MPRA
Munich Personal RePEc Archive

The allocation of attention: theory and evidence

Gabaix, Xavier and Laibson, David Isaac and Moloche, Guillermo and Stephen, Weinberg

Massachussets Institute of Technology, Harvard University, National Bureau of Economic Research

29 August 2003

Online at https://mpra.ub.uni-muenchen.de/47339/
MPRA Paper No. 47339, posted 01 Jun 2013 23:07 UTC

THE ALLOCATION OF ATTENTION: THEORY AND EVIDENCE

Xavier Gabaix*
MIT and NBER

David Laibson
Harvard University and NBER

Guillermo Moloche
NBER

Stephen Weinberg
Harvard University

Current Draft: August 29, 2003

Abstract. A host of recent studies show that attention allocation has important economic consequences. This paper reports the first empirical test of a cost-benefit model of the endogenous allocation of attention. The model assumes that economic agents have finite mental processing speeds and cannot analyze all of the elements in complex problems. The model makes tractable predictions about attention allocation, despite the high level of complexity in our environment. The model successfully predicts the key empirical regularities of attention allocation measured in a choice experiment. In the experiment, decision time is a scarce resource and attention allocation is continuously measured using Mouselab. Subject choices correspond well to the quantitative predictions of the model, which are derived from cost-benefit and option-value principles.

JEL classification: C7, C9, D8.

Keywords: attention, satisficing, Mouselab.

*Email addresses: xgabaix@mit.edu, dlaibson@arrow.fas.harvard.edu, gmoloche@mit.edu, sweinber@kuznets.fas.harvard.edu. We thank Colin Camerer, David Cooper, Miguel Costa-Gomes, Vince Crawford, Andrei Shleifer, Marty Weitzman, three anonymous referees, and seminar participants at Caltech, Harvard, MIT, Stanford, University of California Berkeley, UCLA, University of Montréal, and the Econometric Society. Numerous research assistants helped run the experiment. We owe a particular debt to Rebecca Thornton and Natalia Tsvetkova. We acknowledge financial support from the NSF (Gabaix and Laibson SES-0099025; Weinberg NSF Graduate Research Fellowship) and the Russell Sage Foundation (Gabaix).

1. Introduction

1.1. Attention as a scarce economic resource. Standard economic models conveniently assume that cognition is instantaneous and costless. The current paper develops and tests models that are based on the fact that cognition takes time. Like players in a chess tournament or students taking a test, real decision-makers need time to solve problems and often encounter trade-offs because they work under time pressure. Time almost always has a positive shadow value, whether or not formal time limits exist.

Because time is scarce, decision-makers ignore some elements in decision problems while selectively thinking about others.[1] Such selective thinking exemplifies the *voluntary* (i.e. conscious) and *involuntary* processes that lead to attention allocation.[2]

Economics is often defined as the study of the allocation of scarce resources. The process of attention allocation, with its consequences for decision-making, seems like a natural topic for economic research. Indeed, in recent years many authors have argued that the scarcity of attention has large effects on economic choices. For example, Lynch (1996), Gabaix and Laibson (2001), and Rogers (2001) study the effects of limited attention on consumption dynamics and the equity premium puzzle. Mankiw and Reis (2002), Gumbau (2003), Sims (2003), and Woodford (2002) study the effects on monetary transmission mechanisms. D'Avolio and Nierenberg (2002), Della Vigna and Pollet (2003), Peng and Xiong (2003), and Pollet (2003) study the effects on asset pricing. Finally, Daniel, Hirshleifer, and Teoh (2002) and Hirshleifer, Lim, and Teoh (2003) study the effects on corporate finance, especially corporate disclosure. Some of these papers analyze limited attention from a theoretical perspective. Others analyze limited attention indirectly, by studying its associated market consequences. None of these papers measures attention directly.

The current paper is part of a new literature in experimental economics that empirically studies

[1] Simon (1955) first analyzed decision algorithms — e.g. satisficing — that are based on the idea that calculation is time-consuming or costly. See Kahneman and Tversky (1974) and Gigerenzer *et al.* (1999) for heuristics that simplify problem-solving. See Conlisk (1996) for a framework in which to understand models of thinking costs.

[2] Attention "allocation" can be divided into four categories: involuntary perception (e.g., hearing your name spoken across a noisy room at a cocktail party, see Moray 1959 and Wood and Cowan 1995); involuntary central cognitive operations (e.g. trying to not think about white bears, but doing so anyway, see Wegner 1989); voluntary perception (e.g. searching for a cereal brand on a display shelf, see Yantis 1998); and voluntary central cognitive operations (e.g. planning and control of action, see Kahneman 1973, and Pashler and Johnston 1998). See Pashler (1998) for an excellent overview of research on attention. The model that we analyze in this paper can be applied to all of these categories of attention allocation, but we discuss an application that focuses on conscious attention allocation (the last two categories).

the decision-making process by *directly* measuring attention allocation (Camerer *et al.* 1993, Costa-Gomes, Crawford, and Broseta 2001, Johnson et al 2002, and Costa-Gomes and Crawford 2003). In the current paper, we report the first direct empirical test of an economic cost-benefit model of endogenous attention allocation.

This economic approach contrasts with more descriptive models of attention allocation in the psychology literature (e.g. Payne, Bettman, and Johnson 1993). The economic approach provides a general framework for deriving quantitative attention allocation predictions without introducing extraneous parameters.

1.2. Directed cognition: An economic approach to attention allocation. To implement the economic approach in a computationally tractable way, we apply a cost-benefit model developed by Gabaix and Laibson (2002), which they call the directed cognition model. In this model, time-pressured agents allocate their thinking time according to simple option-value calculations. The agents in this model analyze information that is likely to be useful and ignore information that is likely to be redundant. For example, it is not optimal to continue to analyze a good that is almost surely dominated by an alternative good. It is also not optimal to continue to investigate a good about which you already have nearly perfect information. The directed cognition model calculates the economic option-value of marginal thinking and directs cognition to the mental activity with the highest expected shadow value.

The directed cognition model comes with the advantage of tractability. The directed cognition model *can* be computationally solved even in highly complex settings. To gain this tractability, however, the directed cognition model introduces some *partial* myopia into the option-theoretic calculations, and can thus only approximate the shadow value of marginal thinking.

In this paper we show that this approximation comes at relatively little cost, since the partially myopic calculations in the directed cognition model generate attention allocation choices that approximate the payoffs from a perfectly rational — i.e. infinitely forward-looking — calculation of option values. Because infinite-horizon option value calculations are not generally computationally tractable and because directed cognition is close to those calculations anyway (cf. Appendix C), we focus our analysis on the directed cognition model.

1.3. Perfect rationality: an intractable model in generic high-dimensional problems. Perfect rationality represents a formal (theoretical) solution to any well-posed problem. But perfectly rational solutions cannot be calculated in complex environments and have limited value as modeling benchmarks in such environments. Researchers in artificial intelligence have acknowledged this limitation and taken the extreme step of abandoning perfect rationality as an organizing principle. Among artificial intelligence researchers, the merit of an algorithm is based on its tractability, predictive accuracy, and generality — not by proximity to the predictions of perfect rationality.

In contrast to artificial intelligence researchers, economists use perfect rationality as their primary modeling approach. Perfect rationality continues to be the most useful and disciplined modeling framework for economic analysis. But alternatives to perfect rationality should be active topics of economic study *either* when perfect rationality is a bad predictive model *or* when perfect rationality is not computationally tractable. The latter case applies to our attention allocation analysis.

The complex environment that we study in this paper only partially reflects the even greater complexity of most attention allocation problems. But even in our relatively simple setting perfectly rational option value calculations are not computationally tractable.[3] Hence we focus our analysis on a solvable cost-benefit model — directed cognition — which can easily be applied to a wide range of realistic attention allocation problems. For our application, the merit of the directed cognition model is its tractability, *not* any predictive deviations from the perfectly rational model of attention allocation. We do not study such deviations due to the computational intractability of the perfectly rational model.

1.4. An empirical analysis of the directed cognition model. This paper compares the qualitative and *quantitative* predictions of the directed cognition model to the results of an experiment in which subjects solve a generalized choice problem that represents a very common set of economic decisions: choose one good from a set of N goods. Many economic decisions are special cases of this general problem.

[3] Of course, even when perfect rationality cannot be solved, its solution might still be *partially* characterized and tested. We do not take this approach in the current paper because the predictions of the infinite-horizon model that we have been able to formally derive have been quite weak (e.g., ceteris paribus, analyze information with the highest variance).

In our experiment, decision time is a scarce resource and attention allocation is measured continuously. We make time scarce in two different ways. In one part of the experiment 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 think about each good's attributes before making a final selection of which good to consume.

In a second part of the experiment 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 chooses for 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. As a result, the subjects trade off quality and quantity. Spending more time on any given choice problem raises the quality of the final selection in that problem but reduces the time available to participate in future choice problems with accumulating payoffs.

Throughout the experiment we measure the process of attention allocation. We measure attention allocation within each choice problem (i.e. how is time allocated in thinking about the different goods within a single choice problem before a final selection is made in that choice problem?). In addition, in the open-ended sequence design, we measure attention allocation across choice problems (i.e. how much time does the subject spend on one choice problem before making a final selection in that problem and moving on to the next choice problem?).

Following the lead of other economists (Camerer *et al.* 1993 and Costa-Gomes, Crawford, and Broseta 2001), we use the "Mouselab" programming language to measure subjects' attention allocation.[4] Mouselab tracks subjects' information search during our experiment. 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 contemplating on a second-by-second basis throughout the experiment.[5]

[4]Payne, Bettman, and Johnson (1993) developed the Mouselab language in the 1970's. Mouselab is one of many "process tracing" methods. For example, Payne, Braunstein, and Carrol (1978) elicit mental processes by asking subjects to "think aloud." Russo (1978) records eye movements.

[5]Mouselab has the drawback that it uses an artificial decision environment, but several studies have shown that

In our experiment, subject behavior corresponds well to the predictions of the directed cognition model. Subjects allocate thinking time when the marginal value of such thinking is high, either because competing goods are close in value or because there is a "large" amount of remaining information to be revealed. We demonstrate this in numerous ways. First, we evaluate the pattern of information acquisition within choice problems. Second, in the endogenous time treatment, we evaluate the relationship between economic incentives and subjects' stopping decisions: i.e. when do subjects stop working on one choice problem so they can move on to the next choice problem? We find that the economic model of attention allocation outperforms mechanistic information acquisition models. The marginal value of information turns out to be the most important predictor of attention allocation.

However, the experiment also reveals one robust deviation from the predictions of the directed cognition cost-benefit model. Subject choices are partially predicted by a "boxes heuristic." Specifically, subjects become more and more likely to end analysis of a problem the more boxes they open, holding *fixed* the economic benefits of additional analysis. In this sense, subjects display a partial insensitivity to the particular economic incentives in each problem that they face.

We also test models of heuristic decision-making, but we find little evidence for commonly studied heuristics, including satisficing and elimination by aspects. Instead, our subjects generally allocate their attention consistently with cost-benefit considerations, matching the precise *quantitative* attention-allocation patterns generated by the directed cognition model. Hence, this paper demonstrates that an economic cost-benefit model predicts both the qualitative and quantitative patterns of attention allocation.

Section 2 describes our experimental set-up. Section 3 summarizes an implementable one-parameter attention allocation model (Gabaix and Laibson 2002). Section 4 summarizes the results of our experiment and compares those results to the predictions of our model. Section 5 concludes.

---

the Mouselab environment only minimally distorts final choices over goods/actions (e.g., Costa-Gomes, Crawford, and Broseta 2001 and Costa-Gomes and Crawford 2002). Mouselab's interface does generate "upper-left" and "left-to-right" search biases, which we discuss in subsection (4.4) below.

2. Experimental Set-up

Our experimental design facilitates second-by-second measurement of attention allocation in a basic decision problem, a choice among N goods.

2.1. An N-good game.

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 of a very wide class of choice problems. We call this problem an N-good game, since the N rows conceptually represent N goods. The subject chooses to consume one of these N goods. The columns represent M different attributes.

For example, consider a shopper who has decided to go to Walmart to select and buy a television. The consumer faces a fixed number of television sets at Walmart (N different TV's from which to choose). The television sets have M different attributes — e.g. 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 1000 (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. Columns on the right represent the attributes with the least (utility-metric) variance, like minor clauses in the warranty.[6]

So far our game sounds simple: "Consume the best good (i.e., row)." To measure attention, we make the task harder by masking the contents of boxes in columns 2 through M. Subjects are shown only the box values in column 1.[7] However, a subject can left-click on a masked box in columns 2 through M to unmask the value of that box (Figure 2).

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

[7] We reveal the value of column one because it helps subjects remember which row is which. In addition, revealing column one 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).

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 analyzing at every point in time.[8] Revealing 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 enables us to record which of the $N(M-1)$ masked boxes the subject chooses to unmask. Of course, we also record which row the subject chooses/consumes for his or her final selection.

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

Masked boxes and time pressure capture important aspects of our Walmart shopper's experience. The Walmart shopper selectively attends to information about the attributes of the TV's 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 which she can do in the store if she selects her TV quickly. We explore both of these types of cases in our experiment.

2.2. 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 Walmart 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

[8] When we designed the experiment, we considered but did not adopt a design that permanently keeps boxes open once they had 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 then subjects have the option to quickly — and mechanically — open many boxes and only afterwards analyze their content. Hence, leaving boxes open implied that we lost the ability credibly to infer what the subject is attending to at each point in time.

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 Walmart 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 attention allocation 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, the between-problem attention allocation decision.

2.3. 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.[9] 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.

3. Directed cognition model

The previous section describes an attention allocation problem. In exogenous games, subjects must decide which boxes to unmask before their time runs out. In endogenous games, subjects must jointly decide which boxes to unmask *and* when to move on to the next game.

We want to determine whether economic principles guide subjects' attention allocation choices. We compare our experimental data to the predictions of an attention allocation model proposed by Gabaix and Laibson (2002). This 'directed cognition' model approximates the economic value of attention allocation using simple option value analysis.

[9] Interested readers can view all of our instructions on the web: http://post.economics.harvard.edu/faculty/laibson/papers.html

The option value calculations capture two types of effects. First, when many different choices are being compared and a particular choice gains a large edge over the available alternatives, the option value of continued analysis declines. Second, when marginal analysis yields little new information (i.e. the standard deviation of marginal information is low), the economic value of continued analysis declines. The directed cognition model captures these two effects with a formal search-theoretic framework that makes sharp quantitative predictions about attention allocation choices.

3.1. Summary of the model. The model can be broken down into three iterative steps, which we first summarize and then describe in detail.

Step 1: Calculate the expected economic benefits and costs of different potential cognitive operations. For example, what is the expected economic benefit of unmasking three boxes in the first row?

Step 2: Execute the cognitive 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 one unless time has run out (in exogenous time games) or until the ratio of expected benefit to cost falls below some threshold value (endogenous time games).

We now describe the details of this search algorithm. We first introduce notation and then formally derive an equation to calculate the expected economic benefits referred to in step one.

3.2. Notation. Since our games all have eight rows (goods), we label the rows A, $B, ..., H$. We will use lower case letters — $a, b, ..., h$ — to track a subject's expectations of the values of the respective rows.

Recall that 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 left-most 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 $29 = 23 + 17 + -11$.

3.3. Step 1:. In step 1 the agent calculates the expected economic benefits and costs of different cognitive operations. For our application, a cognitive 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 cognitive operation at a time, we assume that individual cognitive operations themselves can include one *or more* box openings. Multiple box openings increase the amount of information revealed by a single cognitive operator, increase the option value of information revealed by that cognitive operator, and make the (partially myopic) model more forward-looking.[10]

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

We assume that an operator that 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 a box with a mouse, reading the contents of the box, and updating expectations. Such updating includes an *addition* operation as well as two *memory* operations: recalling the prior expectation and rememorizing the updated expectation.

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

$$w(x, \sigma) \equiv \sigma \phi \left(\frac{x}{\sigma} \right) - |x| \, \Phi \left(-\frac{|x|}{\sigma} \right), \tag{1}$$

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

[10] To gain intuition for this effect, consider two goods A and B, with $E(a) = 3/2$ and $E(b) = 0$. Suppose that opening one box reveals one unit of information so that after this information is revealed $E(a') = 5/2$ or $E(a') = 1/2$. Hence, a partially myopic agent won't see the benefit of opening one box, since no matter what happens $E(a') > E(b)$. However, if the agent considers opening two boxes, then there is a chance that $E(a')$ will fall below 0, implying that gathering the information from two boxes would be useful to the agent. Analogous arguments generalize to the case of continuous information densities.

revealed by the cognitive operator. We motivate equation (1) below, but first present an example calculation.

In the game shown in Figure 2, consider a cognitive 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 variation revealed by the cognitive operator is

$$\sigma_O^2 = var\left(\eta_{H2} + \eta_{H3} + \eta_{H4}\right) = (40.8)^2\left(\frac{9}{10} + \frac{8}{10} + \frac{7}{10}\right),$$

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

We now motivate equation (1). 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 cognitive operator O_A^Γ. Executing the cognitive 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 doesn't execute the cognitive operator, her expected payoff will be

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

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

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

This expectation captures the option value generated by being able to pick either row A or row B, *contingent* on the information revealed by cognitive operator O_A^Γ. The value of executing the cognitive operator is the difference between the previous two expressions:

$$E\left[\max\left(a', b\right)\right] - \max\left(a, b\right). \tag{2}$$

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 cognitive operator:

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

Appendix A shows that the value of the cognitive operator is[11]

$$E\left[\max\left(a', b\right)\right] - \max\left(a, b\right) = w\left(a - b, \sigma\right). \tag{3}$$

To help gain intuition about the w function, Figure 3 plots $w\left(x, 1\right)$. The general case can be deduced from the fact that $w\left(x, \sigma\right) = \sigma w\left(x/\sigma, 1\right)$.

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: $w\left(x, \sigma\right)$ is decreasing in $|x|$. Second, the value of a row exploration increases with the variability of the information that will be obtained: $w\left(x, \sigma\right)$ is increasing in σ. In other words, the more information that is likely to be revealed by a row exploration, the more valuable such a path exploration becomes.

3.4. Step 2. Step 2 executes the cognitive 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.

[11] This result assumes Gaussian innovations, which is the density used to generate the games in our experiment.

The subject executes the cognitive operator with the greatest benefit/cost ratio[12],

$$G \equiv \max_{O} \frac{w(x_O, \sigma_O)}{\kappa \Gamma_O}, \tag{4}$$

where κ is the cost of unmasking a single box. Since κ is constant, the subject executes the cognitive operator

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

Appendix B contains an example of such a calculation.

3.5. Step 3. Step 3 is a stopping rule. In games with an exogenous time budget, the subject keeps returning to step one 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.

3.6. 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

$$(750 \text{ cents})/\,(25 \text{ minutes} \cdot 60 \text{ seconds/minute}) = 0.50 \text{ cents/second}.$$

[12] Our model thus gives a crude but compact way to address the "accuracy vs simplicity" trade-off in cognitive processing. See Aragones *et al.* (2003) for a much more sophisticated treatment.

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 explored a 1-parameter version of the directed cognition model, in which the cost of cognition — κ — was 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 open-ended game, not to match the order of search or the distribution of search across games.

3.7. 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 cognitive operator. This assumption adopts the approach taken by Jehiel (1995), who assumes a constrained planning horizon in a game-theory context. Second, the directed cognition 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 directed cognition model avoids the infinite regress problem (i.e. the costs of thinking about thinking about thinking, etc...), by implicitly assuming that solving for O^* is costless to the agent.

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 even computationally. 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.[13]

We return now to the first of the three points listed in the previous paragraph: the perfectly

[13] 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.

rational attention allocation 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 8 rows, one standard deviation of unexplored information in each of the 8 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.[14] By contrast, the directed cognition model is equivalent to 8 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 directed cognition model is only 2 (compared to 17 for the model of perfect rationality).

Although this paper does not attempt to solve the model of perfectly rational attention allocation (or to test it), we can compare the performance of the partially myopic directed cognition model and the performance of the perfectly rational model. Like the directed cognition 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). Appendix C gives lower bounds on the payoffs of the directed cognition model relative to the payoffs of the perfectly rational model. Directed cognition does at least 91% as well as perfect rationality for exogenous time games and at least 71% as well as perfect rationality for endogenous time games.

3.8. Other decision algorithms. In this subsection, we discuss three other models that we compare to the directed cognition model. The first two models are simple, mechanical search rules, which can be parameterized as variants of the satisficing model (Simon 1955). These first two models are just benchmarks, which are presented as objects for comparison, not as serious candidate theories of attention allocation. The third model — Elimination by Aspects (Tversky 1972) — is the leading psychology model of choice from a set of multi-attribute goods.

The *Column model* unmasks boxes column by column, stopping either when time runs out (in

[14] 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. The Gittins index (Gittins 1979, Weitzman 1979, Weitzman and Roberts 1980) 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 only do one thing to a row (an "arm"): access it. 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 in Gittins' framework. We explored several modifications of the Gittins index, but they proved unfruitful at breaking the curse of dimensionality.

games with exogenous time budgets) or stopping according to a satisficing rule (in games with endogenous time budgets). Specifically 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.[15] In endogenous 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 estimated 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).

The *Row model* unmasks boxes row by row, starting with the "best" row and moving to the "worst" row, stopping either when time runs out (in games with exogenous time budgets) or stopping according to a satisficing rule (in games with endogenous time budgets). Specifically the Row model ranks the rows according to their values in column 1. Then the Row model unmasks all the boxes in the best row, then the second best row, etc. 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 $A^{\text{Row model}}$, an aspiration or satisficing level estimated so that the simulations generate an average number of simulated box openings that matches the average number of empirical box openings.

A choice algorithm called *Elimination by Aspects* (EBA) has been widely studied in the psychology literature (see work by Tversky 1972 and Payne, Bettman, and Johnson 1993). We apply this algorithm to our decision framework to analyze games with endogenous time budgets. We use the interpretation that each row is a good with 10 different attributes or "aspects" represented by the ten different boxes of the row. Our EBA application assumes that the agent 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 threshold value A^{EBA}. This elimination continues, stopping at the point where the *next* elimination would eliminate all remaining rows. At this stopping point, we pick the remaining row with the highest estimated value. As above, we estimate A^{EBA} so that the simulations generate an average number of simulated box openings that matches the average

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

number of empirical box openings.

4. Results

4.1. Subject Pool. Our subject pool is comprised of 388 Harvard undergraduates. Two-thirds of the subjects are male (66%). The subjects are distributed relatively evenly over concentrations: 11% math or statistics; 21% natural sciences; 20% humanities; 29% economics; and 20% other social sciences. In a debriefing survey we asked our subjects to report their statistical background. Only 15% report taking an advanced statistics class; 40% report only an introductory level class; 45% report never having taken a statistics course.

Subjects received a mean total payoff of \$29.23, with a standard deviation of \$5.49. Payoffs range from \$13.07 to \$46.69.[16] 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 ranges from 21 to 65.

Payoffs do not systematically vary with experimental experience. To get a sense of learning, we calculate the average payoff $X(k)$ of games played in round $k = 1, ..., 12$ of the exogenous time games. We calculate $X^B(k)$ for the exogenous games played Before the endogenous games and $X^A(k)$ for the exogenous games played After the endogenous games. We estimate a separate regression $X(k) = \alpha + \beta k$ for the Before and After datasets. We find that in both regressions, β is not significantly different from 0, which indicates that there is no significant learning within exogenous time games. Learning would imply $\beta > 0$. Also, the constant α is statistically indistinguishable across the two samples. Playing endogenous time games first does not contribute to any significant learning in exogenous time games. Hence we fail to detect any significant learning in our data.[17]

4.2. Statistical methodology. Throughout our analysis we report bootstrap standard errors for our empirical statistics. The standard errors are calculated by drawing, with replacement, 500 samples of 388 subjects. Hence the standard errors reflect uncertainty arising from the particular

[16] Payoffs show little systematic variation across demographic categories. In light of this, we have chosen to adopt the useful approximation that all subjects have identical strategies. Relaxing this assumption is beyond the scope of the current paper, but we regard such a relaxation as an important future research goal.

[17] See e.g. Camerer (2003), Camerer and Ho (1999) and Erev and Roth (1998) for some examples of problems where learning plays an important role.

sample of 388 subjects that attended our experiments. Our point estimates are the bootstrap means. In the figures discussed below, the confidence intervals are plotted as dotted lines around those point estimates.

We also calculated Monte Carlo means and standard errors for our *model* predictions. The associated standard errors tend to be extremely small, since the strategy is determined by the model. In the relevant figures (4-9), the confidence intervals for the model predictions are visually indistinguishable from the means.[18]

To calculate measures of fit of our models, we generate a bootstrap sample of 388 subjects and calculate a fit statistic that compares the model predictions for that bootstrap sample to the associated data for that bootstrap sample. We repeat this exercise 500 times to generate a bootstrap mean fit statistic and a bootstrap standard error for the fit statistic.

4.3. Games with exogenous time budgets. We begin by analyzing attention allocation patterns in the games with exogenous time budgets. Our analysis focuses on the pattern of box openings across columns and rows. 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 92.6% of the time (s.e. 0.7%[19]). Specifically, when subjects open a previously unopened box in a given row, 92.6% of the time that box has the highest variance of the as-yet-unopened boxes in that row. 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 of the exogenous games that were played. Specifically, each of our 388 subjects played 12 exogenous games,

[18] In Figures 4-9, the width of the associated confidence intervals for the models is on average only one-half of one percent of the value of the means.

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

yielding a total of $388 \times 12 = 4656$ exogenous games played. To generate this total, each subject was assigned a subset of 12 games from a set of 160 unique games. Hence, each of the 160 games was played an average of $4656/160 \sim 30$ times in the exogenous time portion of the experiment.

To calculate the empirical column profile (and all of the profiles that we analyze) we count only the first unmasking of each box. So if a subject unmasks the same box twice, this counts as only one opening. Approximately 90% of box openings are first-time unmaskings. We do not study repeat unmaskings because they are relatively rare in our data and because we are interested in building a simple and parsimonious model. Hence, we only implicitly model memory costs as a reduced form. Memory costs are part of the (time) cost of opening a box and encoding/remembering its value, which includes the cost of mechanically reopening boxes whose values were forgotten.[20]

Figure 4 also plots the theoretical predictions generated by yoked simulations of our model.[21] Specifically, these predictions are calculated by simulating the directed cognition model on the exact set of 4656 games played by the subjects. We simulate the model on each game from this set of 4656 games and instruct the computer to unmask 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 matches the theoretical predictions. This measure is simply the R^2 statistic[22] from the following constrained regression:[23]

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

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

[20] An extension of this framework might consider the case in which the memory technology include memory capacity constraints or depreciation of memories over time.

[21] See footnote 15 for a description of our yoking procedure.

[22] In other words,

$$R'^2 = 1 - \frac{\sum_{col}\left(Boxes(col) - \langle Boxes\rangle - \widehat{Boxes}(col) + \left\langle \widehat{Boxes}\right\rangle\right)^2}{\sum_{col}\left(\widehat{Boxes}(col) - \left\langle \widehat{Boxes}\right\rangle\right)^2}$$

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

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

$\widehat{Boxes}(col)$ 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 1 (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% (s.e. 1.7%), implying a very close match between the data and the predictions of the model.

Figure 5 reports analogous calculations by row. Figure 5 reports the number of boxes opened on average by row, with the rows ranked by their value in column one. (Recall that column one is never masked.) 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% (s.e. 9.2%), implying a poor match between the data and the predictions of the model. The model simulations predict too many unmaskings on the top ranked rows and far too few unmaskings on the bottom ranked rows. The subjects are much less selective than the model. The R'^2 is negative because we constrain the coefficient on simulated boxes to be unity. This is the only bad prediction that the model makes.

Figure 6 reports similar calculations using an alternative way of ordering rows. Figure 6 reports the number of boxes opened on average by row, with the rows ranked by their values at the *end* of each game. For our row predictions the R'^2 statistic is 86.7% (s.e. 1.4%).

4.4. Endogenous games. We repeat the analysis above for the endogenous games. As discussed in section 3, 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' expected earnings per unit time in the endogenous games: $\kappa = 0.49$ cents/box opened (see calibration discussion in section 3). With this calibration, subjects are predicted to open 15.67 boxes per game (s.e. 0.01). In the data, however, subjects open 26.53 boxes per game (s.e. 0.55). 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.

We begin with an evaluation of the declining variance rule. In our endogenous games, subjects follow the declining variance rule 91.0% of the time (s.e. 0.8%). Specifically, when subjects open a previously unopened box in a given row, 91.0% of the time that box has the highest variance of the as-yet-unopened boxes in that row.

Figure 7 reports the average number of boxes unmasked in columns 2-10 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 $\kappa = 0.18$.

The empirical data is calculated by averaging together subject responses on all of the endogenous games that were played. The theoretical predictions are generated by yoked simulations of our model. Specifically, we use the directed cognition model to simulate play of the 10,931 endogenous games that the subjects played. The model generates its own stopping rule (so we no longer yoke the number of box openings).

Figure 7 also reports the associated R'^2 statistic for these column comparisons in the endogenous games. For these endogenous games, the column R'^2 statistic is 96.3% (s.e. 1.4%) for $\kappa = 0.49$ and 73.1% (s.e. 4.2%) for $\kappa = 0.18$.

Figure 8 reports analogous calculations by row for the endogenous games. Figure 8 reports the number of boxes opened on average by row, with the rows ranked by their values in column one. We report the mean number of boxes opened by row for both the subject data and the model predictions. For these endogenous games, the row R'^2 statistics are 85.3% (s.e. 1.3%) for $\kappa = 0.49$ and 64.6% (s.e. 2.4%) for $\kappa = 0.18$.

Figure 9 reports similar calculations using the alternative way of ordering rows. Figure 9 reports the number of boxes opened on average by row, with the rows ranked by their values at the *end* of each game. For these endogenous games, the alternative row R'^2 statistics are 91.8% (s.e. 0.5%) for $\kappa = 0.49$ and 84.5% (s.e. 0.8%) for $\kappa = 0.18$.

These figures show that the model explains a very large fraction of the variation in attention across rows and columns. However, a subset of the results in this section are confounded by an effect that Costa-Gomes, Crawford, and Broseta (2001) have found in their analysis. In particular, subjects who use the Mouselab interface tend to have a bias toward selecting cells in the upper left corner of the screen and transitioning from left to right as they explore the screen.

The up-down component of this bias does *not* affect our results, since our rows are randomly ordered with respect to their respective payoffs. The left-right bias affects only our column results (figures 4 and 7), since information with greater economic relevance is located toward the left-hand side of the screen.

One way to evaluate this column bias would be to flip the presentation of the problem during the experiment so that the rows are labeled on the right and variance declines from right to left. Alternatively, one could rotate the presentation of the problem, swapping columns and rows. Unfortunately, the internal constraints of Mouselab make either of these relabelings impossible.[24] Future work should use a more flexible programming language that facilitates such spatial rearrangements. Finally, we note that neither the up-down nor the left-right biases influence any of our analyses of either row openings (above) or endogenous stopping decisions (below).

4.5. Stopping Decisions in Endogenous Games. Almost all of the analysis above reports *within*-game variation in attention allocation. The analysis above shows that subjects allocate most of their attention to economically relevant columns and rows within a game, matching the patterns predicted by the directed cognition model. Our experimental design also enables us to evaluate how subjects allocate their attention *between* games. In this subsection we focus on several measures of such between-game variation.

First, we compare the empirical variation in boxes opened per game to the predicted variation in boxes opened per game. Most importantly, we ask whether the model can correctly predict which games received the most attention from our subjects. Our experiment utilized 160 unique games, though no single 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 moments 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 elements, since we are analyzing game-specific averages.

We begin by comparing first moments. The empirical (equally weighted) mean of $Boxes(g)$ is 26.53 (s.e. 0.55). By contrast the 0-parameter version of our model (with $\kappa = 0.49$) generates a predicted mean of 15.67 (s.e. 0.01). Hence, unless we pick κ to match the empirical mean (i.e.

[24] However, one could leave the labels on the left and have the unobserved variances decline from right to left. In this design, the displayed cells would have the minimal variance instead of the maximal variance. Such a design would eliminate the old left-right bias, but it would create a new bias by making the row "label" appear on the opposite side of the screen from where search should (theoretically) begin.

$\kappa = 0.18$), our model only crudely approximates the average number of boxes opened per game.

We turn now to second moments. The empirical 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). Moreover, when we set $\kappa = 0.18$ to match the average 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 individual (160) datapoints 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 does a good job predicting the relationship between economic incentives and depth of analysis. Figure 11 plots a non-parametric kernel estimate of the expected number of additional box openings in a game, conditional on the current value of the ratio of the benefit of marginal analysis to the cost of marginal analysis (see eq. 4):

$$G \equiv \max_{O} \frac{w(x_O, \sigma_O)}{\kappa \Gamma_O}.$$

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 which is revealed by mental operator O, and Γ_O is the number of boxes opened by mental operator O.

The solid line represents the subject data. The light dashed line represents the relationship predicted by the model (with $\kappa = 0.18$). The figure also shows bootstrap estimates of the 99% 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 directed cognition 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 endogenous games, generating 330,873 observations.[25] We find that the variables with the greatest *economic* significance are G and *boxes,* with G playing the most important role. The coefficient on G is -0.3660 (with standard error 0.0081); the coefficient on *boxes* is 0.0404 (0.0008); the coefficient on *leader* is 0.0064 (0.0002). 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 1/2 as large and a one-standard-deviation increase in the value of *leader* has an effect less than 1/4 as large.

The logistic regression shows that the economic value of information — G — is by far the most important predictor of the decision to stop searching. However, our subjects are also using other information, 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 reliance 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 Lovallo (1999) and Massey and Wu (2002). These authors find that subjects use

[25] In other terms, we estimate:

$$\text{Probability of continuation } = \frac{\exp(\beta' x)}{1 + \exp(\beta' x)}.$$

where x is a vector of decision-node attributes, and β is the vector of estimated coefficients.

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 attention allocation.

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. Exploring the robustness of this finding is a goal for future research.

4.6. Comparisons with Other Models. The analysis to date focuses on the predictions of the directed cognition model. In this subsection we consider alternative models and evaluate their performance. Table 1 reports R'^2 measures for all of the alternative models summarized in subsection 3.8.

None of these benchmarks does nearly as well as the zero-parameter directed cognition model. For the exogenous games, the directed cognition model has an average R'^2 value of 52.4% (with a standard deviation of 3.4 percentage points). The Column and Row models have respective averages of -18.8% (s.e. 6.4%) and -137.4% (s.e. 13.9%).

For the endogenous games, the zero-parameter directed cognition model has an average R'^2 of 91.2% (s.e. 0.7%). The Column model with a satisficing stopping rule has an average R'^2 of 42.9% (s.e. 0.7%). The Row model with a satisficing stopping rule has an average R'^2 of 55.4% (s.e. 0.4%). The Elimination by Aspects model has an average R'^2 of -6.3% (s.e. 6.2%).

We also evaluate the different models' ability to forecast game-by-game variation in average time allocation.[26] Table 1 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. The Elimination by Aspects model actually has a large negative correlation. By contrast, the zero-parameter directed cognition model generates predictions that are highly correlated (0.66) with the empirical average boxes opened by

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

game.

5. Conclusion

A rapidly growing literature shows that limited attention can play an important role in economic analysis. Most of the papers in that literature have assumed that scarce attention is allocated according to cost-benefit principals. In this paper we develop a tractable framework for such cost-benefit analysis, and we directly *measure* the attention allocation process. We report the first direct empirical test of an economic cost-benefit model of endogenous attention allocation.

Our research was conducted with a stylized choice problem that makes it possible to measure attention allocation with a well-understood laboratory method (i.e. Mouselab). In this sense, our choice problem was a natural starting point for this research program.

The directed cognition model successfully predicts the key empirical regularities of attention allocation measured in our experiment. Moreover, the model makes these predictions without relying on extraneous parameters. Instead, the model is based on cost-benefit principles.

Understanding *how* decision-makers think about complex problems is a relatively new frontier in economic science. We believe that understanding the cognitive processes underlying economic choices will further economists' understanding of the choices themselves. Stripping back decision-making to the level of cognitive processes poses a fundamental set of long-run research goals.

The study of cognitive processes in economic decision-making is still in its infancy, particularly the study of attention allocation. But we are optimistic that such process-based research will ultimately pave the way for a science of decision-making with much greater predictive power than the classical "as if" modeling that treated 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 and Costa-Gomes, Crawford, and Broseta 2001). We look forward to future process-based research that will open the box more completely.

6. Appendix A: Derivation of equation (3)

To gain intuition for equation (3), begin by assuming that $b \geq a$. In this case,

$$\max\left(a', b\right) - \max(a, b) = \begin{cases} 0 & \text{if} \quad b > a + \varepsilon \\ a + \varepsilon - b & \text{if} \quad b \leq a + \varepsilon \end{cases}$$

Because ε is drawn from a Normal$\left(0, \sigma^2\right)$ distribution, integrating over the relevant states yields the right-hand-side of equation (3):

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

We can re-express the option value formula by noting that $x\phi(x) = -\phi'(x)$, so that:

$$\begin{aligned} \int_{|a-b|}^{\infty} (\varepsilon - |a - b|)\, \phi\left(\frac{\varepsilon}{\sigma}\right) \frac{d\varepsilon}{\sigma} &= \int_{|a-b|}^{\infty} \varepsilon\phi\left(\frac{\varepsilon}{\sigma}\right) \frac{d\varepsilon}{\sigma} - |a - b| \left(1 - \Phi\left(\frac{|a-b|}{\sigma}\right)\right) \\ &= \sigma\phi\left(\frac{|a-b|}{\sigma}\right) - |a - b|\, \Phi\left(-\frac{|a-b|}{\sigma}\right) \\ &= w\left(|a - b|, \sigma\right). \end{aligned}$$

Likewise, when $b \leq a$,

$$\max\left(a', b\right) - \max(a, b) = \begin{cases} b - a & \text{if} \quad b > a + \varepsilon \\ \varepsilon & \text{if} \quad b \leq a + \varepsilon \end{cases}$$

Integrating over the relevant states again yields the right-hand-side of equation (3):

$$\begin{aligned} (b-a)\Phi\left(\frac{b-a}{\sigma}\right) + \int_{b-a}^{\infty} \varepsilon\phi\left(\frac{\varepsilon}{\sigma}\right) \frac{d\varepsilon}{\sigma} &= (b-a)\Phi\left(\frac{b-a}{\sigma}\right) + \sigma\phi\left(\frac{b-a}{\sigma}\right) \\ &= w\left(|b - a|, \sigma\right). \end{aligned}$$

7. Appendix B: Example calculation of G

Consider a game where the payoffs in column 1 are drawn from a $N(0, \sigma^2)$ distribution. The variance of payoffs in column n is then $\sigma_n^2 = \sigma^2 (11 - n)/10$, for $n \in \{1...10\}$. Call η_{mn} the payoff of the box located in row m and column n. Using this notation, the current expected value of row

A is

$$a = \sum_{\text{already opened boxes } n \text{ in row } A} \eta_{An}$$

Suppose that at the current time in row A a subject has already opened boxes $\{1, 2, 4, 7, 8\}$ and that boxes $\{3, 5, 9, 10\}$ remain unopened.[27] The current value of row A is

$$a = \eta_{A1} + \eta_{A2} + \eta_{A4} + \eta_{A7} + \eta_{A8}$$

Call $a^{\tilde{}}$ the value of the current best alternative to row A. In most instances $a^{\tilde{}}$ is the value of the leading row, except if A is itself the leading row itself, in which case $a^{\tilde{}}$ is the value of the second leading row.

The relevant cognitive operators for exploring row A are 'open box 3,' 'open boxes 3 and 5,' 'open boxes 3, 5, and 9' and 'open boxes 3, 5, 9, and 10.' Their costs are respectively 1, 2, 3 and 4. So, if $x = a - a^{\tilde{}}$, then the value of the benefit/cost ratio G_A of row A is,

$$\begin{aligned} G_A &= \max\{w(x,\sigma_3)\,, w\left(x, \sqrt{\sigma_3^2+\sigma_5^2}\right)/2, w\left(x, \sqrt{\sigma_3^2+\sigma_5^2+\sigma_9^2}\right)/3, \\ & \quad w\left(x, \sqrt{\sigma_3^2+\sigma_5^2+\sigma_9^2+\sigma_{10}^2}\right)/4\} \\ x &= a - a^{\tilde{}} \end{aligned}$$

and the value of G at the current time is simply the highest possible benefit/cost ratio across all rows:

$$G = \max_n G_n.$$

Finally, note that the highest benefit/cost ratio may be associated with a cognitive operator that opens more than one box. For example, consider a case in which no boxes have been opened in a game in which the leading row and the next best row differ by $3\sigma_1$. Hence, if only one box is opened, a three-sigma event ($p = 0.0013$) is required to generate a rank reversal in the rows. By contrast, if two boxes are opened, the chance of reversal is 11 time greater. Specifically, the standard

[27] This pattern of box openings — $\{1, 2, 4, 7, 8\}$ — would not arise under the model, since the model would never skip a column. However, such skips can arise in experimental play, and hence the directed cognition model must be able to handle such cases if it is used to predict continuation play (e.g., stopping decisions).

deviation of aggregated noise in the two box case is $\sigma = \sqrt{\sigma_1^2 + (0.9)\,\sigma_1^2}$. Since $3\sigma_1/\sqrt{\sigma_1^2 + (0.9)\,\sigma_1^2} = 3/\sqrt{1.9} = 2.2$, we only require a 2.2-sigma event ($p = 0.0146$) in the two-box case to generate a rank reversal. Hence, the two-box cognitive operator may be much more cost effective (when evaluated by the directed cognition algorithm) than the one-box cognitive operator, since the two-box cognitive operator increases the chance of reversal by a factor of 11 but only has twice the cognitive cost.

8. Appendix C: Lower bounds for the performance of the directed cognition algorithm

We wish to evaluate how well directed cognition does compared to the Perfectly Rational solution of a model with calculation costs. Perfect Rationality's complexity makes it prohibitively hard to evaluate. Here we provide bounds on its performance. We will show that directed cognition performs at least 91% as well as Perfect Rationality for exogenous time games and at least 71% as well as Perfect Rationality for endogenous time games.

A useful benchmark is that of Zero Cognition Costs, i.e. the performance of an algorithm that could inspect all the boxes at zero cost. We first evaluate Zero Cognition Costs' performance for a game with standard deviation of the first column σ. Each of the 8 rows will have a value which is a normal with mean 0 and standard deviation σ times $\left(\sum_{j=1}^{10} j/10\right)^{1/2} = 5.5^{1/2}$. Zero Cognition Costs will select the best of those 8 values. The expected value of the maximum of 8 i.i.d. $N(0,1)$ variables is:

$$\nu \simeq 1.423. \tag{5}$$

So the average performance of Zero Cognition Costs will be $5.5^{1/2}\nu\sigma \simeq 3.34\sigma$. We gather this as the following Fact.

Fact: The average payoff V^* of the optimal algorithm will be less than 3.34σ.

$$V^* \leq 3.34\sigma. \tag{6}$$

We use this result to analyze exogenous and endogenous time games.

8.1. Exogenous time games. We simulate the performance of the directed cognition algorithm on a large number of i.i.d. games with the same stochastic structure as in the experiments, and unit standard deviation of the first column. The time allocated is randomly drawn from [10sec, 49sec], as in the experiment. The performance of directed cognition is on average 3.064. Applying inequality (6), the performance of Perfect Rationality is less than 3.34, so directed cognition's relative performance is at least $3.064/3.34 = 0.917$. Hence, directed cognition does 91% as well as Perfect Rationality in exogenous time games.

8.2. Endogenous time games. Monte-Carlo simulations determine the performance of $V^{DC}(\sigma,\kappa)$ in games of standard deviation σ, defined as the expected value of the chosen rows minus κ, the cost of time, times the time taken to explore the games. Call $V^*(\sigma,\kappa)$ the performance of Full Rationality. The relative performance of directed cognition is:

$$\pi(\sigma,\kappa) = \frac{V^{DC}(\sigma,\kappa)}{V^*(\sigma,\kappa)}. \tag{7}$$

By definition $\pi(\sigma,\kappa) \le 1$. We now look for an upper bound on $V^*(\sigma,\kappa)$. We consider a generic game G with a first column of unit variance. The cost of exploring t boxes is κt. V^* is the expected value under the optimal strategy. Call $\eta_{A1},...,\eta_{H1}$ the values of the 8 rows in column 1, and $\xi = \max\{\eta_{A1},...,\eta_{H1}\}$ the highest valued box in the first column. The game G has thus a lower value than the game G' that would have equal values of ξ in the first column: $V^* \le V^{G'}$. The expected value of G' is the expected payoff of the first column, $E[\xi] = \nu \simeq 1.42$, plus the value of a game G'' that has zeroes in the first column. $V^{G'} = V^{G''} + \nu$. So

$$V^* \le V^{G''} + \nu$$

Game G'' has the structure of our games, with zeros in column 1. It has 8 rows. Thus, it has lower value than a game G''' that has an *infinite* number of rows, and the same stochastic structure otherwise: column 1 has boxes of value 0, and column $n \in \{2,...,10\}$ has boxes with variance $(N+1-n)/N$ with $N = 10$ columns. $V^{G''} \le V^{G'''}$. The value $M := V^{G'''}$ of G''' has the property similar to a Gittins index. Indeed, suppose that one is thinking about exploring a new row. Without loss of generality call that row A. One will stop exploring this row after stopping in column τ. Call a_t the value of that row after exploring column 2 through t. So we have:

$$a_t = a_1 + \sum_{n=1}^{t-1} \left(\frac{N-n}{N} \right)^{1/2} \eta_{A,n+1}$$

where the η_{An}'s are i.i.d. standard normals. In game G''' we have $a_1 = 0$. At the end of the exploration of this row, one can either pick the box, which has a value a_τ, or drop the current row, and start the game from scratch, which has a value $M = V^{G'''}$. So the payoff after exploring the row is $\max(a_\tau, M)$. So the value M of game G''' satisfies:

$$M = \max_{\text{stopping time } \tau} E\left[\max(a_\tau, M) - \kappa(\tau - 1) \mid a_1 = 0\right].$$

i.e.

$$0 = \max_{\text{stopping time } \tau} E\left[\max(a_\tau - M, 0) - \kappa(\tau - 1) \mid a_1 = 0\right]. \tag{8}$$

Call $v(a, t, \kappa)$ the value of the game[28] that has explored up to column t, has current payoff a, and that at the end of the search τ yields a payoff $\max(a_\tau, 0)$ minus the cognition costs $\kappa(\tau - t)$. Formally:

$$v(a, t, \kappa) = \max_{\text{stopping time } \tau \geq t} E\left[\max(a_\tau, 0) - \kappa(\tau - t) \mid a_t = a\right]$$

for $a \in \mathbb{R}$, $t \in \{1, ..., 10\}$, $\kappa > 0$. Because of (8), the value of the game $M(\kappa)$ satisfies $v(-M, 1, \kappa) = 0$, and by standard properties of value functions it is the smallest real with that property:

$$M(\kappa) = \min\{x \geq 0 \text{ s.t. } v(-x, 1, \kappa) = 0\}$$

Using (5), we conclude that

$$V^* \leq 1.42 + M(\kappa).$$

for a game of a unit standard deviation. As $V^*(\sigma, \kappa)$ is homogenous of degree 1, the inequality for a general σ is:

$$V^*(\sigma, \kappa) \leq 1.42\sigma + \sigma M(\kappa/\sigma).$$

[28] In terms of Gabaix and Laibson (2002), this is the value function of an "a vs 0" game.

Full Rationality does less well that Zero Cognition Costs, so (6) gives:

$$V^* (\sigma, \kappa) \leq 3.34\sigma$$

We combine those two inequalities using the fact that $\overline{V}^*$ is convex in κ, which comes from the usual replication arguments. This gives:

$$V^* (\sigma, \kappa) \leq \overline{V}^* (\sigma, \kappa)$$

where $\overline{V}^* (\sigma, \kappa)$ is the convex envelope[29] of 3.34σ and $1.42\sigma + \sigma M (\kappa/\sigma)$, seen as functions of κ.

Eq. (7) finally gives

$$\pi (\sigma, \kappa) \geq \underline{\pi} := \inf_{\sigma,\kappa} V^{DC} (\sigma, \kappa) / \overline{V}^* (\sigma, \kappa) \tag{9}$$

We evaluate $\underline{\pi}$ numerically, and find $\underline{\pi} = 0.71$. We conclude that directed cognition does at least 71% as well as Full Rationality in endogenous games.

[29] As we take the lower cord between and 3.34σ and $1.42\sigma + \sigma M (\kappa_0/\sigma)$, this gives, formally:

$$\overline{V}^* (\sigma, \kappa) = \min_{\kappa_0 \geq \kappa} \left(1 - \frac{\kappa}{\kappa_0}\right) 3.34\sigma + \frac{\kappa}{\kappa_0} (1.42\sigma + \sigma M (\kappa_0/\sigma)) .$$

9. REFERENCES

Aragones, E., Gilboa, I., Postlewaite, A., and Schmeidler, D. (2003), "Accuracy vs. Simplicity: A Complex Trade-Off " (Mimeo).

Camerer, C. (2003), *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton: Princeton University Press).

Camerer, C. and Ho, T-H. (1999), "Experience Weighted Attraction Learning in Normal-Form Games", *Econometrica,* 67, 837-874.

Camerer, C., Johnson, E.J., Rymon, T., and Sen, S. (1993), "Cognition and Framing in Sequential Bargaining for Gains and Losses", in K. Binmore, A. Kirman, and P. Tani (eds.) *Frontiers of Game Theory* (Cambridge, MA: MIT Press), 1994, 27-47.

Camerer, C. and Lovallo, D. (1999), "Overconfidence and Excess Entry: an Experimental Approach", *American Economic Review*, March, 89, 306-318.

Conlisk, J. (1996), "Why Bounded Rationality?" *Journal of Economic Literature*, June, *34* (2), 669-700.

Costa-Gomes, M., Crawford, V., and Broseta, B. (2001), "Cognition and Behavior in Normal-Form Games: An Experimental Study", *Econometrica*, 69, 1193-1235.

Costa-Gomes, M. and Crawford, V. (2003), "Cognition and Behavior in Two Person Guessing Games: an Experimental Study" (Mimeo, University of California, San Diego).

Daniel, K., Hirshleifer, D., and Teoh, S.H. (2002), "Investor Psychology in Capital Markets: Evidence and Policy Implications", *Journal of Monetary Economics* 49(1) ,139-209.

D'Avolio, G. and Nierenberg, E. (2003), "Investor Attention: Evidence from the Message Boards" (Mimeo, Harvard).

Della Vigna, S. and Pollet, J. (2003), "Attention, Demographic Changes, and the Stock Market" (Mimeo, University of California at Berkeley).

Erev, I. and Roth, A. (1998), "Predicting How People Play Games", *American Economic Review*, September, 88(2), 848-81.

Gabaix, X. and Laibson, D. (2001), "The 6D Bias and the Equity Premium Puzzle", in B. Bernanke and K. Rogoff (eds.) *NBER Macroeconomics Annual* , 257-312.

Gabaix, X. and Laibson, D. (2002), "Bounded Rationality and Directed Cognition," (Mimeo, MIT and Harvard).

Gabaix, X. and Laibson, D. (2003), "Some Industrial Organization with Boundedly Rational Consumers", (Mimeo, MIT and Harvard).

Gigerenzer, G., Todd, P. and the ABC Research Group (1999), *Simple Heuristics that Make Us Smart* (Oxford: Oxford University Press).

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

Gumbau, F. (2003), "Inflation Persistence and Learning in General Equilibrium" (Mimeo, Harvard).

Hirshleifer, D., Lim, S.S., and Teoh, S.H. (2003), "Disclosure to a Credulous Audience: The Role of Limited Attention" (Mimeo).

Jehiel, P. (1995), "Limited Horizon Forecast in Repeated Alternate Games", *Journal of Economic Theory*, 67, 497-519.

Johnson, E.J., Camerer, C., Sen, S., and Rymon, T. (2002), "Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining", *Journal of Economic Theory*, May 2002, 104(1), 16-47.

Kahneman, D. (1973), *Attention and Effort* (New Jersey: Prentice Hall Inc.).

Kahneman, D. and Tversky, A. (1974), "Judgment Under Uncertainty: Heuristics and Biases", *Science, 185, 1124-1131.*

Lynch, A. (1996), "Decision Frequency and Synchronization Across Agents: Implications for Aggregate Consumption and Equity Returns", *Journal of Finance* 51(4): 1479-97.

Mankiw, N.G. and Reis, R. (2002), "Sticky Information Versus Sticky Prices: a Proposal to Replace the New Keynesian Phillips Curve", *Quarterly Journal of Economics* 117(4), 1295-1328.

Massey, B. C. and Wu, G. (2002), "Detecting Regime Shifts: The Psychology of Over- and Under-reaction" (Mimeo, University of Chicago Graduate School of Business).

Moray, N. (1959), "Attention in Dichotic Listening: Affective Cues and the Influence of Instructions", *Quarterly Journal of Experimental Psychology*, 11, 56-60.

Pashler, H. (ed) (1998), *Attention*. (UK: Psychology Press).

Pashler, H. and Johnston, J.C. (1998), "Attentional Limitations in Dual Task Performance," in H. Pashler (ed) *Attention* (UK: Psychology Press), 155-190.

Payne, J.W., Bettman, J.R, and Johnson, E.J. (1993), *The Adaptive Decision Maker* (Cambridge: Cambridge University Press).

Payne, J.W), Braunstein, M.L., and Carroll, J.S. (1978), "Exploring Predecisional Behavior: an Alternative Approach to Decision Research", *Organizational Behavior and Human Performance*, 22, 17-44.

Peng, L. and Xiong, W. (2003) "Capacity Constrained Learning and Asset Price Comovement", (Mimeo, Princeton).

Pollet, J. (2003), "Predicting Asset Returns with Expected Oil Price Changes" (Mimeo, Harvard).

Rogers L.C.G. (2001) "The relaxed investor and parameter uncertainty." *Finance and Stochastics* 5(2): 131-54.

Russo, J. E. (1978), "Eye Fixations Can Save the World: a Critical Evaluation and a Comparison between Eye Fixations and Other Information Processing Methodologies", *Memory and Cognition*, 3(3), 267-276.

Simon, H. (1955), "A Behavioral Model of Rational Choice", *Quarterly Journal of Economics*, Feb, 69(1), 99-118.

Sims, C. (2003), "Implications of Rational Inattention", *Journal of Monetary Economics*, 50(3), 665-690.

Tversky, A. (1972), "Elimination by Aspects: A Theory of Choice", *Psychological Review,* 79, 281-299.

Wegner, D.M. (1989), *White Bears and Other Unwanted Thoughts: Suppression, Obsession, and the Psychology of Mental Control* (New York: Viking/Penguin).

Weitzman, M. (1979), "Optimal Search for the Best Alternative", *Econometrica,* 47(3), 641-54.

Weitzman, M. and Roberts, K. (1980), "On a General Approach to Search and Information Gathering", MIT Working Paper 263.

Wood, N. and Cowan, N. (1995), "The Cocktail Party Phenomenon Revisited: How Frequent Are Attention Shifts to One's Name in an Irrelevant Auditory Channel", *Journal of Experimental Psychology: Learning Memory and Cognition*, 21(1), 255-60.

Woodford, M. (2002),"Imperfect Common Knowledge and the Effects of Monetary Policy", in P. Aghion, J. Stiglitz, and M. Woodford (eds.) *Knowledge, Information, and Expectations in Modern Macroeconomics: in Honor of Edmund S. Phelps* (Princeton: Princeton University Press).

Yantis, S. (1998), "Control of Visual Attention," in Harold Pashler (ed.) *Attention* (UK: Psychology Press), 223-256.

Table 1: Evaluation of Alternative Models

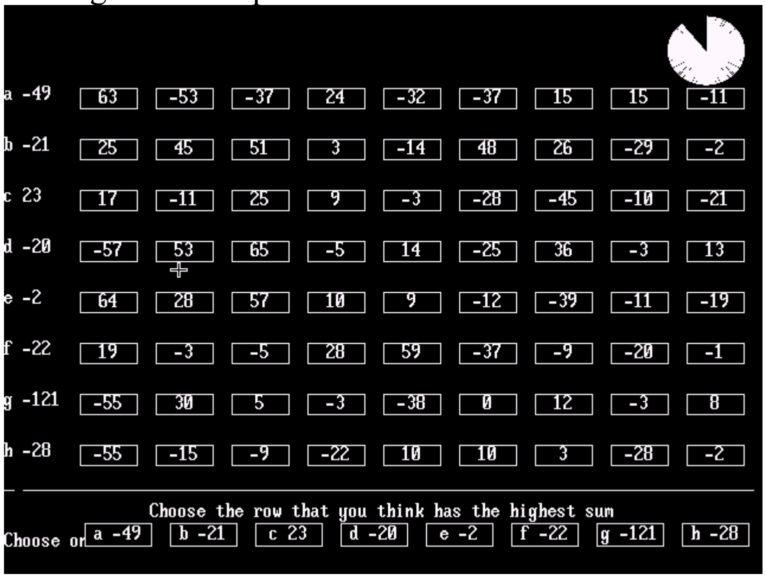
	Games with Exogenous Time Budgets			Games with Endogenous Time Budgets					
	Directed Cognition	Column Model	Row Model	**Directed Cognition**	Directed Cognition	Column Model Satisficing	Row Model Satisficing	Directed Cognition Satisficing	Elimination by Aspects
Free parameters?	**No**	No	No	**No**	$\kappa = 0.18$	S = 42	S = 43	S = 45	A = -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%)	(0.9%)	(1.4%)	(4.2%)	(1.8%)	(0.9%)	(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%)	(0.5%)	(0.4%)	(1.3%)	(2.4%)	(0.4%)	(0.7%)	(1.0%)	(0.7%)
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%)	(0.4%)	(1.5%)	(0.5%)	(0.8%)	(0.3%)	(0.6%)	(0.8%)	(0.8%)
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%)	**(0.7%)**	(1.7%)	(0.7%)	(0.4%)	(0.6%)	(6.2%)
Correlation with Empirical Number of Boxes per Game	**---**	---	---	**0.66**	0.61	0.00	0.02	0.03	-0.39
	---	---	---	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)

Bootstrapped 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 co-efficient on the predictions is fixed to equal 1.

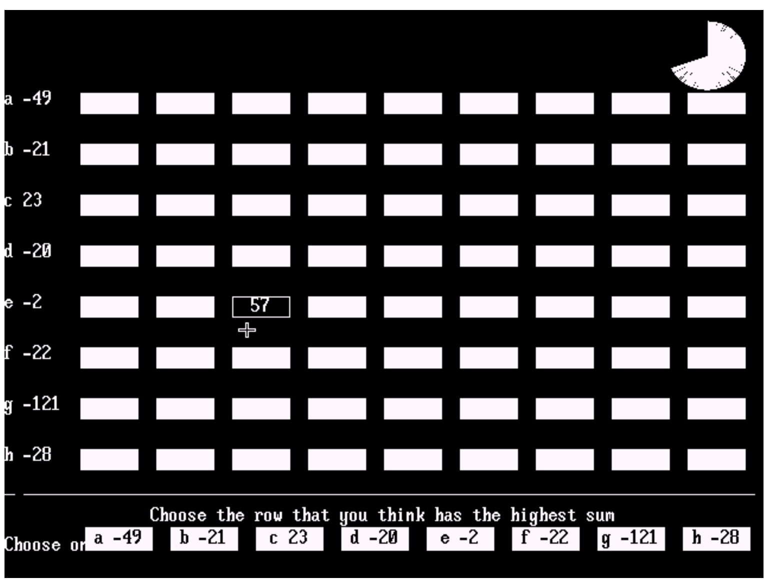
Correlation refers to the correlation between the number of boxes opened by subjects in each of the 160 games and the number of box openings predicted by the models.

Figure 1: Sample Game with All Values Unmasked



In the actual experiment, the subjects see only the value of 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% confidence interval of –60 to 60 cents.

Figure 2: Sample Game with Values Concealed



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.

Figure 3: w(x,σ) for σ = 1

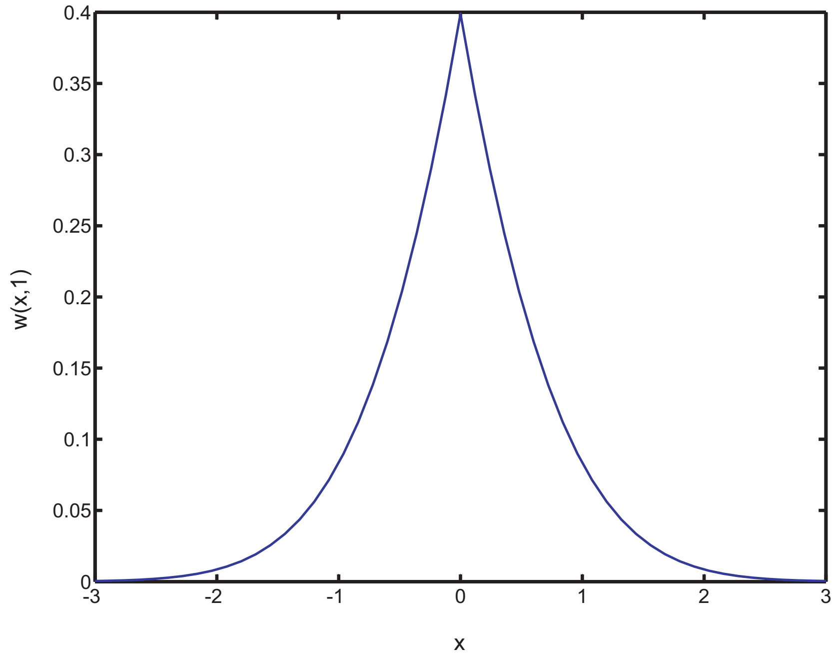


Figure 3 plots the expected benefit from continued search on an alternative if the difference between the value of the searched alternative and its best alternative is x, and the standard deviation of the information gained is $\sigma = 1$ as defined in Equation 3. This w function is homogeneous of degree one.

Figure 4: Column Profiles for Games with Exogenous Time Budgets

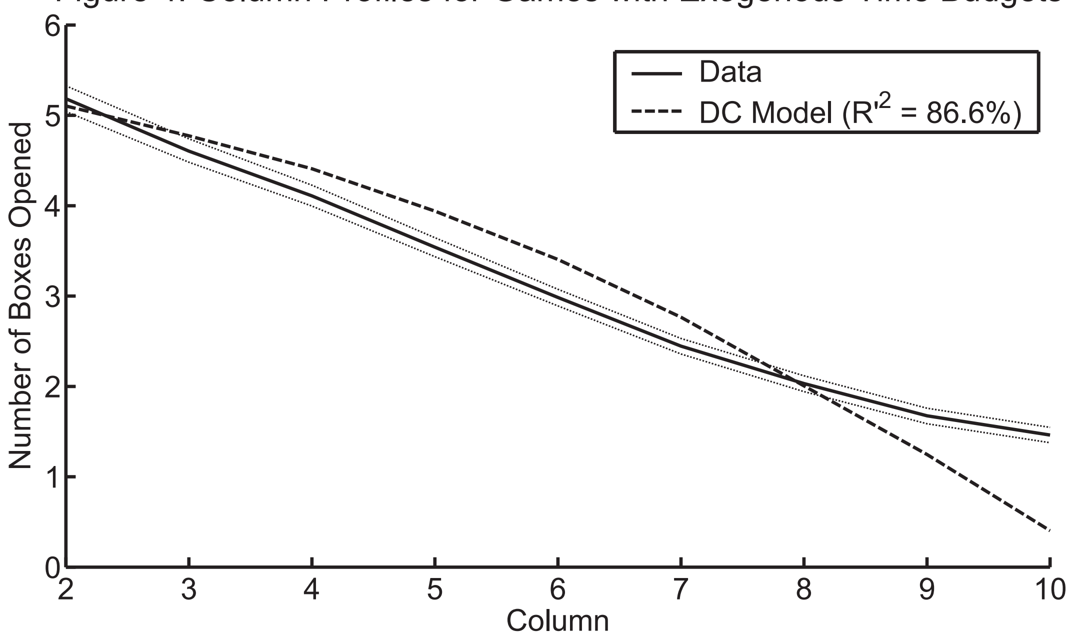


Figure 4 plots the mean number of boxes opened in each column by subjects and by the directed cognition model, for games in which time budgets are imposed exogenously. Dotted lines show the bootstrapped 95% confidence intervals for the data.

Figure 5: Row Profiles for Games with Exogenous Time Budgets

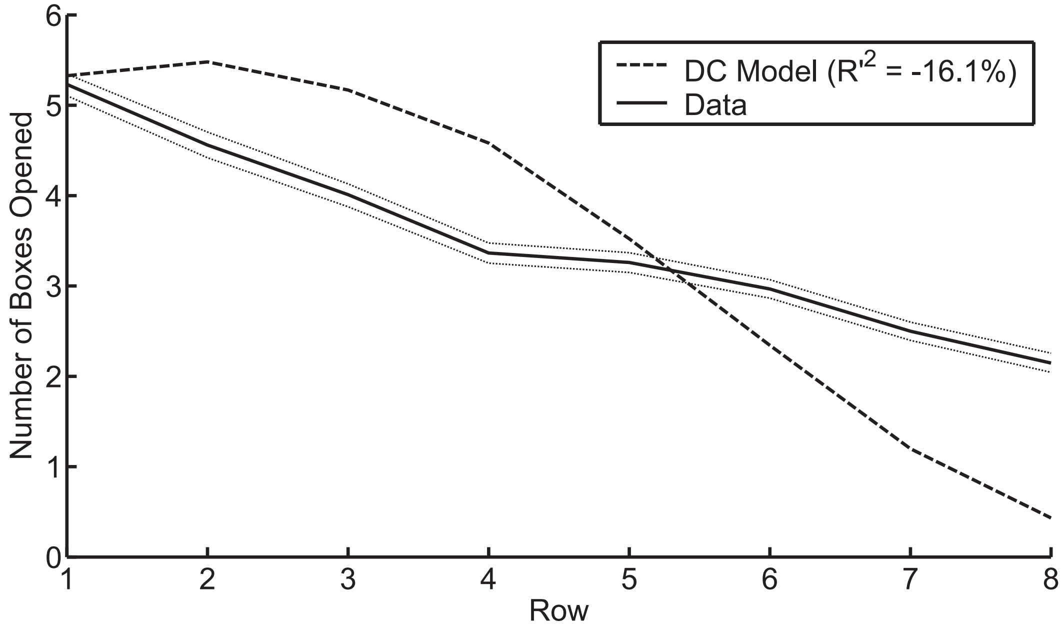


Figure 5 plots the mean number of boxes opened in each row by subjects and by the directed cognition model, for games in which time budgets are imposed exogenously. Dotted lines show the bootstrapped 95% confidence intervals for the data. The rows are ordered according to the values in the first column.

Figure 6: Row Profiles (alternative ordering) for Games with Exogenous Time Budgets

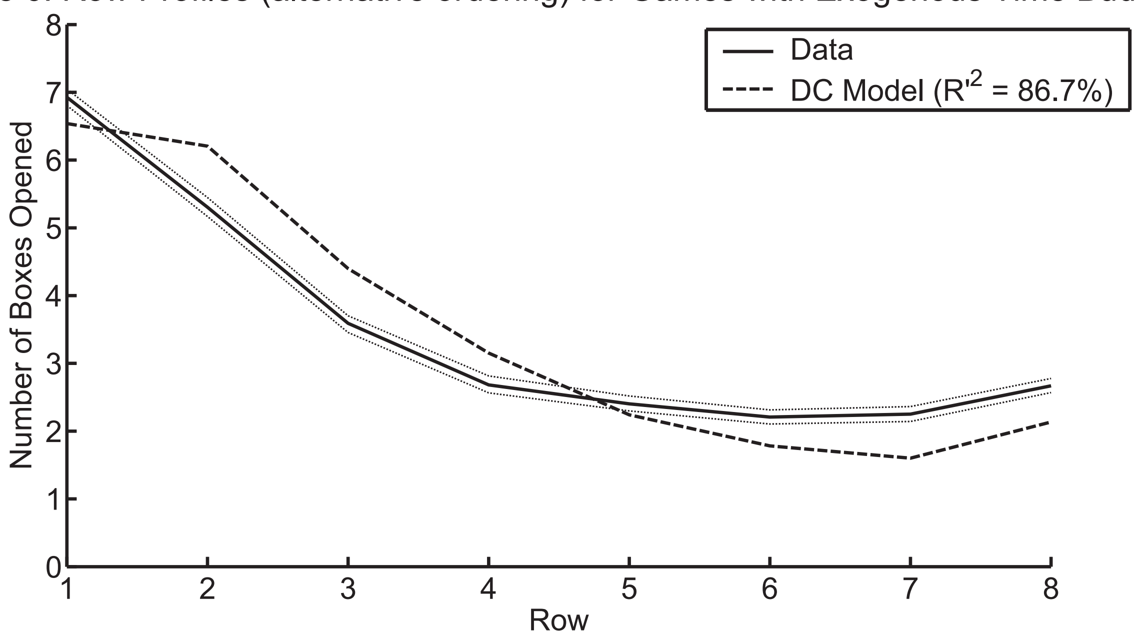


Figure 6 plots the mean number of boxes opened in each row by subjects and by the directed cognition model, for games in which time budgets are imposed exogenously. Dotted lines show the bootstrapped 95% confidence intervals for the data. The rows are ordered according to all the information available to subjects after they have concluded their search, just prior to making a choice.

Figure 7: Column Profiles for Games with Endogenous Time Budgets

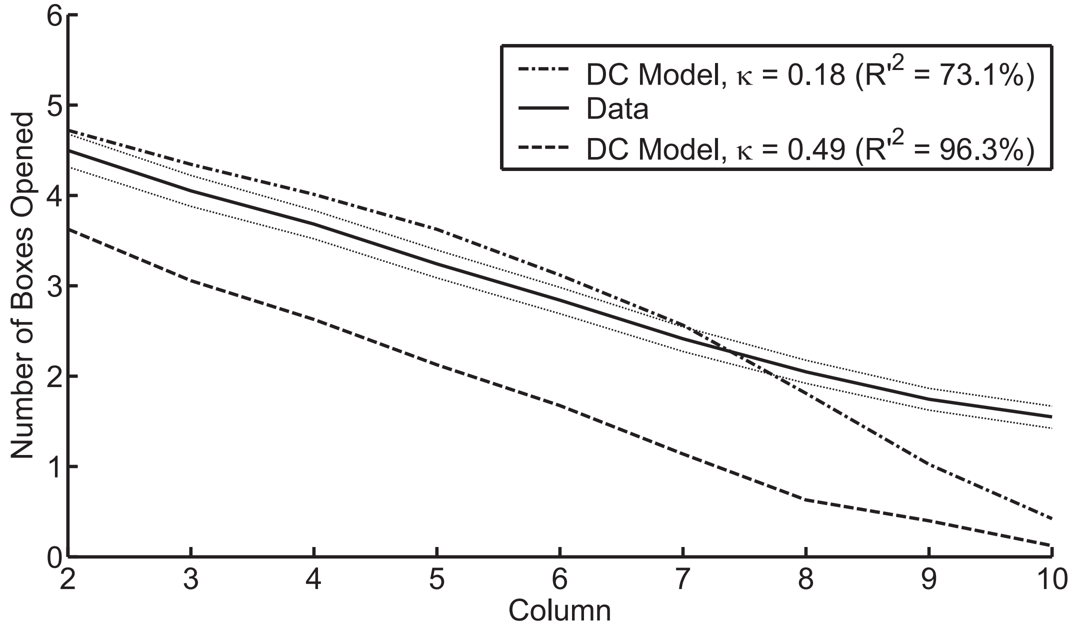


Figure 7 plots the mean number of boxes opened in each column by subjects and by both calibrations of the directed cognition model, for games in which agents choose how much time to allocate from a fixed time budget. Dotted lines show the bootstrapped 95% confidence intervals for the data.

Figure 8: Row Profiles for Games with Endogenous Time Budgets

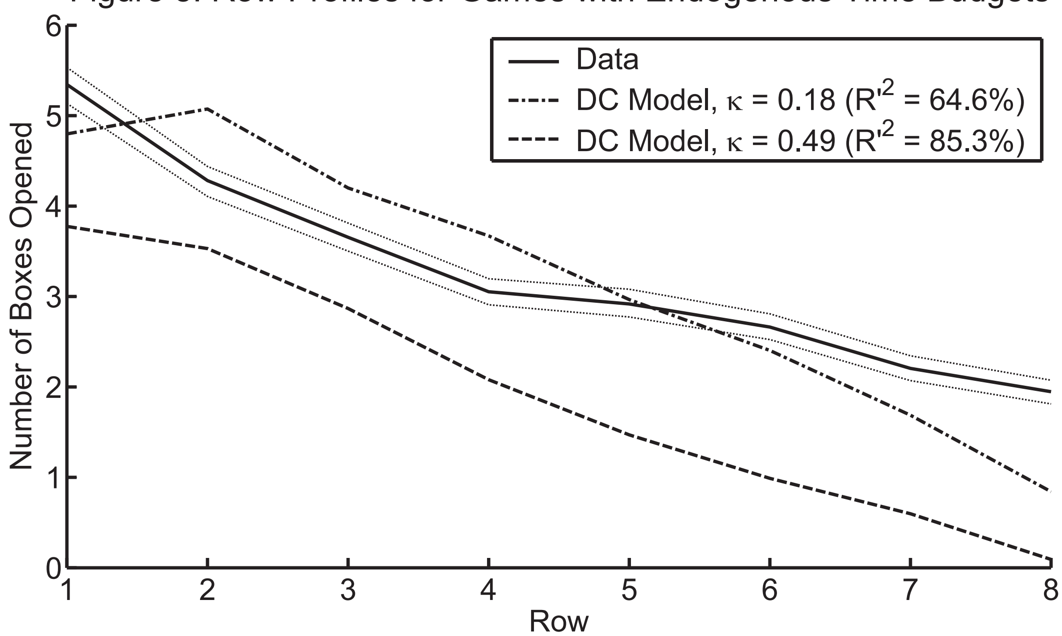


Figure 8 plots the mean number of boxes opened in each row by subjects and by both calibrations of the directed cognition model, for games in which agents choose how much time to allocate from a fixed time budget. Dotted lines show the bootstrapped 95% confidence intervals for the data. The rows are ordered according to the values in the initial column.

Figure 9: Row Profiles (Alternative Ordering) for Games with Endogenous Time Budgets

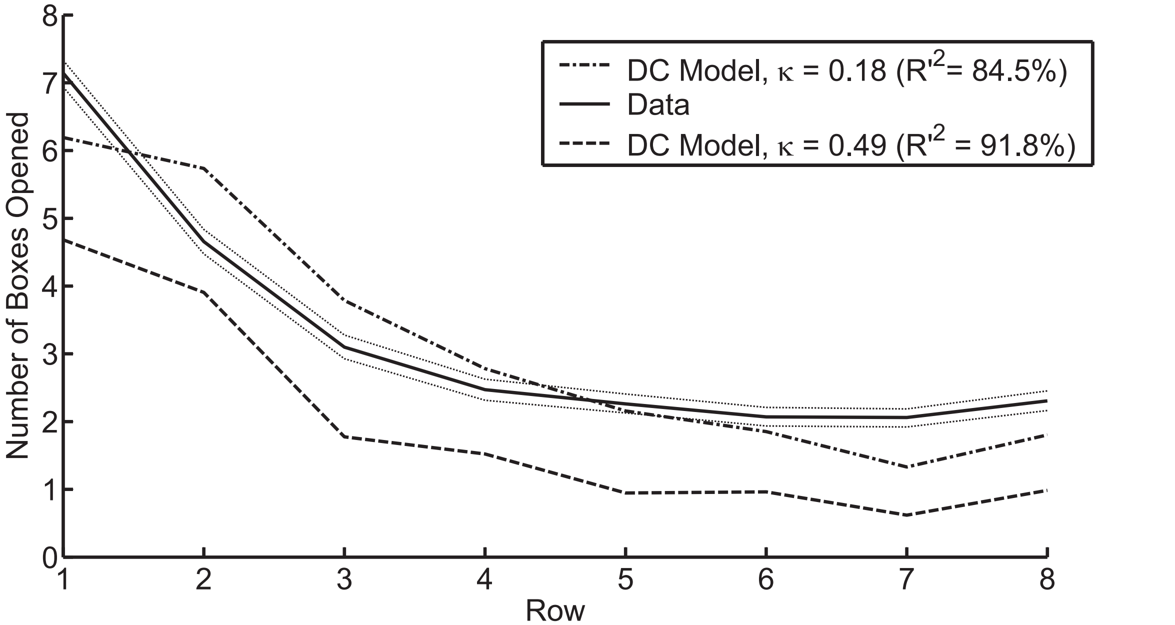


Figure 9 plots the mean number of boxes opened in each row by subjects and by both calibrations of the directed cognition model, for games in which agents choose how much time to allocate from a fixed time budget. Dotted lines show the bootstrapped 95% confidence intervals for the data. The rows are ordered according to all the information available to subjects after they have concluded their search, just prior to making a choice.

Figure 10: Boxes Opened by Game

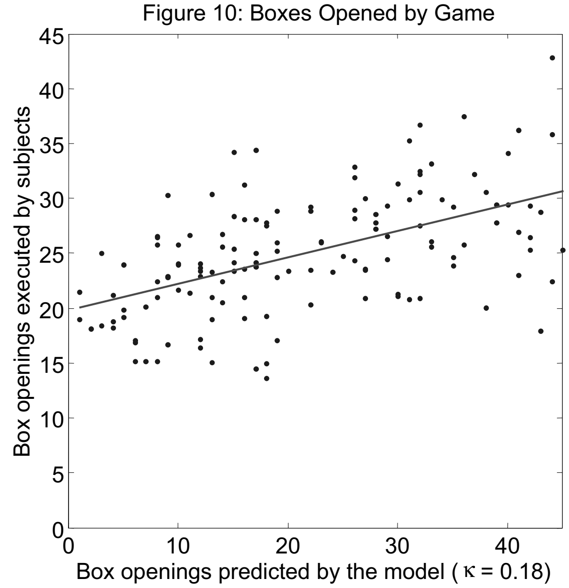


Figure 10 takes each of the 160 game-types played by subjects and compares the number of boxes opened by subjects to the number of box openings predicted by the model, for the games in which agents choose their cross-game time allocations. The correlation between the model and the data is .61.

Figure 11: Expected Measure of Economic Gain "G" vs. Expected Remaining Time

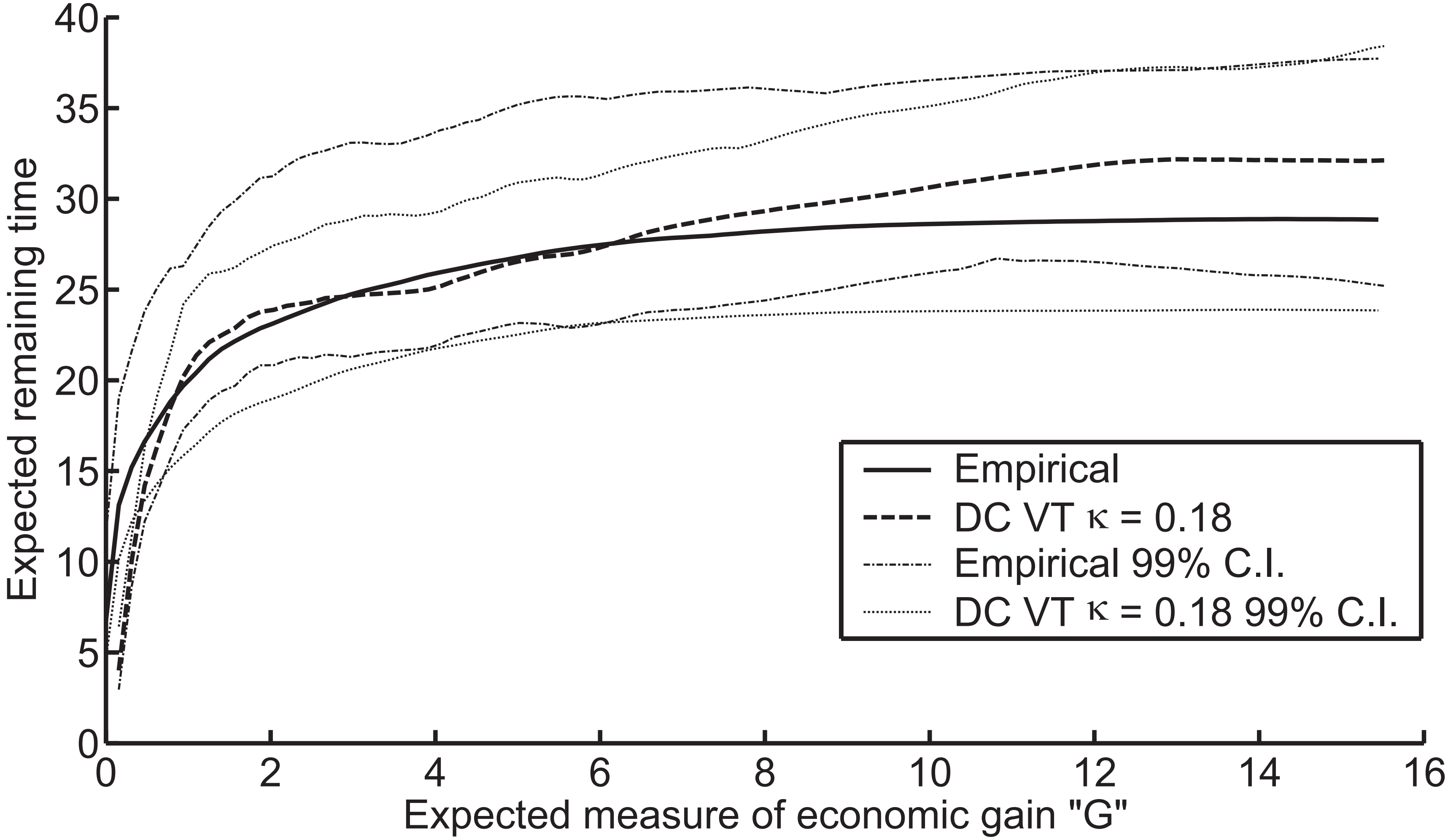


Figure 11 plots non-parametric (kernel) estimates of the expected number of additional box openings in a game, conditional on the current "G" value. "G" is the benefit-to-cost ratio of marginal analysis in the directed cognition model, as defined in equation (4).