Searching for the Sunk Cost Fallacy

By Daniel Friedman*, Kai Pommerenke, Rajan Lukose,
Garrett Milam and Bernardo Huberman

July 2004

Abstract

We seek to isolate in the laboratory factors that encourage and discourage the sunk cost fallacy. Subjects play a computer game in which they decide whether to keep digging for treasure on an island or to sink a cost (which will turn out to be either high or low) to move to another island. The research hypothesis is that subjects will stay longer on islands that were more costly to find. Nine treatment variables are considered, e.g. alternative visual displays, whether the treasure value of an island is shown on arrival or discovered by trial and error, and alternative parameters for sunk costs. The data reveal a surprisingly small and erratic sunk cost effect that is generally insensitive to the proposed psychological drivers.

Acknowledgements

The National Science Foundation funded this work under grant IIS-9986651. It builds on intellectual contributions by undergraduate researchers Dan Goldwasser and Mike McCaffrey, graduate research assistant Ralf Hepp, and programmers Nitai Farmer and Kannan Goundan. We also benefited from comments by numerous colleagues including Joshua Aizenman, Dan Ariely, Rachel Croson, Eric Eyster, Rob Fairlie, Federico Ravenna, Nirvikar Singh, and Donald Wittman. The usual caveat applies.

***Corresponding author.**

Daniel Friedman	(831) 459-4981 (fax 459-5900)
Professor of Economics	http://cash.ucsc.edu/Friedman/
212 Social Sciences I	dan@ucsc.edu

University of California, Santa Cruz 95064

1. Introduction

A cost is *sunk* when it cannot be recovered. Once a cost is sunk, it has no effect on the incremental payoffs of future decisions, and therefore plays no role in rational choice. Indeed, sunk costs play no role in any outcome-oriented process, rational or otherwise.

Economists generally assume that people are rational, and pride themselves on internal consistency. It is therefore remarkable that they devote so much class time and so many textbook pages to teaching undergraduate and MBA students to ignore sunk costs. A favorite anecdote is a concertgoer who realizes after the first five minutes that the show is horrible, but sticks around just to "get his money's worth" from his purchase of admission ticket. The textbooks explain that expenditures on the ticket are sunk and thus irrelevant. The student is warned never in her career to commit the *sunk cost fallacy* of taking some otherwise undesirable action simply because of a sunk cost, e.g., investing another million dollars on an unprofitable product line just because ten million has already been invested.

So who is right, the economist model builders who assume rationality, or the economist teachers and textbook writers who think that hard work is needed to stamp out the fallacy? How widespread really is the sunk cost fallacy? Our own interest in these questions began with a practical issue in e-commerce: is it true, as claimed by several observers, that people stay at a website longer when it takes longer to download?

Our investigation encountered many surprises. The first is the slenderness of published evidence for the fallacy. We will show in the next section that it is easy to rationalize the supposedly fallacious choices featured in most studies and anecdotes.

We therefore devised a direct laboratory test inspired by the e-commerce issue. We put subjects in front of a computer screen, present them sequentially with "islands" that contain various amounts of "buried treasure", and grant them a limited number of mouse clicks for uncovering the treasure. To get to a new island the subject must sink a cost that will turn out to be either high or low. The sunk cost fallacy is present if subjects expend more of their click budget in high cost islands than in low cost islands.

The next surprise was the difficulty in demonstrating the fallacy. Our initial treatments produced essentially the same distribution of clicks on low and high cost islands. Following advice of colleagues, we tried many new treatments and created a design capable of detecting very small effects. The most recent data confirm a sunk cost effect, but it is much smaller and more erratic than we had expected. Variables representing rational choice are much more powerful in explaining the data than any of the psychological variables we have investigated.

The next section summarizes existing empirical evidence on sunk costs, emphasizing gaps in the literature. We then describe our experiment, sketch optimal choice, obtain testable hypotheses, and present the results. Appendix A derives optimal search for our task, and Appendix B reproduces the instructions to subjects.

2. Existing Evidence

Recent news stories suggest that the sunk cost fallacy exists on a grand scale, but at the same time they underscore ambiguities. Was the US government's final decision to invade Iraq in March 2003 a sunk cost fallacy? Iraq's dictator seemed ready to agree to very intrusive weapons inspections and to the placement of tens of thousands of NATO troops, meeting the stated goals of the US. On the other hand, in readying the attack, the US had already spent tens of billions of dollars and disrupted the lives of more than one hundred thousand soldiers. Some commentators argued that these sunk costs precluded calling off the invasion. Several other interpretations are equally plausible, however. For instance, US policymakers may have believed that a last minute cancellation of the invasion would hurt their credibility.

The loss of space shuttle Columbia in February 2003 brought to mind numerous previous decisions to continue the NASA's shuttle program. From its inception in the 1970s, the shuttle was criticized as extremely cost ineffective and dangerous. Yet each time its supporters pointed to the lives and dollars already spent as a reason for continuing the program, and so far Congress has always agreed (Economist, 2003). Again there are other interpretations. For example, the incentives facing NASA managers (and Congress) may not push them towards safe and cost effective space programs; the interests of contractors and other clienteles may be more urgent. And admitting a huge mistake might not be good for their future careers or for their mentors' place in history.

Psychologists have studied the fallacy for several decades (e.g., Staw, 1976; Bazerman, 1986, chapter 4; see also Thaler, 1980), often referring to it as "irrational escalation of commitment." The underlying mechanism mentioned in older papers is cognitive dissonance (Festinger, 1957) or self-justification (Aronson, 1968), but more recent discussions often tie it to prospect theory (Kahneman and Tversky, 1979), specifically to a fixed reference point and loss aversion.

Most of the evidence consists of responses to hypothetical survey questions. For example, subjects are asked to imagine that they have spent \$50 on a ticket for event A and \$100 on a ticket for event B (e.g., A and B are ski weekends in Wisconsin and Michigan). They are also to imagine that they really prefer A to B, that they have just discovered that the events are mutually exclusive, and that the tickets have no salvage value. When asked which event they would then choose, about half the subjects select the more expensive but less preferred alternative B. The psychologist authors urge the interpretation that their subjects respond more strongly to the \$50 difference in sunk costs than to the "true" preferences. Other interpretations of such evidence are that subjects attend more to their actual homegrown preferences between A and B, or to the impression they make on the person asking the question.

Psychologists report that the sunk cost effect increases in the size of the hypothetical sunk cost, especially in proportional terms (e.g., Garland and Newport, 1991). The effect is very sensitive to framing, and seems reduced by emphasizing the salience of the additional costs (e.g., Northcraft and Neale, 1986; Tan and Yates, 1995). Posing hypothetical questions as whether to grant an additional bank loan for continuing a project, Garland and Conlan (1998) find that sunk costs are less important than whether the additional loan will allow project completion. Their interpretation is that the goal of project completion psychologically displaces the profit goal. Taking the survey responses at face value, however, an alternative explanation is that the subjects have a better intuition for the value of real options than the experimenters. Refusing the additional loan to complete a project would extinguish the wait option for the project, and might hurt the bank's reputation.

Eyster (2002) expresses a consensus view that "the most convincing single experiment comes from Arkes and Blumer (1985)," experiment 2. In this field experiment, 20 randomly selected buyers were given a small (\$2) discount, 20 others a large discount (\$7, almost half the price), and 20 others no discount, on season tickets to the campus theatre. After excluding 10% of the subjects who bought tickets as couples, the authors report that the no-discount group used more tickets than either discount group in the first half of the theater season (p<.05). This is consistent with the sunk cost fallacy, but the evidence is not as strong as one might hope. The reported significance levels apparently assume that (apart from the excluded couples) all attendance choices are independent. The authors do not explain why they divided the season in half, nor do they report the significance levels for the entire season (or first quarter, etc.). The data show no significant difference between the small and large discount groups in the first half season nor among any of the groups in the second half season. We are not aware of any replication of this field experiment.

The animal behavior literature reports a controversy dating back to Trivers (1972) on the "Concorde effect," an allusion to continuing government subsidies of the uneconomic supersonic passenger plane. Arkes and Ayton (1999) conclude that "there are no unambiguous instances" of the sunk cost fallacy among animals, or even human children. They argue that adult humans commit the fallacy by misapplying the "don't waste" rule.

There are only a few relevant laboratory experiments using salient payments. Phillips, Battalio, and Kogut (1991) found that increases in the sunk price of a lottery ticket led to increased valuations by one quarter of the subjects, but to decreased valuations by another quarter and no response by half. With opportunities to learn in a market setting, very few subjects responded to the sunk price. Meyer (1993) reported that about half of his subjects bid more relative to a benchmark when an auction entry fee became larger. However, questions remain about his theoretical benchmark, the symmetric equilibrium bidding function for a different kind of auction than used in the experiment. We are also aware of an unpublished study by Offerman and Potters (2001) that shows sunk costs can facilitate coordination, and an inconclusive study by Elliott and Curme (1998).

Some non-experimental field evidence suggests the sunk cost fallacy. Camerer and Weber (1999) confirm Staw and Hoang's (1995) observation that first year professional basketball players who are drafted earlier (and thus, by the nature of the draft system, represent larger sunk costs) get more playing time, conditional on measured performance. Of course, it is hard to completely rule out other explanations based on unobserved components of performance or the coaches' Bayesian priors. Barron et al. (2001) find that US firms are significantly more likely to terminate projects following the departure of top managers. This might reflect the new managers' insensitivity to costs sunk by their predecessors, or it might simply reflect two aspects of the same broad realignment decision.

Do Internet users respond to sunk time costs? Manley and Seltzer (1997) report that after a particular website imposed an access charge, the remaining users stayed longer. A rival explanation to the sunk cost fallacy is selection bias: the users with shortest stays when the site was free are those who stopped coming when they had to pay. Klein et al (1999) report that users stick around longer on their site after encountering delays while playing a game, but again selection bias is a possible alternative explanation. The issue is important in e-commerce because "stickier" sites earn more advertising revenue. Schwartz (1999) reports that managers of the free Wall Street Journal site deliberately slowed the login process in the belief that users would then stay longer. One of us (Lukose) took a sample of 2000 user logs from a website and found a significant positive correlation between residence time at the site and download latency. One alternative explanation is unobserved congestion on the web, and users may have been responding more to expected future time costs than to time costs already sunk. Also, good sites may be more popular because they are good, leading to a) congestion and b) more time spent on the site.

To summarize, there are at least two distinct psychological mechanisms that might create an irrational regard for sunk costs. Self-justification (or cognitive dissonance) induces people who have sunk resources into an unprofitable activity to irrationally revise their beliefs about the profitability of an additional investment, in order to avoid the unpleasant acknowledgment that they made a mistake and wasted the sunk resources. Loss aversion (with respect to a reference point fixed before the costs were sunk) might induce people to choose an additional investment whose incremental return has negative expected value but still has some chance of allowing a positive return on the overall investment.

There are also several possible rational explanations for an apparent concern with sunk costs. Maintaining a reputation for finishing what you start may have sufficient value to compensate for the expected loss on an additional investment. The "real option" value (e.g., Dixit and Pindyck, 1994) of continuing a project also may offset an expected loss. Agency problems in organizations may make it personally better for a manager to continue an unprofitable project than to cancel it and take the heat from its supporters (e.g., Milgrom and Roberts, 1992).

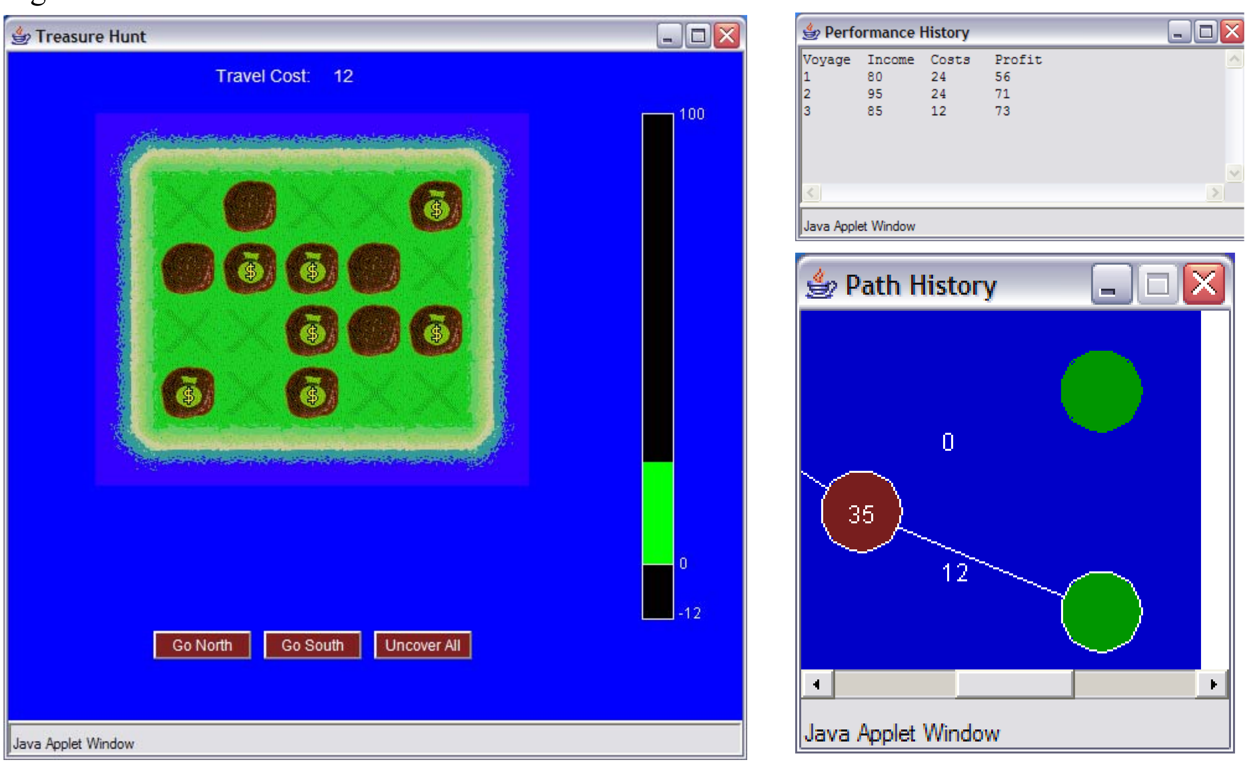
The available evidence is remarkably ambiguous. Besides confounding the various rational and irrational explanations, the studies often are unable to control for unobserved Bayesian priors, selection biases, etc. Clearly there is room for a new laboratory experiment that eliminates the rational explanations and the unobservable factors, and that allows alternative psychological explanations to demonstrate their explanatory power.

3. Experiment Design

Subjects play a computerized treasure hunt game in which they visit a sequence of islands. In the baseline treatment, each island has 20 sites the subject can "dig up" by clicking the mouse and she gets 5 points each time she clicks a site with buried treasure. The "voyage" ends when the subject exhausts a fixed click budget, e.g. 200 clicks. Budget permitting, she can click as many of the 20 sites as she wants before "sailing North or South" to the next island.

Leaving for a new island involves a sunk cost. The subject is told that because of unpredictable weather at sea, "your cost (in points) of reaching the next island is either high or low [e.g., is either c_H = 12 points or c_L = 0 points with equal probability, and the] … amount of buried treasure on an island is not affected by the cost of getting there." Figure 1 shows the user interface.

Figure 1



The number of treasures buried on each island is determined by an i.i.d. uniform random draw from consecutive integers {L,…,U} with $0 \le L < U \le 20$, e.g., L=2 and U=18. Subjects are told all relevant parameters, e.g., c_L, c_H , L and H. The pay rate, e.g., two cents per point, is posted on the board. Subjects are paid for one voyage, selected at the end of the experiment by a public random device.

We report results for nine variants on the baseline treatment:[1]

1. Displaying upon arrival the number of treasures on the island (Show Island Value, SIV=y).
2. Requiring subjects to click all sites or none (Require Complete Uncovering, RCU=y), used only in conjunction with SIV=y.
3. A click budget different from 200, e.g. Nclicks=100.
4. In the baseline, subjects choose whether to sail North or South, but these choices have no effect on the distribution of sunk costs and island values. In the treatment 'Choose Next Island' (CNI) = a, the distributions differ in ways known to the subjects, e.g. high sunk costs and more buried treasure are more likely when sailing North. At the other extreme, in the treatment CNI = no, the subject chooses only when to leave, not whether to go North or South.
5. In the baseline, after choosing North or South, the subject sees the cost of her own choice and also the cost of the direction not chosen. In the alternative treatment (Show Other Island Cost, SOIC=n), she sees only her own cost. Treatments 4 and 5 are intended to manipulate the saliency of self-justification.
6. When traveling to the next island, subjects experience a time delay proportional to the cost in points. In the baseline, the proportion is 'Cost Pause' (CP) = 0.8 seconds per point; alternatives include CP = 0, 0.4 and 1.6.
7. In the baseline, the screen displays a thermometer-like graph of net cumulative points earned on the island. It starts out in the red (as the negative of c_L or c_H, as the case may be) and turns green when it reaches positive territory as treasures are found. In the alternative treatment (Thermometer Displayed, TD = no), this part of the graphical display is suppressed. Treatments 6 and 7 are intended to manipulate the saliency of loss aversion.

[1] Pilot experiments not reported here explored two additional treatments: more formal instructions, and sunk costs incurred as clicks rather than as points. Neither treatment had a discernable effect on responsiveness to sunk costs.

8. Alternative choices to the baseline parameters c_L = 2 points, c_H = 18 points, L= 2 sites and U=18 sites are often used. Some colleagues have suggested that the sunk cost effect will be stronger when $c_L = 0$.
9. The baseline treatment draws the probability of hitting a treasure with no replacement (Replace=n) from the discrete uniform distribution with endpoints $0 \le L < U \le 20$. The alternative treatment (Replace=y) draws a probability p for each island independently from a continuous uniform distribution on $[l, u] \subset [0, 1]$, and each click on that island has independent probability p of hitting treasure. Sometimes p is displayed on arrival to the island (Show Hit Probability, SHP=y). We shall soon see that the replacement treatment has important consequences for rational choice.

One other design feature should be mentioned. Except in treatment CNI=a, subjects are paired so that for each subject who reaches a given island at high cost, there is another subject who reaches the same island at low cost. This pairing reduces experimental error, and it is feasible (except in CNI=a) because the random sequences of travel costs and of hits and misses on each island are drawn in advance.

Table A1 in the Appendix A summarizes the design of the 36 sessions analyzed below.

4. Search Theory and Testable Hypotheses.

The benchmark of optimal behavior will strengthen the data analysis. It turns out that treasure hunt game is not easily solved; indeed it took us several tries over a period of months to get it right. There are several cases, depending on whether the island value is displayed (SIV), whether subjects may click sites one at a time (RCU), and whether replacement is used in probabilities of hitting treasure (Replace). Appendix A collects the analytical results, which can be summarized as follows.

Case 1: SIV=y, RCU=y, i.e., the island value v (or hit probability p) is shown on arrival and the click choice is all-or-none, as in approximately 8% of the voyages listed in Table A1. This case requires only a minor extension of classic search theory, and the optimal search is characterized by a reservation value R. Budget permitting, the optimal strategy is to click out the island (expend 20 clicks) if $v \ge R$, and immediately to sink the cost and move on if $v \le R$. Appendix A shows that for the uniform distributions on $[L, U]$ used in the experiment, the reservation value is closely approximated by $R = bU - \sqrt{2cb(U-L)}$. For example, with the default parameters (expected sunk cost $c = (c_L + c_H)/2 = (2 + 18)/2 = 10$, treasure value b=5, and the number of treasures per island between L=2 and U=18) we get R =50. Then a rational player would click out the islands showing values v=50 and above and would pass on the islands showing values 45 and below. For Replace=y, set $v = 20bp = 100p$; e.g., in the example, click out islands showing $p = 0.50 = 50\%$ and above.

Case 2: SIV=y, RCU=n, approximately 42% of the voyages. This case is identical to case 1, except that the player has the option to leave the island after clicking some but not all sites. With the number of clicks divisible by 20 and with replacement (Replace=y), the option has no value: if it is optimal to click once on a given island, then it is optimal to click 20 times, so the rule is the same as in Case 1. For Replace=n, however, the option is valuable. For a player who gets sufficiently lucky on the first several clicks, the remaining sites are not worth clicking because his luck must "catch up" to the displayed island value. Hence optimal behavior is more complicated and is computed using techniques similar to those discussed for case 4b below.

Case 3. SIV=n, RCU=y. The player must decide all-or-none whether to click out an island about which he can obtain no information. This case is uninteresting and is not used in the experiment.

Case 4. SIV=n, RCU=n, approximately 50% of the voyages. Even when it does not hit treasure, each click has an information value because the player can update his estimate of the number of remaining treasures. This would seem to give more scope for self-justification and perhaps strengthen the sunk cost effect. There are two subcases. (a) When Replace=y, we obtain essentially a click-by-click reservation value $h^*(n)$ for the minimum number of hits in the first n clicks required to justify staying on an island. This value also depends on the size of the remaining click budget. (b) When Replace=n, the optimal policy is even more complicated because the catch-up effect (mentioned in case 2 above) opposes the effect of information updating.[2] Usually the optimal choice (for a given number of clicks on an island and given

[2] The authors were confused on this point for quite a while. An example may provide intuition. Suppose that you know initially that between 2 and 18 of the 20 sites contain treasure, and the first 4 clicks do not hit anything. The 16 remaining sites must therefore contain between 2 and 16 treasures, increasing the probability of hitting treasure on the next click. This catch-up effect sometimes dominates the Bayesian updating effect that negatively skews the posterior distribution. The treatment Replace=y eliminates the catch-up effect and simplifies the analysis.

number of clicks remaining) is to leave if the number of hits falls into a middle range and otherwise to click once more, but in a few cases the optimal strategy is still more complicated.

Using the optimality computations, we can determine for each decision on each voyage whether it was optimal and, if not, the loss in expected value. The data analysis uses the following classification for each decision.

- Impatient: the subject left the island when it was optimal to stay, thus incurring a loss $x > 0$ in expected earnings; or
- Stubborn: the subject clicked another site on the island when it was optimal to leave, thus incurring a loss $y > 0$ in expected earnings; or
- Optimal: no alternative action would generate higher expected earnings, and $x = y = 0$.

Needless to say, subjects will not always choose optimally. Their task is computationally very challenging, and even highly intelligent subjects unaffected by biases will occasionally make impatient or stubborn choices. The purpose of the classification is to better characterize behavior. For example, suppose that on average subjects stay longer on high cost islands. The sunk cost fallacy is confirmed if this arises from stubbornness on high cost islands and approximate optimality on low cost islands. However, if we find optimality on high cost islands but impatience on low cost islands, then the data reflect other departures from rationality.

The following testable hypotheses will guide our examination of the data.

H1.0 The average number of clicks on each island is the same for players who reached it with a low sunk cost as for players who reached it with a high sunk cost.

H1.R The research hypothesis, a one-sided alternative to this null hypothesis, is that the average number of clicks is higher for high sunk cost players.

In the cases (1 and 2) where the island value is displayed on arrival, we expect a stronger sunk cost effect for “close-calls,” the islands whose value is in the vicinity of the optimal reservation value R. The sunk cost fallacy should have less impact when islands are obviously worth clicking out or obviously better to skip.

The three-way classification of choices permits a more refined test:

H2.0 Impatient choices and stubborn choices have the same distribution on islands reached with high sunk cost as on islands reached with low sunk cost.

H2.R Impatient choices are less frequent and stubborn choices are more frequent on islands reached with high sunk cost than on islands reached with low sunk cost.

It is reasonable to say that costly mistakes are more meaningful than mistakes that incur negligible losses. Hence we make the same comparison in the payoff domain:

H3.0 The loss of expected earnings due to impatient choices (IL=$\sum x$) and the loss of expected earnings due to stubborn choices (SL=$\sum y$) have the same distribution on islands reached with high sunk cost as on islands reached with low sunk cost.

H3.R Average IL is smaller and average SL is larger on islands reached with high sunk cost than on islands reached with low sunk cost.

Hypotheses H1-H3 focus on the sunk cost effect, but the experiment design encourages a more detailed dissection of treatment effects and individual choice. Each click by each subject gives us a value of the indicator variable for staying: $Z = 0$ if the player sinks a cost to go to the next island, and $Z = 1$ if the player clicks another site on the current island. This dependent variable is explained in a logit regression by variables representing rational and psychological motives, other treatment variables, and their interactions. The null hypothesis is:

H4.0 Estimated coefficients in the logit regressions will be large and significantly positive for variables representing rational motives for staying, and will be insignificant for the dummy (indicator variable) for high sunk cost and for its interactions with other variables.

H4.R The dummy variable for high sunk cost and its interactions with several of the treatment variables will be significant. In particular, self-justification theory suggests positive interactions with treatments 4 and 5, especially in case 4, and loss aversion suggests positive interactions with treatments 6 and 7.

5. Results

First consider H1, the most direct test of the sunk cost fallacy. For each island we calculate the average number of clicks under high cost c_H minus average clicks under low cost c_L, so positive differences represent a sunk cost effect. Differences are averaged across islands using weights proportional to the number of players reaching the island. The overall click difference is -.17, indicating a slight reverse sunk cost effect. The value is, however, not significantly different from zero (by a paired t-test value of –1.63). Table 1 below splits the data by case (recall that case 3 is uninteresting and not used) and reports a significant sunk cost effect only in case 4, again a reverse effect. Panel B of the table separates out the non-obvious islands, defined as those within displayed value within 15 of the reservation value R. Contrary to our conjecture, the sunk cost effect is not stronger for such islands; indeed, Panel B shows a stronger (but still not significant) reverse sunk cost effect for Medium island values in case 1. Thus direct tests of H1 fail to establish a sunk cost effect.

Table 1: Weighted Average Click Difference

Panel A: By case

Case	W. avg. click difference	t-value	nobs
1: SIV=y, RCU=y	-0.25	-0.75	966
2: SIV=y, RCU=n	0.07	0.47	4449
4: SIV=n, RCU=n	-0.45	-2.56	3609

Panel B: By case and island value

Case	Island value	W. avg. click difference	t-value	nobs
1: SIV=y, RCU=y	Low	-0.01	-0.02	437
1: SIV=y, RCU=y	Medium	-0.63	-0.78	317
1: SIV=y, RCU=y	High	-0.21	-1.05	212
2: SIV=y, RCU=n	Low	0.16	0.65	2014
2: SIV=y, RCU=n	Medium	0	-0.01	1684
2: SIV=y, RCU=n	High	-0.03	-0.19	751

Note: Weighted average click difference is the difference between the mean number of clicks in the high cost group and the low cost group on each island, with islands weighted by the number of subjects in the smaller group. Reported t-values compare the weighted average click difference to zero. Low (resp. high) island values in Panel B are those less than R-15 (resp. greater than R+15), where R is the reservation value defined in section 4. Panel B excludes Case 4 observations since the High/Medium/Low classification is not well defined when the island value is not displayed.

We turn now to a finer grained examination of hypotheses H2 to H4. Table 2 shows that, despite the complexity of the calculation, from 65 percent (in Case 2) to 80 percent (in Case 4) of all choices are optimal. Stubborn choices account for most of the departures. This choice asymmetry is not surprising: on a given island one can be stubborn many times but impatient at most once.

Table 2: Choices by Case and Cost

Panel A: By case (all islands)

Case	Cost	Choices (% of total) Optimal	Stubborn	Impatient	Chi-square (nobs)	p-value
1: SIV=y, RCU=y	Low	78.5	21.4	0.2	2.35	0.31
1: SIV=y, RCU=y	High	77.4	21.9	0.8	(1193)	
2: SIV=y, RCU=n	Low	64.8	34.6	0.6	0.72	0.70
2: SIV=y, RCU=n	High	65.1	34.3	0.6	(81,292)	
4: SIV=n, RCU=n	Low	80.0	19.0	1.0	5.01	0.08
4: SIV=n, RCU=n	High	79.4	19.6	1.0	(82,843)	

Panel B: By case for medium island values

Case	Cost	Optimal	Stubborn	Impatient	Chi-square (nobs)	p-value
1: SIV=y, RCU=y	Low	63.4	36.1	0.5	2.06	0.36
1: SIV=y, RCU=y	High	59.1	39.1	1.9	(409)	
2: SIV=y, RCU=n	Low	68.7	30.3	0.9	6.34	0.04
2: SIV=y, RCU=n	High	69.9	29.2	0.9	(39,790)	

Table 2 also shows that in cases 1 and 4 stubborn choices are more frequent and optimal choices are less frequent on high cost islands than on low cost islands, consistent with H2.R. However, the shift is small and is reversed in case 2, and no consistent picture emerges for impatient choices. A chi-square test indicates the shifts are insignificant except perhaps in case 4 (significant at a marginal 8% level). Panel B restricts the analysis to medium island values (R-15 ≤ island value ≤ R+15), and here the choice shift becomes significant for Case 2 at the 4% level, but in a direction inconsistent with the sunk cost hypothesis H2.R.

Now consider the payoff domain. Cases 2 and 4 allow the calculation of value gained or lost on each click. Define total potential value (TPV) as the sum of the absolute difference in expected profit between immediately leaving the island and staying for another click. Otherwise put (see Appendix A for details), TPV= actual value gained + SL + IL. Table 3 shows that subjects overall lose less than 7% of TPV. There really is an asymmetry in that 3-6% of TPV is lost due to stubbornness but less than 1% due to impatience. The data support H3.R for case 2. The difference in value lost due to stubbornness between the low and high cost groups is small (0.3%) but highly significant (t=-4.1, p<0.0001). In case 4, however, stubborn losses and impatient losses are the same on high and low cost islands, so we can't reject the null hypothesis.

Table 3: Value Gained or Lost by Cost

Case	Cost	Gain	Value as % of TPV Stubborn loss	Impatient loss	Nobs
2: SIV=y, RCU=n	Low	93.6	6.0	0.4	38,090
2: SIV=y, RCU=n	High	93.4	6.3	0.4	43,292
4: SIV=n, RCU=n	Low	96.3	3.0	0.7	39,889
4: SIV=n, RCU=n	High	96.3	3.0	0.7	42,954

Case 2 (SIV=y, RCU=n) without replacement (replace=n) permits sharp tests of two well-known heuristics. According to the "win-stay, lose-shift" heuristic (e.g., Eyster, 2003), subjects who just experienced success are more likely to stay on an island. Optimality predicts exactly the opposite because in Case 2 the catch-up effect (described in the previous section) is not offset by an update effect. To test these opposing predictions, and to refine the estimates of the sunk cost effect, we ran the logit regressions reported in Table 4.

Table 4: Logistic Regression for Win-Stay, Lose-Switch Heuristic

Case 2 data. Number of Observations = 74,952

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
Intercept	3.48	0.05	6303	<.0001
Cost	0.09	0.04	3.9	0.05
Stay surplus	0.36	0.01	971	<.0001
Last-click-successful dummy	-0.70	0.05	227	<.0001

Notes: Dependent variable Z=1 if subject clicks once more, Z=0 if subject leaves the island. The interaction term cost*Last-click-successful is insignificant when included in the regression.

The second line indicates that higher sunk cost (cost=1 if $c = c_H$ and =0 otherwise) increases the log odds of staying by 0.09, significant at the 5% level, consistent with the sunk cost effect and hypothesis H4.R. The third line shows a strong impact of the difference between the value of clicking and the value of leaving the island (stay surplus), e.g., increasing stay surplus by 2 treasures or 10 points increases the log odds by 3.6, consistent with noisy rational search. The last line investigates the heuristic. The dummy variable Last-click-successful is set to one if the previous click hit treasure, and to zero otherwise. The line indicates that the rational catch-up effect dominates the win-stay, lose-switch heuristic.

A second heuristic is based on loss aversion. Eyster (private conversation) predicts that subjects whose cumulative earnings on a given island are still negative (i.e., have not yet covered the sunk cost) are more likely to stay. The variable Cumulative loss is the minimum of zero and the cumulative earnings on that island. The next to last line in panel A of Table 5 indicates that larger cumulative losses (more negative values) tend to decrease the probability of staying on an island in Case 2, contrary to loss aversion. On the other hand, the results are consistent with loss aversion in Case 4, where subjects do not have information about the value of an island. In either case, the sum of the last two coefficients indicates that the effect disappears when sunk costs are high. The Cumulative loss variable has range [-c, 0], so its absolute value remains small when $c=c_L$. Hence the heuristic has little impact overall. Alternative specifications, e.g. a binary (dummy) version of the Cumulative loss variable, give roughly similar results.

Table 5: Logistic Regression for Loss Aversion

Panel A: Case 2. Number of Observations = 81,382

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
Intercept	2.85	0.03	9581	<<.0001
Cost	0.34	0.05	56.7	<.0001
Stay surplus	0.35	0.01	1249	<<.0001
Cumulative loss	0.79	0.06	166	<.0001
Cost*Cum. loss	-0.72	0.06	134	<.0001

Panel B: Case 4. Number of Observations = 82,843

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
Intercept	2.84	0.03	8092	<<.0001
Cost	-0.01	0.05	0.05	0.82
Stay surplus	0.32	0.01	1096	<<.0001
Cumulative loss	-0.74	0.18	16.8	<.0001
Cost*Cum. loss	0.73	0.18	16.3	<.0001

Notes: Dependent variable Z=1 if subject clicks once more, Z=0 if subject leaves the island.

Table 6 reports the most sensitive tests of Hypothesis 4. The Stay surplus coefficients in all panels confirm the huge impact of rational considerations. The second line of Panel A reports the best evidence we have found for the sunk cost effect: although relatively small, the Cost coefficient estimate has the predicted sign and is significant at the 1% level. The main effects for the self-justification treatments are both significant at the 5% level (indeed SOIC at the 0.01% level). Allowing the subject to choose North or South, and including information on the cost of the route not taken, both seem to increase the tendency to stay on the current island.

The treatment predictions in Hypothesis 4 concern interactions rather than main effects, and Panels B and C report the test results. Panel C covers case 4 and hence should give the hypotheses their best shot. Here the CNI interaction with Cost is significant at the 5% level but has the wrong sign. The other self-justification interaction, with SOIC, has the predicted sign but it is barely significant at the 10% level. The loss-aversion interactions both are highly significant and have the predicted sign. In Panel B (case 2), the CNI interaction coefficient flips to the predicted sign and the SOIC interaction coefficient keeps the predicted sign and becomes significant. The loss-aversion interactions both flip to sign, although only the TD (the thermometer display) retains significance. All these variables become insignificant when all main effects and interactions are included in the same logistic regression.

Additional tests, omitted here, find no evidence that offering an asymmetric choice between North and South (CNI=a) increases the sunk cost effect. Indeed, when such a choice is offered, the effect is slightly stronger for those who choose South, contrary to prediction; the difference however at best is marginally significant. Also, increasing the contrast between high and low sunk costs (with or without a corresponding change in time delay, CP) seems perversely to reduce the sunk cost effect, but again the impact is not significant at the 5% level. Reducing the lower sunk cost to $c_L = 0$ has no detectable incremental impact.

Table 6: Logistic Regressions for Decision to Stay

Panel A: Main Effects, Cases 2 and 4 combined. Number of Observations = 164,225

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
Intercept	2.34	0.12	397.	<.0001
Cost	0.07	0.03	6.2	0.01
Stay surplus	0.35	0.01	2785.	<.0001
Choose Next Island	0.16	0.08	4.2	0.04
Show Other Island Cost	0.25	0.06	19.1	<.0001
Cost Pause	0.00	0.00	0.13	0.72
Thermometer Displayed	0.03	0.04	0.42	0.52

Panel B: Interactions, Case 2. Number of Observations=81,382

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
Intercept	2.74	0.03	10620.	<.0001
Cost	-0.89	0.24	14.0	0.0002
Stay surplus	0.37	0.01	1398.	<.0001
Cost*CNI	0.55	0.14	16.1	<.0001
Cost*SOIC	0.79	0.20	14.9	0.0001
Cost*CP	-0.00013	0.00008	2.33	0.13
Cost*TDn	-0.24	0.10	5.71	0.02

Panel C: Interactions, Case 4. Number of Observations=82,843

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr>ChiSq
Intercept	2.86	0.03	8369.	<.0001
Cost	-0.37	0.29	1.65	0.20
Stay surplus	0.32	0.01	1113.	<.0001
Cost*CNI	-0.40	0.20	3.95	0.047
Cost*SOIC	0.14	0.09	2.74	0.098
Cost*CP	0.0006	0.0002	10.3	0.001
Cost*TD	0.17	0.08	4.91	0.027

Note: Dependent variable Z=1 if subject clicks once more, Z=0 if subject leaves the island. The dummy variables CNI, SOIC and TD are 1 for default value y of the corresponding treatments and are 0 when the treatments have value n. All interaction terms between cost and treatment variables are insignificant when jointly included in the Panel A regression.

6. Discussion

The experiment seeks to isolate the famous but elusive sunk cost fallacy. The treasure hunt task is simple to understand but very difficult to master fully. The matched pair design can detect even very small effects, and numerous treatments enable us to explore, in a variety of contexts, the proposed psychological drivers of the fallacy.

The results can be summarized as follows.

1. Subjects' choices are surprisingly consistent with optimal search behavior. A large majority of choices are optimal, and actual losses in expected payoff represent less than 7% of possible losses. Losses due to stubbornness are larger than losses due to impatience, probably because subjects can be stubborn many times on a given island, but impatient only once.
2. There is evidence for the sunk cost fallacy. Stubborn errors are more frequent when sunk costs are high in Case 1 (all-or-none choice, island value displayed on arrival) and Case 4 (click-by-click choice, island value not displayed), and on average stubborn losses are larger in the remaining case (2, click-by-click choice, value not displayed). The relevant logit regressions indicate that subjects are more likely to stay on islands with higher sunk cost.
3. The effect is surprisingly small and inconsistent. The simple comparison (click difference) indicates a small reverse effect, and so do several variants of the stubborn error and logit specifications. Probably the strongest evidence for the fallacy is the main effect for Cost in Panel A of Table 6. The relevant coefficient is significant at the 1% level and has the right sign, but implies a rather small effect: even a player who would stay with probability 0.50 low sunk cost would stay with probability $1/(1+\exp(0.07)) \approx 0.52$ with high cost.
4. The treatments intended to manipulate the psychological drivers of the fallacy also have rather small and inconsistent impact. Contrary to conjecture, the variables manipulating self-justification work best in Case 2, while in Case 4 they either have the wrong sign or are hardly significant. On the other hand, the variables manipulating loss aversion have the wrong sign (and one is insignificant) in Case 2 but work better in Case 4.

The results reported here arose from an extensive design search. We had expected to quickly find a substantial sunk cost effect, but did not. With helpful advice from many colleagues, we tried a succession of treatments. The best we could come up with were the small and inconsistent results just described. In sum, we were unable to find sunk-cost tasks and treatments that reliably lead subjects to substantial departures from rational behavior. The challenge thus remains for future investigators.

Bibliography

Arkes, Hal R. and Peter Ayton (1999). "The Sunk Cost and Concorde Effects: Are Humans Less Rational Than Lower Animals?" *Psychological Bulletin*, Vol.125, No.5, 591-600.

Arkes, Hal R. and Catherine Blumer(1985). "The Psychology of Sunk Cost," *Organizational Behavior and Human Decision Processes*, 35, 124-140.

Aronson, Elliot (1968). "Dissonance Theory: Progress and Problems." In R. Abelson *et al*, eds, *Theories of Cognitive Consistency: A Sourcebook*, Rand McNally and Co, pp5-27.

Barron, John M.; Chulkov, Dmitriy V. and Glen R. Waddell (2001). "The Escalation Phenomenon and Executive Turnover: Theory and Evidence." Mimeo, Purdue University, July 2001.

Bazerman, Max H. (1986). *Judgment in Managerial Decision Making* 2nd ed. Chapter 4. John Wiley & Sons, Inc.

Camerer, Colin F. and Roberto A. Weber (1999). "The Econometrics and Behavioral Economics of Escalation of Commitment: A Re-examination of Staw and Hoang's NBA data," *Journal of Economic Behavior and Organization*, 39 (1), 59-82.

Devine, Kevin and Priscilla O'Clock (1995), "The Effect of Sunk Costs and Opportunity Costs on a Subjective Capital Allocation Decision," *The Mid-Atlantic Journal of Business*, 31(1), 25-38.

Dixit, Avinash and Robert H. Pindyck (1994). *Investment Under Uncertainty*. Princeton University Press.

Economist (2003). "The space shuttle: Old, unsafe and costly." *Economist* Magazine, Aug. 28, 2003, pp. 77.

Elliott, Steven R. and Michael Curme (1998). "Understanding the Sunk Cost Effect: An Experimental Approach," Mimeo, Miami University.

Eyster, Erik (2002). "Rationalizing the Past: A Taste for Consistency," Nuffield College, Oxford, Job-Market Paper.

Festinger, Leon (1957). *A Theory of Cognitive Dissonance*. Stanford: Stanford University Press.

Friedman, Daniel and Shyam Sunder (1994) *Experimental Methods: A Primer for Economists*. Cambridge University Press. New York, NY. 1994.

Garland, Howard and Donald E. Conlan (1998). “Too Close to Quit: The Role of Project Completion in Maintaining Commitment,” *Journal of Applied Social Psychology*, 28(22), 2025-2048.

Garland, Howard and Stephanie Newport (1991). “Effects of Absolute and Relative Sunk Costs on the Decision to Persist with a Course of Action,” *Organizational Behavior and Human Decision Processes*, 48, 55-69.

Klein, Jonathan; Youngme Moon, and Rosalind W. Picard (1999). “This Computer Responds to User Frustration,” MIT Media Laboratory, June 17, 1999 Vision and Modeling Group Technical Report #502.

Kahneman, Daniel and Amos Tversky (1979). “Prospect Theory: An Analysis of Decision Under Risk,” *Econometrica*, 47(2), 263-291.

Manley, S. and M. Seltzer (1997). “Web Facts and Fantasy.” *Proceedings of the 1997 USENIX Symposium on Internet Technologies and Systems*, Monterey, CA, December 1997.

Meyer, Donald J. (1993). “First Price Auctions with Entry: An Experimental Investigation,” *The Quarterly Journal of Economics and Finance*, 33(2), 107-122.

Milgrom, P. and J. Roberts (1992). *Economics, Organization, and Management.* Prentice-Hall, Englewood Cliffs: NJ.

Northcraft, Gregory B. and Margaret A. Neale (1986). “Opportunity Costs and the Framing of Resource Allocation Decisions,” *Organization Behavior and Human Decision Processes*, 37, 348-356.

Offerman, Theo and Jan Potters (2001). “Does Auctioning of Entry Licenses Induce Collusion? An Experimental Study,” Working paper, University of Amsterdam.

Parayre, Roch (1995). “The Strategic Implications of Sunk Costs: A Behavioral Perspective,” *Journal of Economic Behavior and Organization*, 28, 417-442.

Phillips, Owen R., Raymond C. Battalio, and Carl A. Kogut (1991). “Sunk and Opportunity Costs in Valuation and Bidding,” *Southern Economic Journal*, 58(1), 112-128.

Rabin, Mathew (1998). “Psychology and Economics,” *Journal of Economic Literature*, 36(1), 11-46.

Schwartz, Evan. I. (1999). “Digital Darwinism: 7 Breakthrough Business Strategies for Surviving in the Cutthroat Web Economy”, by Evan I. Schwartz, Broadway Books; (June 8, 1999)

Staw, B. M. (1976). “Knee deep in the big muddy: A Study of Escalating Commitment to a Chosen Course of Action.” *Organizational Behavior and Human Performance*, 16, 27-44.

Staw, B. M. and Hoang, H. (1995). “ Sunk Costs in the NBA: Why Draft Order Affects Playing Time and Survival in Professional Basketball.” *Administrative Science Quarterly*, 40, 474-494.

Tan, Hun-Tong and Frank J. Yates (1995). “Sunk Cost Effects: The Influences of Instruction and Future Return Estimates,” *Organizational Behavior and Human Decision Processes*, 63(3), 311-319.

Thaler, Richard (1980).“Toward a Positive Theory of Consumer Choice,” *Journal of Economic Behavior and Organization*, 1, 39-60.

Trivers, Robert L. (1972). “Parental Investment and Sexual Selection,” In B. Campbell (Ed.), *Sexual Selection and the Descent of Man*, 1871-1971(pp. 136-179). Chicago: Aldine.

Table A1: Sessions and Treatments

Date	# of complete voyages	# periods w/ cost diffs	Low cost (c_L)	High cost (c_H)	Island values, hit prob.	Click budget	Stakes (cents/ point)	CP	SIV, SHP	RCU	TD	SOIC	CNI	# of people	Exper-ience
8/1/02	52	11	2	18	35-65	100	4	TMT	No	No	Yes	Yes	yes	4	no
8/7/02	60	16	2	18	5-95	100	4	TMT	No	No	Yes	Yes	yes	4	no
8/15/02	16	0	2	18	20-80	100	4	800	No	No	Yes	Yes	TMT	2	no
8/20/02	57	11	2	18	20-80	100	4	800	TMT	TMT	Yes	Yes	yes	4	yes
8/21/02	26	7	2	18	20-80	200	2	800	No	No	Yes	Yes	TMT	4	no
8/22/02	14	4	2	18	20-80	200	2	800	Yes	No	Yes	Yes	TMT	4	no
9/5/02	15	4	1	13	10-90	200	2	800	Yes	No	Yes	Yes	TMT	4	no
9/11/02	30	5	1	13	10-90	200	2	800	Yes	Yes	Yes	Yes	TMT	6	yes
9/12/02	13	4	1	13	10-90	200	2	800	No	No	Yes	Yes	TMT	3	yes
1/24/03	46	9	1	13	20-80	200	2	800	TMT	No	Yes	Yes	Yes	6	no
1/31/03	36	9	1	13	10-90	200	2	800	TMT	No	Yes	Yes	Yes	4	no
2/4/03	29	0	1	13	10-90*	200	2	800	TMT	No	Yes	Yes	Yes	4	no
2/5/03	46	8	1	13	10-90*	200	2	800	TMT	No	Yes	Yes	Yes	6	yes
2/19/03	22	0	1	13	10-90*	200	2	800	TMT	No	Yes	Yes	Yes	4	yes
5/19/03	38	5	1	13	10-90*	200	2	800	TMT	No	Yes	Yes	Yes	8	no
5/23/03	21	4	1	13	10-90*	200	2	800	TMT	No	Yes	Yes	Yes	6	no
6/2/03	18	6	1	13	10-90*	200	2	800	TMT	No	Yes	Yes	Yes	4	yes
8/14/03	62	12	1	13	10-90*	200	4	800	No	No	Yes	No	Yes	6	no
8/21/03	76	15	0	10	10-90*	200	4	TMT	Yes	TMT	Yes	Yes	Yes	6	no
10/2/03	23	7	1	13	.1-.9	200	2	800	Yes	No	Yes	Yes	Yes	4	no
10/3/03	24	5	1	13	.1-.9	200	2	800	Yes	No	Yes	Yes	Yes	6	no
10/10/03	34	5	1	13	.1-.9	200	2	800	TMT	No	Yes	Yes	Yes	6	yes
10/23/03	54	9	0	12	.1-.9	200	2	800	TMT	No	Yes	Yes	Yes	8	both
11/6/03	24	5	0	12	.1-.9	200	2	800	TMT	No	Yes	Yes	Yes	6	both
11/14/03	31	6	0	12	.1-.9	200	2	800	TMT	No	Yes	Yes	Yes	6	no
12/4/03	44	12	0	12	.1-.9	200	1	800	TMT	No	Yes	Yes	Yes	4	yes
12/4/03	3	1	0	12	.1-.9	100	5	800	No	No	Yes	Yes	Yes	4	yes
1/23/04	27	5	0	12	.1-.9	200	2	800	TMT	No	TMT	Yes	Yes	8	no
1/23/04	8	1	0	12	.1-.9	100	4	800	No	No	Yes	Yes	Yes	8	no
1/28/04	34	8	0	12	.1-.9	200	2	800	TMT	No	TMT	Yes	Yes	7	no
2/4/04	34	9	0	12	.1-.9	200	1	800	TMT	No	TMT	Yes	Yes	4	no
2/4/04	4	1	0	12	.1-.9	100	4	800	No	No	Yes	Yes	Yes	4	no
2/11/04	58	7	0	12	.1-.9	200	1	800	TMT	No	TMT	Yes	Yes	12	no
2/11/04	12	1	0	12	.1-.9	100	4	800	No	No	Yes	Yes	Yes	12	no
2/25/04	17	4	0	12	.1-.9	200	1	800	Yes	TMT	TMT	TMT	Yes	7	no
2/25/04	6	1	0	12	.1-.9	100	5	800	No	No	Yes	Yes	Yes	7	no

Notes: The third column reports the cases where completed voyages had complementary cost structures. All sessions use treatments r = 5 points per treasure, AutoDig=yes, while other treatments vary as indicated; TMT (Treatment) indicates variation within session. Prob($c=c_H$|North)=prob($c=c_H$|South)=0.5, except for sessions with *, for which prob($c=c_H$|North)=0.7 and prob($c=c_H$|South)=0.4. Island values|North=Island values|South, except for *, for which the entry gives Island values|North, and Island values|South=10-70 (10-80 on 2/19/03).
Pilot experiments conducted before 8/1/02 are not listed and were excluded from the data analysis because different instructions were used and the data format is incompatible with later formats. Similarly, three sessions, run on 8/14/02, 2/12/03 and 2/14/03, were excluded because of technical problems with the software during the session.

Appendix A: Computation of Rational Decisions

June 6, 2004

The classic economic model of sequential job search (e.g., Lippman and McCall, 1976) gives insight into our treasure search task. In the classic model, a job seeker can always pay a cost $c \geq 0$ to receive another job offer $y \in [0, \infty)$, assumed an iid random variable with known distribution function F. The search terminates as soon as an offer $y = x$ is accepted, and the payoff is $x - nc$, where n is the number of offers purchased. In the simplest version, the job seeker is risk-neutral, there is no discounting (time lags are negligible, as in our experiment), and there is no bound on the number of offers that can be purchased.

With a current offer x in hand, the job seeker maximizes expected value V defined recursively by the Bellman equation

$$V(x) = \max\{x, -c + EV(y)\}. \tag{1}$$

It is well-known that this problem has a unique solution using a reservation price R. That is, the solution of the form: accept the most recent offer x if $x \geq R$ and otherwise pay c for another offer drawn from the distribution F. The reservation price R is determined from (1) by equating the value of the current offer $x = R$ to the value of continuing an optimal search $-c + EV(y) = -c + \int(\max\{y, R\})dF(y) = -c + R\int_0^R dF(y) + \int_R^\infty y dF(y) = -c + R + \int_R^\infty (y - R)dF(y)$. Cancelling R from both sides of the equation and simplifying slightly we get the marginal condition

$$c = H(R), \text{where } H(z) = \int_z^\infty (y - z)dF(y). \tag{2}$$

That is, R equates the incremental cost of search c to its incremental expected benefit $H(R)$. If F has a positive density over its support $[L, U] \subset [0, \infty)$, it is easily checked that the function H is strictly decreasing from $H(L) = Ey - L$ to $H(U) = 0$. Then (2) has a unique solution $R = H^{-1}(c) > 0$ for any search cost $c \in (0, Ey - L)$.

The classic problem can be adapted to a finite horizon. If only m more draws are possible, then the value function depends on m as well as x and the solution reservation price decreases as m decreases.

Case 1: Value Displayed, Uncover All or None

Our treasure search problem at first glance looks like the classic finite horizon problem, but it turns out to be a bit different. Consider first the case where each island value is known upon arrival and one must uncover all sites on the island or none. Without further loss of generality, normalize so that one click uncovers all the sites and the initial click budget is $Y > 0$. In the experiment the standard click budget is $Y = 200/20 = 10$ with this normalization.

Since the click budget is separate from earnings, there is no limit on the number of islands that can be visited and skipped. Hence the analogy is to

Y classic job searchers, each with an infinite horizon. The solution is to click islands whose displayed value W is at least R, and otherwise to skip to the next island, until the click budget is exhausted. (This assumes that the travel cost is not exorbitant. If $c \geq Ey = (U+L)/2$, a situation never seen in the experiment, then the player should quit playing rather than skip.)

To compute R, first recall that the (expected) travel cost is $c = (c_L + c_H)/2 > 0$, and that the number of treasures is uniformly distributed between $L \geq 0$ and $U \leq n_{max} = 20$ with value $b = 5$ value per treasure. The continuous uniform distribution function for value then is $F(x) = (x - bL)/b(U-L)$ for $x \in [bL, bU]$, with $F(x) = 0$ for $x < bL$ and $F(x) = 1$ for $x > bH$. Then for $z \in [bL, bU]$ we have

$$H(z) = \int_z^{bU} (y-z)dF(y) = (bU - bL)^{-1} \int_z^{bU} (y-z)dy = \frac{(bU-z)^2}{2b(U-L)}. \qquad (3)$$

This function H is continuous and is linear decreasing (slope=-1) for $z < bL$, quadratic decreasing for $z \in [bL, bU]$ and is 0 for $z > bU$. Solving $c = H(R)$ we obtain

$$R = bU - \sqrt{2bc(U-L)}. \qquad (4)$$

It follows that R decreases from $R = bU$ when $c = 0$ to $R = bL$ when $c = Ey - bL = b(U-L)/2$. For $c \in [b(U-L)/2, b(U+L)/2]$ every island should be clicked, but for $c \geq b(U+L)/2 = Ey$ the search should be abandoned. For $0 \leq c \leq b(U-L)/2$, the expected value of moving to the next island is R before sinking the cost and is $R + c$ after arrival.

The reservation value from (4) is not exact when, as in the experiment, the number of treasures on an island must be an integer. The uniform distribution then has discrete support $\{bL, b(L+1), ..., bU\}$ with equal mass $1/(U-L+1)$ at each point. For $z \in [bL, bU]$, write $z = b(L + i_z + r_z)$, where i_z is the unique integer between 0 and $U - L$ such that the residue r_z is in $[0, 1)$. The H-function for this discrete uniform distribution is $H_D(z) = \int_z^\infty (y-z)dF(y) = \frac{b}{U-L+1}\left[(1-r_z) + (2-r_z) + ... + (U-L-i_z-r_z)\right]$. Sum the series to get

$$H_D(z) = \frac{b(U-L-i_z)(U-L-i_z+1-2r_z)}{2(U-L+1)}. \qquad (5)$$

Comparing (3) and (5), or just noting that the distributions have the same support, one can see that $H_D(z) = H(z) = 0$ for $z \geq bU$ and that $H_D(z) = H(z)$ for $z \leq bL$; in particular $H_D(bL) = H(bL) = b(U-L)/2$. Moreover, the slope of H_D increases in $U - L + 1$ equal steps from -1 at $z = bL$ to 0 at $z = bU$, while the slope of H increases linearly from -1 to 0 over the same interval. Thus H_D is a continuous, piecewise linear approximation of the quadratic function H, and (4) closely approximates the exact reservation value when $U - L$ is reasonably large, as in the experiment.

Case 2: Value Displayed, Uncovering Discretionary

Now consider the decision problem when the player can choose click by click whether to skip to the next island. Recall that there are two ways of specifying island value. With *replacement*, each click hits treasure with independent constant probability p. When p is displayed on arrival and the number of remaining clicks is evenly divisible by the number of n_{max} of sites per island, then the decision problem is equivalent to the problem in the non-discretionary case (under the maintained hypothesis of risk neutrality). The expected island value is $W = bn_{max}p = 100p$ and the optimal strategy is to click all $n_{max} = 20$ sites on the present island if p is at least $r = R/100$, and otherwise to click none. The logic and computation of R are exactly as in the previous case.

The problem is considerably more complicated in the other subcase, *no replacement*. In this case the island value $W = bw$ is displayed on arrival. For the first click, the hit probability is $p_o = w/n_{max}$, but it changes after that. The hit probability p given h hits out of $n < n_{max}$ clicks so far on the current island is the number of remaining treasures divided by the number of remaining sites,

$$p = \frac{w - h}{n_{max} - n}. \tag{6}$$

Hence the hit probability p typically rises after a miss and declines after a hit. This *catch-up effect* can cause an initially attractive value of $p = p_o$ to become quite unattractive after a hot streak. Thus the player may rationally click some but not all sites on an island. Likewise, if the click budget is not evenly divisible by $n_{max} = 20$, then a detailed analysis is again necessary. The material below on Bellman equations covers these subcases.

Case 3: Value Not Displayed, Uncover All or None

The player has no basis for distinguishing one island from another on arrival and can't sample. Thus there is never a reason to skip an island; that only increases cost without increasing expected revenue. As in case 1, the player should quit if cost exceeds average island value. Otherwise he should dig up every island in order until the budget is exhausted. This case is trivial and we don't use it in the experiment.

Case 4: Value Not Displayed, Uncovering Discretionary

This case is intricate, due to the *update effect*: after each click, a player should use Bayes theorem to update his estimate of the island value. As in case 2, there are two subcases. With no replacement (Case 4b) the catchup effect opposes the update effect and is stronger in the extreme cases (very few hits or very few misses), but is weaker in other cases. Consequently the optimal strategy here cannot be expressed in terms of a reservation price. Replacement (case 4a) eliminates the catchup effect. Here the optimal search is characterized by a reservation price (in terms of hits and clicks so far on the island) that reflects the information value of another click as well as the Bayes posterior expected values.

In both subcases, a key computation is $p(h, n)$, the posterior probability that the next click on the current island will hit treasure, given h hits out of n clicks so far on the current island. We are given the prior distribution

$f(p)$ with support contained in $[0,1]$. Of course, $p(0,0)$ is simply the prior mean $\int_0^1 pf(p)dp \equiv \overline{p}$. In the experiment, subjects are told the maximum U and minimum L numbers of treasures and that the distribution is uniform, so $\overline{p} = (U+L)/(2n_{max}) = (U+L)/40$.

By Bayes theorem, the posterior density $f(p|h,n)$ of the hit probability given (h,n) is the likelihood of (h,n) times the prior probability $f(p)$ and normalized so that the expression integrates to 1.0. The desired posterior probability $p(h,n)$ is the expectation $\int_0^1 xf(x|h,n)dx$.

With replacement as in case 4a, the likelihood is the binomial expression $\binom{n}{h}x^h(1-x)^{n-h}$. With a continuous uniform prior supported on $[l,u] \subset [0,1]$ we therefore have

$$p(h,n) = \frac{\int_l^u x^{h+1}(1-x)^{n-h}dx}{\int_l^u x^h(1-x)^{n-h}dx}. \tag{7}$$

When $u = 1$ and $l = 0$, we can integrate by parts repeatedly and the surface terms (i.e., $x^i(1-x)^j$ evaluated at 0 and 1) vanish, yielding $p(h,n) = \frac{h+1}{n+2}$.

In the experiment, the distributions are discrete uniform. Equation (7) gives a close approximation for $u = U/n_{max}$ and $l = L/n_{max}$. The exact expression replaces the integrals by sums over $t = L, ..., U$ and replaces x by $x_t = t/n_{max}$, viz.,

$$p(h,n) = \frac{\sum_{t=L}^U x_t^{h+1}(1-x_t)^{n-h}}{\sum_{t=L}^U x_t^h(1-x_t)^{n-h}}. \tag{8}$$

In the no-replacement case 4b, the likelihood is hypergeometric instead of binomial. The likelihood that there are exactly $t \leq U \leq 20$ treasures on the island, given that h were found on the first $n \leq n_{max} = 20$ tries, is

$$p(h,n|t) = \frac{\binom{n}{h}\frac{t!}{(t-h)!}\frac{(20-t)!}{(20-t-(n-h))!}}{\frac{20!}{(20-n)!}} \tag{9}$$

if $h \leq t \leq u$ and otherwise is 0. In the expression for the Bayesian posterior probability, both denominator (the normalizing constant) and numerator contain the binomial coefficient $\binom{n}{h}$, the expression $\frac{20!}{(20-n)!}$, and the constant prior probability $1/(U-L+1)$. Hence these expressions cancel and we obtain the exact posterior distribution

$$f(t|h,n) = \frac{G(t|h,n)}{\sum_{s=L}^U G(s|h,n)}, \text{where } G(t|h,n) = \frac{t!}{(t-h)!}\frac{(20-t)!}{(20-t-(n-h))!} \tag{10}$$

for $h, L \leq t \leq U$ and $h \leq n$. Finally, the desired exact probability is the expectation of the remaining number of treasures (without replacement) divided by the remaining number of sites,

$$p(h,n) = \sum_{t=max\{L,h\}}^{U} \frac{t-h}{20-n} f(t|h,n). \tag{11}$$

Bellman Equations
We now are prepared to derive solutions for cases 2b, 4a and 4b. The approach is the same in each case: we write out the Bellman equation for optimal decision, insert appropriate boundary values and state transitions, and compute the values and contingent decisions by backward induction on the number of clicks remaining.

The Bellman equations take the following form.

$$\begin{aligned} V(l,s) &= \max\{C(l,s), S(l), 0\}, && (12)\\ C(l,s) &= E(x|s) + EV(l-1,s'), && (13)\\ S(l) &= -c + V(l,s_o). && (14) \end{aligned}$$

The first line says that the value, i.e., the expected payoff over the rest of the voyage given l clicks remaining and state s, is the maximum obtainable from three options: clicking on the present island (C), skipping to the next island (S), or quitting immediately (0). The second line defines the click value recursively as the expected payoff from the next click $E(x|s)$, plus the expected value of continuing the voyage with one less click, taking into account the transition from the current state s to a new state s'. The third line defines the skip value as the value of starting on a new island (state s_o) less expected travel cost; note that it depends only on the number l of clicks remaining, and not on the current state s.

General boundary conditions include

1. $V(0,s) = 0$, i.e., the game is over when zero clicks remain; and
2. $E(x|s_m) = -\infty\ \forall s_m$ such that $n = n_{max}$, i.e., only $n = n_{max}$ clicks are permitted on each island.

In case 4a, relevant state s is (h,n), the number of hits and clicks so far on the current island. Using $p = p(h,n)$ from equation (8), the click value for $l > 0$ clicks remaining and $n < n_{max}$ is

$$\begin{aligned} C(l,s) &= E(x|s) + EV(l-1,s') && (15)\\ &= pb + pV(l-1,h+1,n+1)) + (1-p)V(l-1,h,n+1). && (16) \end{aligned}$$

The skip value here is

$$S(l) = -c + V(l,s_o) = -c + \max\{0,\ C(l,0,0)\} \qquad (17)$$

where $s_o = (0,0)$ refers to the state on arrival at a new island, 0 hits on 0 clicks. The last term uses only the click value, because skipping at $s_o = (0,0)$ sinks the travel cost without improving prospects and thus is dominated by clicking or by quitting. Recall that the skip value is constant across states, while the click value obviously is increasing in h for given n. Hence the optimal decision typically is of the form: Click iff $h \geq h^*(n)$, for some reservation value function $h^*(n)$.

Case 4b, no replacement, is the same as case 4a except that the expectations in the click value use the probability $p = p(h, n)$ defined in equation (11). For reasons noted earlier, the probability here is not monotone in h for given n, and therefore the optimal decision cannot be characterized by a reservation value.

Recall that in Case 2a the island value $bw \in \{bL, ..., bU\}$ is observed, so the relevant state now is $s = (w, h, n)$. Hence in this case the expression $V(l, s_o)$ in equation (14) expands to $E_w V(l, w, 0, 0)$. The click value is the same as in Case 4a except that the hit probability now comes from equation (6). The skip value is more complicated because new islands with low w should be skipped immediately. The skip value satisfies

$$S(l) = -c + V(l, s_o) = -c + E_w V(l, w, 0, 0), \tag{18}$$

where the value of arriving at a new island of value bw satisfies

$$V(l, w, 0, 0) = \max\{C(l, w, 0, 0), S(l), 0\}. \tag{19}$$

The difficulty is that equations (18) and (19) do not tell us directly whether to click or skip on a new island; the skip value $S(l)$ in (18) also enters the right hand side of (19).

To work it out, recall that the skip value to be determined is independent of the new island value w, while the click value $k_w = C(l, w, 0, 0)$ is increasing in w because it takes an expectation using probability $p = w/20$. Hence there is some threshold $w^*(l)$ such that optimally one clicks at least once on an island iff the displayed $w \geq w^*(l)$. The tentative skip value $T(w_o)$ is the value obtained using an arbitrary threshold $w_o \in \{L, ..., U\}$. By definition, $T(w_o) = -c + \frac{1}{U-L+1}\left[\sum_{w=L}^{w_o-1} T(w_o) + \sum_{w=w_o}^{U} k_w\right]$, so

$$T(w_o) = \frac{\sum_{w=w_o}^{U} k_w - (U - L + 1)c}{U - w_o + 1}. \tag{20}$$

The optimal threshold is the smallest number of treasures for which the tentative click surplus $K_w = k_w - T(w)$ is positive, i.e., $w^*(l) = \min\{w : K_w \geq 0\}$, and the true skip value is $S(l) = T(w^*(l))$.

The solution is straightforward to compute and well behaved because K_w is an increasing sequence in w that is positive for $w = U$. To see this, use (20) to write $K_w = k_w - \overline{k_{w+}} + b_w c$. The term $k_w - \overline{k_{w+}}$ is increasing because k_w increases faster than its (upper) average $\overline{k_{w+}} = \frac{\sum_{v=w}^{U} k_v}{U-w+1}$. The cost coefficient $b_w = \frac{U-L+1}{U-w+1}$ is also increasing in w. Clearly $K_U = 0 + (U - L + 1)c > 0$.

One last subcase remains. When the number of clicks remaining on arrival at a new island is not divisible by $n_{max} = 20$, then Case 2b no longer reduces to Case 1. For example, one should not skip to the next island when the value of a new island is slightly below R when only 10 clicks remain, because the marginal benefit of skipping to a new island is depressed but the marginal cost is not. By backward induction, this complication also affects choices when 30 clicks remain on arrival to a new island, etc. Although every subject in the experiment begins every voyage with a multiple of 20 clicks, most subjects depart from optimal strategy at some point and find themselves in the situation just described.

We compute the (henceforth) optimal strategy in this subcase essentially the same way as in Case 2a. Use the displayed hit probability p to compute the click values k_p. Note that the possible values of p are discrete multiples of 0.01 and, since $b = 5$, they can be indexed $i = 100p_i$, where $i = 5L, 5L+1, ..., 5U$. Thus in this case, equations (18) and (19) reduce to

$$S(l) = -c + E_p V(l, p, 0, 0) = -c + \sum_{i=5L}^{5U} \max\{k_{p_i}, S(l), 0\}. \tag{21}$$

Now (21)can be solved to yield

$$S(l) = \frac{\sum_{i=i^*}^{5U} k_{p_i} - (5U - 5L + 1)c}{5U - i^* + 1}, \tag{22}$$

where i^* is the threshold index, for which k_{p_i} first exceeds $S(l)$.

Matlab Code

The Matlab code below implements the Bellman equation approach just outlined. For each case, each l and each state s it computes the click value, the skip value, and their difference z. Each choice i by each subject in the experiment then is compared to the optimal choice. It is an error ($i \in \mathcal{E}$) if it is not optimal. The decision errors then are tabulated and their costs ($SL = -\sum_{i\in\mathcal{E}} \min\{z_i, 0\}$ and $IL = \sum_{i\in\mathcal{E}} \max\{z_i, 0\}$) are summed.

The code departs from the conventions above in a few minor respects. It uses i instead of l as the index for clicks remaining, since the letter l and the number 1 are indistinguishable in Matlab. Also, since the index for a Matlab array has to start at 1 and cannot start at 0, the actual number of clicks remaining is $i-1$, not i. Similarly, the actual number of hits on the island is $h-1$, and the number of clicks on the island is $n-1$.

Matlab Code

```
% This script is for cases 2b, 4a and 4b;

tic;                    % start the timer;

% Experimental parameters
clicks = 100;           % initial click budget;
L = 2;                  % minimum number of treasures per island;
U = 18;                 % maximum number of treasures per island;
nmax = 20;              % maximum number of clicks on each island;
b = 5;                  % value in points of each treasure;
cbar = 10;              % expected travel cost;
Case='2b';              % 2b, 4a (replacement), 4b (no replacement);


% Sizing and initializing the data structure (the set of V(.) tables)
if Case=='2b'
   IVs=U-L+1;           % number of different island values;
   S=2;                 % used for sizing the staysurplus variable;
else
   IVs=1;               % Cases 4a and 4b don't show island value;
   S=1;                 % used for sizing the staysurplus variable;
end

V = cell(clicks+1,IVs);           % clicks+1 as V{1,w} is for 0 clicks;
for w=1:IVs                       % w= # of treasures on the island;
   V{1,w}=zeros(nmax+1, nmax+1);  % value=0 if no clicks remain
end

firstClickValue=zeros(IVs,1);
staysurplus = cell(clicks*IVs,S);
w_star=zeros(clicks);
Not_satisfied=0;        % =max(# clicks) when condition was not met (2b);


% Calculate the value function for each value of i (=clicks remaining +1);
for i = 2:(clicks+1)              % i=2 means 1 click remaining;
   ['***************** clicks remaining= ' num2str(i-1) '*****************']

   % Calculate the skip value;
   if Case=='4a' | Case=='4b'
      p0=(L+U)/(2*nmax);
      skipValue = -cbar + b * p0 + p0 * V{i-1}(2,2) + (1-p0) * V{i-1}(1,2);
   end
   if Case=='2b'
      % Calculate the initial click values for (n,h)=(0,0);
      for j=1:IVs       % j is an index covering all island values, but starting at 1, not L;
         w=j+L-1;       % Number of treasures on island;
         p0=w/nmax;     % Initial probability of finding a treasure;
         firstClickValue(j)=p0*b+p0*V{i-1,j}(2,2)+(1-p0)*V{i-1,j}(1,2);
      end

      % Calculate all possible skip values and identify the true one;
      satisfied=0;
      w0=L;
      while satisfied==0 & w0<=U
```

```
            skipValue = -(U-L+1)*cbar/(U-w0+1);     % assuming for now that w*=w0;
            for w=w0:U
               j=w-L+1;
               skipValue = skipValue + firstClickValue(j)/(U-w0+1);
            end
            if firstClickValue(w0-L+1) >= skipValue % checking if assumption was justified;
               satisfied=1;
               w_star(i-1)=w0;                    % w*=w0 is correct;
               ['w*=' num2str(w0) ', skip value: ' num2str(skipValue)]
            end
            %['w*=' num2str(w0)) ', k(w*-1)=' num2str(firstClickValue(w0-L)) ', skip value=' num2str(skipValue0) ',
k(w*)=' num2str(firstClickValue(w0-L+1))]
            w0=w0+1; % if assumption wasn't justified, try again with higher value for w0;
         end
         if satisfied==0
            Not_satisfied=i-1;
         end
      end

      % Calculate the value function;
      for j=1:IVs                        % j=1 for cases 4a and 4b;
         w=j+L-1;                        % Number of treasures on island;
         staysurplus{(i-2)*IVs+j,S}=w; % last column shows island value (w);

         % Boundary condition (all sites uncovered, n=20)
         for h = 1:(nmax+1)
            V{i,j}(h, nmax+1) = max(skipValue,0);
         end

         % Calculate the value function for n=0-19
         p=-1;
         staysurplus{(i-2)*IVs+j,1}=zeros(nmax,nmax);
         for n = 1:nmax      % n=1 means 0 clicks spent on the island so far;
            for h = 1:n

               % Calculate the prob. of hitting a treasure on the next click;
               if Case=='2b'
                  p = prob2b(h-1,n-1,w);         % calls .m-file 'prob2b'
               end
               if Case=='4a'
                  p = prob4a(h-1,n-1,L,U,nmax);  % calls .m-file 'prob4a'
               end
               if Case=='4b'
                  p = prob4b(h-1,n-1,L,U);       % calls .m-file 'prob4b'
               end

               % Calculating the click value and the value function;
               if p>=0     % p=-1 for impossible combinations of h,n,w,L,U;
                  clickValue=b*p+p*V{i-1,j}(h+1,n+1)+(1-p)*V{i-1,j}(h, n+1);
                  V{i,j}(h, n) = max(max(clickValue, skipValue), 0);
                  staysurplus{(i-2)*IVs+j,1}(h, n) = clickValue - skipValue;
               end
            end
         end
      end
end
```

```
if Case=='2b'
   plot(w_star)
   xlabel('Clicks remaining');
   ylabel('w*');
end
Not_satisfied
toc            % stop the timer;

% Saving the staysurplus (clickValue - skipValue);

tic;
% File name: island value range, cbar, Case;
name=['staysurplus_',num2str(L*b),num2str(U*b),'_',num2str(cbar),'_',Case,'.txt'];
output=zeros(clicks*IVs*nmax,nmax+S-1);
for i = 1:clicks*IVs
   output((i-1)*nmax+1:i*nmax,1:nmax)=staysurplus{i,1};     % n0-19;
   if Case=='2b'
      output((i-1)*nmax+1:i*nmax,nmax+1)=staysurplus{i,2};  % island value;
   end
end
%dlmwrite(name, output, ' ');
['Output saved to file']
toc
```