al. (1978) elicit mental processes by asking subjects to "think aloud." J. Edward Russo (1978) records eye movements.
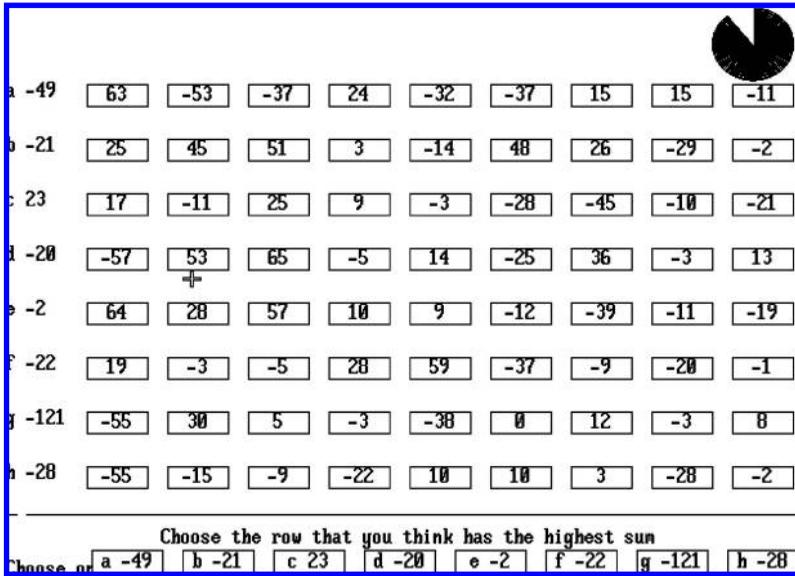
FIGURE 1. SAMPLE GAME WITH ALL VALUES UNMASKED

Notes: In the actual experiment, the subjects see the value of only one box at a time (see Figure 2). Values in each column are drawn independently from a normal distribution with the same variance, with variances declining linearly from left to right. In this sample, the left-most column is generated with a standard deviation of 30.6 cents, which is explained to subjects as a 95-percent confidence interval of −60 to 60 cents.

second-by-second basis throughout the experiment.[8]

A. *The Details of an N-Good Choice Task*

An N-good game is an N-row by M-column matrix of boxes (Figure 1). Each box contains a random payoff (in units of cents) generated with normal density and zero mean. After analyzing an N-good game, the subject makes a final selection and "consumes" a single row from that game. The subject is paid the sum of the boxes in the consumed row.

Consuming a row represents an abstraction from a very wide class of choice problems. We call this problem an N-good game, since

the N rows conceptually represent N goods. The columns represent M different attributes.

For example, consider a shopper who has decided to go to Wal-Mart to select and buy a television. The consumer faces a fixed number of television sets at Wal-Mart (N different TV's from which to choose). The television sets have M different attributes—size, price, remote control, warranty, etc. By analogy, the N TV's are the rows of Figure 1, and the M attributes (in a utility metric) appear in the M columns of each row.

In our experiment, the importance or variability of the attributes declines as the columns move from left to right. In particular, the variance decrements across columns equal one-tenth of the variance in column one. For example, if the variance used to generate column one is 1,000 (squared cents), then the variance for column 2 is 900, and so on, ending with a variance for column 10 of 100. So columns on the left represent the attributes with the most (utility-metric) variance, like screen size or price in our TV example.

---

[8] Mouselab has the drawback that it uses an artificial decision environment, but several studies have shown that the Mouselab environment distorts final choices over goods/actions only minimally (e.g., Costa-Gomes et al., 2001; Costa-Gomes and Crawford, forthcoming). Mouselab's interface does generate "upper-left" and "left-to-right" search biases, which we discuss in Section IIB below.
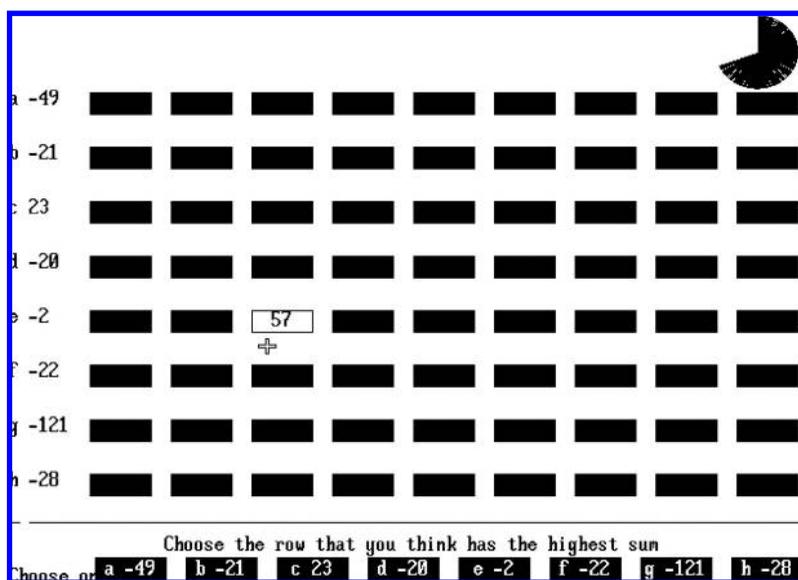
FIGURE 2. SAMPLE GAME WITH VALUES CONCEALED

Notes: This is how a sample game would appear to subjects, with values concealed by boxes. Subjects can use the mouse to open one box at a time. In this game the subject faces a set time limit; the clock in the upper-right corner reveals the fraction of time remaining.

Columns on the right represent the attributes with the least (utility-metric) variance, like minor clauses in the warranty.[9]

So far our game sounds simple: "Consume the best good (i.e., row)." To measure information acquisition, we mask the contents of boxes in columns 2 through M. Subjects are shown only the box values in column 1.[10] A subject can, however, left-click on a masked box in columns 2 through M to unmask the value of that box (Figure 2).

Only *one* box from columns 2 through M can be unmasked at a time. This procedure enables us to record exactly what information the subject is observing at every point in time.[11] Re-

vealing the contents of a box does not imply that the subject consumes that box. Note, too, that if a row is picked for consumption, then all boxes in that row are consumed, whether or not they have been previously unmasked.

We introduce time pressure, so that subjects will not be able to unmask—or will not choose to unmask—all of the boxes in the game. Mouselab records which of the $N(M - 1)$ masked boxes the subjects unmask. Of course, we also record which rows the subjects choose/consume.

We study a setting that reflects realistic—i.e., high—levels of decision complexity. This complexity forces subjects to confront tradeoffs. Real consumers in real markets frequently face decisions that are much more complex.

Masked boxes and time pressure capture important aspects of our Wal-Mart shopper's experience. The Wal-Mart shopper selectively attends to information about the attributes of the

---

[9] In our experiment, all of the attributes have been demeaned.

[10] We reveal the value of column 1 because it helps subjects remember which row is which. In addition, revealing column 1 initializes the game by breaking the eight-way tie that would exist if subjects began with the expectation that all rows had the same value (zero).

[11] When we designed the experiment, we considered but did not adopt a design that permanently keeps boxes open once they have been selected by the subject. This alternative approach has the advantage that subjects face a reduced memory burden. On the other hand, if boxes stay open

permanently, subjects have the option to quickly—and mechanically—open many boxes, and only afterward analyze their content. Hence, leaving boxes open implies that we lose the ability credibly to infer the subject's attention at each point in time.

TVs among which she is picking. The shopper may also face some time pressure, either because she has a fixed amount of time to buy a TV, or because she has other tasks she can do in the store if she selects her TV quickly. We explore both types of cases in our experiment.

B. *Games with Exogenous and Endogenous Time Budgets*

In our experiment, subjects play two different types of *N*-good games: games with exogenous time budgets and games with endogenous time budgets. We will refer to these as "exogenous" and "endogenous" games.

For each exogenous game a game-specific time budget is generated from the uniform distribution over the interval [10 seconds, 49 seconds]. A clock shows the subject the amount of time remaining for each exogenous time game (see clock in Figure 2). This is the case of a Wal-Mart shopper with a fixed amount of time to buy a good.

In endogenous games, subjects have a fixed budget of time—25 minutes—in which to play as many different *N*-good games as they choose. In this design, adjacent *N*-good games are separated by 20-second buffer screens, which count toward the total budget of 25 minutes. Subjects are free to spend as little or as much time as they want on each game, so time spent on each game becomes an endogenous choice variable. This is the case of a Wal-Mart shopper who can move on to other purchases if she selects her TV quickly.

We study both exogenous time games and endogenous time games because these two classes of problems commonly arise in the real world and any cognition model should be able to handle both situations robustly. Both types of problems enable us to study within-problem attention allocation decisions. In addition, the endogenous time games provide a natural framework for studying stopping rules, i.e., the decision to allocate less analysis to the current game, freeing time for subsequent games.

C. *Experimental Logistics*

Subjects receive printed instructions explaining the structure of an *N*-good game and the setup for the exogenous and endogenous games. Subjects are then given a laptop on which they read instructions that explain the Mouselab interface. Subjects play three test games, which do not count toward their payoffs.

Then subjects play 12 games with separate exogenous time budgets. Finally, subjects play a set of endogenous games with a joint 25-minute time budget. For half of the subjects we reverse the order of the exogenous and endogenous games. At the end of the experiment, subjects answer demographic and debriefing questions.

Subjects are paid the cumulative sum of all rows that they consume. After every game, feedback reports the running cumulative value of the consumed rows.

III. Application of the Directed Cognition Model

Application of the directed cognition model can be broken down into three iterative steps, which we first summarize and then describe in detail.

Step 1: Using a myopic planning horizon, calculate the expected economic benefits and costs of different potential search operations. Specifically, evaluate each incremental search operation *as if* it were the last search operation.[12]

Step 2: Execute the search operation with the highest ratio of expected benefit to cost. For example, if exploration of the "next" two boxes in the sixth row has the highest expected ratio of benefit to cost, then unmask those two boxes.

Step 3: Return to step 1 unless time has run out (in exogenous time games) or until the ratio of expected benefit to cost falls below some threshold value (in endogenous time games).

A. *Notation*

Since our games all have eight rows (goods), we label the rows *A*, *B*, ... , *H*. We use lower-case

---

[12] We imagine that this step recruits both conscious and unconscious mental processing and do not take a position on their relative contributions. Daniel M. Wegner (2002) and Thomas Gilovich et al. (2002) argue that unconscious/automated mental processes play a central role in many behavioral choices. Much of the activity of the limbic system appears to be unavailable to consciousness, but nevertheless is critical for reward valuation, approach/avoidance, and attention allocation.

letters—*a*, *b*, ... , *h*—to track a subject's expectations of the values of the respective rows.

The subject knows the values of all boxes in column 1 when the game begins. Thus, at the beginning of the game, the row expectations will equal the value of the payoff in the leftmost cell of each row. For example, if row *C* has a 23 in its first cell, then at time zero $c = 23$. If the subject unmasks the second and third cells in row *C*, revealing cell values of 17 and -11, then c would be updated to $29 = 23 + 17 + -11$. We now describe the three steps in detail.

B. *Step 1*

In step 1 the agent calculates the expected economic benefits and costs of different search operations. For our application, a search operation is a partial or complete unmasking/analysis of boxes in a particular row of the matrix. Such an operation enables the decision maker to improve her forecast of the expected value of that row. In our notation, O_A^Γ represents the operation "open Γ additional boxes in row *A*." Because tractability concerns lead us to limit agents' planning horizons to only a single search operation at a time, we assume that individual search operations themselves can include one *or more* box openings, though multiple box openings must be sequential and in the same row. Multiple box openings increase the amount of information revealed by a single search operator, increase the option value of information revealed by that search operator, and make the (partially myopic) model more forward-looking.[13]

The operator O_A^Γ selects the boxes to be opened using a maximal-information rule. In other words, the O_A^Γ operator would select the Γ unopened boxes (in row *A*) that have the highest variances (i.e., with the most information). In our game, this corresponds with left to right box

[13] To gain intuition for this effect, consider two goods *A* and *B*, with $a = \frac{3}{2}$ and $b = 0$. Suppose that each of two remaining boxes in row A can take the value 1 or -1. After the information in a single box has been revealed, a' will equal $\frac{5}{2}$ or $\frac{1}{2}$. Hence, a partially myopic agent will not see the benefit of opening one box, since no matter what happens, $a' > b$. If, however, the agent considers opening both remaining boxes, there is a chance that a' will fall below 0, implying that gathering the information from the two boxes would be useful to the agent.
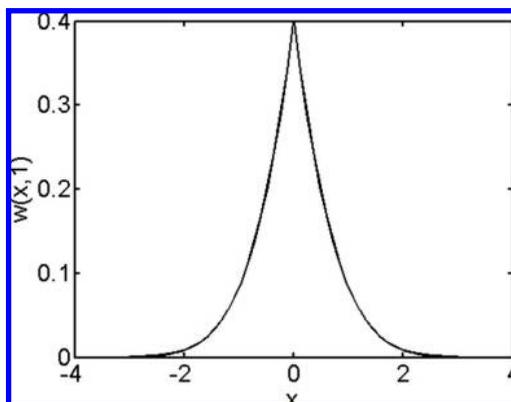

FIGURE 3. $w(x, \sigma)$ FOR $\sigma = 1$

Notes: This figure plots the expected benefit from continued search on an alternative if the difference between the value of the searched alternative and the best of the other alternatives is *x*, and the standard deviation of the information gained is $\sigma = 1$ as defined in equation (3). This *w* function is homogeneous of degree one.

openings (skipping any boxes that may have been opened already).

We assume that an operator who opens Γ boxes has cost $\Gamma \cdot \kappa$, where κ is the cost of unmasking a single box. We take this cost to include many components, including the time involved in opening the box with the mouse, reading the contents of the box, and updating expectations.

The expected benefit (i.e., option value) of a search operation is given by

$$(3) \qquad w(x, \sigma) \equiv \sigma\phi\left(\frac{x}{\sigma}\right) - |x|\Phi\left(-\frac{|x|}{\sigma}\right),$$

where ϕ represents the standard normal density function, Φ represents the associated cumulative distribution function, *x* is the estimated value gap between the row that is under consideration and its next best alternative, and σ is the standard deviation of the payoff information that would be revealed by the search operator. Figure 3 plots this benefit function. We motivate equation (3) below, but first present an example calculation.

In the game shown in Figure 2, consider a search operator O_H^3 that explores three boxes in row *H*. The initial expected value of *H* is $h = -28$. The best current alternative is row *C*, which has a current payoff of $c = 23$. So the

estimated value gap between H and the best alternative is

$$x_O = |h - c| = 51.$$

A box in column n will reveal a payoff η_{Hn} with variance $(40.8)^2(1 - n/10)$, and the updated value of H after the three boxes have been opened will be

$$h' = -28 + \eta_{H2} + \eta_{H3} + \eta_{H4}.$$

Hence the variance of the information revealed by the search operator is

$$\sigma_O^2 = var(\eta_{H2} + \eta_{H3} + \eta_{H4})$$

$$= (40.8)^2\left(\frac{9}{10} + \frac{8}{10} + \frac{7}{10}\right),$$

i.e., $\sigma_O = 63.2$. So the benefit of the search operator is $w(x_O, \sigma_O) = 7.5$, and its cost is $\Gamma_O \cdot \kappa = 3\kappa$.

We now motivate equation (3). To fix ideas, consider a new game. Suppose that the decision maker is analyzing row A and will then immediately use that information to choose a row. Assume that row B would be the leading row if row A were eliminated, so row B is the next best alternative to row A.

The agent is considering learning more about row A by executing a search operator O_A^Γ. Executing the search operator will enable the agent to update the expected payoff of row A from a to $a' = a + \varepsilon$, where ε is the sum of the values in the Γ newly unmasked boxes in row A.

If the agent does not execute the search operator, her expected payoff will be

$$\max(a, b).$$

If the agent plans to execute the search operator, her expected payoff will be

$$E[\max(a', b)].$$

This expectation captures the option value generated by being able to pick either row A or row

B, *contingent* on the information revealed by search operator O_A^Γ. The value of executing the search operator is the difference between the previous two expressions:

$$(4) \qquad E[\max(a', b)] - \max(a, b).$$

This value can be represented with a simple expression. Let σ represent the standard deviation of the change in the estimate resulting from applying the search operator

$$\sigma^2 = E(a' - a)^2.$$

The value of the search operator is[14]

$$(5) \quad E[\max(a', b)] - \max(a, b) = w(a - b, \sigma).$$

To develop intuition for the w function, Figure 3 plots $w(x, 1)$. In the general case, $w(x, \sigma) = \sigma w(x/\sigma, 1)$.

The option value framework captures two fundamental comparative statics. First, the value of a row exploration decreases the larger the gap between the active row and the next best row: $w(x, \sigma)$ is decreasing in $|x|$. Second, the value of a row exploration increases with the variability of the information that will be obtained: $w(x, \sigma)$ is increasing in σ. In other words, the more information that is likely to be

---

[14] This result assumes Gaussian innovations, which is the density used to generate the games in our experiment. To derive equation (5), begin by assuming that $b \geq a$. In this case, $\Delta := \max(a', b) - \max(a, b) = \max(a + \varepsilon - b, 0)$. Because ε is drawn from a Normal$(0, \sigma^2)$ distribution, we get the right-hand side of equation (5):

$$\Delta = E[\max(a', b)] - \max(a, b)] = \int_{b-a}^{\infty} (a + \varepsilon - b)\phi\left(\frac{\varepsilon}{\sigma}\right)\frac{d\varepsilon}{\sigma}$$

$$= \int_{|a-b|}^{\infty} (\varepsilon - |a - b|)\phi\left(\frac{\varepsilon}{\sigma}\right)\frac{d\varepsilon}{\sigma}$$

$$= \int_{|a-b|}^{\infty} \varepsilon\phi\left(\frac{\varepsilon}{\sigma}\right)\frac{d\varepsilon}{\sigma} - |a - b|\left(1 - \Phi\left(\frac{|a - b|}{\sigma}\right)\right)$$

$$= \sigma\phi\left(\frac{|a - b|}{\sigma}\right) - |a - b|\Phi\left(-\frac{|a - b|}{\sigma}\right) \text{as } x\phi(x) = -\phi'(x)$$

so $\Delta = w(|a - b|, \sigma)$. The calculation is similar when $b \leq a$.

revealed by a row exploration, the more valuable such an exploration becomes.

C. Step 2

Step 2 executes the search operation with the highest ratio of expected benefit to cost. Recall that the expected benefit of an operator is given by the $w(x, \sigma)$ function and that the implementation cost of an operator is proportional to the number of boxes that it unmasks. The subject executes the search operator with the greatest benefit/cost ratio,[15]

$$(6) \qquad G \equiv \max_O \frac{w(x_O, \sigma_O)}{\Gamma_O},$$

where κ is the cost of unmasking a single box.[16] Since κ is constant, the subject executes the search operator

$$O^* = \arg \max_O \frac{w(x_O, \sigma_O)}{\Gamma_O}.$$

The Web appendix to this article contains an example of such a calculation.

D. Step 3

Step 3 is a stopping rule. In games with an exogenous time budget, the subject keeps returning to step 1 until time runs out. In games with an endogenous time budget, the subject keeps returning until G falls below the marginal value of time, which must be calibrated.

---

[15] We postulate that the agent picks the operator O that maximizes the ratio $w(x_O, \sigma_O)/\Gamma_O$ rather than the difference $w(x_O, \sigma_O) - \kappa \Gamma_O$, where κ is a marginal cost of time. The ratio criterion does not require the subject to calculate κ. Hence, the ratio formulation is more frugal than the difference formulation. In the simple experiment of Section I, using a ratio or difference formulation leads to identical predictions, as both risky projects require the same search cost. In general, however, the ratio and difference approaches yield different predictions. When the cost of time is known, the difference formulation is preferable.

[16] Our model thus gives a crude but compact way to address the "accuracy versus simplicity" trade-off in cognitive processing. See Enriqueta Aragones et al. (2005).

E. Calibration of the Model

We use two different methods to calibrate the *marginal* value of time during the *endogenous time games.* First, we estimate the marginal value of time as perceived by our subjects. Advertisements for the experiment implied that subjects would be paid about $20 for their participation, which would take about an hour. In addition, subjects were told that the experiment would be divided into two halves, and that they were guaranteed a $5 show-up fee.

Using this information, we calculate the subjects' anticipated marginal payoff per unit time during games with endogenous time budgets. This marginal payoff per unit time is the relevant opportunity cost of time during the endogenous time games. Since subjects were promised $5 of guaranteed payoffs, their expected *marginal* payoff for their choices during the experiment was about $15. Dividing this in half implies an expectation of about $7.50 of marginal payoffs for the endogenous time games. Since the endogenous time games were budgeted to take 25 minutes, which was known to the subjects, the perceived marginal payoff per second of time in the experiment was

$$\frac{750 \text{ cents}}{25 \text{ minutes} \cdot 60 \text{ seconds/minute}}$$

$$= 0.50 \text{ cents/second.}$$

Since subjects took on average 0.98 seconds to open each box, we end up with an implied marginal shadow cost per box opening of

$$(0.50 \text{ cents/second})(0.98 \text{ seconds/box})$$

$$= 0.49 \text{ cents/box.}$$

We also explore a one-parameter version of the directed cognition model, in which the cost of cognition—κ—is chosen to make the model partially fit the data. Calibrating the model so it matches the average number of boxes explored in the endogenous games implies $\kappa = 0.18$ cents/box. Here κ is chosen only to match the average amount of search per endogenous time game, not to match the order of search or the distribution of search across games.

F. *Conceptual Issues*

This model is easy to analyze and is computationally tractable, implying that it can be empirically tested. The simplicity of the model follows from three special assumptions. First, the model assumes that the agent calculates only a partially myopic expected gain from executing each search operator. This assumption adopts the approach taken by Philippe Jehiel (1995), who assumes a constrained planning horizon in a game-theory context. Second, the DC model assumes that the agent uses a fixed positive shadow value of time. This shadow value of time enables the agent to trade off current opportunities with future opportunities. Third, the DC model avoids the infinite regress problem (i.e., the costs of thinking about thinking about thinking, etc.) by using a myopic option value calculation assumed to be costless to the agent. In some evolutionary relevant domains, the brain may have evolved specialized algorithms to solve the infinite regress problem (e.g., intuition). But given the "evolutionary irrelevance" of the experimental task, it is unlikely that sophisticated evolved solutions would be applicable to the current problem. Hence, a crude myopic solution—like the one we adopt—may be reasonable. Our assumption that the myopic option-value calculation is costless is not realistic, however, and is made only with modeling convenience in mind.

Without some version of these three simplifying assumptions, the model would not be useful in practice. Without some partial myopia (i.e., a limited evaluation horizon for option value calculations), the problem could not be solved either analytically or computationally.[17] Without the positive shadow value of time, the agent would not be able to trade off her current activity with unspecified future activities and would never finish an endogenous time game without first (counterfactually) opening up all of the boxes. Finally, without eliminating cognition costs at some primitive stage of reasoning, maximization models are not well defined.[18]

We return now to the first of the three points listed in the previous paragraph: the perfectly rational search model is not solvable in our context. An exact solution of the perfect rationality model requires the calculation of a value function with 17 state variables: one expected value for each of the eight rows, one standard deviation of unexplored information in each of the eight rows, and finally the time remaining in the game. This dynamic programming problem in $\mathbb{R}^{17}$ suffers from the curse of dimensionality and would overwhelm modern supercomputers.[19] By contrast, the directed cognition model is equivalent to eight completely *separable* problems, each of which has only two state variables: x, the difference between the current expected value of the row and the current expected value of the next best alternative row; and σ, the standard deviation of unexplored information in the row. So the "dimensionality" of the DC model is only 2 (compared to 17 for the model of perfect rationality).

We can compare the performance of the partially myopic DC model and the performance of the perfectly rational model. Like the DC model, the perfectly rational model assumes that examining a new box is costly and that calculating the optimal search *strategy* is costless (analogous to our assumption that solving for O^* is costless). The Web Appendix gives lower bounds on the payoffs of the DC model relative to the payoffs of the perfectly rational model. DC does at least 91 percent as well as perfect rationality for exogenous time games

---

[17] Gabaix and Laibson (2006) propose that consumer myopia explains a series of market phenomena.

[18] See Conlisk (1996) for a description of the infinite regress problem and an explanation of why it plagues all decision cost models. We follow Conlisk in advocating exogenous truncation of the infinite regress of thinking.

[19] Approximating algorithms could be developed, but after consulting with experts in operations research, we concluded that existing approximation algorithms cannot be used without a prohibitive computational burden. Our problem is discrete (eight rows, ten boxes), and so the state-space is large (seven continuous variables for the relative current payoffs, plus eight discrete variables for the depth of operations), and there is no clear way to simplify the state-space. The standard simplifying tools (Dimitri P. Bertsekas and John N. Tsitsiklis, 1996) do not apply here. The Gittins index (Gittins, 1979; Weitzman, 1979) does not apply here either, for much the same reason it does not apply to most dynamic problems. In Gittins's framework, it is crucial that one can do only one thing to a row (i.e., an "arm") at a particular point in time. In contrast, in our game, a subject can do more than one thing with a row. She can explore it further, or take it and end the game. Hence, our game does not fit into Gittins's framework. We explored several modifications of the Gittins index, but they proved unfruitful at breaking the curse of dimensionality.

and at least 71 percent as well as perfect rationality for endogenous time games. In other words, using the DC algorithm instead of perfect rationality would reduce payoffs by no more than 8 percent and 29 percent, respectively, in the exogenous and endogenous time games.[20]

G. *Other Decision Algorithms*

In this subsection, we describe several naive benchmarks which we compare to the directed cognition model. These benchmarks are motivated by economists' interest in determining whether other simple models can organize our experimental data.[21] The benchmarks are also related to prototypical models in the psychology literature.

Column Model.—The *column model* unmasks all the boxes in column 2 (top to bottom), then in column 3, ... , etc. In exogenous games, this column-by-column unmasking continues until the simulation has explored the same number of boxes as a "yoked" subject.[22] In endogenous time games, the unmasking continues until a row has been revealed with an estimated value greater than or equal to $A^{\text{Column model}}$, an aspiration or satisficing level. This aspiration level (and those discussed below) are all chosen so that the simulations generate an average number of simulated box openings that matches the average number of empirical box openings (26 boxes per game).

Row Model.—The *row model* unmasks boxes row by row, sequencing the unmasking according to the values in column 1. In exogenous games, this row-by-row unmasking continues until the simulation has explored the same number of boxes as a yoked subject (see previous footnote). In endogenous games, the unmasking continues

until a row has been revealed with an estimated value greater than or equal to $A^{\text{Row model}}$.

Directed Cognition with Satisficing.—In endogenous games, we also analyze a DC model with a satisficing stopping rule. In this model, the DC model is iterated until a row has been revealed with an estimated value greater than or equal to $A^{\text{DC model}}$.

Elimination by Aspects.—A choice algorithm called *Elimination by Aspects* (EBA) has been widely studied in the psychology literature (e.g., Tversky, 1972; and Payne et al., 1993). We use EBA to analyze games with endogenous time budgets. Each row is a good with ten different attributes or "aspects" represented by the ten different boxes of the row. The EBA algorithm proceeds aspect by aspect (i.e., column by column) from left to right, eliminating goods (i.e., rows) with an aspect that falls below some aspiration value A^{EBA}. This elimination continues, stopping at the point where the *next* elimination would eliminate all remaining rows. At this stopping point, EBA picks the remaining row with the highest estimated value.

IV. Results

Our 388 subjects[23] received a mean total payoff of $29.23, with a standard deviation of $5.49. Payoffs ranged from $13.07 to $46.69.[24] All subjects played 12 games with exogenous times. On average, subjects chose to play 28.7 games under the endogenous time limit, with a standard deviation of 7.9. The number of games played (endogenous and exogenous games combined) ranged from 21 to 65. Our AER Web Appendix provides details about (negative) tests for subject learning effects, reports evidence that subjects have imperfect memory, and describes our (standard) bootstrap methodology for calculating standard errors. The Appendix also contains the experimental protocol and a Web-based simulation of the actual experiment.

---

[20] Of course, these bounds do not apply to different classes of games. It is possible to construct specific games in which the losses approach 100 percent. Partially rational algorithms will always fail in special circumstances that are designed to exploit the algorithm.

[21] See Camerer and Teck-Hua Ho (1999), Ido Erev and Alvin E. Roth (1998); and Gerd Gigerenzer et al. (1999).

[22] In such a yoking, the simulation is tied to a particular subject. If the empirical subject opens N boxes in game g, the yoked simulation opens N boxes in game g.

[23] Subjects are Harvard University undergraduates. Of them, 55 percent report having taken at least one statistics course.

[24] Payoffs do not vary with subject demographics (e.g., statistical coursework). We adopt the simplifying approximation that all subjects have identical strategies.
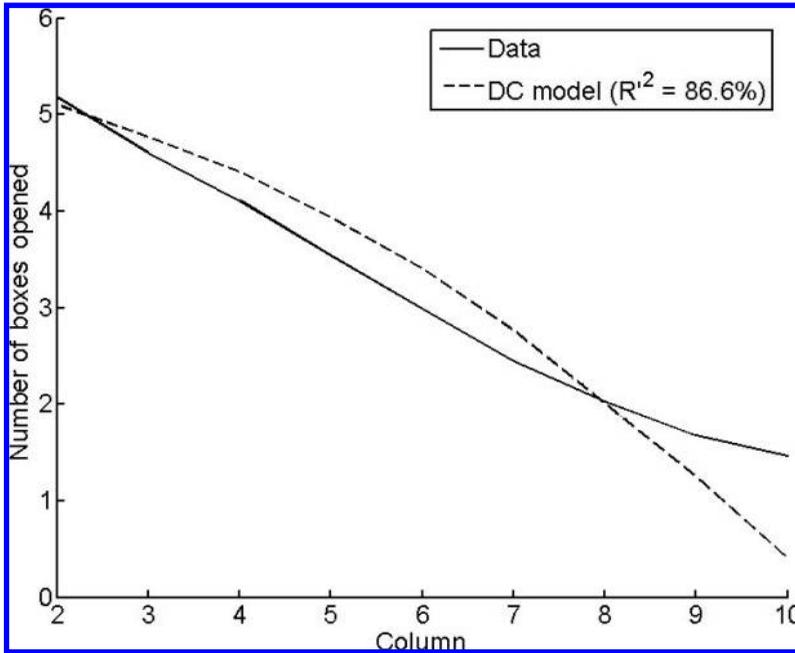
FIGURE 4. COLUMN PROFILES FOR GAMES WITH EXOGENOUS TIME BUDGETS

Notes: This figure plots the mean number of boxes opened in each column by subjects and by the DC model, for games in which time budgets are imposed exogenously. Dotted lines show the bootstrapped 95-percent confidence intervals for the data.

A. *Games with Exogenous Time Budgets*

We compare the empirical patterns of box openings to the patterns of box openings predicted by the directed cognition model. We begin with a trivial prediction of our theory: subjects should always open boxes from left to right, following a declining variance rule. In our experimental data, subjects follow the declining variance rule 91.0 percent of the time (s.e. 0.8 percent).[25] Specifically, when subjects open a previously unopened box in a given row, 91.0 percent of the time that box has the highest variance of the as-yet-unopened boxes in that row.[26] For reasons that we explain below, such left-to-right box openings may arise because of spatial biases instead of the information processing reasons implied by our theory.

Now we consider the pattern of search across columns and rows. Figure 4 reports the average number of boxes opened in columns 2–10. We report the average number of boxes unmasked, column by column, for both the subject data and the model predictions.

The empirical profile is calculated by averaging together subject responses on all the exogenous games that were played. Specifically, each of our 388 subjects played 12 exogenous games, yielding a total of $388 \times 12 = 4,656$ exogenous games played. Each subject was assigned a subset of 12 games from a set of 160 unique games. Hence, each of the 160 games was played about $4,656/160 \simeq 30$ times in the exogenous time portion of the experiment.

Figure 4 also plots the theoretical predictions generated by yoked simulations of our model. Specifically, these predictions are calculated by simulating the DC model on the exact set of 4,656 games played by the subjects. We simulate the model on each game from this set of 4,656 games and instruct the computer to un-
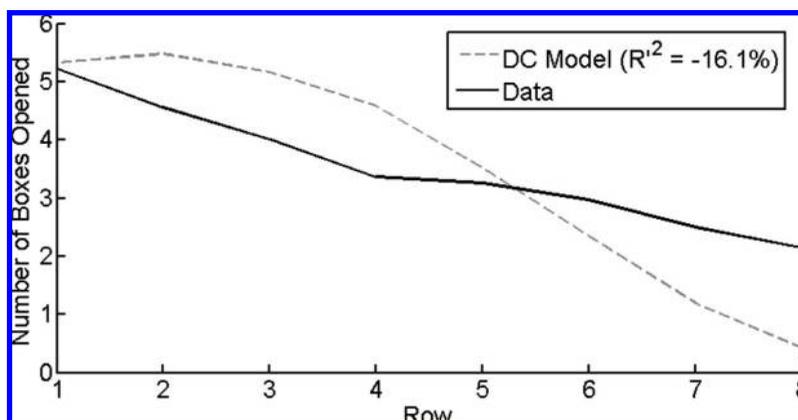
---

[25] The units throughout refer to percentage points, not to a percentage of the point estimate.

[26] In our endogenous games, subjects follow the declining variance rule 92.6 percent of the time (s.e. 0.7 percent).

FIGURE 5. ROW PROFILES FOR GAMES WITH EXOGENOUS TIME BUDGETS

Notes: This figure plots the mean number of boxes opened in each row by subjects and by the DC model for games in which time budgets are imposed exogenously. Dotted lines show the bootstrapped 95-percent confidence intervals for the data. The rows are ordered according to the values in the first column.

mask the same number of boxes that was unmasked by the subject who played each respective game.

The analysis compares the particular boxes opened by the subject to the particular boxes opened by the yoked simulation of the model. Figure 4 reports an R'^2 measure, which captures the extent to which the empirical data match the theoretical predictions. This measure is simply the R^2 statistic[27] from the following *constrained* regression:[28]

$$Boxes(col) = \text{constant} + \widehat{Boxes}(col) + \varepsilon(col).$$

Here $Boxes(col)$ represents the *empirical* average number of boxes unmasked in column *col*

and $\widehat{Boxes}(col)$ represents the *simulated* average number of boxes opened in column *col*. Note that *col* varies from 2 to 10, since the boxes in column 1 are always unmasked. This R'^2 statistic is bounded below by $-\infty$ (since the coefficient on $\widehat{Boxes}(col)$ is constrained equal to unity) and bounded above by one (a perfect fit). Intuitively, the R'^2 statistic represents the fraction of squared deviations around the mean explained by the model. For the column predictions, the R'^2 statistic is 86.6 percent (s.e. 1.7 percent), implying a very close match between the data and the predictions of the model.

Figure 5 reports analogous calculations by row and reports the number of boxes opened on average by row, with the rows ranked by their value in column one. We report the number of boxes opened on average by row for both the subject data and the model predictions. As above, the model predictions are calculated using yoked simulations.

Figure 5 also reports an R'^2 measure analogous to the one described above. The only difference is that now the variable of interest is $Boxes(row)$, the empirical average number of boxes opened in row *row*. For our row predictions our R'^2 measure is -16.1 percent (s.e. 9.2 percent), implying a poor match between the data and the predictions of the model. The data

---

[27] In other words,

$$R'^2 = 1$$

$$-\frac{\sum_{col}(Boxes(col) - \langle Boxes \rangle - \widehat{Boxes}(col) + \langle \widehat{Boxes} \rangle)^2}{\sum_{col}(\widehat{Boxes}(col) - \langle \widehat{Boxes} \rangle)^2},$$

where $\langle \cdot \rangle$ represents empirical means.

[28] In this section of the paper, the constant is redundant, since the dependent variable has the same mean as the independent variable. In the next subsection, however, we will consider cases in which this equivalence does not hold, necessitating the presence of the constant.

show that subjects unmask at least two to three boxes on average in each row. The model predicts an average, however, of between zero and one unmaskings in the rows that have the two lowest values in the first column (rows 7 and 8 in the figure). The subjects are less selective than the model. The R'^2 is negative because we constrain the coefficient on simulated boxes to be unity. This figure captures the most problematic predictions that the model makes.

Figure 6 reports similar calculations using an alternative way of ordering rows. It reports the number of boxes opened on average by row, with the rows ranked by their revealed values at the *end* of search in each game. In other words, we rank the rows based on the boxes that have been revealed during the course of information acquisition in each game. This ranking rule produces very different patterns than the ranking rule based on the column 1 values (i.e., in Figure 5).[29] For this alternative ranking, the model predictions have an R'^2 statistic of 86.7 percent (s.e. 1.4 percent).

We also report a performance measure that is based on individual-level data. We ask whether the model can predict the next box opening at the time of a row change. By conditioning on a row change, we eliminate left-to-right biases.[30] At the time of a row change, the DC algorithm correctly predicts the next box opening 37 percent of the time.

We also calculate the DC ranking of rows to which subjects switch. The rankings are normalized so that an index score of zero represents the worst DC ranking and an index score of one represents the best DC ranking.[31] We find that rows to which subjects switch have a mean DC rank index of 0.69 and a median rank index of 0.80, implying that the DC model places a relatively high search value on the rows to which subjects actually do switch.

Finally, we determine whether the DC algorithm successfully predicts final consumption choices using only ex ante information about each game (i.e., conditioning only on the information in the 80 boxes of each game and not using any of the subject choices). We simulate the DC algorithm on each game. We find that the final rows consumed by our subjects match the rows chosen by the DC algorithm 54 percent of the time in exogenous time games (a random algorithm would have a 13-percent success rate). Likewise, the rows picked by subjects match one of the top two rows chosen by the DC algorithm 73 percent of the time (25 percent for a random algorithm). Finally, the rows picked by subjects match one of the top three rows chosen by the DC algorithm 82 percent of the time (38 percent for a random algorithm).

B. *Games with Endogenous Time Budgets*

We repeat the analysis above for the endogenous games. As discussed in Section III, we consider two variants of the directed cognition model when analyzing the endogenous games. We calibrate one variant by exogenously setting κ to match the subjects' anticipated earnings per unit time in the endogenous games: $\kappa = 0.49$ cents/box opened (see calibration discussion in Section III). With this calibration, subjects are predicted to open 15.57 boxes per game (s.e. 0.01). In the data, however, subjects open 26.06 boxes per game (s.e. 0.57). To match this frequency of box opening, we consider a second calibration with $\kappa = 0.18$. With this lower level of κ, the model opens the empirically "right" number of boxes.

Figure 7 reports the average number of boxes

---

[29] Consider an illustrative example in which all subjects randomly pick one (and only one) row, and then explore the nine covered boxes in that row. Then the profile of box openings in Figure 5 would be flat and equal to 9/8 for every row (each row has a 1/8 chance of being chosen for 9 openings). By contrast, the profile in Figure 6 would be U-shaped with peaks slightly below 9/2 for both the best row and the worst row. Intuitively, the row that was randomly chosen for exploration is likely to end up being either the best or worst row when the ranking is made *including the nine boxes that are opened in the course of search;* opening up nine boxes increases variance! So in any particular game, the extreme rows have about a 1/2 chance of having nine box openings and about a 1/2 chance of having zero box openings, implying about 9/2 box openings on average for the row with the best ranking and the row with the worst ranking.

[30] When we do not condition on row changes, the model predicts the next box opening 14 percent of the time. This percentage is biased down by the left-to-right bias in behavioral data. Models based on cost-benefit analysis will not reproduce the left-to-right search pattern, unless the model assumes that left-to-right movements are less costly than other search operations.

[31] The formula for the rank index is $(rank - 1)/(N - 1)$, where N is the number of rows to which the subject could switch without implementing a repeat box opening and *rank* is the DC rank (1 to N) of the row to which the subject switches (higher ranks being better).
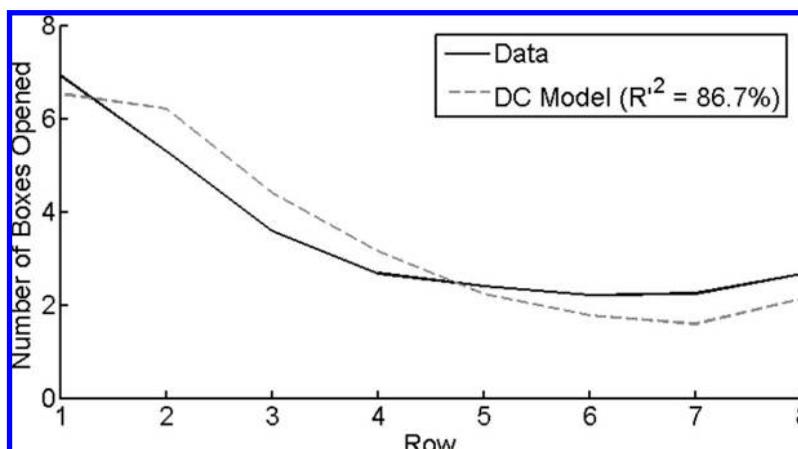
FIGURE 6. ROW PROFILES (ALTERNATIVE ORDERING) FOR GAMES
WITH EXOGENOUS TIME BUDGETS

Notes: This figure plots the mean number of boxes opened in each row by subjects and by the
DC model, for games in which time budgets are imposed exogenously. Dotted lines show the
bootstrapped 95-percent confidence intervals for the data. The rows are ordered according to
all the information available to subjects after they have concluded their search, just prior to
making a choice.

unmasked in columns 2 to 10 in the endogenous
games. We report the average number of boxes
unmasked by column for the subject data and
for the model predictions with $\kappa = 0.49$ and
$\kappa = 0.18$. To generate the theoretical predic-
tions, we use the directed cognition model to
simulate play of the 11,124 endogenous games
that the subjects actually played. The model
generates its own stopping rule (so we no longer
yoke the specific number of box openings). For
these endogenous games, the column R'^2 statistic
is 96.3 percent (s.e. 1.4 percent) for $\kappa = 0.49$
and 73.1 percent (s.e. 4.2 percent) for $\kappa = 0.18$.

Figure 8 reports the number of boxes opened
on average by row, with the rows ranked by
their values in column one. The row R'^2 statis-
tics are 85.3 percent (s.e. 1.3 percent) for $\kappa =
0.49$ and 64.6 percent (s.e. 2.4 percent) for $\kappa =
0.18$. Figure 9 reports similar calculations using
the alternative way of ordering rows, with the
rows ranked by their values at the *end* of each
game. The alternative row R'^2 statistics are 91.8
percent (s.e. 0.5 percent) for $\kappa = 0.49$ and 84.5
percent (s.e. 0.8 percent) for $\kappa = 0.18$.

These figures show that the model explains a
large fraction of the variation in attention across
rows and columns. Some of these results are
confounded, however, by a subject bias that
Costa-Gomes et al. (2001) have identified. In

particular, subjects who use the Mouselab inter-
face tend to have a bias toward selecting cells in
the upper-left corner of the screen and transi-
tioning from left to right as they explore the
screen. The left-right bias affects our column
results (Figures 4 and 7), since information with
greater economic relevance is located toward
the left-hand side of the screen in our
experiments.[32]

We also report the model's ability to predict
the next box opening at the time of a row
change. By conditioning on a row change, we
eliminate left-to-right biases. At the time of a
row change, the DC algorithm correctly predicts
the next box opening 38 percent of the time.

We also calculate the DC ranking of rows to
which subjects switch. The rankings are nor-
malized between zero (low) and one (high) as
explained above. We find that rows to which
subjects switch have a mean DC rank index of
0.70 and a median rank index of 0.83.

Finally, we determine whether the DC algo-
rithm successfully predicts final consumption

---

[32] The up-down bias does not affect our results, since our
rows are randomly ordered with respect to their respective
payoffs. Neither the up-down nor the left-right biases influ-
ence our analyses of either row openings (above) or endog-
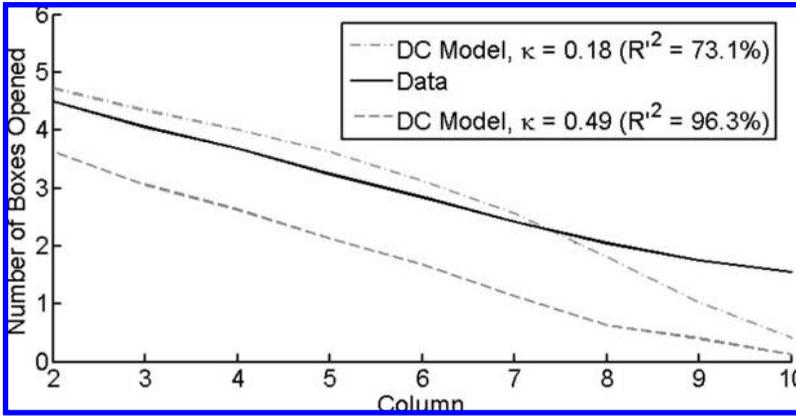enous stopping decisions (below).

FIGURE 7. COLUMN PROFILES FOR GAMES WITH ENDOGENOUS TIME BUDGETS

Notes: This figure plots the mean number of boxes opened in each column by subjects and by both calibrations of the DC model, for games in which agents choose how much time to allocate from a fixed time budget. Dotted lines show the bootstrapped 95-percent confidence intervals for the data.

FIGURE 8. ROW PROFILES FOR GAMES WITH ENDOGENOUS TIME BUDGETS

Notes: This figure plots the mean number of boxes opened in each row by subjects and by both calibrations of the DC model, for games in which agents choose how much time to allocate from a fixed time budget. Dotted lines show the bootstrapped 95-percent confidence intervals for the data. The rows are ordered according to the values in the initial column.
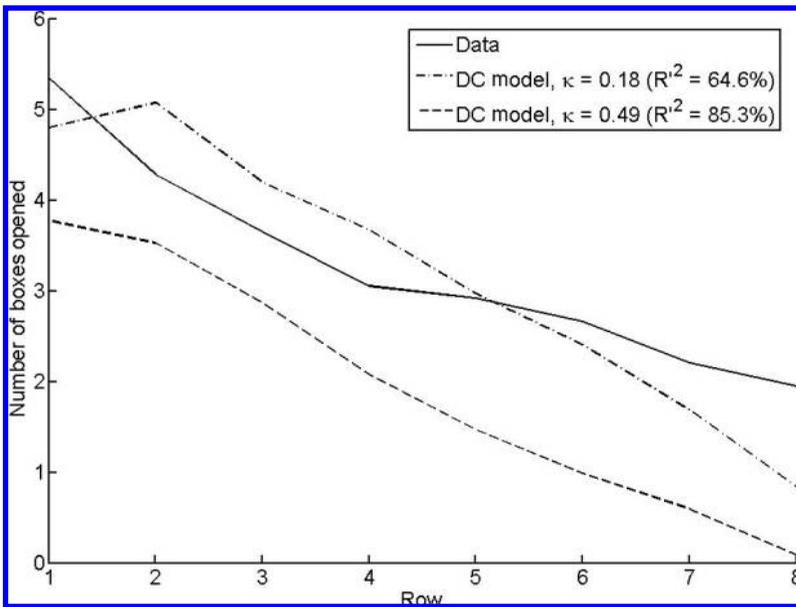
choices using only ex ante information. We simulate the DC algorithm with $\kappa = 0.18$. The final rows consumed by our subjects match the rows chosen by the DC algorithm 57 percent of the time. Likewise, the rows picked by subjects match one of the top two (three) rows chosen by the DC algorithm 76 percent (85 percent) of the time.
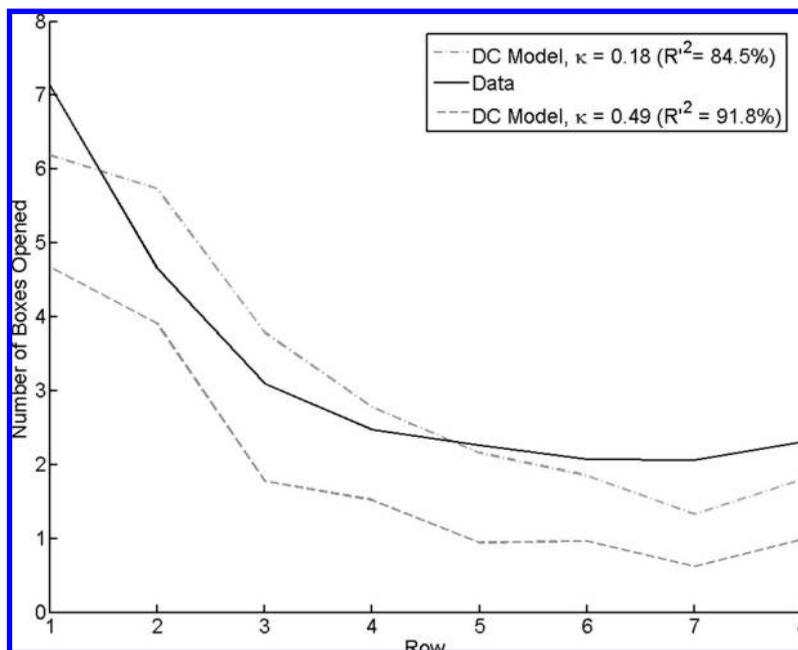
FIGURE 9. ROW PROFILES (ALTERNATIVE ORDERING) FOR GAMES
WITH ENDOGENOUS TIME BUDGETS

Notes: This figure plots the mean number of boxes opened in each row by subjects and by both
calibrations of the DC model, for games in which agents choose how much time to allocate
from a fixed time budget. Dotted lines show the bootstrapped 95-percent confidence intervals
for the data. The rows are ordered according to all the information available to subjects after
they have concluded their search, just prior to making a choice.

C. *Stopping Decisions in Endogenous Games*

Almost all of the analysis above reports *within-*
game variation in information acquisition. The
analysis above shows that subjects allocate most
of their attention to economically relevant col-
umns and rows within a game, matching the pat-
terns predicted by the directed cognition model.
Our experimental design also enables us to eval-
uate how subjects allocate search time *between*
games. In this subsection we focus on several
measures of such between-game variation.

Most importantly, we ask whether the model
can correctly predict which games received the
most attention from our subjects. Our experi-
ment utilized 160 unique games, though no sin-
gle subject played all 160 games. Let $Boxes(g)$
represent the average number of boxes opened
by subjects who played game g. Let $\widehat{Boxes}(g)$
represent the average number of boxes opened
by the model when playing game g. In this

subsection we analyze the first and second mo-
ments of the empirical sample $\{Boxes(g)\}_{g=1}^{160}$
and the simulated sample $\{\widehat{Boxes}(g)\}_{g=1}^{160}$. Note
that these respective vectors each have 160 el-
ements, since we are analyzing game-specific
averages.

We begin by comparing first moments. The
empirical (equally weighted) mean of $Boxes(g)$
is 26.06 (s.e. 0.57). By contrast, the 0-parameter
version of our model (with $\kappa = 0.49$) generates
a predicted mean of 15.57 (s.e. 0.01). Hence,
unless we pick κ to match the empirical mean
(i.e., $\kappa = 0.18$), our model only crudely approx-
imates the average number of boxes opened per
game.

We turn now to second moments. The empir-
ical standard deviation of $Boxes(g)$ is 6.32 (s.e.
0.14), while the 0-parameter version of our
model (with $\kappa = 0.49$) generates a predicted
standard deviation of 12.05 (s.e. 0.03). More-
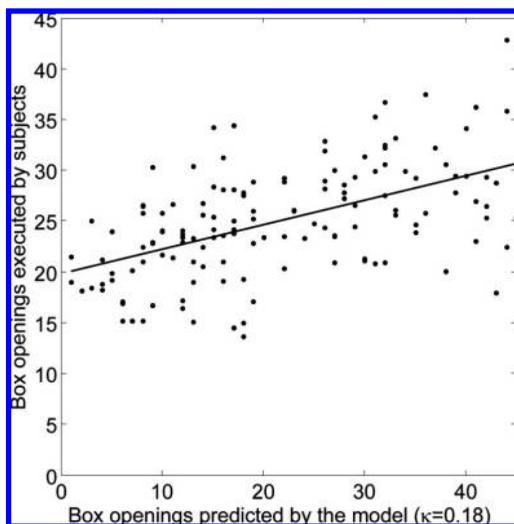over, when we set $\kappa = 0.18$ to match the aver-

FIGURE 10. BOXES OPENED BY GAME

Notes: This figure takes each of the 160 game-types played by subjects and compares the number of boxes opened by subjects to the number of box openings predicted by the model, for the games in which agents choose their cross-game time allocations. The correlation between the model and the data is 0.61.

age boxes per game, the standard deviation rises to 15.85 (s.e. 0.03). The relatively high standard deviations from the simulations reflect the model's sophisticated search strategy. The model continuously adjusts its search strategies in response to instantaneous variation in the economic incentives that the subjects face. By contrast, the subjects are less sensitive to high frequency variation in economic incentives.

Despite these shortcomings, the model successfully predicts the pattern of empirical variation in the number of boxes opened in each game. The correlation between $Boxes(g)$ and $\widehat{Boxes}(g)$ is 0.66 (s.e. 0.02) when we set $\kappa = 0.49$. Similarly, the correlation is 0.61 (s.e. 0.02) when we set $\kappa = 0.18$. See Figure 10 for a plot of the 160 data points for the $\kappa = 0.18$ case. These high correlations imply that the model does a good job predicting which games the subjects will analyze most thoroughly.

The model also predicts the statistical relationship between economic incentives and depth of analysis. Figure 11 reports the relationship between G—the myopic estimate of the

marginal benefit of incremental search[33]—and the average amount of time the subject (or algorithm) continues to play the game. The relationship is calculated (using a nonparametric kernel estimator) for both simulated data and the experimental data. To construct the simulated relationship, we use the model to "play" each of our games and then use the simulated move-by-move data to estimate the relationship between G and the amount of simulated time remaining. To construct the experimental relationship, we use the experimental move-by-move data to estimate the relationship between G and the empirical amount of time remaining.

The dashed line represents the relationship estimated with the simulated data (with $\kappa = 0.18$). The solid line represents the relationship estimated with the experimental data. The figure also shows bootstrap estimates of the 95-percent confidence intervals. For most levels of G (the benefit-cost ratio), the model's predictions are close to the pattern in the subject data. Subjects do more analysis (i.e., open up more boxes) when the economic incentives to do so are high. Moreover, the functional form of this relationship roughly matches the form predicted by the theory.

We also evaluate the DC model by asking whether G predicts when subjects decide to stop working on the current game and move on to the next game. We run a stopping logit (1 = stop, 0 = continue) with explanatory variables that include the measure of the economic value of continued search (or G), the number of different boxes opened to date in the current game (*boxes*), the expected value in cents of the leading row in the current N-good game (*leader*), and subject fixed effects. Note that satisficing models predict that G and *boxes* should not have predictive power in this logit regression, but that a higher value of the leader variable will increase the stopping probability.

Each observation for this logit is a decision-node (i.e., a choice over mouse clicks) in our

---

[33] See equation (6): $G \equiv \max_O w(x_O, \sigma_O)/(\kappa \Gamma_O)$. Recall that x_O is the gap between the expected value of the current row and the expected value of the next best row, σ_O is the standard deviation of the information revealed by mental operator O, and Γ_O is the number of boxes opened by mental operator O.
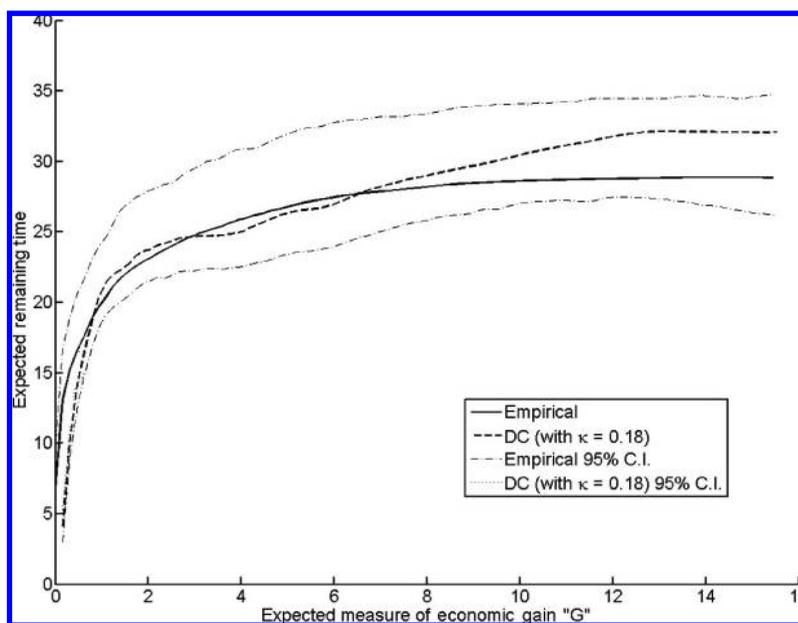
FIGURE 11. EXPECTED MEASURE OF ECONOMIC GAIN "*G*" VERSUS EXPECTED REMAINING TIME

Notes: This figure plots nonparametric (kernel) estimates of the expected number of additional box openings in a game, conditional on the current "*G*" value. "*G*" is the benefit-to-cost ratio of marginal analysis in the DC model, as defined in equation (6).

endogenous games, generating 330,873 observations.[34] The coefficient on G is -0.3660 (standard error 0.0081 and t-stat of 45); the coefficient on *boxes* is 0.0404 (standard error 0.0008 and t-stat of 50); the coefficient on *leader* is 0.0064 (standard error 0.0002 and t-stat of 32). At the mean values of the explanatory variables, a one-standard-deviation reduction in the value of G more than doubles the probability of stopping. A one-standard-deviation increase in the value of *boxes* has an effect roughly ½ as large and a one-standard-deviation increase in the value of *leader* has an effect less than ¼ as large. Although *boxes* has the greatest t-stat, G is the most important predictor of the decision to stop searching (variation in G has a bigger impact than variation in *boxes*). Incremental logit regressions

also confirm that G has more predictive power than *boxes*.[35]

The logistic analysis shows that the economic value of information—G—is the most important predictor of the decision to stop searching. Our subjects are also using other information, however, as shown most importantly by the strong predictive power of the *boxes* variable. If subjects place partial weight on the "*boxes* heuristic"—i.e., increasing the propensity to stop analyzing the current *N*-good game as more and more boxes are opened—they will be less likely to spend a long time on any one game. For an unsophisticated player who can only imperfectly calculate G, the *boxes* heuristic is a useful additional decision input. We view our subjects' partial reli-

[34] In other terms, we estimate: probability of continuation $= \exp(\beta'x)/[1 + \exp(\beta'x)]$, where x is a vector of decision-node attributes, and β is the vector of estimated coefficients.

[35] When a stopping logit is estimated with only a constant, $\ln L = -48{,}026$, where L is the likelihood function. When a stopping logit is estimated with a constant and *boxes,* $\ln L = -47{,}403$. When a stopping logit is estimated with a constant and G, $\ln L = -45{,}353$. When a stopping logit is estimated with a constant, *boxes* and G, $\ln L = -45{,}264$. Hence, including G in the model increases $\ln L$ four times more than including *boxes*.

TABLE 8—EVALUATION OF ALTERNATIVE MODELS

	Games with exogenous time budgets			Games with endogenous time budgets					
	Directed cognition	Column model	Row model	**Directed cognition**	**Directed cognition**	Column model satisficing	Row model satisficing	Directed cognition satisficing	Elimination by aspects
Fitted parameter	**None**	None	None	**None**	$\kappa = 0.18$	$A^{\text{Column}} = 42$	$A^{\text{Row}} = 43$	$A^{\text{DC}} = 45$	$A^{\text{EBA}} = -6$
R'^2 for column profile	**86.6%**	−107.7%	39.3%	**96.3%**	**73.1%**	79.7%	34.5%	75.3%	−70.4%
	(1.7%)	(18.7%)	(0.93%)	(1.4%)	(4.2%)	(1.8%)	(1.0%)	(1.5%)	(18.4%)
R'^2 for row profile	**−16.1%**	26.2%	−540.6%	**85.3%**	**64.6%**	25.1%	88.6%	65.6%	−31.1%
	(9.2%)	(0.55%)	(0.40%)	(1.3%)	(2.4%)	(0.44%)	(0.69%)	(1.0%)	(0.66%)
R'^2 for alt. row profile	**86.7%**	25.0%	89.1%	**91.8%**	**84.5%**	23.8%	43.0%	53.3%	82.5%
	(1.4%)	(0.42%)	(1.5%)	(0.55%)	(0.83%)	(0.39%)	(0.59%)	(0.77%)	(0.79%)
Average R'^2	**52.4%**	−18.8%	−137.4%	**91.2%**	**74.1%**	42.9%	55.4%	64.7%	−6.3%
	(3.4%)	(6.4%)	(13.9%)	(0.67%)	(1.7%)	(0.73%)	(0.41%)	(0.62%)	(6.2%)
Correlation with empirical number of boxes	—	—	—	**0.66**	**0.61**	−0.005	0.020	0.026	−0.39
	—	—	—	(0.19)	(0.018)	(0.024)	(0.023)	(0.024)	(0.021)
Agreement with final choices	**53.8%**	51.2%	52.9%	**55.7%**	**57.4%**	50.7%	47.1%	49.0%	47.1%
	(0.73%)	(0.73%)	(0.73%)	(0.47%)	(0.47%)	(0.47%)	(0.47%)	(0.47%)	(0.47%)

Notes: Standard errors in parentheses. R'^2 refers to the R^2 statistic from a regression of the empirical number of boxes opened by subjects in each row or column on a constant, plus the number of boxes predicted by the models, where the coefficient on the predictions is fixed to equal 1. "Correlation with empirical number of boxes" refers to the correlation between the number of boxes opened by subjects in each of the 160 games and the number of openings predicted by the models. "Agreement with final choices" records the percentage of the games played by subjects in which their final choices match the predictions of each model, when the model is simulated with no knowledge of subject box openings.

ance on *boxes* as an example of sensible—perhaps constrained optimal—decision-making.

The predictive power of the *boxes* variables supports experimental research on "system neglect" by Camerer and Dan Lovallo (1999) and Cade Massey and George Wu (2005). These authors find that subjects use sensible rules, but fail to adjust those rules adequately to the particular problem at hand. The *boxes* heuristic is a good general rule, but it is not the first-best rule since it neglects the idiosyncratic incentives generated by each specific *N*-good game. The *boxes* heuristic is a type of imperfectly sophisticated search. The *boxes* heuristic is also consistent with findings in the sequential search literature that subjects are often overly sensitive to the total search costs paid (Carl Kogut, 1990; Philip Moon and Andrew Martin, 1990; Joep Sonnemans, 1998).

Finally, our experiment reveals another bias that we have not emphasized because we are uncertain about its generalizability to other classes of games. Specifically, our subjects allocated too much time per game in the endogenous games. Subject payoffs would have been higher if they had generally collected less information in each game, thereby enabling them to

play more games. This result is contrary to the sequential search literature, in which subjects often search less than predicted (Kogut, 1990; Sonnemans, 1998). Exploring the robustness of this finding is a goal for future research.

D. *Comparisons with Other Models*

Table 8 reports R'^2 measures for the alternative models summarized in Section IIIG. For the exogenous games, the DC model has an average R'^2 value of 52.4 percent (with a standard deviation of 3.4 percentage points). The column and row models have respective averages of −18.8 percent (s.e. 6.4 percent) and −137.4 percent (s.e. 13.9 percent). For the endogenous games, the zero-parameter version of the DC model has an average R'^2 of 91.2 percent (s.e. 0.7 percent). The column model with a satisficing stopping rule has an average R'^2 of 42.9 percent (s.e. 0.7 percent). The row model with a satisficing stopping rule has an average R'^2 of 55.4 percent (s.e. 0.4 percent). The elimination by aspects model has an average R'^2 of −6.3 percent (s.e. 6.2 percent).

We also evaluate the different models' ability to forecast game-by-game variation in average

time allocation.[36] Table 8 shows that all of the satisficing models yield effectively no correlation between the game-by-game average number of box openings predicted by these models and the game-by-game average number of box openings in the data.[37] By contrast, the zero-parameter directed cognition model generates predictions that are highly correlated (0.66) with the empirical average boxes opened per game.

V. Conclusion

Economists assume that agents act as if they were choosing cognitive operations optimally. We have experimentally analyzed a parsimonious cognition model based on partially myopic cost-benefit calculations: the DC model. The partial myopia makes the model computationally tractable, and our experimental results demonstrate that the partial myopia improves the model's empirical fit.

Our first experiment studies a class of simple decision problems. These problems admit a rational search solution, do not pose spatial biases, and do not require subject memorization. In these simple problems, we also assume an exogenous financial search cost. We find that the DC model successfully explains the patterns of information acquisition. Indeed, when the DC model and the fully rational model make different predictions, the DC model does a *better* job of matching the laboratory evidence.

Our second set of experiments studies a highly complex choice problem. The DC model can be used to study such complex settings because the model simplifies analysis (for the subject and the researcher). In contrast, the rational actor model does not admit a computa-

tionally tractable solution in such problems. In these experiments, subjects analyzed eight goods, each of which had nine hidden attributes that could be revealed with some effort, though without financial cost. Subjects decided which attributes to reveal before choosing to consume one of the eight goods. Subjects faced endogenous cognition costs because they were under time pressure. We created this time pressure in two ways: first, by giving subjects a fixed amount of time for each game, and, second, by giving subjects a fixed amount of time for a set of games. The DC model successfully predicts the aggregate empirical regularities of information acquisition in this experiment, including the pattern of information acquisition within each game and the pattern of information acquisition across games. This second set of experiments illustrated the applicability of the DC problem to highly complex settings in which the classical rational actor model is analytically and computationally intractable.

Understanding how decision makers think about problems is a relatively new frontier in economic science. Stripping back decision-making to the level of cognitive processes poses fundamental research challenges. In this paper we have studied the information acquisition process. This is one example of a wider set of attention allocation processes. Actors allocate attention to acquire new information and to analyze already available information (e.g., Christopher A. Sims, 2003).

The study of cognitive processes in economic decision-making is still in its infancy. But we are optimistic that such process-based research may pave the way for a science of decision-making with much greater predictive power than the classical "as if" modeling that treats cognition as a closed black box. This paper joins a growing body of work that pries open that box (e.g., Camerer et al., 1993; Costa-Gomes et al., 2001). We look forward to future process-based research that will open the box more completely.

---

[36] Recall that this analysis uses the average number of boxes opened per game for the 160 unique games in the dataset. These data are compiled from the endogenous time segment of the experiment.

[37] We do, however, find some support for a satisficing framework. A higher payoff in the current decision problem leads to a slightly higher intensity of search in subsequent problems. For example, a $10 higher realized payoff in the current game leads to 0.16 additional box openings in the subsequent game (controlling for subject fixed effects). A satisficing model (with ongoing updates of the satisficing threshold) would predict a relationship with this sign.

REFERENCES

Aragones, Enriqueta; Gilboa, Itzhak; Postlewaite, Andrew and Schmeidler, David. "Fact-Free Learning." *American Economic Review*, 2005, *95*(5), pp. 1355–68.

Bertsekas, Dimitri P. and Tsitsiklis, John N. *Neuro-dynamic programming*. Belmont, MA: Athena Scientific, 1996.

Camerer, Colin and Ho, Teck-Hua. "Experience-Weighted Attraction Learning in Normal Form Games." *Econometrica*, 1999, *67*(4), pp. 827–74.

Camerer, Colin F.; Johnson, Eric; Rymon, Talia and Sen, Senkar. "Cognition and Framing in Sequential Bargaining for Gains and Losses," in Ken Binmore, Alan Kirman, and Piero Tani, eds., *Frontiers of game theory*. Cambridge, MA: MIT Press, 1993, pp. 27–47.

Camerer, Colin F. and Lovallo, Dan. "Overconfidence and Excess Entry: An Experimental Approach." *American Economic Review*, 1999, *89*(1), pp. 306–18.

Conlisk, John. "Why Bounded Rationality?" *Journal of Economic Literature*, 1996, *34*(2), pp. 669–700.

Costa-Gomes, Miguel A. and Crawford, Vincent P. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study." *American Economic Review* (forthcoming).

Costa-Gomes, Miguel A.; Crawford, Vincent P. and Broseta, Bruno. "Cognition and Behavior in Normal-Form Games: An Experimental Study." *Econometrica*, 2001, *69*(5), pp. 1193–1235.

Erev, Ido and Roth, Alvin E. "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria." *American Economic Review*, 1998, *88*(4), pp. 848–81.

Gabaix, Xavier and Laibson, David. "Bounded Rationality and Directed Cognition." Unpublished Paper, 2005.

Gabaix, Xavier and Laibson, David. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics*, 2006, *121*(2), pp. 505–40.

Gigerenzer, Gerd; Todd, Peter and ABC Research Group. *Simple heuristics that make us smart*. Oxford: Oxford University Press, 1999.

Gilovich, Thomas; Griffin, Dale and Kahneman, Daniel. *Heuristics and biases: The psychology of intuitive judgment*. Cambridge: Cambridge University Press, 2002.

Gittins, John C. "Bandit Processes and Dynamic Allocation Indices." *Journal of the Royal Statistical Society, Series B*, 1979, *41*(2), pp. 148–77.

Jehiel, Philippe. "Limited Horizon Forecast in Repeated Alternate Games." *Journal of Economic Theory*, 1995, *67*(2), pp. 497–519.

Johnson, Eric J.; Camerer, Colin; Sen, Sankar and Rymon, Talia. "Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining." *Journal of Economic Theory*, 2002, *104*(1), pp. 16–47.

Kogut, Carl A. "Consumer Search Behavior and Sunk Costs." *Journal of Economic Behavior and Organization*, 1990, *14*(3), pp. 381–92.

Lipman, Barton L. "How to Decide How to Decide How To … : Modeling Limited Rationality." *Econometrica*, 1991, *59*(4), pp. 1105–25.

Massey, Cade and Wu, George. "Detecting Regime Shifts: The Psychology of Under- and Over-Reaction." *Management Science*, 2005, *51*(6), pp. 932–47.

Moon, Philip and Martin, Andrew. "Better Heuristics for Economic Search: Experimental and Simulation Evidence." *Journal of Behavioral Decision Making*, 1990, *3*(3), pp. 175–93.

Payne, John W.; Bettman, James R. and Johnson, Eric J. *The adaptive decision maker*. Cambridge: Cambridge University Press, 1993.

Payne, John W.; Braunstein, Myron L. and Carroll, John S. "Exploring Predecisional Behavior: An Alternative Approach to Decision Research." *Organizational Behavior and Human Performance*, 1978, *22*(1), pp. 17–44.

Russo, J. Edward. "Eye Fixations Can Save the World: A Critical Evaluation and a Comparison between Eye Fixations and Other Information Processing Methodologies," in H. Keith Hunt, ed., *Advances in consumer research*. Vol. 5. Ann Arbor: Association for Consumer Research, 1978, pp. 561–70.

Simon, Herbert A. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics*, 1955, *69*(1), pp. 99–118.

Sims, Christopher A. "Implications of Rational Inattention." *Journal of Monetary Economics*, 2003, *50*(3), pp. 665–90.

Sonnemans, Joep. "Strategies of Search." *Journal of Economic Behavior and Organization*, 1998, *35*(3), pp. 309–32.

Tversky, Amos. "Elimination by Aspects: A Theory of Choice." *Psychological Review*, 1972, *79*(4), pp. 281–99.

Tversky, Amos and Kahneman, Daniel. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty*, 1992, *5*(4), pp. 297–323.

Wegner, Daniel M. *The illusion of conscious will*. Cambridge, MA: MIT Press, 2002.

Weitzman, Martin L. "Optimal Search for the Best Alternative." *Econometrica*, 1979, *47*(3), pp. 641–54.