Monty Hall's Three Doors: Construction and Deconstruction of a Choice Anomaly

Daniel Friedman

Monty Hall's Three Doors: Construction and Deconstruction of a Choice Anomaly

By DANIEL FRIEDMAN *

Sometimes people make decisions that seem inconsistent with rational choice theory. We have a "choice anomaly" when such decisions are systematic and well documented. From a few isolated examples such as the Maurice Allais (1953) paradox and the probability matching puzzle of William K. Estes (1954), the set of anomalies expanded dramatically in the last two decades, especially following the work of Daniel Kahneman and Amos Tversky (e.g., 1979). By now the empirical literature offers dozens of interrelated anomalies documented in hundreds of articles and surveys (e.g., Colin F. Camerer, 1995).

The problem is not just theoretical. Economists' traditional analytical tools are based squarely on rational choice theory, so anomalies provoke debates about the practical value of these tools. The present paper offers some new laboratory evidence bearing on the anomalies debates. It focuses on a particular choice task, chosen (after a search) for its robustness and persistence.

I. Construction

As recounted by Barry Nalebuff (1987), host Monty Hall of the once-popular TV game

show "Let's Make A Deal" asked his final guest of the day to choose one of three doors (or curtains). One door led to the "grand prize" such as a new car and the other two doors led to "zonks" or worthless prizes such as goats. After the guest chose a door, Monty always opened one of the other two doors to reveal a zonk and always offered the guest the opportunity to switch her choice to the remaining unopened door. The stylized fact is that very few guests accepted the opportunity to switch.

Nonswitching is anomalous because in the game just described the probability of winning is 1/3 for nonswitchers and 2/3 for switchers, as explained in popular articles by Steven Selvin (1975a, b), Nalebuff (1987), Marilyn vos Savant (1990), and Leonard Gillman (1992). Forgoing the incremental probability of 1/3 on more than $10,000 in incremental prize value represents a serious loss to a nonswitching guest. Indeed, it might seem that millions of dollars were left on the table because the game ran in the United States daily from 1963–1968 and later weekly until as late as 1986, with over 3,800 shows altogether (Alex McNeil, 1991 pp. 369–70).

But the stylized fact is not quite what it seems. In reply to my recent questions by phone and e-mail, Monty Hall says that to the best of his recollection, he rarely offered guests the switch option and cannot say how often it was accepted (see also John Tierney, 1991). His longtime production assistant, Carol Andrews, asserts that Monty never offered the switch option to the final guest. Nalebuff says he personally recalls watching the show and seeing the switch option offered (as do I), but he cannot recall "whether Monty offered this option all the time and whether or not his making the offer was at all connected to whether you picked the right door."

Recovered memories are unreliable (Elizabeth F. Loftus, 1997) so one might wish to extract evidence from tapes of Monty Hall's show. Unfortunately there are two problems. First, the tapes are not available. Inquiries to Andrews and to the Museums of Broadcasting in New York and Los Angeles disclose available tapes for at most five shows, hardly an adequate sample. Second, the game show environment was poorly controlled. Some of the difficulties can be seen in the work of Andrew Metrick (1995), who analyzed tapes from another popular TV game show, "Jeopardy." Metrick finds that people do not always best respond in a simultaneous move game ("Final Jeopardy"). However, his "empirical best response" assumes that each player faces the overall empirical distribution of play. Actual players may have well-founded beliefs that their specific opponents differ from average, and they may have idiosyncratic preferences (e.g., looking timid while losing is worse than going for it but still losing). Hence, despite careful work, Metrick's failure rate may be overstated. The three-door game poses additional problems. Perhaps guests were victims of Monty Hall's unique personality or their own "room temperature IQ's" (Hal Erickson, 1989 p. 230). Even worse, if Monty offered the option to switch only when the guest's initial choice was correct, then nonswitching would be optimal after all.

Existing academic literature includes at least two studies inspired by Monty Hall's game show, but none to my knowledge that actually tests the three-door choice task. Ruma Falk (1992) analyzes conceptually some related judgment tasks. In order to test regret theory, Thomas Gilovich et al. (1995) conduct laboratory experiments that disguise an unwinnable task as the three-door task.

A. Method

The essential ingredients of the three-door task are an initial risky choice, the revelation of information concerning a rejected choice, and the opportunity to reconsider the initial choice. My experiment uses these ingredients in a manner parallel to the stylized game show.

One hundred four subjects were recruited, mainly from large introductory classes in bi-

ology, organic chemistry, and politics at the University of California, Santa Cruz (UCSC) and at Cabrillo Community College. Each subject entered a quiet room and sat at a table opposite the conductor with no other subjects present. After reading the instructions, each subject completed a series of ten trials (or "periods"). In each trial, the subject initially picked one of three face-down cards. Then the conductor turned over a nonprize card that the subject did not choose and offered her the opportunity to switch to the other face-down card. Finally both face-down cards were turned over, one of which was the prize card. Each trial the subject earned 40 cents when her final choice turned out to be the prize card and 10 cents otherwise. The instructions (available on request) briefly and unambiguously explained the procedure.

B. Results

Table 1 summarizes the results. Subjects' relevant decision each period is binary, either to switch (the rational choice) or to remain with the card initially chosen (the irrational choice). UCSC Humanities (and Arts) students switched most often (in 34.4 percent of 250 trials) and UCSC Natural Science majors switched least often (in 23.3 percent of 270 trials). Subjects who took only lower-division statistics courses switched less often than those with more statistics or none; apparently here, as elsewhere, a little knowledge is a dangerous thing! Female subjects were slightly more rational (31.8 percent) than males (24.9 percent). Of the 104 subjects tested, only 6 switched more than half the time. (One of the 6 subjects was already familiar with the task, having read and remembered vos Savant's [1990] column in *Parade* magazine; no other subject admitted to prior knowledge even at the exit interview after final payment.)

Figure 1 shows the time trend. The switch rate started out extremely low (less than 10 percent in the first trial) and increased fairly steadily over the next several trials. But it stagnates at about 40 percent after the sixth trial and actually declines in the last trial to about 30 percent. The overall switch percentage is 28.7 percent.

TABLE 1—CHOICES AND EARNINGS IN RUN1

Pool	Nobs	Switch rate percent	Mean earnings
School			
UCSC—Lower division	420	26.7	$2.35
UCSC—Upper division	340	32.0	$2.30
Other schools	280	27.1	$2.31
Major			
Natural Sciences	270	23.3	$2.32
Social Sciences	360	30.6	$2.30
Humanities/Arts	250	34.4	$2.37
Undeclared	160	23.8	$2.37
Sex			
Male	490	24.9	$2.34
Female	550	31.8	$2.32
Previous statistics class			
None	620	31.0	$2.38
Lower division	340	24.2	$2.32
Upper division	80	28.8	$1.98
Age			
17–19	440	27.3	$2.32
29–23	440	31.3	$2.35
24+	160	24.4	$2.31

C. Discussion

The three-door task is now a true anomaly. The data are not hypothetical: more than 100 real people left lots of real money on the table in a controlled laboratory setting. The observed behavior is not "approximately" rational; most of the people most of the time made the irrational choice when the rational choice was just as convenient. And with ten trials, each laboratory subject had the opportunity to become familiar with the task and to find the rational decision.

How might this anomaly be explained? A straightforward response to any choice anomaly is to accommodate it by relaxing some of the standard axioms of subjective expected utility and Bayesian decision theory. Such responses to previous anomalies has produced a vast theoretical literature on generalized expected utility; for an introduction see, for example, Mark J. Machina (1987) or Camerer

and Martin W. Weber (1992). Unfortunately, many of the leading theories still forbid selection of a stochastically dominated action and so cannot accommodate the new data. Generalizations that can accommodate dominated choices are unlikely to generate sharp enough predictions to be useful in applications. For present purposes, perhaps the best way to proceed is to review briefly some related anomalies.

First, the three-door anomaly may be related to the "gambler's fallacy" or more generally, the "illusion of control" (Camerer, 1995). A subject may believe that, despite our best attempts to randomize and a careful explanation of our attempts, he can somehow intuit which face-down card is the prize card and so he may regard his initial choice as the most likely. Revealing the nonprize status of a card he already believed less likely (especially if he thought it second most likely) may then only confirm his initial belief. Such beliefs conceivably were rational in the TV game show, but clearly are irrational in the trials just described because 35.0 percent of the choices to remain won the prize, while 70.3 percent of the choices to switch were winners.

Second, the anomaly may be related to the "nonrational escalation of commitment" featured in Max H. Bazerman (1990). Having made his initial choice, the subject is reluctant to change it; he sees persistence as a virtue and flip-flopping as a vice. Economists might regard this as a cousin of the sunk cost fallacy: the subject rationally should regard his initial decision as sunk and of no relevance to his decision between the last two cards, but he may be psychologically unable to let bygones be bygones. Such reluctance to switch has been called the "endowment effect" in a slightly different context, e.g., Jack L. Knetsch (1989). As noted before, reluctance to switch might be rational in the TV game show if Monty offered the switch option more often when the initial choice was correct, but it definitely is irrational in the experiment where the switch option is always offered.

Third, subjects often misapply Bayesian updating in various ways (Camerer, 1995) and may have done so here. With a prior probability of 1/3 for each card, a random information process (i.e., equally likely a priori to
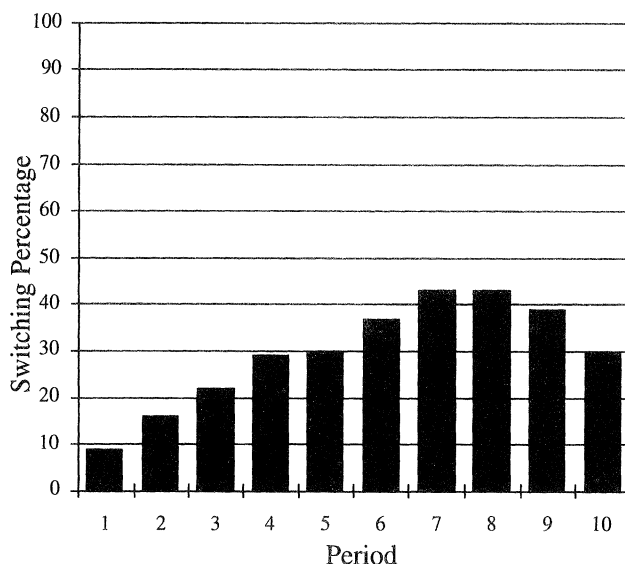
FIGURE 1. SWITCHING PERCENTAGE IN RUN1

reveal any one of the cards) that happens to reveal a nonprize, nonchosen card will yield Bayesian posterior probabilities of 1/2 for each of the nonchosen cards. That is, in accord with most people's intuition, Laplace's principle of insufficient reason applies to the posterior as well as to the prior probabilities. In this case there would be no advantage (or disadvantage) to switching. However, the actual information process is not random; the conductor *must* reveal a nonprize, nonchosen card, so the Bayesian posterior probability remains at 1/3 for the initially chosen card and moves to 2/3 for the other card. Such a nonrandom process is outside most people's experience, so they may be unaware of the advantage of switching to the card with the 2/3 win probability.

A fourth anomaly may be worth mentioning. Probability matching behavior, documented in numerous studies such as Estes (1954), consists of not always choosing the most likely alternative (as rationality would dictate), but rather choosing each alternative in proportion to its likelihood. Like the Bayesian failure just discussed, probability matching is not potentially a complete explanation of the three-door anomaly because it would predict a 50-percent switch rate for confused Bayesians

and a 67-percent switch rate for good Bayesians. Probability matching may partially account for the rarity of subjects who always switch. The underlying reasons for probability matching are unclear (it seems to be regarded as innate in much of the psychological literature) but Sidney Siegel (1961) demonstrates that its effects diminish when salient cash rewards become more intense.

One conclusion seems clear. The three-door task now deserves a place among the leading choice anomalies. The laboratory evidence strongly corroborates anecdotal evidence (e.g., Massimo Piattelli-Palmarini, 1994).[1] Indeed, I am not aware of any anomaly that has produced stronger departures from rationality in a controlled laboratory environment. The ultimate reasons for its strength remain unclear, but a proximate reason is that the task

---

[1] This popular book on "mental tunnels," i.e., choice anomalies, refers to the three-door task as a "supertunnel ... the best and most striking instance" (p. 7). Despite the fact that he offers only anecdotal evidence, Piattelli-Palmarini devotes his last chapter ("Grand Finale") to the three-door task and emphasizes the Bayesian update failure.

ties together several classic anomalies, including the gambler's fallacy, commitment escalation and the endowment effect, Bayesian updating failures, and probability matching. It would appear difficult to accommodate the three-door anomaly into a useful generalization of expected utility theory.

II. Deconstruction

The anomalous laboratory evidence confronts economists with at least two hard questions. Why are people so irrational? When (if ever) can people become more rational in tasks of this sort? In an attempt to find answers to these questions, I had each subject in the experiment continue with a second part.

A. Method

After completing Run1, the ten trials described in the previous section, each subject was paid accumulated earnings (between $1.00 and $4.00) in cash and asked to complete a second set of 12 or 15 trials, called Run2. The baseline procedures for Run2 were exactly as in Run1 with the addition of a non-salient $3.00 cash payment. Most subjects in Run2 received one or more of the following alternative treatments, chosen to encourage more effective learning.

1. Intense incentives (referred to below as *Intense*).—The prize card was worth +$1.00 each trial and the other two cards were worth −$0.50. The initial $3.00 cash payment was put on the table at the beginning of Run2. Each trial $1.00 was added or $0.50 subtracted from the pile; at the end of the run, the subject received the final pile.[2] Note that each period the expected value for remaining was $(1/3)\$1 + (2/3)(-\$0.50) = \$0.00$ and for switching was $(2/3)\$1 + (1/3)(-\$0.50) = \$0.50$ with intense incentives, versus $0.20 and $0.30 respectively for remaining and switching with the baseline incentives.

2. Track record (*Track*).—Each subject wrote down the results each period on a record sheet that included her own Run2 cumulative cash earnings as well as the cumulative earnings for the strategy "always remain" and for the strategy "always switch." Usually (but not invariably) the "always switch" strategy did best.

3. Written advice (*Advice*).—Before the first period of Run2, a subject in this treatment was handed a page with two paragraphs in random order. One paragraph recommended always switching and explained succinctly why switching improves the odds. The other paragraph, written in a similar style, recommended always remaining and tried to evoke the "escalation" and "gambler's fallacy" justifications. The paragraphs, reproduced in the Appendix, were not presented separately to avoid confounding demand effects with learning through written advice. Questionnaires reveal that subjects initially found the paragraphs equally persuasive.

4. Comparative results (*Compare*).—The broad results for the first 40 subjects were compiled and summarized for some later subjects. After the sixth period in Run2, each subject in this treatment received a four-line statement pointing out that 62.3 percent of the switch choices won the prize versus 30.5 percent of the remain choices. The summary also revealed that the majority of choices were to remain. The conductor believes that all subjects in this condition found the summary credible.

The Run2 data comprise 1,407 observations of the binary variables choice and outcome, gathered from 103 subjects. The subject who had prior knowledge of the task is excluded from the analysis to follow because his choices (almost all to switch) might be incorrectly attributed to the treatments he received (*Intense* and *Advice*) rather than to his prior knowledge. The data include the actual choices of the seven subjects dismissed early because of bankruptcy, but of course exclude the choices they would have made in later periods had they not been dismissed (about four periods per bankrupt subject). As explained in the Appendix, the design mixes the treatments in a balanced fashion to permit unbiased estimation of the direct and pairwise impacts of the treatments.

---

[2] Seven subjects experienced bankruptcy (i.e., exhausted the cash pile before the last trial), most of them in trial 10 or later, at which point they were dismissed and received no further payment. A referee notes correctly that the *Intense* treatment combines higher stakes with possible asymmetric responses to gains and losses.
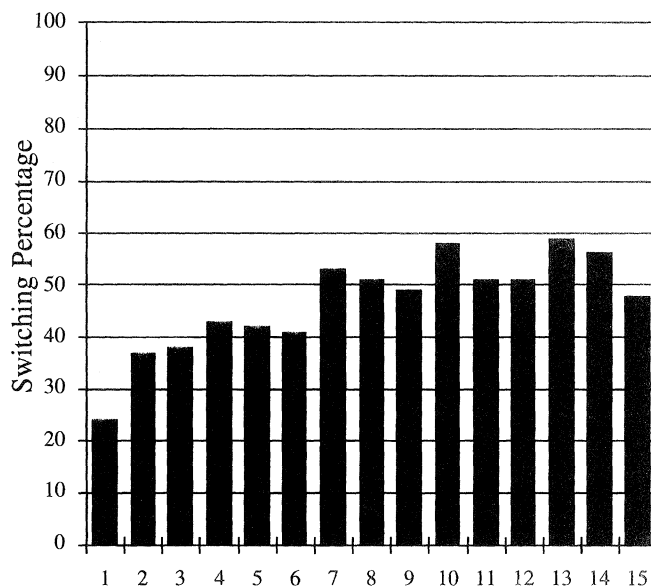
FIGURE 2. SWITCHING PERCENTAGE IN RUN2

B. Results

Surprisingly, the switch rate in Run2 did not begin where it left off in Run1. Figure 2 shows that the overall switch rate was less than 25 percent in the first period, but soon rose above 40 percent and remained mostly in the 40–50-percent range after period 4. It rose to 55 percent in period 10 and trailed off somewhat in the last few periods.

Table 2 summarizes the impact of the treatments. Overall, 46.0 percent of the 1,407 choices were to switch. The switch rate rose from 39.8 percent in the earlier periods (1–7) to 52.6 percent in the later periods (8–15). The increase is highly significant (at least the 0.0001 level) according to Fisher's exact test (e.g., William J. Conover, 1980). Intense incentives did not increase the switch rate, mainly because in early periods only 38 percent of subjects with intense incentives chose to switch. (By luck of the draw, many of the subjects with intense incentives did better, or would have done better, by not switching in the first two or three periods, and it took a while for the impact of early experience with intense incentives to die out.)

The other three treatments increased the frequency of rational behavior. The impact of *Track* was not quite significant at the 0.10 level in either early or late periods separately, but was significant at the 0.05 level overall according to the Fisher exact test. *Advice* had no significant effect in early periods (or overall) but was significant at the 10-percent level in later periods; perhaps it took some time for subjects to figure out which part of the advice jibed with experience. By design, *Compare* can have no effect on periods 1–6, so it is no surprise that it is insignificant in early periods. It is highly significant in later periods and overall. The combination of *Advice* and *Compare* may be especially powerful. Switch rates in later periods remained above 50 percent for the 20 subjects with this combination and exceeded 70 percent in several periods.

Table 3 presents a finer-grained analysis. The dependent variable all regressions is binary choice, with value 1 if the subject switched and 0 if the subject remained that period. The controls for nuisance variables include a constant (here the same for all subjects, but in the Appendix allowed to vary across subjects) and the trend variable *Period*

TABLE 2—SWITCH RATES BY TREATMENT

		Nobs	Percent All periods	Percent Periods 1–7	Percent Periods 8–15
0	Overall	1,407	46.0	39.8	52.6
	(*p*-value)				(0.000)
1	*Intense*	791	43.9	37.9	51.3
	Not *Intense*	616	48.7	42.7	54.0
	(*p*-value)		(0.968)	(0.915)	(0.782)
2	*Track*	647	48.0	41.9	54.8
	No *Track*	760	43.4	36.9	49.8
	(*p*-value)		(0.048)	(0.101)	(0.113)
3	*Advice*	804	47.4	40.6	55.0
	No *Advice*	603	44.1	38.7	49.5
	(*p*-value)		(0.122)	(0.327)	(0.087)
4	*Compare*	647	50.3	40.2	59.4
	No *Compare*	760	42.4	39.5	46.0
	(*p*-value)		(0.001)	(0.458)	(0.000)

Notes: Numbers of observations (Nobs) are almost evenly divided between Periods 1–7 (overall 51.6 percent) and Periods 8–15 (overall 48.4 percent of Nobs). Reported *p*-values are the probability that Fisher's exact test incorrectly rejects the null hypothesis of no effect in favor of the one-sided alternative hypothesis that the treatment (or period number in item 0) has a positive effect on switch rates. The four treatments (*Intense, Track, Advice,* and *Compare*) are defined in the text.

(the period number in Run1 or in Run2). The explanatory variables include dummies for each of the four treatments and for the interaction of interest, *Advice *Compare*. Two other explanatory variables are suggested by various learning theories (e.g., Drew Fudenberg and David Levine, 1998). Fictitious play (or reinforcement learning) suggests the variable *Switchbonus*, defined as the earnings from always switching minus earnings from always remaining, cumulated from the first period in Run1 to the most recent period. Cournot behavior (or directional learning, e.g., Reinhard Selten and Joachim Buchta, 1995) suggests *Switchwon*, a dummy variable equal to 1 if and only if switching would have won the prize in the most recent period. For completeness, the table also considers the interactions of the continuous learning variable with *Advice* and *Compare*.

The ordinary least-squares (OLS) estimates confirm the Table 2 findings on the four treatment variables, but do not confirm the conjecture that *Advice* and *Compare* are more powerful in tandem; the relevant interaction near the bottom of Table 3 is insignificant and

negative. The significant positive coefficient estimates for *Switchbonus* and *Period* suggest that subjects do learn from personal experience and from time on task. The coefficient estimates for *Switchwon* and the interactions raise the possibility of weak recency effects and subadditive impacts of the treatments with personal experience.

Two caveats are in order. First, the significance levels in both tables assume independent observations, but in fact successive observations from any given subject share a common history and therefore are unlikely to be truly independent. Second, the dependent variable in the regressions has limited (in fact binary) range, so the linear model underlying the OLS estimates is inappropriate. Orley Ashenfelter et al (1992) copes with analogous econometric problems arising in pairwise bargaining experiments by supplementing the OLS estimates with probit estimates, with and without fixed pair effects. In the Appendix, I report probit estimates with and without random individual subject effects; the estimates confirm and extend the inferences drawn from Table 3.

TABLE 3—OLS REGRESSIONS FOR SWITCH RATE

Variable	Run1	t-statistic	Run2	t-statistic
Constant	0.125 (0.029)	4.24	0.191 (0.043)	4.42
Period	0.017 (0.005)	3.18	0.012 (0.003)	3.71
Switchbonus (cumulative earnings for switching − not switching)	0.116 (0.024)	4.85	0.030 (0.008)	3.95
Switchwon (=1 if switch won in previous period)	0.039 (0.038)	1.03	0.114 (0.030)	3.83
Intense (sensitive earnings)	—	—	−0.089 (0.029)	−3.05
Track (write out cumulative earnings for all strategies)	—	—	0.104 (0.039)	2.63
Advice (read conflicting explanations)	—	—	0.127 (0.050)	2.52
Compare (told other subjects' performance)	—	—	0.080 (0.043)	1.87
Switchbonus* Track	—	—	−0.017 (0.010)	−1.63
Switchbonus* Advice	—	—	−0.014 (0.009)	−1.55
Advice* Compare	—	—	−0.045 (0.059)	−0.77
Adjusted R^2	0.061	—	0.053	—
NOBs	1,040	—	1,407	—

Notes: The dependent variable is 1 in periods the subject chose to switch and is zero in other periods. The table reports ordinary least-squares coefficient estimates (and standard errors) and student t-statistics.

The significantly positive point estimates for *Switchbonus* have the important implication that switching will usually be chosen when sufficiently favorable evidence accumulates. Indeed, the probit estimates in the Appendix imply that subjects would choose switch more than 90 percent of the time after about $T = 60$ Run2 trials with intense incentives. The calculation is as follows. Set $0.90 = N(z)$, or $z \approx 1.3$, for the unit normal cumulative distribution function N with argument $z = \text{const} + \text{coeff} \times Switchbonus$. Recall that with intense incentives the expected value of *Switchbonus* is $0.50 \times T$, and solve for T. With const $= -0.863 - 0.238 \approx 1.1$ and coeff $= 0.084$ from the last column of Table

A2 the solution is $T \approx 57$. I rely on this indirect evidence because running 60 trials by hand seems infeasible, and even on the computer there is the problem of credibly precommitting the prize card position before the subject decides whether to switch. Of course, there is also the usual boredom problem. Solutions to these problems are possible (e.g., multisession by-hand experiments with sequestered subjects) but not easy.

C. *Discussion*

Evidently people are capable of learning to overcome the three-door anomaly. Most of

the people made the rational choice most of the time in the more favorable conditions. The rationality rate sometimes exceeded the classic "probability matching" benchmark of 2/3 for transparent risky choices with the same odds. Coefficient estimates imply that, given a sufficiently favorable track record, subjects would make the rational choice almost all the time.

The three-door task is not uniquely easy to learn. To the contrary, the evidence in Section I shows that it is exceptionally difficult. If people are not hardwired to behave irrationally here, then perhaps they are capable of learning to behave rationally in other tasks. Indeed, it now seems reasonable to make the following.

Assertion.—Every choice "anomaly" can be greatly diminished or entirely eliminated in appropriately structured learning environments.

Gerd Gigerenzer (1991, and numerous later studies) offers many supporting examples; I welcome readers' suggestions of possible counterexamples to the assertion.

If the assertion is correct then there is no such thing as an anomaly in the traditional sense of stable behavior that is inconsistent with rationality. There are only pseudo-anomalies describing transient behavior before the learning process has been completed.

The argument can be taken a step further. The contrapositive form of the assertion says that the absence of adequate learning opportunities is a necessary condition for observing pseudo-anomalies. A moment's reflection on sufficient conditions leads to the following.

RECIPE FOR PSEUDO-ANOMALIES
Ingredients
1. A common and useful rule of thumb, e.g., procrastinate when anticipating important new information.
2. A laboratory environment that evokes the rule of thumb when it is inappropriate, e.g., a lab environment where new information is never important.
3. Subjects inexperienced in the laboratory environment.
4. Standard laboratory procedures.
5. Standard statistical techniques and standard rhetoric.

Directions
Put subjects in the laboratory environment. Run them through the standard procedures, carefully excluding any opportunities for them to learn from their experience or from imitating more successful subjects. Tabulate and analyze the results statistically. Write up the results using rhetoric appropriate for your target social science journal. This recipe can serve for one or more publications and may produce a new entry in the pantheon of choice anomalies.

I do not claim that previous authors have consciously used the recipe, but it may give some insight into even the best empirical anomalies papers. Take Craig Fox and Tversky (1995), for example, the first anomaly paper I encountered after writing down the recipe. The underlying rule of thumb is to avoid decisions when ill informed, especially to avoid dealing with better-informed people with divergent interests. Such rules definitely are useful in most field settings and can be derived rigorously from expected utility theory under appropriate assumptions (e.g., Jack Hirshleifer and John G. Riley, 1992 pp. 204ff, 307ff). Using variations on the famous Daniel Ellsberg (1961) paradox, the laboratory environment evokes comparisons to better-informed people and to tasks about which the subject has better information. It is inappropriate to avoid decisions in the laboratory environment because better information will not be forthcoming nor will the subject actually contract with better-informed people. Even though the authors use mild financial incentives at one point, the lab environments all are one shot with no opportunities to learn from own (or others') experience. Standard statistical techniques demonstrate that subjects avoid the ill-informed tasks when they are compared to better-informed tasks or better-informed people. The rhetoric is a bit subdued but appropriate for an economics journal; the authors say that they ". . . provide evidence against the descriptive validity of expected utility theory" (p. 586) that may be consistent with some generalizations of that theory.

Does the recipe work when followed consciously? The main problem is in finding rules of thumb that have not already been used. For

example, I discovered that the procrastination rule of thumb has already been thoroughly and imaginatively studied in Tversky and Eldar Shafir (1992). In Section I of this paper I used instead a combination of the less-exploited rules, "don't flip-flop" and "insufficient reason." Opportunities to use the recipe still remain for social scientists with imagination and a knowledge of the literature.

What are the theoretical implications? Optical illusions arising from misleading visual cues are interesting but do not imply the need to modify the theory of optics. Likewise, irrational choices arising from incomplete learning do not imply the need to modify standard choice theory.

III. Conclusion

Some readers will rightfully object to the facile tone of the preceding subsection. Despite the use of strong treatments, many subjects continued to make lots of irrational choices. The three-door anomaly declined substantially but did not actually disappear in 15 periods.

In my view, the real message to economists is neither "ignore anomalies because they will eventually disappear" nor "rewrite choice theory to accommodate the evidence." Rather, economists should focus their energies on two sorts of questions. First, which learning environments encourage or discourage specific kinds of anomalies? Second, which institutions are sensitive to anomalous choice behavior? Anomalies can be ignored when the economic situation of interest (for example, a simple competitive market) involves an institution insensitive to anomalous behavior or when the institution allows effective learning. Otherwise, economists should take anomalies seriously, not by relaxing axioms of choice theory, but rather by modelling the specific institutional sensitivity and/or the incomplete learning.

Exploration of the three-door anomaly is far from complete. It surely would be useful to develop the technology to run each subject for 30 or more trials and to fit individual learning models to the individual data. With or without such technology, new treatments could isolate the underlying reasons for the initial strength of the anomaly and the reasons for its persistence. One promising angle is to try to distance the subject from the initial choice among the three doors in order to reduce the strength of commitment and/or the illusion of control. Relevant treatments might include initial choice by a random device or by a second subject or by the same subject at an earlier time. Another angle is to manipulate the subject's initial choice by beginning with a more transparent task such as the million-door problem suggested in vos Savant (1990).

A third angle is to put the three-door task in a richer institutional framework such as a market in choice opportunities. A subject could switch, remain, or sell the choice or buy additional choices at a price set exogenously or in a double-auction market. Economists might predict that the market price would quickly move to the maximal expected value ($0.50 with intense incentives) minus a small risk premium, with only switchers buying and all nonswitchers selling. Psychologists (along with some other experimentalists and observers of human nature) might predict a core of stubborn nonswitching nonselling subjects, even when prices are above the confused Bayesian expected value ($0.25 with intense incentives). Such experiments, in tandem with the serious modelling of learning processes within economic institutions, hold the promise of ending the now-sterile debate on anomalies (see also L. Jonathan Cohen, 1981; Robin M. Hogarth and Melvin W. Reder, 1987; Lola Lopes, 1990; Vernon L. Smith, 1991) and creating a productive new research agenda on rationality.

APPENDIX: RUN2 DETAILS

Advice treatment text.—Please read the [following] two pieces of advice on how to earn money in this experiment. The pieces disagree on what you should do; you must make up your own mind.

Advice S1: You have 1 chance in 3 of picking the Prize card initially. If you did pick it initially you will win the prize if you Remain. You have 2 chances in 3 of not picking the Prize card initially. If you did not pick it initially you will win the Prize if you Switch. So, the overall chance of winning is 2 chances in

TABLE A1—NUMBER OF SUBJECTS IN ALTERNATIVE RUN2 TREATMENTS

| | Intense = 1 | | Intense = 0 | | |
	Track = 1	Track = 0	Track = 1	Track = 0	Total
Advice = 1					
Compare = 1	10	0	0	10	20
Compare = 0	20	10 + 1*	10	0	41
Advice = 0					
Compare = 1	0	11	11	1	23
Compare-0	10	0	0	10	20
TOTAL	40	22	21	21	104

Notes: The four binary treatment variables (Intense, Track, Advice, and Compare) are defined in the text. The eight cells with ten (or 11) subjects comprise a half-fractional factorial design that confounds Compare with the three-way interaction of the other variables. The asterisk (*) identifies an individual subject excluded from subsequent analysis because of prior knowledge.

3 if you Switch and only 1 in 3 chances if you Remain. Therefore, you will do the best in the long run if you always Switch.

Advice R1: The conductor may try to distract you by offers to Switch, but these offers are made for his/her own reasons and are not necessarily in your interest. When only 2 cards remain face down, you have at least 1 chance in 2 of picking the Prize card and Switching will not improve your chances. You should not let yourself be distracted. Therefore, you should always Remain with your initial choice and never Switch.

Design.—Table A1 shows that 80 subjects were run in conditions that comprise a half-fractional factorial design in the four treatments with ten subjects per cell. Twenty other subjects were run in the condition Track and Intense only; the remaining three subjects were run in miscellaneous conditions. The fractional factorial design allows estimation of the direct impact of each treatment and the two-way interactions but not the three-way interactions; see, e.g., George Box et al., 1978).

Probit estimation.—Table A2 reports maximum likelihood probit coefficient estimates using the Limdep 7.0 package on all data. Logit regressions, with and without fixed effects, produced very similar but less complete results; the software could not handle the full unbalanced panel in the logit regression corresponding to the last column in Table A2. See William H. Greene (1990) for a full discus-

sion of the relation between logit, probit, and standard regression techniques and a discussion of random and fixed effects.

The first two columns of Table A2 report the results for Run1. Random effects (individual subject differences in the basic tendency to switch) do appear to be present, as indicated by the diagnostic statistic Rho (significant at better than the 0.1-percent level) and by the improved log-likelihood reported at the bottom of the second column. The Switchwon coefficient changes sign but is insignificant either way. The other coefficients (and their significance levels) are quite consistent in both columns. The significantly negative constant indicates a strong tendency to remain rather than switch, all else equal. The significantly positive Period coefficient indicates an upward trend in the switch rate. Learning from accumulated experience has an even stronger and more significant impact as indicated by the Switchbonus coefficient.

The Run2 results in the remaining columns include the impact of the treatments (none of which were used in Run1) and tell a somewhat more complex story. The full specification in third column indicates that the average bias against switching is less than in Run1 but still highly significant. The −0.814 constant indicates a basic tendency to switch in about $N(-0.814) = 21$ percent of trials, where N is the unit normal cumulative distribution function. Trend as measured by the Period

TABLE A2—PROBIT ESTIMATES (AND p-VALUES) FOR SWITCH RATE

Variable	Run1	Run1	Run2	Run2
Constant	−1.090	−1.166	−0.814	−0.863
	(0.000)	(0.000)	(0.000)	(0.001)
Period	0.055	0.064	0.032	0.043
	(0.001)	(0.001)	(0.000)	(0.000)
Switchbonus	0.325	0.385	0.082	0.084
	(0.000)	(0.000)	(0.000)	(0.001)
Switchwon	0.106	−0.146	0.293	0.131
	(0.344)	(0.250)	(0.000)	(0.137)
Intense	—	—	−0.234	−0.238
			(0.002)	(0.109)
Track	—	—	0.276	0.286
			(0.009)	(0.117)
Advice	—	—	0.337	0.380
			(0.012)	(0.253)
Compare	—	—	0.208	0.229
			(0.069)	(0.411)
Switchbonus* Track	—	—	−0.045	−0.052
			(0.099)	(0.249)
Switchbonus* Advice	—	—	−0.038	−0.034
			(0.113)	(0.395)
Advice*Compare	—	—	−0.116	−0.087
			(0.458)	(0.799)
Random effects	No	Yes	No	Yes
Rho	—	0.384	—	0.453
		(0.000)		(0.000)
NOBs	1,040	1,040	1,407	1,407
−Log-likelihood	589.9	581.3	927.8	897.5

Notes: The dependent variable is 1 in periods the subject chose to swtich and is 0 in other periods. The table reports maximum likelihood probit coefficient estimates with and without random effects (and corresponding p-values). Rho is the standard Hausman test statistic for the presence of random effects.

coefficient is weaker than in Run1 but still highly significant. The fictitious play variable *Switchbonus* also remains highly significant. The Cournot variable *Switchwon* has a significantly positive coefficient here. Since *Switchbonus* already captures the most relevant positive aspects of intense incentives, the significantly negative *Intense* coefficient should be thought of as measuring a residual impact. (The incentives for most subjects in Run2 are five times more intense than in Run1. If subjects respond to the stronger incentives but less than proportionately, then we would see a negative *Intense* coefficient and/or a smaller *Switchbonus* coefficient in Run2, as indeed we do.) *Track* and *Advice* have positive and significant coefficients. The *Compare* coefficient is marginally significantly positive (at the $p = $ 6.9-percent level), but the a priori most important interactions are less significant.

The last column allows for random effects. The effects are highly significant and substantially improve the fit. The most important impacts are that the *Switchwon* dummy and all the interactions become insignificant, the treatments become less significant, and the learning proxies *Switchbonus* and *Period* become the most dominant explanatory variables.

REFERENCES

Allais, Maurice. "Le Comportement de l'homme Rationel Devant le Risque, Critique des Postulates et Axiomes de l'ecole Americaine." *Econometrica*, October 1953, *21*(4), pp. 503–46.

Ashenfelter, Orley; Currie, Janet; Farber, Henry S. and Spiegel, Matthew. "An Experimental Comparison of Dispute Rates in Alternative Arbitration Systems." *Econometrica*, November 1992, *60*(6), pp. 1407–33.

Bazerman, Max H. *Judgment in managerial decision making*. New York: Wiley, 1990.

Box, George; Hunter, William and Hunter, J. Stuart. *Statistics for experimenters*. New York: Wiley, 1978.

Camerer, Colin F. "Individual Decision Making," in John H. Kagel and Alvin E. Roth, eds., *The handbook of experimental economics*. Princeton, NJ: Princeton University Press, 1995, pp. 587–703.

Camerer, Colin F. and Weber, Martin W. "Recent Developments in Modelling Preferences: Uncertainty and Ambiguity." *Journal of Risk and Uncertainty*, October 1992, *5*(4), pp. 325–70.

Cohen, L. Jonathan. "Can Human Irrationality Be Experimentally Demonstrated?" *Behavior and Brain Sciences*, September 1981, *4*(3), pp. 317–70.

Conover, William J. *Practical nonparametric statistics*, 2nd Ed. New York: Wiley, 1980.

Ellsberg, Daniel. "Risk, Ambiguity, and the Savage Axioms." *Quarterly Journal of Economics*, November 1961, *75*(4), pp. 643–69.

Erickson, Hal. *Syndicated television: The first forty years, 1947–1987*. Jefferson, NC: McFarland, 1989.

Estes, William K. "Individual Behavior in Uncertain Situations: An Interpretation in Terms of Statistical Association Theory," in Robert M. Thrall, C. H. Coombs, and R. L. Davis, eds., *Decision processes*. New York: Wiley, 1954, pp. 127–37.

Falk, Ruma. "A Closer Look at the Probabilities of the Notorious Three Prisoners." *Cognition*, June 1992, *43*(3), pp. 197–223.

Fox, Craig and Tversky, Amos. "Ambiguity Aversion and Comparative Ignorance." *Quarterly Journal of Economics*, August 1995, *110*(3), pp. 585–603.

Fudenberg, Drew and Levine, David. *The theory of learning in games*. Cambridge, MA: MIT Press, 1998.

Gillman, Leonard. "The Car and the Goats." *American Mathematical Monthly*, January 1992, *99*(1), pp. 3–7.

Gigerenzer, Gerd. "How to Make Cognitive Illusions Disappear: Beyond 'Heuristics and Biases'." *European Review of Social Psychology*, March 1991, *2*(1), pp. 83–115.

Gilovich, Thomas; Medved, Victoria Husted and Chen, Serena. "Commission, Omission, and Dissonance Reduction: Coping with Regret in the 'Monty Hall' Problem." *Personality and Social Psychology Bulletin*, February 1995, *21*(2), pp. 182–90.

Greene, William H. *Econometric analysis*. New York: MacMillan, 1990.

Hirshleifer, Jack and Riley, John G. *The analytics of uncertainty and information*. Cambridge: Cambridge University Press, 1992.

Hogarth, Robin M. and Reder, Melvin W., eds. *Rational choice: The contrast between economics and psychology*. Chicago: University of Chicago Press, 1987.

Kahneman, Daniel and Tversky, Amos. "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica*, March 1979, *47*(2), pp. 263–91.

Knetsch, Jack L. "The Endowment Effect and Evidence of Nonreversible Indifference Curves." *American Economic Review*, December 1989, *79*(5), pp. 507–21.

Loftus, Elizabeth F. "Repressed Memory Accusations: Devastated Families and Devastated Patients." *Applied Cognitive Psychology*, February 1997, *11*(1), pp. 25–30.

Lopes, Lola. "The Rhetoric of Irrationality." *Theory and Psychology*, February 1990, *1*(1), pp. 65–82.

Machina, Mark J. "Choice Under Uncertainty: Problems Solved and Unsolved." *Journal*

of Economic Perspectives, Summer 1987, 1(1), pp. 121–54.

McNeil, Alex. Total television: A comprehensive guide to programming from 1948 to the present, 3rd Ed. New York: Penguin, 1991.

Metrick, Andrew. "A Natural Experiment in 'Jeopardy!' " American Economic Review, March 1995, 85(1), pp. 240–53.

Nalebuff, Barry. "Puzzles." Journal of Economic Perspectives, Summer 1987, 1(1), pp. 157–63.

Piattelli-Palmarini, Massimo. Inevitable illusions. New York: Wiley, 1994.

Selten, Reinhard and Buchta, Joachim. "Experimental Sealed Bid First Price Auctions with Directly Observed Bid Functions." Discussion Paper B-270, University of Bonn, Germany, 1995.

Selvin, Steven. "Letters." American Statistician, February 1975a, 29(2), p. 67.

_____ . "Letters." American Statistician, May 1975b, 29(3), p. 134.

Smith, Vernon L. "Rational Choice: The Contrast Between Economics and Psychology." Journal of Political Economy, August 1991, 99(4), pp. 877–97.

Siegel, Sidney. "Decision Making and Learning Under Varying Conditions of Reinforcement." Annals of the New York Academy of Sciences, November 1961, 89(5), pp. 766–83.

Tierney, John. "Behind Monty Hall's Doors: Puzzle, Debate and Answer?" New York Times, July 21, 1991, Sec. 1, p. 1.

Tversky, Amos and Shafir, Eldar. "Choice Under Conflict: The Dynamics of Deferred Decision." Psychological Science, November 1992, 3(6), pp. 358–61.

vos Savant, Marilyn. "Ask Marilyn." Parade, September 8, 1990; December 2, 1990; February 17, 1991.