

# Appendix to “Information Acquisition: Experimental Analysis of a Boundedly Rational Model”

XAVIER GABAIX<sup>‡</sup><sup>◇</sup>, DAVID LAIBSON<sup>†</sup><sup>◇</sup>, GUILLERMO MOLOCHE<sup>◇</sup>, STEPHEN WEINBERG<sup>†</sup>

<sup>†</sup>HARVARD UNIVERSITY, <sup>‡</sup>MIT, <sup>◇</sup>NBER

October 17, 2005

ABSTRACT. This Appendix provides supplementary results for the experiment as well as the experimental instructions.

## 1. SIMPLE EXPERIMENT

**1.1. Predictions about net payoffs.** We compare empirical outcomes and model predictions for the net payoffs in each game – i.e. the project payoff net of search cost. We calculate the Euclidean distance between a model’s predicted distribution of outcomes and the empirical distribution of outcomes.

For each game  $g = A...E$ , we call  $q_g^M \in \mathbb{R}^7$  the vector of probabilities predicted by model  $M = GW, DC$ , and  $q_g^e$  the vector of empirical probabilities. We define the Euclidean distance between the model and the empirical predictions:

$$Q_g^M = \|q_g^M - q_g^e\|^2 = \sum_{\text{outcome } i} (q_{gi}^M - q_{gi}^e)^2.$$

$$Q_{A-E}^M = \frac{1}{5} \sum_{g=A}^E Q_g^M$$

$$\Delta_g = Q_g^{DC} - Q_g^{GW} \text{ for } g = A...E, A - E$$

A bad performance corresponds to a large distance  $Q$ . We use a standard bootstrap methodology to calculate standard errors.

[Insert Table A about here]

Table A reports the results. In each of the games A-E, DC performs better than GW, but the difference is statistically insignificant ( $t = -1.6$ ).

**1.2. Predictions about project choices.** We compare empirical outcomes and model predictions for the chosen projects in each game – i.e. the investment project that is ultimately taken in the game. We calculate the Euclidean distance between a model’s predicted distribution of outcomes and the empirical distribution of outcomes. Again, we study games  $g = A...E$ .

Call  $H$  the high expected value project,  $L$  the low expected value project, and  $S$  the safe project (with sure-thing payoff \$1). Let  $p$  represent a project’s probability of success. For instance, in Project A,  $p_H = 0.76$ ,  $p_L = 0.09$ . DC recommends exploring  $H$  first, stopping if successful, and otherwise exploring  $L$ . GW recommends exploring  $L$  first, stopping if successful, and otherwise exploring  $H$ . Call  $q_H$  the probability that project  $H$  will be selected in the end – i.e. explored, and, if successful, chosen. The predictions are:

$$\begin{aligned} (q_H, q_L, q_S)^{DC} &= (p_H, (1 - p_H)p_L, (1 - p_H)(1 - p_L)) \\ (q_H, q_L, q_S)^{GW} &= ((1 - p_L)p_H, p_L, (1 - p_H)(1 - p_L)) \end{aligned}$$

We call  $(q_H, q_L, q_S)^e$  the empirical probabilities. We reports two kinds of statistics: an aggregate statistic, and individual statistics. We estimate all the standard errors by bootstrapping over the subjects.

The aggregate statistic is the quadratic distance  $Q$  between the model and the experimental probabilities:

$$Q_g^M = \sum_{i \in \{H, L, S\}} (q_i^M - q_i^e)^2 \text{ in game } g. \quad (1)$$

We form the difference between the distances of DC and GW:

$$\Delta_g = Q_g^{DC} - Q_g^{GW} \quad (2)$$

$$\Delta_{A-E} = \frac{1}{5} \sum_{g=A...E} \Delta_g. \quad (3)$$

If  $\Delta_{A-E}$  is large and positive, then DC predicts choices less well than GW. Table B reports that  $\Delta_{A-E} \cdot \Delta_{A-E}$  is not statistically significant from 0 (see the bottom-right corner of the table). This means that, if one just looks at outcomes, DC and GW are performing equally well. This is why, of course, the paper looks at the actual search strategy. There is less randomness for the search

strategies, and hence more statistical power.

We now turn to the performance of individual probability measures. For instance, take  $q_H$ , in a given game. We want to test if  $(q_H^e - q_H^{DC})^2 > (q_H^e - q_H^{GW})^2$ . To do so, we form:

$$\begin{aligned} L &= (q_H^e - q_H^{DC})^2 - (q_H^e - q_H^{GW})^2 \\ &= (q_H^{DC} - q_H^{GW}) \left( \frac{q_H^{DC} + q_H^{GW}}{2} - q_H^e \right). \end{aligned}$$

So our test statistics associated with  $L > 0$  is the  $t$  ratio:

$$t = \text{sign}(q_H^{DC} - q_H^{GW}) \cdot \frac{\left( \frac{q_H^{DC} + q_H^{GW}}{2} - q_H^e \right)}{\text{standard error on } q_H^e}. \quad (4)$$

DC makes a worse prediction when  $L$  is significantly greater than 0, i.e. when  $t > 1.95$ . We reports those  $t$  ratios in the Table. Of course, when the models make the same predictions (in games F-J),  $L = t = 0$ , so we do not report  $t$  for those games.

**[Insert Table B about here]**

As always, a size adjustment is needed when interpreting one extremal number out of many jointly reported statistics. For instance, the threshold for the maximum of the absolute value of 5 random Gaussians is 2.56, not 1.95. (It is 2.32 for the 1-sided test.) So, the high ratio  $t = 2.1$  for game B should be measured with the cutoff of 2.56, as game  $B$  is chosen to be the game with the highest  $t$  ratio. (The aggregate measure  $\Delta_{A-E}$  is free of this selection bias.)

Finally, and perhaps most importantly, 16% of subjects check the low-probability option even *after* they have already found that the high-probability option is a winner. Such an irrational<sup>1</sup> search strategy is not predicted by either GW or DC. However in this experiment such a strategy generates investment project outcomes that reproduce the predictions of the rational GW model. This additional path to the GW choice provides a critical reason to focus on joint tests of search strategies and project choices, instead of focusing on project choices alone.

---

<sup>1</sup>It is more sensible to check the option with the bigger prize first (the low probability option). This ordering saves search costs when the lottery with the bigger prize turns out to be a winner.

## 2. COMPLEX EXPERIMENT

**2.1. Example calculation of  $G$ .** Consider a game in which the payoffs in column 1 are drawn from a  $N(0, \sigma^2)$  distribution. The variance of payoffs in column  $n$  is  $\sigma_n^2 = \sigma^2 (11 - n) / 10$ , for  $n \in \{1 \dots 10\}$ . Call  $\eta_{mn}$  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.<sup>2</sup> The current value of row  $A$  is

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

Call  $\tilde{a}$  the value of the current best alternative to row  $A$ . In most instances  $\tilde{a}$  is the value of the leading row. If  $A$  is the leading row itself,  $\tilde{a}$  is the value of the second leading row.

The relevant search 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  $x = a - \tilde{a}$ , then the value of the benefit/cost ratio  $G_A$  of row  $A$  is,

$$G_A = \max\{w(x, \sigma_3), w\left(x, \sqrt{\sigma_3^2 + \sigma_5^2}\right) / 2, w\left(x, \sqrt{\sigma_3^2 + \sigma_5^2 + \sigma_9^2}\right) / 3, w\left(x, \sqrt{\sigma_3^2 + \sigma_5^2 + \sigma_9^2 + \sigma_{10}^2}\right) / 4\}$$

with  $x = a - \tilde{a}$ , 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 search 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,

---

<sup>2</sup>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).

if two boxes are opened, the chance of reversal is 11 time greater. The standard 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 search operator may be much more cost effective (when evaluated by the directed cognition algorithm) than the one-box search operator, since the two-box search operator increases the chance of reversal by a factor of 11 but only has twice the search cost.

**2.2. Lower bounds for the performance of the directed cognition algorithm.** We wish to evaluate the payoff consequences of following directed cognition. We compare the DC payoffs to the payoffs that would be realized under Perfect Rationality. Perfect Rationality’s complexity makes it prohibitively hard to evaluate. Here we provide bounds on its performance. 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  $\sigma$ . Each of the 8 rows will have a value which is normal with mean 0 and standard deviation  $\sigma$  times  $\sqrt{\sum_{j=1}^{10} j/10} = \sqrt{5.5}$ . 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:  $\nu \simeq 1.423$ . So the average performance of Zero Cognition Costs will be  $5.5^{1/2}\nu\sigma \simeq 3.34\sigma$ .

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

$$V^* \leq 3.34\sigma. \quad (5)$$

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

**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 (5), 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 92% as well as

Perfect Rationality in exogenous time games.

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

$$\pi(\sigma, \kappa) = \frac{V^{DC}(\sigma, \kappa)}{V^*(\sigma, \kappa)}. \quad (6)$$

By definition  $\pi(\sigma, \kappa) \leq 1$ . 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  $\kappa t$ .  $V^*$  is the expected value under the optimal strategy. Call  $\eta_{A1}, \dots, \eta_{H1}$  the values of the 8 rows in column 1, and  $\xi = \max\{\eta_{A1}, \dots, \eta_{H1}\}$  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  $\xi$  in the first column:  $V^* \leq V^{G'}$ . 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  $V^* \leq V^{G''} + \nu$ .

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  $n \in \{2, \dots, 10\}$  has boxes with variance  $(N+1-n)/N$  with  $N = 10$  columns.  $V^{G''} \leq V^{G'''}$ . The value  $M := V^{G'''}$  of  $G'''$  has the property similar to a Gittins index. 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  $\tau$ . 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  $\eta_{An}$ 's are i.i.d. standard normals. In game  $G'''$  we have  $a_1 = 0$ . After exploring this row, one can either pick the box, which has a value  $a_\tau$ , 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  $\max(a_\tau, M)$ . The value  $M$  of game  $G'''$  satisfies:  $M = \max_{\text{stopping time } \tau} E[\max(a_\tau, M) - \kappa(\tau - 1) \mid a_1 = 0]$ , i.e.

$$0 = \max_{\text{stopping time } \tau} E[\max(a_\tau - M, 0) - \kappa(\tau - 1) \mid a_1 = 0]. \quad (7)$$

Call  $v(a, t, \kappa)$  the value of the game<sup>3</sup> that has explored up to column  $t$ , has current payoff  $a$ , and that at the end of the search  $\tau$  yields a payoff  $\max(a_\tau, 0)$  minus the cognition costs  $\kappa(\tau - t)$ . Formally:

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

for  $a \in \mathbb{R}$ ,  $t \in \{1, \dots, 10\}$ ,  $\kappa > 0$ . Because of (7), 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  $\nu \simeq 1.423$ , 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  $\sigma$  is:  $V^*(\sigma, \kappa) \leq 1.42\sigma + \sigma M(\kappa/\sigma)$ .

Full Rationality does less well than Zero Cognition Costs, so (5) gives:  $V^*(\sigma, \kappa) \leq 3.34\sigma$ .

We combine those two inequalities using the fact that  $V^*$  is convex in  $\kappa$ , which comes from the usual replication arguments. This gives  $V^*(\sigma, \kappa) \leq \bar{V}^*(\sigma, \kappa)$ , where  $\bar{V}^*(\sigma, \kappa)$  is the convex envelope<sup>4</sup> of  $3.34\sigma$  and  $1.42\sigma + \sigma M(\kappa/\sigma)$ , seen as functions of  $\kappa$  on  $\kappa \geq 0$ .

Eq. 6 finally gives

$$\pi(\sigma, \kappa) \geq \underline{\pi} := \inf_{\sigma, \kappa} V^{DC}(\sigma, \kappa) / \bar{V}^*(\sigma, \kappa) \quad (8)$$

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.

**2.3. Learning effects.** Payoffs do not systematically vary with experimental experience. To document this fact, we calculate the average payoff  $X(k)$  of games played in round  $k = 1, \dots, 12$  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. Learning would imply  $\beta > 0$ . We find that in both regressions  $\beta$  is not significantly different from 0. Also, the constant  $\alpha$  is statistically indistinguishable across the two samples. In addition, playing endogenous

<sup>3</sup>In terms of Gabaix and Laibson (2005a), this is the value function of an “ $a$  vs 0” game.

<sup>4</sup>As we take the lower cord between  $3.34\sigma$  and  $1.42\sigma + \sigma M(\kappa_0/\sigma)$ , this gives, formally:

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

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.

**2.4. Memory limitations.** Our data reveal that subjects have memory limitations. Repeat unmaskings represent 10% of box openings. In addition, only 69% of final row choices in exogenous time games match the row with the greatest revealed value; the analogous number is 78% for endogenous time games.<sup>5</sup> In the interest of parsimony, we do not explicitly incorporate imperfect memory into our model.<sup>6</sup> In all of our analyses, we count only the first unmasking of each box.

**2.5. Statistical methodology.** 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 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 calculate Monte Carlo means and standard errors for our *model* predictions. The associated standard errors tend to be extremely small, because 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.<sup>7</sup>

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.

**2.6. Depth at which a row is explored.** We evaluate DC predictions about the depth of continuous row exploration immediately after each empirical row switch. Specifically, we consider

---

<sup>5</sup>For exogenous time games, 85% of row choices correspond to one of the two rows with the greatest revealed value, and 91% of row choices correspond to one of the three rows with greatest revealed value. For endogenous time games 90% of row choices correspond to one of the two rows with the greatest revealed value, and 95% of row choices correspond to one of the three rows with greatest revealed value.

<sup>6</sup>We 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 reopening boxes whose values have been forgotten. An extension of our framework might consider the case in which the technology includes memory capacity constraints or depreciation of memories over time.

<sup>7</sup>In Figures 4-9, the width of the confidence intervals for the models is on average one-half of one percent of the value of the means.



every event in which a subject has just switched rows. We ask the DC algorithm to predict how far this subject will search in the new row before again switching rows. For each event, we difference the predicted DC depth and the actual empirical depth. We find that the mode of this difference is 0. Nevertheless, such exact matches are achieved only 18% of the time. Moreover, 63% of the time, the DC depth prediction is at least two boxes away from the empirical outcome.

Table A. Frequency of net payoffs obtained at the end of each game  
Theoretical predictions, and Empirical values

Game	Euclidean Distances					
	GW	DC	Delta	s.e.	t	
A	0.373	0.194	-0.179	0.109	-1.65	
B	0.369	0.222	-0.147	0.105	-1.40	
C	0.252	0.251	-0.002	0.101	-0.02	
D	0.452	0.190	-0.262	0.119	-2.20	
E	0.337	0.332	-0.004	0.119	-0.04	
F	0.059	0.059	0			
G	0.069	0.069	0			
H	0.032	0.032	0			
I	0.037	0.037	0			
J	0.055	0.055	0			
Mean A-E	0.357	0.238	-0.119	0.074	-1.60	
Mean F-J	0.051	0.051				
Mean A-J	0.204	0.144	-0.059	0.037	-1.60	

"DC distance" is the sum of the squared distance in probabilities between the DC prediction and the empirical choice probabilities.

Delta=DC distance - GW distance

The mean Euclidean distances of the DC and GW models are not statistically different.

Table B. Frequency of a subject selecting the high Expected Value, low expected value, or riskless project  
Theoretical predictions, and Empirical values

Game	High EV Project					Low EV project					Riskless Project				Euclidean Distances				
	GW	DC	Empirical	s.e.	t	GW	DC	Empirical	s.e.	t	GW	DC	Empirical	s.e.	GW	DC	Delta	s.e.	t
A	0.692	0.760	0.729	0.039	-0.07	0.090	0.022	0.062	0.021	0.30	0.218	0.218	0.209	0.036	0.0022	0.0027	0.0005	0.0071	0.06
B	0.703	0.790	0.659	0.041	2.14	0.110	0.023	0.101	0.027	1.27	0.187	0.187	0.240	0.037	0.0049	0.0261	0.0212	0.0100	2.12
C	0.655	0.720	0.698	0.041	-0.25	0.090	0.025	0.054	0.020	-0.17	0.255	0.255	0.248	0.038	0.0031	0.0014	-0.0017	0.0066	-0.26
D	0.713	0.810	0.814	0.033	-1.59	0.120	0.023	0.062	0.021	-0.45	0.167	0.167	0.124	0.029	0.0155	0.0034	-0.0120	0.0095	-1.27
E	0.748	0.850	0.721	0.039	2.00	0.120	0.018	0.047	0.018	-1.25	0.132	0.132	0.233	0.037	0.0162	0.0276	0.0113	0.0101	1.12
F	0.480	0.480	0.395	0.043		0.385	0.385	0.341	0.042		0.135	0.135	0.264	0.039	0.0256	0.0256	0		
G	0.340	0.340	0.279	0.039		0.462	0.462	0.481	0.044		0.198	0.198	0.240	0.038	0.0058	0.0058	0		
H	0.520	0.520	0.450	0.044		0.355	0.355	0.403	0.043		0.125	0.125	0.147	0.031	0.0078	0.0078	0		
I	0.390	0.390	0.357	0.043		0.427	0.427	0.434	0.044		0.183	0.183	0.209	0.036	0.0019	0.0019	0		
J	0.850	0.850	0.829	0.033		0.000	0.000	0.016	0.011		0.150	0.150	0.155	0.031	0.0007	0.0007	0		
Mean A-E	0.702	0.786	0.724	0.019	1.05	0.106	0.022	0.065	0.010	0.10	0.192	0.192	0.211	0.018	0.0084	0.0122	0.0038	0.0043	0.89
Mean F-J	0.516	0.516	0.462	0.021		0.326	0.326	0.335	0.020		0.158	0.158	0.203	0.019	0.0083	0.0083	0		
Mean A-J	0.609	0.651	0.593	0.015	1.05	0.216	0.174	0.200	0.011	0.10	0.175	0.175	0.207	0.014	0.0084	0.0103	0.0019	0.0021	0.89

"DC distance" is the sum of the squared distance in probabilities between the DC prediction and the empirical choice probabilities.

Delta=DC distance - GW distance

The mean Euclidean distance of the DC and GW models are not statistically different.

## Experiment 1 Materials

### Computer Instructions

Thanks for participating in this experiment. During the experiment you will be asked to choose among groups of "lottery tickets." Tickets have different probabilities of paying off. Consider the following set of tickets:

Ticket 1:	50% chance of \$20	Winner: Unknown	<u>Find Out Whether Ticket is a Winner</u>
Ticket 2:	75% chance of \$10	Winner: Unknown	<u>Find Out Whether Ticket is a Winner</u>
Ticket 3:	100% chance of \$1	Winner: <b>Yes</b>	<b><u>Take Ticket -- worth \$1</u></b>

Here, ticket 1 has a 50% chance of paying \$20. To be more precise, it pays off \$20 with a 50% chance and pays off \$0 with a 50% chance. Similarly, ticket 2 pays off \$10 with a 75% chance and \$0 with a 25% chance. Ticket 3 always pays \$1.

For \$1 you can find out whether an "unknown" ticket is a winner or not by clicking on "Find Out Whether Ticket is a Winner" in the far right hand column. You may investigate as many tickets as you wish, at a cost of \$1 each.

At any point you may stop learning and end the round by choosing ONE ticket. Click on "Take Ticket" in the far right hand column. The "Take Ticket" option will be available only for tickets which have been revealed to be winners. Once you select a ticket, we will calculate your winnings for that round as the value of the selected ticket minus \$1 for each "unknown" ticket that you investigated. We will tell you your winnings for the round and then reset the screen for the next round.

You will start with a practice round that illustrates the process. The experiment will then begin. The experiment will consist of 10 games (in addition to the practice round). You will get to keep your winnings (net of learning costs) in 3 of the 10 games selected at random.

# BEGIN

## Practice Round

Ticket 1:	50% chance of \$20	Winner: Unknown	<a href="#">Find Out Whether Ticket is a Winner</a>
Ticket 2:	75% chance of \$10	Winner: Unknown	<a href="#">Find Out Whether Ticket is a Winner</a>
Ticket 3:	100% chance of \$1	Winner: <b>Yes</b>	<a href="#">Take Ticket -- worth \$1</a>

Clicking on "Find Out Whether Ticket is a Winner" costs \$1 and displays whether the ticket has won or not.

Clicking on "Take Ticket" ends the round and computes your net winnings for the round.



## Round 1 of 10

Ticket 1:	74% chance of \$11	Winner: Unknown	<u>Find Out Whether Ticket is a Winner</u>
Ticket 2:	52% chance of \$18	Winner: Unknown	<u>Find Out Whether Ticket is a Winner</u>
Ticket 3:	100% chance of \$1	Winner: <b>Yes</b>	<u><b>Take Ticket -- worth \$1</b></u>

Clicking on "Find Out Whether Ticket is a Winner" costs \$1 and displays whether the ticket has won or not.

Clicking on "Take Ticket" ends the round and computes your net winnings for the round.

## Round 2 of 10

Ticket 1:	9% chance of \$21	Winner: Unknown	<a href="#">Find Out Whether Ticket is a Winner</a>
Ticket 2:	76% chance of \$10	Winner: Unknown	<a href="#">Find Out Whether Ticket is a Winner</a>
Ticket 3:	100% chance of \$1	Winner: <b>Yes</b>	<a href="#">Take Ticket -- worth \$1</a>

Clicking on "Find Out Whether Ticket is a Winner" costs \$1 and displays whether the ticket has won or not.

Clicking on "Take Ticket" ends the round and computes your net winnings for the round.

## Round 3 of 10

Ticket 1:	48% chance of \$22	Winner: Unknown	<a href="#">Find Out Whether Ticket is a Winner</a>
Ticket 2:	100% chance of \$1	Winner: <b>Yes</b>	<a href="#">Take Ticket -- worth \$1</a>
Ticket 3:	74% chance of \$11	Winner: Unknown	<a href="#">Find Out Whether Ticket is a Winner</a>

Clicking on "Find Out Whether Ticket is a Winner" costs \$1 and displays whether the ticket has won or not.

Clicking on "Take Ticket" ends the round and computes your net winnings for the round.



## Round 4 of 10

Ticket 1:	79% chance of \$10	Winner: Unknown	<a href="#">Find Out Whether Ticket is a Winner</a>
Ticket 2:	100% chance of \$1	Winner: <b>Yes</b>	<a href="#">Take Ticket -- worth \$1</a>
Ticket 3:	11% chance of \$19	Winner: Unknown	<a href="#">Find Out Whether Ticket is a Winner</a>

Clicking on "Find Out Whether Ticket is a Winner" costs \$1 and displays whether the ticket has won or not.

Clicking on "Take Ticket" ends the round and computes your net winnings for the round.

## Round 5 of 10

Ticket 1:	100% chance of \$1	Winner: <b>Yes</b>	<b><u>Take Ticket -- worth \$1</u></b>
Ticket 2:	34% chance of \$24	Winner: Unknown	<u>Find Out Whether Ticket is a Winner</u>
Ticket 3:	70% chance of \$9	Winner: Unknown	<u>Find Out Whether Ticket is a Winner</u>

Clicking on "Find Out Whether Ticket is a Winner" costs \$1 and displays whether the ticket has won or not.

Clicking on "Take Ticket" ends the round and computes your net winnings for the round.

## Round 6 of 10

Ticket 1:	70% chance of \$9	Winner: Unknown	<a href="#">Find Out Whether Ticket is a Winner</a>
Ticket 2:	100% chance of \$1	Winner: <b>Yes</b>	<a href="#">Take Ticket -- worth \$1</a>
Ticket 3:	39% chance of \$25	Winner: Unknown	<a href="#">Find Out Whether Ticket is a Winner</a>

Clicking on "Find Out Whether Ticket is a Winner" costs \$1 and displays whether the ticket has won or not.

Clicking on "Take Ticket" ends the round and computes your net winnings for the round.

## Round 7 of 10

Ticket 1:	100% chance of \$1	Winner: <b>Yes</b>	<b><u>Take Ticket -- worth \$1</u></b>
Ticket 2:	12% chance of \$20	Winner: Unknown	<u>Find Out Whether Ticket is a Winner</u>
Ticket 3:	85% chance of \$12	Winner: Unknown	<u>Find Out Whether Ticket is a Winner</u>

Clicking on "Find Out Whether Ticket is a Winner" costs \$1 and displays whether the ticket has won or not.

Clicking on "Take Ticket" ends the round and computes your net winnings for the round.



## Round 8 of 10

Ticket 1:	85% chance of \$8	Winner: Unknown	<u>Find Out Whether Ticket is a Winner</u>
Ticket 2:	9% chance of \$10	Winner: Unknown	<u>Find Out Whether Ticket is a Winner</u>
Ticket 3:	100% chance of \$1	Winner: <b>Yes</b>	<u><b>Take Ticket -- worth \$1</b></u>

Clicking on "Find Out Whether Ticket is a Winner" costs \$1 and displays whether the ticket has won or not.

Clicking on "Take Ticket" ends the round and computes your net winnings for the round.

## Round 9 of 10

Ticket 1:	100% chance of \$1	Winner: <b>Yes</b>	<b><u>Take Ticket -- worth \$1</u></b>
Ticket 2:	12% chance of \$18	Winner: Unknown	<u>Find Out Whether Ticket is a Winner</u>
Ticket 3:	81% chance of \$10	Winner: Unknown	<u>Find Out Whether Ticket is a Winner</u>

Clicking on "Find Out Whether Ticket is a Winner" costs \$1 and displays whether the ticket has won or not.

Clicking on "Take Ticket" ends the round and computes your net winnings for the round.

## Round 10 of 10

Ticket 1:	9% chance of \$23	Winner: Unknown	<u>Find Out Whether Ticket is a Winner</u>
Ticket 2:	100% chance of \$1	Winner: <b>Yes</b>	<u><b>Take Ticket -- worth \$1</b></u>
Ticket 3:	72% chance of \$13	Winner: Unknown	<u>Find Out Whether Ticket is a Winner</u>

Clicking on "Find Out Whether Ticket is a Winner" costs \$1 and displays whether the ticket has won or not.

Clicking on "Take Ticket" ends the round and computes your net winnings for the round.



# Thank You

We have randomly selected that you will receive your payout from games 5,2,4.  
Your total winnings are \$27

Please answer the following questions.

Were the instructions in this experiment clear?

Yes: ☐ No: ☐

What do you think this experiment was about?

Please describe your strategy in these games:

What is your age?

Are you in high school or college?

High School: ☐ College: ☐

Have you taken any statistics?

Yes: ☐ No: ☐

Have you taken any economics, decision theory, or game theory courses?

Yes: ☐ No: ☐

Submit Form

Total winnings are \$27



## Experiment 2 Materials

### Written Instructions

Please take 5-6 minutes to read these instructions. These written instructions describe the decisions you will be making in this experiment. You will also see a computer demo on a laptop computer that will explain how the computer works. After the demo, you will spend approximately 35 minutes making decisions on the laptop.

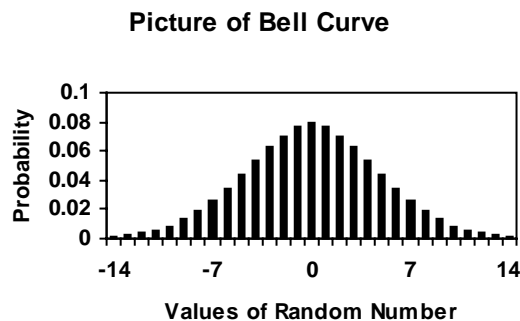
In this experiment, we are studying how people make decisions under time limits. We will ask you to solve several problems where the choices you make can increase (or decrease) the amount of money that you will earn in this experiment. You will have a limited amount of time for each problem. Each problem will consist of eight rows of ten random numbers. Boxes will cover nine of the random numbers in each row; you can see these numbers by clicking on them with the mouse. You will always be able to see the first number in each row. For the sake of these instructions, we will show you pictures where the boxes have all been opened; the computer demo will show you how the boxes work. With all the boxes opened, two rows might look like this:

a	-15	-28	18	21	-18	-24	7	21	-14	2
b	57	-24	-11	48	1	4	18	18	-19	9

We will pay you the sum of the row you pick (in pennies), so your best choice will be the row with the highest sum. We will discuss payment in more detail at the end of these instructions.

#### I. Random Numbers

You may find it useful to understand how we generate the random numbers in the problems you will solve. We use bell curves to generate these numbers.



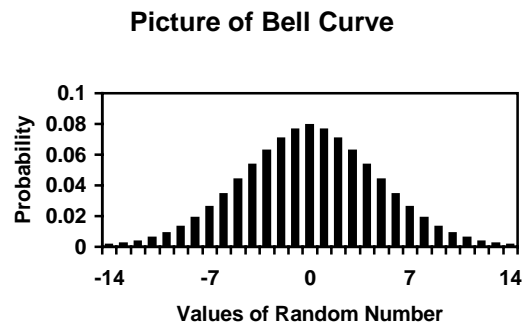
A bell curve, also called a “normal distribution,” describes the way that random numbers are spread around some central value. In this experiment, the random numbers cluster around 0. A bell curve has three key properties:

- The higher the level of the bell curve for a particular number, the more likely we are to draw that number.
- There is an equal chance of getting a positive value or a negative value.

- The curve is highest at zero and then slopes downward on either side of zero, so the bell curve is more likely to generate numbers the closer they are to zero. For example, 1 should be more likely than 11.

How do we generate random numbers from bell curves? The bars show the percentage of the time that a random number takes a given integer value. There's only a small chance of getting any single number. For example, if we generated 100,000 random numbers from this bar graph, 7.6% of them should be  $-1$ , and 4.3% of them should be 5. For a smaller number of draws, these percentages may be only approximations.

Here is another look at the same bell curve.



## II. Ranges of Random Numbers

Because we are unlikely to generate any single number, it is easier to think about ranges of numbers. If we generate lots of numbers from this bell curve, 95% of the numbers should fall between  $-10$  and  $10$ .

For example, the following twelve numbers were generated from this bell curve:

-1  
12  
6  
-9  
1  
-4  
2  
2  
-5  
0  
5  
-7

Notice that six of the numbers are positive; five of the numbers are negative; and one of the numbers is 0. Eleven of the numbers are between  $-10$  and  $10$ , the 95% range. The lone exception is the 12.

For each problem, we will generate random numbers from different bell curves. **We will describe these different bell curves to you by giving you the 95% range, the range within which we expect about 95% of the random numbers to fall.** Remember that while about 95% of the numbers will be

within the 95% range, most of the numbers will actually be considerably closer to 0. In the few cases where a number falls outside the 95% range, it is most likely to fall very close to the range.

Let's recap the main points you need to remember:

**1) The wider the 95% range is, the more likely you are to see very big numbers, either positive or negative.**

**2) The random numbers are just as likely to be positive as negative.**

Please keep these facts in mind. We will now describe the problems you will encounter in this experiment.

### III. The Problems

Each problem will consist of 8 rows of 10 random numbers. For example, a problem might look like the following, if we open all the boxes so you can see all the numbers:

a	-15	-28	18	21	-18	-24	7	21	-14	2
b	57	-24	-11	48	1	4	10	10	-19	9
c	-13	-31	-11	-40	-29	-6	10	8	-25	9
d	39	14	14	-10	30	1	-19	-13	-28	-7
e	-23	15	31	41	5	-24	-30	6	6	1
f	31	-24	25	1	-14	-15	8	-5	-5	5
g	20	81	11	-19	-20	1	-51	-5	-5	8
h	9	8	26	22	29	11	-24	-26	6	-12

You want to choose the row you think has the highest sum.

There are 10 columns in each problem; in the example above, column 1 contains the numbers -15, 57, -13, 39, -23, etc. The numbers in column 1 are not in boxes; numbers in the other columns are in boxes. In the actual problems, you will need to click on the boxes with your mouse to see those numbers; we will show you how the boxes work in the computer demo.

**We generate each column from its own bell curve with its own 95% range.** Within a single row, such as row (a), every number is generated from a different bell curve.

The leftmost column (-15, 57, -13, 39, etc) comes from a bell curve that has a 95% range of -90 to 90. The next column (-28, -24, -31, 14, etc) comes from another bell curve, which has a slightly narrower 95% range of -86 to 86. We thus expect numbers from the second column, on average, to be a little closer to 0. The 95% ranges narrow gradually from left to right, so the rightmost column (2, 9,

9, -7, etc) comes from a tenth bell curve, which has a 95% range of -25 to 25. **This means the leftmost numbers are most likely to be far from 0, and as you move from left to right the numbers should tend to be closer to 0.**

To review, if you look down the columns of the problem, every number is generated from the same bell curve. As you look left-to-right along the rows of the problem, every number is generated from a different bell curve, going from widest to narrowest. **Before each problem, you will be told the 95% range for the bell curves generating the leftmost and rightmost random numbers in that problem.**

Feel free to skip the following paragraph, provided for those who prefer a more technical explanation:

All you need to understand is that the ranges narrow gradually and smoothly from left to right. In case you are interested and have taken a statistics class, we will tell you the actual process used to generate these ranges. The bounds on the 95% confidence intervals we report are 1.96 times the standard deviation used for the bell curve; the statistics are based on the squares of the standard deviations, or the variances. Moving from left to right the variances fall linearly in increments of 1/10ths. Therefore, the squares of the bounds fall linearly. For example, if the leftmost range is -10 to 10, then its bound is 10 and the square is 100. The square of the bound of the second range is 9/10 times 100, or 90. The actual range comes from the square root of this, or 9.5. Moving to the 3<sup>rd</sup> column, we take 8/10 of the bound of the first range 100, and then take the square root, for a bound of 8.9. (If this paragraph is confusing, don't worry about it.)

We have randomly generated all the problems as described above; furthermore, we have also randomized the order of the problems, and even the order of the different parts of the experiment. You do not need to spend any mental energy looking for patterns in the problems or other tricks. Economics has a professional standard against deceiving our subjects in any way. The editors of our journals require us to be completely straightforward with you.

#### IV. Payment

You will begin the experiment with a payment account of \$5. **After each problem, we will add the sum of the row you selected to your account (in pennies).** Remember that numbers in the first column, though not covered by boxes, still count towards your payment. Suppose the example problem were the first problem in the experiment. At the beginning of the problem, you have your starting payment account of \$5.00. In the example problem, row (a) has a sum of -30. If you selected row (a), your account would decrease from \$5 to \$4.70. You would then move on to the next problem with an account worth \$4.70. Row (b) has a sum of 85. If you selected row (b) instead, your account would increase from \$5 to \$5.85, and you would start the next problem with an account worth \$5.85. **At the end of the experiment we will record the value of your payment account and send you a check for that amount within 72 hours.** So suppose you played a total of 50 problems and in each problem you chose a row with a sum of 50 (worth \$.50). Then your final account would be \$.50 times 50, plus the initial \$5, so you would receive a check for \$30. To recap: Your payoff in each problem will be the sum of the boxes in the row you choose, and we will pay you \$5 plus the sum of your payoffs from every problem you finish.

There is one friendly exception to this calculation. We guarantee you a payment of \$5 if you finish the experiment, even if the total payoff from your choices would be less than \$5. We believe if you try hard you can earn \$10 to \$30. You may choose to leave at any time.

You will now practice on a computer demo. Please pay close attention to the demo instructions, which explain the various twists and turns of our user interface. **Please raise your hand for an experimental assistant.**

# Computer Instructions

Subjects received the following instructions on a laptop computer.

## DECISION MAKING EXPERIMENT

You will now begin the computer portion of the experiment.

You should have just finished reading a written set of instructions that discusses bell-curves. If not, please raise your hand and notify an assistant.

Before beginning, you will first be guided through some screens to help you get comfortable with this software.

[next screen]

Remember, you will be shown 8 rows of 10 numbers. Your goal is to choose the row that you think has the highest sum.

To record how you approach these problems, we are using a computer program that will cover the numbers with boxes. Only the left-most numbers in each row will not be covered and will always be visible. To look at the numbers covered by boxes, you must use your mouse to click on the box that you want to see.

Each box covers one number.

[next screen]

To open a box, move the cursor into the box and click the LEFT mouse button.

To close a box, click on the RIGHT mouse button. You do not need to have the cursor in a box to close it. Only one box can be open at a time.

After deciding which row you think has the highest sum, close all of the boxes by right clicking, and then make a choice by clicking on the corresponding button at the bottom of the screen. You MUST close all the boxes before making a choice. You will then be asked to confirm your selection. The buttons at the bottom of the screen are labeled with the letter and initial value for each row.

[next screen]

Please remember the following main points:

1. To open a box -> left click with your mouse  
To close a box -> right click with your mouse

NOTE: your cursor can be anywhere on the screen to close a box

2. When you are ready to choose a row with the highest sum, first close all of the open boxes by right clicking. Then click on the starting value of the row you believe to have the highest sum and confirm your choice.

3. You will be asked to confirm your choice. To do this, click on the long horizontal button on the bottom of the screen.

4. You cannot make a choice unless all of the boxes are closed.  
Remember to right click to close any open box.

The next screen is a sample for you to practice opening and closing the boxes.

THE NEXT SCREEN IS PRACTICE #1

[The next screen is a demo.]

[Note: this text applies to subjects who play the Exogenous Time segment first, followed by the Endogenous Time segment. This ordering was randomly determined for each subject.]

In Part I of the experiment, you will have a fixed amount of time for each problem. You will have a different amount of time for each game. There will be a round clock in the top right corner of the screen that will show you how much time you have left for each game.

When the clock runs out, all the boxes will close and you will hear a beep indicating that your time has expired. After you hear the beep and are prompted to make a selection, choose the row you think has the highest sum.

If you make a choice before the time has expired, you will be returned to the screen for the remaining amount of time.

WARNING: If there are any boxes open, the clock will temporarily pause  
...but the time will continue to countdown!!!

[next screen]

The next screen will let you practice making a selection with a time restriction.

You will have 1 1/2 minutes to practice. (This is more time than you will have in the actual problems)

You must wait for the entire 1 1/2 minutes before making a selection. If you make a choice before the time has expired, you will be returned to the screen for the remaining amount of time.

## THE NEXT SCREEN IS PRACTICE #2

[The next screen is a demo.]

There will be a 20-second delay between each problem. During these delays, you will see a "buffer screen" instead of a math problem.

During the buffer screen you will be told:

1. The 95% ranges of the left-most columns and right-most columns.
2. The time limit of the next screen
3. How much money you have earned

You must wait for 20 seconds during the buffer screen. After the time has expired and you hear a beep, you will be able to advance to the next problem.

[next screen]

The next practice will involve a 20 second buffer screen, followed by a 45 second fixed time problem.

Remember, you must wait for the time to expire before making your choice.

## THE NEXT SCREEN IS PRACTICE #3

[The next screen is a demo.]

You will now begin Part I of the experiment.

You will have 12 problems with randomly selected time limits. For these problems, you must use the entire time before making your selection.

Remember:

- Look on the buffer screens to know how much time you will have for the next game
- You must wait for the time to expire before making your choice

- The higher the sum of the rows you choose, the more money you will earn in this experiment.

GOOD LUCK!!!!

THE NEXT PROBLEMS WILL ALL COUNT TOWARDS YOUR PAYOFF!!!

[The subjects would then begin the experiment. At the end, they take the following survey.]

You have now completed the experiment!

Thank you for your time.

Please take a few minutes to answer a few questions.

[next screen]

What would you say best describes your current concentration or expected concentration at Harvard?

- a. Economics
- b. History or History and Literature
- c. Humanities
- d. Language
- e. Math, Statistics, or Applied Math
- f. Psychology
- g. Other Natural Sciences
- h. Other Social Sciences

[next screen]

Which of the following years do you expect to graduate from Harvard?

(If you are advanced standing, please choose the year you expect to finish.)

- a. 2000
- b. 2001
- c. 2002
- d. 2003
- e. 2004
- f. 2005



g. Other

[next screen]

Which of the following best describes your class?

- a. First Year
- b. Sophomore
- c. Junior
- d. Senior

[next screen]

How old are you?

- a. under 18
- b. 18
- c. 19
- d. 20
- e. 21
- f. 22
- g. over 22

[next screen]

How would you best describe your gender?

- a. Female
- b. Male

[next screen]

We would like to know about your statistical background. Please give us your best assessment of your previous coursework

Pick the HIGHEST number that applies to you.

- a. I have never taken a statistics course.
- b. I took a statistics class in high school.
- c. I took an introductory statistics course such as statistics 100 or statistics 104, or I took the AP statistics test in high school.

d. I have taken a higher-level statistics course such as statistics 110 or higher, economics 1123, or government 1000.

e. I have taken more than one advanced level statistics courses.

[next screen]

Please give your best estimate of your statistical ability, compared to all Harvard undergraduates.

a. I have below average skills in statistics for a Harvard undergraduate.

b. I have about average skills in statistics for a Harvard undergraduate.

c. I have above average skills in statistics for a Harvard undergraduate

[next screen]

Please raise your hand for assistance.

Thank you.

[An experimental assistant must enter a special code to access the next screen.]

Please open the sealed envelope given to you at the beginning of the experiment. Please sit at a desk without a computer and take a few moments to answer the questions found inside the envelope. Your answers will be helpful for us to study how you approached the problems in this experiment.

Final score:

[The subject's final score (i.e. total payment) would now be displayed.]

# Questionnaire

Before leaving, subjects filled out the following written questionnaire.

Subject #:

## Questionnaire

1. Please describe how you approached the problem of finding the best rows in the problems you just analyzed. Your detailed response to this question will be very helpful to us. Please feel free to use the back of this form if needed.
2. Has any part of this experiment been confusing? Please explain.