THE JOURNAL OF FINANCE • VOL. LIX, NO. 3 • JUNE 2004

Are Judgment Errors Reflected in Market Prices and Allocations? Experimental Evidence Based on the Monty Hall Problem

BRIAN D. KLUGER and STEVE B. WYATT*

ABSTRACT

The question of whether individual judgment errors survive in market equilibrium is an issue that naturally lends itself to experimental analysis. Here, the Monty Hall problem is used to detect probability judgment errors both in a cohort of individuals and in a market setting. When all subjects in a cohort made probability judgment errors, market prices also reflected the error. However, competition among two bias-free subjects was sufficient to drive prices to error-free levels. Thus, heterogeneity in behavior can be an important factor in asset pricing, and further, it may take few bias-free traders to make asset prices bias-free.

In recent years, behavioral researchers have challenged the joint hypotheses of efficient markets and the idea that average (excess) returns are solely rewards for risk, suggesting that these are deficient for interpreting capital market data (De Bondt and Thaler (1985), Shleifer (2000), Hirshleifer (2001)). They argue that individual cognitive errors are useful in interpreting marketwide anomalies, presupposing that individual judgment errors aggregate into similar marketwide errors. Other researchers instead maintain that individual judgment errors do not materially affect markets in equilibrium. Rubinstein (2001, p. 26) summarizes: "Some adherents of behavioral finance begin sensibly enough with the results of convincing experiments that show human beings are irrational in certain specific systemic ways. But then comes the hand waving as they try to extend the results to the much more complex, long-lasting, repetitive and subtle environment of the market. This extension requires a big leap of faith."

This paper offers direct experimental evidence about the link between individuals' judgment errors and market equilibrium. We perform two sets of experiments: the first establishes the absence or presence of judgment errors in a cohort of individuals. In the second experiment we change the context to a market setting. We can then observe the extent to which judgment errors observed in the individual experiments are reflected in market prices and allocations. We further vary the types of trading opportunities to test some conjectures about how judgment errors could be eliminated in markets.

*Both authors are at the University of Cincinnati. We are grateful to Lucy Ackert, Ann Gillette, Rick Green (the editor), Brian Hatch, Chuck Schnitzlein, Steve Slezak, David Stolin, seminar participants at McMaster University and Stockholm Institute for Financial Research, and an anonymous referee for helpful suggestions.

We focus on a probability judgment error inspired by the Monty Hall problem. Because the Monty Hall problem commonly leads to very specific errors, the presence or absence of judgment errors can be readily detected in individuals and markets. The commonly modeled biases (e.g., overconfidence, conservatism, or representativeness) lead to a range of errors. For example, with conservatism, individuals overweight prior information, but the degree of overweighting is imprecise and may vary across individuals. This imprecision makes it difficult to form a well-specified alternative hypothesis that can be rejected as easily as the hypothesis that probability judgments are error free.

We do not argue that the Monty Hall problem should be a basis for behavioral models or that it is a common problem in financial decision making.[1] Rather, we chose the Monty Hall problem as a tool to study probability judgment errors in general because of the precise predictions associated with this error. Further, using the Monty Hall problem also allows a design where the decision problem is identical in both the individual and the market experiments.

The importance of probability judgment errors for understanding asset pricing and for explaining market anomalies is a matter of intense debate.[2] Despite the importance of this question, only a few studies have directly examined the link between judgment errors and prices using experimental economics methods. Duh and Sunder (1986), Camerer (1987), Ganguly, Kagel, and Moser (1994), and Anderson and Sunder (1995) focus on whether deviations from Bayesian predictions are present in experimental market prices, but present inconclusive evidence on the question. Bayes' rule, however, is not an ideal tool to study probability judgment errors in markets because a behavioral alternative to Bayes' rule is not well specified. By contrast, using the Monty Hall problem provides precise predictions for both the absence and presence of judgment errors.

Our experiments suggest several general conclusions. First, in markets where all participants misjudge probabilities, market prices also reflect this misjudgment. Even after repeated trials, subjects do not learn to avoid probability mistakes even if there are significant rewards for doing so. Individual probability misjudgment aggregates into market prices as maintained by behavioral models. However, when markets are made up of both traders who do and do not misjudge probabilities, we find the outcome is reversed. In cases where two or more subjects correctly estimated probabilities in the individual trials, these participants dominated trading in the corresponding market experiments. Market prices reflected correct probability beliefs. Thus, our data suggest that heterogeneity in judgment errors by investors is an important determinant of asset market equilibrium.

[1] Crack (2000, pp. 201–203), however, reports that the Monty Hall problem is often used by Wall Street firms as a screening device to evaluate job candidates.

[2] See Arrow (1982), Blume and Easley (1982), Russell and Thaler (1985), DeLong et al. (1990), Camerer and Ho (1996), Shleifer and Vishny (1997), Fama (1998), Daniel, Hirshleifer, and Subrahmanyam (1998, 2001), Gervais and Odean (2001), and Bossaerts (2002) for discussions of forces (e.g., arbitrage, learning, and competition) that may or may not cause markets to aggregate individuals' judgment errors.

The rest of the paper is organized as follows. Section I describes the Monty Hall probability judgment error. We present the experimental design in Section II, and the experimental results in Section III. Section IV concludes.

I. The Monty Hall Judgment Error

The classic presentation of the Monty Hall problem is based on stylized facts from the very popular 1960s TV game show, "Let's Make A Deal." Game show host Monty Hall presents a final contestant with a choice among three doors (or curtains). Behind one of these doors is a valuable prize, but behind the other two doors is a worthless prize. After the contestant picks a door, Monty shows the contestant one of the remaining doors revealing a worthless prize. Then Monty offers the contestant the opportunity to switch doors, that is, to choose the prize behind the other remaining doors.

As long as Monty does not decide whether or not to offer the option to switch doors based on the contestants' initial selection, the contestant should always switch doors. The probability that the prize is behind the original door is one-third, but switching to the unrevealed door has a two-thirds probability of winning the prize.[3] However, when faced with this problem, many individuals appear to believe that the odds of winning are even regardless of the door chosen, so they do not switch or switch too infrequently.

The Monty Hall problem has been extensively studied by Friedman (1998) in a series of individual decision experiments. He reports an average switch rate of about 30% over his entire sample and trials, and further finds that subjects require a large number of trials before they learn (if ever) to correct their decision errors.[4] Because the Monty Hall problem is such a pervasive

[3] To see why the odds are one-third and two-thirds, consider the following explanation. The objective probability of picking the winning door is one-third at the outset and the probability that the winning prize is behind one of the other two doors is two-thirds. Since it is common knowledge that Monty has precommitted to reveal a nonselected losing door after the contestant has chosen, the only new information contained in this revelation is *which* of the two nonselected doors does not have a prize. Since we know that only one door has a valuable prize, it will always be possible to reveal a losing prize behind at least one of the two remaining doors. The prior probability was one-third for the selected door and two-thirds for both of the unselected doors. When the nonselected door is revealed as a loser, the probability that the prize is behind an unselected door remains two-thirds. Since the door revealed as a loser has a zero probability of winning, the remaining nonselected door must have a two-thirds probability of winning. After Monty reveals the losing door, it is also common knowledge that he will always allow the contestant to switch from their original choice (with a one-third chance of winning) to the remaining door (with a two-thirds chance of winning). Since Monty is precluded by the common knowledge assumption from offering the opportunity to switch conditional on the initial door selected, the offer to switch contains no information. Under these circumstances, we know switching has twice the expected payoff as not switching. Monty is effectively offering both of the remaining doors to the contestant in exchange for the original pick.

[4] Friedman (1998) reports that with very intense rewards for winning and intense losses, it would take about 60 trials before his subjects would be expected to switch over 90% of the time.

judgment error, it provides a powerful test to observe if these same errors are fully impounded into market prices.

II. Experimental Design

A. Experimental Overview

In each experimental session, a cohort of six traders first participates in a sequence of individual decision-making experiments, and then participates in a sequence of experimental asset markets. The individual decision-making experiment measures the existence and the frequency of the Monty Hall judgment error for each participant. The second part of the experimental session consists of a sequence of experimental markets, which are designed to study the same judgment decision when subjects have the opportunity to trade decision rights. We accomplish this by embedding the Monty Hall problem into the decision rights associated with traded assets.

In total, we conducted 12 sessions, each with a unique cohort of six subjects. A typical session (including time spent on instructions) lasted approximately three hours. Participants were all students at the University of Cincinnati. Payouts per session averaged approximately $25 U.S. dollars per participant, paid in cash at the end of the experiment.

B. Individual Judgment Experiments

Our individual experiments employ a design very similar to Friedman (1998). Instead of three doors, a cohort of subjects were shown three cards, marked with the suits clubs, diamonds, and hearts that had the same numerical value (2, 3, 4, etc.). At the start of the trial, each participant received an asset certificate. The certificates were either club, diamond, or heart certificates. Only one type of certificate was issued in each trial, so all participants always had the same type of certificate. For exposition purposes, suppose that heart certificates are the type issued.

Participants were then shown three playing cards. The cards were shuffled and randomly placed face down on a mat in squares labeled one through three. Participants were then shown a bingo cage with three balls, labeled one through three. A ball was then randomly selected, and the card in the position corresponding to the number announced was set aside, still face down.[5] The participants were told that (because they have a heart certificate) they would receive

[5] This double randomization method for choosing the initial card helps to ensure that participants truly believe that the initial selection is indeed random. Choosing the initial card randomly (as opposed to letting the subjects choose) allows us to separate probability assessment errors from errors due to the illusion of control fallacy. Camerer (1995) suggests that individuals may fail to revise beliefs according to objective probability laws because they seem to believe they can control chance events. Thus their initial choices are better than objective odds would indicate. This is related to overconfidence in general. Since our subjects do not select the initial card, they cannot fall prey to this particular fallacy, and so any failure to switch must be related to errors in probability assessment.

$1 if the card set aside were a heart, and nothing otherwise. However, they are also told that later they will have a chance to switch. They will be able to convert from the original card to one of the remaining cards.

The participants are then shown one of the two remaining cards. Participants are told that the card revealed will always be a losing card, so if heart certificates were initially issued, the revealed card must be either a club or a diamond. The instructor then reveals a losing card. For illustrative purposes, suppose a diamond is revealed.

After this revelation, subjects face the decision to convert from heart to club certificates. If they do not convert, then they will receive $1 if the set-aside card is a heart and nothing otherwise. If they convert, then they will receive $1 if the set-aside card is a club and nothing otherwise. The participants privately mark whether they want to convert their asset certificates. The set-aside card is then revealed, and subjects earn $1 if the set-aside card has the same suit as their asset certificate.

The sequence of events for each trial is illustrated in Figure 1. This sequence was repeated 12 times per cohort, with the order of the certificate type rotated across trials and sessions.

We chose to frame the judgment problem in this manner so that the decision faced in the individual experiments would be the same as the conversion decision in the subsequent market trials. Therefore, any observed differences can only be attributed to the opportunity to trade decision rights.

C. Judgment Errors in Market Experiments

After the individual decision-making trials, the same cohort participates in 12 market trials. Participants are initially endowed with three certificates. Procedures and certificate payoffs are identical to those described in the individual experiments. Market trials consist of two types of auctions: Second-price sealed bid auctions and oral double auctions.

In the second-price auction, participants submit bids for one additional share of the asset. This forces the participants to explicitly value the asset. Further, each second-price auction guarantees a transaction, and the price is made public. Although six participants are generally sufficient to generate some trading volume, it is possible that with unusually reticent participants, no shares will trade. The second-price auctions guarantee one transaction both before and after the losing card is revealed.[6]

A two-minute oral double auction is then conducted. At all times, participants can observe active bids and asks, transaction prices, and their share and cash balances. Short sales were not permitted, and participants had a sufficient cash balance so that margin purchases were not necessary.

[6] The mechanics as well as the optimal strategy for the second-price sealed bid auction are discussed in the instructions. In case of ties, the buyer is chosen randomly among bidders with the highest bid.

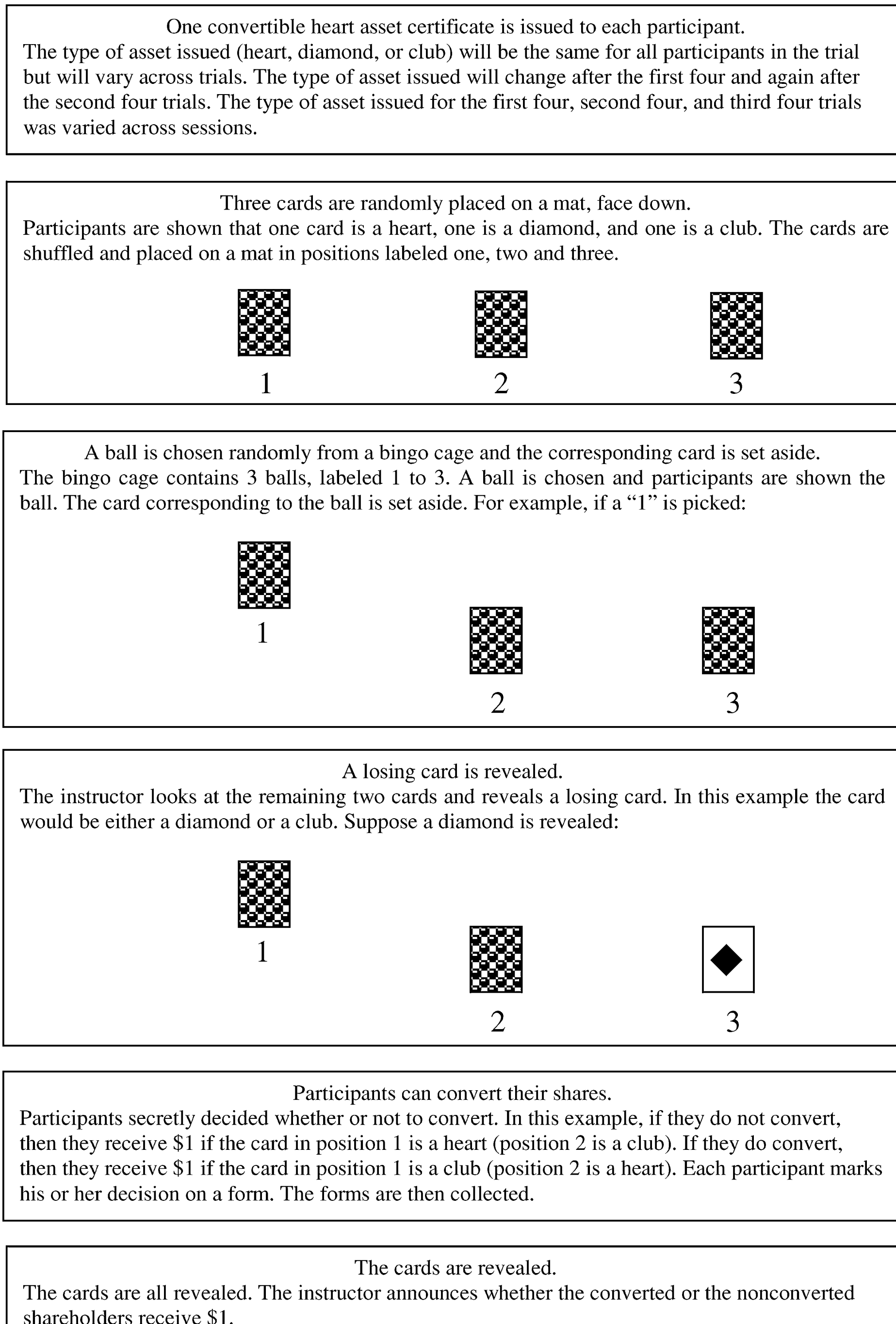

Figure 1. Sequence of events for a sample trial of the individual decision-making task. Each session consists of twelve trials.

After two minutes, the auction is stopped and a losing card is revealed. Again, the procedure is exactly the same as described in the individual decision-making task. As before, all participants know that a losing card is going to be revealed. The double auction is then reopened for two more minutes. Lastly, a second-price sealed bid auction is conducted for an additional share.

D. Conversion Treatments

Three types of conversion treatments were conducted in each session. In the P sessions, at the start of each trial, the instructor announced whether the shares were nonconvertible, all-or-none convertible, or partial convertible. In the B sessions, partial-convertible shares were not used. The partial-convertible treatment was replaced by a treatment with both nonconvertible and all-or-none convertible shares.

If shares were nonconvertible, then participants did not have the opportunity to "switch doors" (convert). Therefore, if the shares issued at the start were, for example, nonconvertible diamond shares, then participants would receive $1 per share if the card chosen (by the bingo draw) was a diamond, and nothing otherwise. The expected value of these shares was 33 cents, since there was a one-third chance of the randomly chosen card matching the designated suit.

If shares were all-or-none convertible, then participants had the opportunity to convert shares just before the set-aside card was revealed. The participants' choice was to convert either all of their shares or none of their shares. The objective expected value of these shares is 67 cents if converted, but 33 cents if not converted. The conversion decision here is exactly the same as in the individual decision-making task, except that the amount of money at risk here is proportional to the individual's share balance.

Partial-convertible shares differed from the all-or-none convertibles in that shareholders could convert some of their shares if they desired. The expected value of the shares is the same as in the all-or-none case. However, here shareholders could hedge. By converting one certificate and not converting a second certificate, the shareholder ensures a payoff of $1. Thus, share prices persisting at levels below 50 cents would represent an arbitrage opportunity.

In the B sessions, three of the participants received nonconvertible shares, and three received convertible shares. We alternated the type of share being traded, switching either after a transaction occurred or after a 15-second period of inactivity (where no new bids or asks were made). As in the partial-convertible treatment, shareholders could hedge. By holding pairs of convertible and nonconvertible certificates (and converting), the shareholder ensures a payoff of $1 per pair. Therefore, the sum of the convertible price and the nonconvertible price should equal $1.

The sequence of events in each market trial is diagrammed in Figure 2.

The experimental design is summarized in Table I.

We replicated our basic study design 12 times across 12 different cohorts. Six sessions were P sessions and six were B sessions.

Three asset certificates are issued to each participant.
The type of asset issued (heart, diamond, or club) will be the same for all participants in the trial but will vary across trials. The assets will be either nonconvertible, all-or-none convertible, or partial-convertible.

Three cards are randomly placed on a mat, face down.
Participants are shown that one card is a heart, one is a diamond, and one is a club. The cards are shuffled and placed on a mat in positions labeled one, two, and three.

A ball is chosen randomly from a bingo cage and the corresponding card is set aside.
The bingo cage contains 3 balls, labeled 1 to 3. A ball is chosen and participants are shown the ball. The card corresponding to the ball is set aside.

An additional asset certificate is sold in a second-price sealed-bid auction.
Participants submit bids to purchase an additional asset certificate from the "house." The asset certificate is identical to assets issued at the start of the trial. If the treatment is both-arbitrage, then two auctions are simultaneously conducted, one for nonconvertible shares and one for convertibles.

Participants can trade shares in an oral double auction.
Participants can buy shares from or sell shares to each other via an oral double auction. The auction lasts for two minutes.

A losing card is revealed.
The instructor looks at the remaining two cards and reveals a losing card. In this example the card would be either a diamond or a club.

Participants can trade shares in another oral double auction.
Once more, participants can buy shares from or sell shares to each other via an oral double auction. The auction lasts for two minutes.

An additional asset certificate is sold in a second-price sealed-bid auction.
Participants submit bids to purchase an additional asset certificate from the "house." The asset certificate is identical to assets issued at the start of the trial.

Nonconvertible Shares

All-or-None Convertible Shares

Each participant secretly decides whether or not to convert. The choice is to convert either all or none of his or her shares.

Partial-Convertible Shares

Each participant secretly decides whether or not to convert. The participant may convert all, some, or none of his or her shares.

Both Nonconvertible and All-or-None Convertible Shares

Each participant secretly decides whether or not to convert all or none of his or her convertible shares.

The cards are revealed.
The cards are all revealed. The instructor announces the liquidation value of all shares.
Participants' profit positions are updated.

Figure 2. Sequence of events for a trial of the market. Each session consists of twelve trials.

Table I

Design for a Session of the Experimental Market

In the no-conversion treatment, only nonconvertible assets were issued. In the all-or-none conversion treatment, only convertible assets were issued, and participants were required to convert all their shares or none of their shares. In the partial-conversion treatment, only convertible assets were issued, but participants were able to convert all, some, or none of their shares. In the both treatment, both nonconvertible and convertible assets were issued and traded. In that treatment, participants with convertibles were required to convert all or none of their convertible shares.

Trial	P Session Treatment	B Session Treatment
1	No conversion	No conversion
2	All-or-none conversion	All-or-none conversion
3	Partial conversion	Both
4	All-or-none conversion	All-or-none conversion
5	Partial conversion	Both
6	No conversion	No conversion
7	Partial conversion	Both
8	No conversion	No conversion
9	All-or-none conversion	All-or-none conversion
10	No conversion	No conversion
11	All-or-none conversion	All-or-none conversion
12	Partial conversion	Both

III. Experimental Results

A. Individual Experiments

In the individual experiments, converting is a dominant strategy and we should observe a 100% conversion rate. Based on the parameters of our individual experiments, if subjects never converted, the expected cost of deviating from the optimal strategy (of always converting) was $4 per subject or $24 per cohort. In the market experiments, the maximum expected cost was $80 for a cohort of six subjects. On a per-subject basis, a zero conversion rate has an expected cost of over $17 across both experiments and all replications. Since the decision to convert or not convert was equally costly in effort (a checkmark in a box), then any deviation from 100% conversion is evidence that subjects are making an error in probability judgment. This is the only formal hypothesis using individual decision data. The expected cost of this error is linear in conversion rate, so we can use the observed rate of conversion in the individual experiments to estimate how far their actions deviate from perfect rationality.

Conversion rates for the individual experiment are presented in Table II. Only 7 of 72 subjects converted their shares every time in the individual experiments.[7] Four others never converted. The overall conversion rate was

[7] Note that 100% conversion does not necessarily indicate that the subject properly assesses the probabilities of winning the dollar payoff. For example, if the subject believes that there is a one-half probability of winning regardless of the conversion choice, he or she might choose any conversion rate.

Table II

Conversion Rates for the Individual Experiments

The conversion rates are presented for the first four trials, the middle four trials, and for the last four trials. The overall conversion rate is the rate for all twelve trials.

Session	Trader	Trials 1–4	Trials 5–8	Trials 9–12	Overall	Session	Trader	Trials 1–4	Trials 5–8	Trials 9–12	Overall
P1	1	50%	50%	50%	50%	B1	1	0%	0%	0%	0%
P1	2	50%	25%	0%	25%	B1	2	0%	50%	0%	17%
P1	3	75%	50%	75%	67%	B1	3	50%	75%	50%	58%
P1	4	75%	75%	25%	58%	B1	4	50%	50%	0%	33%
P1	5	0%	0%	0%	0%	B1	5	50%	0%	50%	33%
P1	6	25%	50%	25%	33%	B1	6	100%	100%	100%	100%
P2	1	50%	25%	75%	50%	B2	1	50%	75%	25%	50%
P2	2	50%	50%	50%	50%	B2	2	75%	25%	25%	42%
P2	3	100%	50%	50%	67%	B2	3	25%	75%	75%	58%
P2	4	25%	75%	25%	42%	B2	4	50%	100%	100%	83%
P2	5	75%	75%	25%	58%	B2	5	100%	100%	100%	100%
P2	6	75%	100%	75%	83%	B2	6	0%	50%	100%	50%
P3	1	75%	25%	50%	50%	B3	1	100%	100%	100%	100%
P3	2	50%	75%	75%	67%	B3	2	50%	25%	50%	42%
P3	3	0%	0%	0%	0%	B3	3	100%	100%	100%	100%
P3	4	75%	50%	100%	75%	B3	4	0%	50%	75%	42%
P3	5	25%	50%	50%	42%	B3	5	75%	50%	25%	50%
P3	6	100%	100%	100%	100%	B3	6	25%	25%	25%	25%
P4	1	100%	100%	100%	100%	B4	1	100%	100%	50%	83%
P4	2	25%	50%	25%	33%	B4	2	50%	75%	75%	67%
P4	3	100%	100%	100%	100%	B4	3	50%	50%	100%	67%
P4	4	100%	75%	100%	92%	B4	4	50%	50%	50%	50%
P4	5	25%	25%	100%	50%	B4	5	50%	50%	50%	50%
P4	6	0%	50%	50%	33%	B4	6	25%	25%	100%	50%
P5	1	50%	50%	25%	42%	B5	1	25%	75%	75%	58%
P5	2	0%	75%	75%	50%	B5	2	25%	50%	50%	42%
P5	3	75%	50%	75%	67%	B5	3	0%	50%	100%	50%
P5	4	75%	0%	100%	58%	B5	4	0%	50%	100%	50%
P5	5	25%	50%	25%	33%	B5	5	75%	75%	100%	83%
P5	6	25%	50%	50%	42%	B5	6	0%	75%	100%	58%
P6	1	25%	50%	25%	33%	B6	1	75%	50%	25%	50%
P6	2	100%	25%	25%	50%	B6	2	75%	75%	50%	67%
P6	3	25%	75%	50%	50%	B6	3	75%	75%	50%	67%
P6	4	0%	0%	25%	8%	B6	4	0%	25%	25%	17%
P6	5	0%	0%	0%	0%	B6	5	75%	50%	50%	58%
P6	6	25%	25%	75%	42%	B6	6	75%	75%	50%	67%
	Avg.	49%	49%	52%	50%		Avg.	48%	59%	61%	56%

53%. This rate is higher than the nearly 30% reported by Friedman (1998), but still sufficient to demonstrate that most of our subjects fall prey to the Monty Hall probability judgment error.[8]

[8] The difference may be due to the illusion of control issue mentioned above. In our experiments, subjects do not choose the initial card, so they may be less attached to their initial choice and therefore be more willing to switch.

The individual decision conversion rates are also reported over the course of the individual trials in Table II. The conversion rates for the first four, the second four, and the final four trials are reported for each subject. Average conversion rates increased slightly in both the P and B sessions, consistent with slight learning. The hypothesis that learning occurred can be formally tested using a Wilcoxon signed-rank test:

H_o: The number of conversions in the last four trials equals the number of conversions in the first four trials.

H_a: The number of conversions in the last four trials is greater than the number of conversions in the first four trials.

On average, subjects converted 1.93 times out of four conversion decisions (48%) in the first four trials and 2.26 times (57%) during the last four trials. The Wilcoxon test statistic is 123 and the associated one-tailed p-value is 0.066. This result indicates that there may be a slight amount of learning over the course of the individual trials. However, if there is learning, it is slow. Even after repeated trials, the majority of subjects do not learn to avoid the Monty Hall judgment errors, even if there are significant rewards for doing so.

B. Market Prices and Judgment Errors

The market experiment provides a rich environment to study the aggregation of judgment errors in equilibrium. Without probability judgment errors, we would expect the prices to be consistent with the true probability of payoffs. Since the objective probability of receiving the payoff is two-thirds for the convertible shares (given that they are converted) and one-third for the nonconvertible shares, the expected payoff for the convertible asset is 67 cents and 33 cents for the nonconvertible asset. If the participants maximize expected utility, and do not make probability judgment errors, we would expect the ratio of the convertible to nonconvertible prices to be 2.

The Monty Hall problem allows us to make some predictions about individual errors in belief formation that may be transmitted to market outcomes. Unlike other judgment biases that are vague in magnitude, we can conjecture exactly how bias arises and what this implies for market prices.

The overwhelming majority of subjects who fall prey to the Monty Hall judgment error believe that after one of the three possible outcomes is eliminated, the two remaining choices are equally likely. Converting or not converting both have a 50% chance of winning the dollar; therefore the subjective expected payoff for the convertible asset is 50 cents and it makes no difference whether or not the conversion rights are actually exercised. This error forms the basis for our behavioral alternative hypothesis.

Since this problem is presented as a conjunction between elimination of a possibility (revealing the losing card) and an opportunity to convert, we do not know if both conditions are necessary to elicit the Monty Hall judgment error.

The first part of this conjunction is the elimination of a possibility. Because a losing card is always revealed, individuals may distribute total probability evenly over the remaining two possibilities. They will then form a subjective belief that the original and the remaining card have an equal chance of winning. We call this line of false reasoning the "possibilities and probability matching error" (PPME) because subjects appear to match possibilities with probability. Only one of two cards can win and the probabilities are judged to be equally likely. Both the convertible and the nonconvertible assets are judged to earn the dollar half of the time, so the mistaken expected payoff would be 50 cents for all types of asset certificates. If these errors aggregate in equilibrium, we would predict that the prices of convertible and nonconvertible certificates would be the same (a relative price of 1).

However, if both the opportunity to switch and the elimination of a possibility (revealing the losing card) are necessary for subjects to make the Monty Hall error, then we will observe different subjective valuations for the convertible and nonconvertible shares. If individuals form their beliefs over choices and not possibilities, then the nonconvertible certificate will be valued based on the original choice of one in three, but the convertible certificates will be valued based on a choice of one in two. This implies a subjective probability of one-third for the nonconvertible certificate and a subjective probability of one-half for the convertible certificate. The mistaken expected payoffs are 33 cents for the nonconvertible asset and 50 cents for the convertible asset.

We call this line of false reasoning the "choice and probability matching error" (CPME) hypothesis. Under this hypothesis, if the error aggregates in equilibrium, the prices of the convertible and nonconvertible certificates would trade at a ratio of the subjective probabilities, a ratio of 1.5. Note that the nonconvertible certificate appears to be correctly priced, but this is only because the mental mistake of focusing on choices and not on possibilities leads by accident to the correct judgment. This probability assessment is not correct because it implies that the total subjective probability for convertible and nonconvertible certificates is $1/3 + 1/2 = 5/6$, not 1.

Markets may also be made up of some individuals subject to the "possibility and probability matching" error and others subject to the "choice and probability matching" error. In this circumstance, both groups would agree on the same (mistaken) expected payoff of the convertible certificate but disagree about the expected payoff of the nonconvertible certificate. As a result, in aggregate, the convertible to nonconvertible price ratio would be somewhere between 1 and 1.5. There should also be a greater trading volume for the nonconvertible certificate relative to the convertible certificate because of the divergent opinion about the value of the convertible asset. We label this the "mixed error" hypothesis.

The hypotheses are summarized in Table III. The benchmark prices and price ratios provide a means to test how individual judgment errors can affect market prices.

Table III

Expected Payoffs Conditional on the Hypothesized Probability Judgment Behavior by Treatment

The price ratio is the predicted ratio of the price of the convertible asset divided by the predicted price of the nonconvertible asset.

Hypothesis	No Conversion	All-or-none Conversion	Partial Conversion	Price Ratio
Rational pricing	33	67	67	2.0
PPME	50	50	50	1.0
CPME	33	50	50	1.5
Mixed error	33–50	50	50	1.25

C. Market Experiments

Prices for the market experiments are presented in Table IV and the ratios of the average convertible to nonconvertible prices are contained in Table V.

Table VI classifies each cohort's pricing according to whether they more closely followed rational pricing or one of the alternative probability judgment error hypotheses. The absolute deviation between the price and the expected payoff (based on each behavioral hypothesis) is averaged over all transactions during each trial for the session. The session is then classified according to which behavioral hypothesis produces the minimum average absolute deviation.

Tables IV and V illustrate that only one of the P session cohorts reached rational prices. The P4 session had an average nonconvertible price of 34 cents. The average all-or-none convertible price was 61 cents and the average partial-convertible price was 67 cents. In the other P sessions, we sometimes observe an average nonconvertible price near 33 cents or an average convertible price near 67 cents, but not both in the same session. In the remaining P sessions, judgment errors are reflected in the prices of either nonconvertible asset certificates, convertible asset certificates, or both. In the B sessions, two cohorts, B3 and B5, had prices close to their rational levels. Prices in the remainder of the B sessions diverged from rational levels, for either one or both types of assets.

The average conversion rates for the all-or-none treatment of the sessions are also reported in Table VI. These rates illustrate a correspondence between rational pricing and allocational efficiency (100% conversion). The conversion rates for the sessions classified as rational ranged from 81 to 90%. The average conversion rates for the remaining sessions ranged from 33 to 70%.

We next examine in more detail the sessions in which the prices were rational, and look at how the market corrected the probability bias. One common argument for rational prices is that rational traders drive out irrational traders in the price formation process. Our design allows us to test this hypothesis.

Table IV

Average Prices for the Market Trials by Session and by Treatment

The prices are averages of transaction prices during the double auctions as well as the transaction prices in the second-price sealed bid auctions. P sessions contain the partial conversion treatment. B sessions contain the both assets treatment.

	Treatment	Average Price											
Trial	Session	P1	P2	P3	P4	P5	P6	B1	B2	B3	B4	B5	B6
1	No conv	50.57	23.25	48.13	27.27	52.50	34.60	33.60	50.67	41.00	45.00	40.80	46.75
6	No conv	49.88	25.67	44.57	37.56	51.67	30.50	33.00	51.25	41.33	47.40	29.17	52.86
8	No conv	46.48	22.67	35.33	36.60	52.86	29.00	30.33	50.40	31.75	43.67	26.60	58.67
10	No conv	50.73	22.75	28.13	33.27	40.83	30.83	30.33	49.63	30.22	30.75	25.83	56.00
Avg.	No conv	49.42	23.59	39.04	33.68	49.47	31.23	31.82	50.49	36.08	41.71	30.60	53.57
2	All or none	50.56	29.50	45.83	54.80	44.83	36.43	34.14	54.71	49.00	52.67	59.29	53.44
4	All or none	50.00	25.00	51.38	64.00	52.00	35.50	34.00	51.40	62.10	54.00	68.80	58.50
9	All or none	51.08	22.75	41.80	64.36	56.83	40.00	27.67	51.30	62.08	39.75	71.20	69.64
11	All or none	50.30	26.33	40.13	60.22	58.43	47.60	28.00	49.30	62.09	31.67	72.30	66.56
Avg.	All or none	50.49	25.90	44.79	60.85	53.02	39.88	30.95	51.68	58.82	44.52	67.90	62.04
3	Partial	52.45	31.40	55.00	66.50	40.29	48.67						
5	Partial	52.18	33.50	64.00	65.67	55.67	53.40						
7	Partial	52.44	36.67	65.00	69.31	68.40	51.75						
12	Partial	52.50	38.14	70.40	65.38	70.90	55.00						
Avg.	Partial	52.39	34.93	63.60	66.72	58.82	52.21						

3	Both	Nonconvertible asset	40.43	48.20	40.83	49.00	45.00	60.80
3	Both	Convertible asset	43.29	54.67	60.38	57.20	69.00	70.83
3	Both	Convertible plus nonconvertible asset	83.72	102.87	101.21	106.20	114.00	131.63
5	Both	Nonconvertible asset	45.60	50.83	51.40	49.50	35.00	65.17
5	Both	Convertible asset	45.83	50.00	66.40	50.75	71.00	71.33
5	Both	Convertible plus nonconvertible asset	91.43	100.83	117.80	100.25	106.00	136.50
7	Both	Nonconvertible asset	38.71	50.86	42.86	49.25	33.67	59.43
7	Both	Convertible asset	45.20	51.57	63.00	49.20	74.00	73.60
7	Both	Convertible plus nonconvertible asset	83.91	102.43	105.86	98.45	107.67	133.03
12	Both	Nonconvertible asset	42.00	48.75	39.20	36.67	22.80	56.00
12	Both	Convertible asset	41.43	51.43	68.71	38.25	76.00	72.80
12	Both	Convertible plus nonconvertible asset	83.43	100.18	107.91	74.92	98.80	128.80
Avg.	Both	Nonconvertible asset	41.66	49.66	43.57	46.11	34.12	60.35
Avg.	Both	Convertible asset	43.94	51.92	64.62	48.85	72.50	72.14
Avg.	Both	Convertible plus nonconvertible asset	85.60	101.58	108.20	94.96	106.62	132.49

Table V

Ratio of Average Prices of Convertible Assets to Nonconvertible Assets

All-or-none to nonconvertible is the average price in the all-or-none conversion treatment divided by the average price in the noconversion treatment. The partial to nonconvertible is the ratio of the price in the partial conversion treatment to the no-conversion treatment price. The both convertible to nonconvertible is the ratio of convertible to nonconvertible prices in the both treatment.

Session	All-or-none to Nonconvertible	Partial to Nonconvertible	Session	All-or-none to Nonconvertible	Both Convertible to Nonconvertible
P1	1.02	1.06	B1	0.97	1.05
P2	1.10	1.48	B2	1.02	1.05
P3	1.15	1.63	B3	1.63	1.48
P4	1.81	1.98	B4	1.07	1.06
P5	1.07	1.19	B5	2.22	2.13
P6	1.28	1.67	B6	1.16	1.20

Table VI

The Average Absolute Deviation between the Actual Transaction Prices and the Hypothesized Prices

Each cohort is classified according to the lowest absolute deviations. The average conversion rate for the all-or-none treatment (calculated as a percentage of all shares converted) is also reported.

	Average Absolute Deviations from Behavioral Hypothesis					
Cohort	Rational Pricing	PPME	CPME	Mixed Error	Classification	Conversion Rate for the All-or-none Treatment
P1	15.79	1.20	7.10	4.09	PPME	33%
P2	28.36	19.96	15.86	17.74	CPME	45%
P3	13.30	10.20	9.32	8.96	Mixed error	50%
P4	4.33	14.85	10.62	12.20	Rational pricing	89%
P5	12.32	9.71	12.75	10.23	PPME	70%
P6	14.56	10.45	5.54	7.67	CPME	39%
B1	18.86	13.89	10.21	10.84	CPME	41%
B2	15.97	1.81	9.37	5.32	PPME	50%
B3	8.01	12.80	11.05	11.20	Rational pricing	81%
B4	15.93	8.20	9.96	8.03	Mixed error	56%
B5	6.59	19.00	13.69	15.71	Rational pricing	90%
B6	14.88	12.33	19.60	15.64	PPME	66%

Because we observe the results from the individual experiments, we can rate our subjects by how far from perfect rationality their prior decisions were and observe in our market experiments if the most rational individuals drive the price formation process to levels predicted by rational asset pricing.

Tables VII and VIII present the conversion rates for both the individual experiments and the market experiments. To minimize possible effects due to learning during the individual trials, individual conversion rates reported here are from the last eight trials of the individual task. Table VII shows that the

Table VII

P Session Conversion Rates for Both the Individual Experiments and the Market Experiments

Individual Experiment conversion rates are based on the last eight trials. For the all-or-none treatment, the conversion rate based on the number of shares converted and based on the number of trials is shown. For the partial treatment, the share conversion rate is shown, counting all shares and again counting only unhedged shares. All pairs consisting of one converted share plus one unconverted share were not considered for the unhedged conversion frequency and rate.

		Individual Task		All-or-none Treatment				Partial Treatment			
Session	Trader	Trials Converted	%	Trials Converted	%	Shares Converted	%	Shares Converted	%	Unhedged Converted	%
P1	1	4	50%	0	0%	0	0%	0	0%	0	0%
P1	2	1	13%	0	0%	0	0%	4	18%	2	11%
P1	3	5	63%	3	75%	17	81%	5	71%	4	80%
P1	4	4	50%	0	0%	0	0%	16	50%	0	–
P1	5	0	0%	2	50%	6	67%	3	75%	2	100%
P1	6	3	38%	1	33%	3	38%	4	57%	2	67%
P1 total		17	35%	6	32%	26	33%	32	40%	10	28%
P2	1	4	50%	1	25%	4	29%	5	33%	1	14%
P2	2	4	50%	2	50%	5	36%	17	71%	10	100%
P2	3	4	50%	2	50%	6	50%	3	75%	2	100%
P2	4	4	50%	0	0%	0	0%	4	40%	0	0%
P2	5	4	50%	2	50%	5	45%	11	61%	4	100%
P2	6	7	88%	4	100%	16	100%	9	100%	9	100%
P2 total		27	56%	11	46%	36	45%	49	61%	26	76%
P3	1	3	43%	3	75%	17	77%	6	46%	2	40%
P3	2	6	75%	4	100%	11	100%	11	55%	4	67%
P3	3	0	0%	0	0%	0	0%	5	50%	0	–
P3	4	6	50%	0	0%	0	0%	3	75%	2	100%
P3	5	4	50%	2	50%	10	50%	8	47%	2	40%
P3	6	8	100%	2	100%	2	100%	8	50%	0	–
P3 total		27	56%	11	50%	40	50%	41	51%	10	56%

Table VII—*Continued*

		Individual Task		All-or-none Treatment				Partial Treatment			
Session	Trader	Trials Converted	%	Trials Converted	%	Shares Converted	%	Shares Converted	%	Unhedged Converted	%
P4	1	8	100%	3	75%	18	82%	10	77%	7	100%
P4	2	3	38%	2	67%	11	79%	11	61%	8	67%
P4	3	8	100%	3	100%	9	100%	0	–	0	–
P4	4	7	88%	4	100%	19	100%	18	82%	11	100%
P4	5	5	63%	4	100%	13	100%	17	74%	10	91%
P4	6	4	50%	0	0%	0	0%	1	25%	0	0%
P4 total		35	73%	16	76%	70	88%	57	71%	36	84%
P5	1	3	50%	2	50%	11	69%	2	50%	1	50%
P5	2	6	75%	3	100%	10	100%	14	100%	14	100%
P5	3	5	63%	3	75%	14	74%	16	67%	8	100%
P5	4	4	50%	2	50%	5	42%	8	67%	5	71%
P5	5	3	38%	3	75%	9	75%	4	31%	1	14%
P5	6	4	50%	3	75%	7	64%	5	38%	1	20%
P5 total		25	52%	16	70%	56	70%	49	61%	30	70%
P6	1	3	38%	2	50%	4	36%	9	56%	4	67%
P6	2	2	25%	3	75%	8	80%	7	64%	6	67%
P6	3	5	63%	2	50%	5	36%	10	59%	3	100%
P6	4	1	13%	0	0%	0	0%	2	15%	0	0%
P6	5	0	0%	2	50%	3	30%	11	73%	9	82%
P6	6	4	50%	1	25%	11	50%	4	50%	0	–
P6 total		15	31%	10	42%	31	39%	43	54%	22	58%

Table VIII

B Session Conversion Rates for Both the Individual Experiments and the Market Experiments

Individual experiment conversion rates are based on the last eight trials. For the market treatment, the conversion rate is calculated based on the number of shares converted and again based on the number of trials where the subject converted.

		Individual Task		All-or-none Treatment				Both Treatment			
Session	Trader	Trials Converted	%	Trials Converted	%	Shares Converted	%	Trials Converted	%	Shares Converted	%
B1	1	0	0%	1	33%	1	17%	4	100%	12	100%
B1	2	2	25%	3	75%	5	56%	4	100%	8	100%
B1	3	5	63%	3	100%	14	100%	2	100%	2	100%
B1	4	2	25%	1	25%	5	33%	2	67%	2	67%
B1	5	2	25%	2	50%	8	53%	2	67%	2	50%
B1	6	8	100%	0	0%	0	0%	4	100%	15	100%
B1 total		19	40%	10	45%	33	41%	18	90%	41	93%
B2	1	4	50%	1	33%	4	31%	2	100%	4	100%
B2	2	2	25%	1	50%	2	40%	0	0%	0	0%
B2	3	6	75%	3	75%	6	50%	0	0%	0	0%
B2	4	8	100%	3	75%	16	76%	3	100%	11	100%
B2	5	8	100%	4	100%	11	100%	4	100%	11	100%
B2	6	6	75%	1	25%	1	7%	0	0%	0	0%
B2 total		34	71%	13	62%	40	50%	9	50%	26	59%
B3	1	8	100%	4	100%	37	100%	3	100%	8	100%
B3	2	3	38%	2	50%	5	50%	4	100%	8	100%
B3	3	8	100%	4	100%	5	100%	1	50%	1	33%
B3	4	5	63%	1	100%	2	100%	3	75%	6	67%
B3	5	3	38%	2	50%	15	60%	4	100%	12	100%
B3	6	2	25%	1	100%	1	100%	4	100%	4	100%
B3 total		29	60%	14	78%	65	81%	19	90%	39	89%

Table VIII—*Continued*

		Individual Task		All-or-none Treatment				Partial Treatment			
Session	Trader	Trials Converted	%	Trials Converted	%	Shares Converted	%	Trials Converted	%	Shares Converted	%
B4	1	6	75%	3	75%	11	85%	4	100%	7	100%
B4	2	6	75%	2	50%	7	50%	3	75%	8	80%
B4	3	6	75%	3	75%	10	77%	3	75%	8	73%
B4	4	4	50%	3	75%	9	75%	2	67%	4	67%
B4	5	4	50%	2	50%	8	53%	3	100%	7	100%
B4	6	5	63%	0	0%	0	0%	1	33%	1	33%
B4 total		31	65%	13	54%	45	56%	16	76%	34	77%
B5	1	6	75%	2	50%	4	44%	2	100%	4	100%
B5	2	4	50%	3	75%	10	83%	4	100%	7	100%
B5	3	6	75%	3	75%	6	86%	3	100%	4	100%
B5	4	6	75%	4	100%	12	100%	3	100%	6	100%
B5	5	7	88%	4	100%	30	100%	4	100%	14	100%
B5	6	7	88%	3	100%	10	100%	4	100%	9	100%
B5 total		36	75%	19	83%	72	90%	20	100%	44	100%
B6	1	3	38%	1	25%	3	21%	2	100%	4	100%
B6	2	5	63%	3	75%	20	74%	1	50%	3	43%
B6	3	5	63%	2	100%	5	100%	2	100%	4	100%
B6	4	2	25%	–	–	–	–	0	0%	0	0%
B6	5	4	50%	3	75%	12	80%	3	100%	7	100%
B6	6	5	63%	4	100%	13	100%	3	75%	17	81%
B6 total		24	50%	13	59%	53	66%	11	79%	35	80%

P4 cohort included two subjects who did not make the Monty Hall error during the individual experiment. Table VIII shows similar results for the B3 and B5 sessions. Two traders had conversion rates (over the last eight individual trials) of 100% in session B3 and 88% in session B5. If these high conversion rates signify that these subjects correctly assessed the probabilities in the Monty Hall problem, then our results show that even if most traders make probability judgment errors, the presence of some rational traders can make prices rational.

Tables VII and VIII also show that some sessions included subjects with high conversion rates in the individual experiments, but did not display rational pricing in the corresponding market experiments. Sessions P2, P3, B1, and B2 all included at least one participant with a conversion rate of 88% (seven of eight) or greater during the individual trials. But these sessions did not show rational pricing in the corresponding market trials. This may or may not mean that these subjects assessed probabilities correctly. Low conversion rates imply that subjects are making judgment errors, but high conversion rates in the individual experiments do not necessarily indicate that subjects properly assess the probabilities of winning the dollar payoff.

Table IX examines the relation between the number of participants displaying judgment errors and whether the market outcome reflected judgment errors. Panel A matches the number of subjects in a session with a conversion rate of 88% or more over the last eight individual trials to whether or not prices in the corresponding market session were rational. In Panel B, the calculations are repeated, counting the number of subjects who converted 88% or more in the last eight individual trials as well as 88% or more of their shares in the all-or-none convertible treatment of the market sessions. Subjects with high conversion rates in the individual trials but low conversion rates in the market trials probably did not assess the probabilities correctly in either portion of the session. In these cases, the market prices were not rational, so either these subjects did not have the correct beliefs, or if they did, the inclusion of one rational subject was not sufficient to achieve rational pricing.

The hypothesis that rational pricing is associated with the existence of at least two rational traders is formally tested:

- H_o: The frequency of rational asset pricing in the market experiment is unrelated to the number of rational traders in the individual experiment.
- H_a: The frequency of rational asset pricing in the market experiment is greater if there are at least two rational traders in the individual experiment.

Using Fisher's exact test, the null hypothesis is rejected with a probability value of 0.0045.[9]

[9] Chi-square statistics are also reported in Table IX. However, this statistical test is controversial when the contingency table has average cell counts of less than 5. Fisher's exact test can be used with 2×2 contingency tables.

Table IX

Table of Frequencies of Sessions Classified According to Whether Market Prices Reflected Judgment Errors

In Panel A the frequencies are also classified according to the number of participants with a conversion rate of 88% or more (seven or more of eight) in the last eight trials of the corresponding individual sessions. In Panel B the frequencies are also classified according to the number of participants with both a conversion rate of 88% or more (seven or more of eight) in the last eight trials of the corresponding individual sessions and a conversion rate of 88% or more shares converted in the all-or-none treatment of the market experiments.

Number of Participants with 88% or More Conversion Rate	Rational Pricing	Prices Reflecting Judgment Error	Total
	Panel A		
None	0	5	5
One	0	3	3
Two	2	1	3
Three	1	0	1
Total	3	9	–
Likelihood ratio chi-square = 9.677 Probability = 0.02			
	Panel B		
Less than two	0	9	9
Two or more	3	0	3
Total	3	9	–
Likelihood ratio chi-square = 13.496 Probability = 0.0002		Fisher's exact test probability = 0.0045	

Other arguments for rational prices have been advanced. A second argument hypothesizes that markets may enhance the learning process. The early experimental economics literature shows that under fairly simple information structures, market prices converge to prices predicted by rational pricing models after repetitions of a stationary environment (see Sunder (1995)). Markets may be inherently self-correcting even if individual traders exhibit judgment errors. If so, judgment errors are eventually eliminated under a wide range of initial conditions. We do not find strong support for this explanation. We never observed rational prices (a convertible/nonconvertible ratio of two) except in the instance where at least two subjects did not make the error in the individual experiment. When all subjects made the Monty Hall error individually, prices were not rational.

D. Arbitrage

Another argument for how markets can correct judgment errors relies on arbitrage. In the partial-conversion treatment where individuals can select the number of certificates to convert, asset certificates should sell for at least 50 cents. Otherwise an arbitrage profit could be obtained from purchasing two shares and converting one. This treatment allows for limited arbitrage opportunities. In the partial-conversion treatment, we did observe (in five of six P sessions) a price greater than 50 cents. However, the markets with this treatment do not have rational pricing. The partial-arbitrage treatment alone was not sufficient to eliminate the judgment error.

In the B sessions, the partial-arbitrage treatment was replaced with the both-arbitrage treatment, and subjects could hold both convertible and nonconvertible assets. Here, participants could ensure a payoff of $1 by holding one share of each type of asset. Prices for the assets during this treatment are shown in Table IV. In most sessions, we report that the no-arbitrage condition (when the prices of the assets sum to exactly $1) is not met. This is probably because our market structure did not allow participants to make risk-free profits, since both assets could not be traded simultaneously, and because short selling was not permitted. To purchase one of each asset required two transactions, and given the sequence of our double auctions, the price of the second asset would not be known at the time the first was purchased. Despite this, the no-arbitrage condition should hold approximately and does in most of the sessions. The asset prices in session B6 add to 132 cents and in B1 add to 86 cents. In other sessions, the sum of the prices differs from $1 by less than 9 cents.

Prices in the both-arbitrage treatment were higher than in the other treatments in the corresponding cohort. The nonconvertible asset price in the both-arbitrage treatment was higher than the nonconvertible asset price in the no-conversion treatment for the same cohort. Similarly, the convertible asset price in the both-arbitrage treatment was higher than the convertible asset price in the all-or-none conversion treatment (see Table IV). The higher prices likely occur because aggregate risk is reduced in the both-arbitrage treatment. In fact, aggregate risk is zero if all conversion rights are exercised.

Aside from the pricing effects, the precise effect of the both treatment on probability judgment errors is not clear. However, arbitrage either in the partial-arbitrage or the both-arbitrage treatments did not automatically correct probability judgment errors.

E. Conversion Behavior and Allocative Efficiency

In our experimental markets, failure to exercise conversion rights is inefficient. Tables VII and VIII contain conversion rates for both the individual experiment and the market trials. Most of the subjects who had high conversion rates in the individual trials continued to have high rates in the market trials. Furthermore, some subjects learned to convert their shares more frequently.

Table X presents the results of Wilcoxon signed-rank tests with the null hypothesis:

H_o: The percentage of shares converted in the all-or-none treatment of the market experiment equals the percentage of shares converted in the last eight trials of the individual experiment.

Table X

Wilcoxon Signed Rank Tests Comparing the Conversion Rates in the Individual Experiments to the Conversion Rates in the Market Experiments

H_o: The percentage of shares converted in the all-or-none treatments of the market experiment equals the percentage of shares converted in the last eight trials of the individual experiment.

H_a: The percentage of shares converted in the all-or-none treatments of the market experiment is greater than the percentage of shares converted in the last eight trials of the individual experiment.

	N	Mean Difference	Wilcoxon Statistic	p-Value
Panel A: Testing Across Subjects				
All subjects	71	0.030	189	0.099
P session subjects	36	0.014	20.5	0.347
B session subjects	35	0.047	77	0.076
Subjects in rational sessions (P4, B3, B5)	18	0.123	32	0.051
Subjects in sessions where prices reflect Judgment error	53	0.003	51.5	0.312
Panel B: Testing Across Sessions				
All sessions	12	0.037	11	0.212
P sessions	6	0.039	2.5	0.344
B sessions	6	0.034	3.5	0.281
Rational sessions (P4, B3, B5)	3	0.172	3	0.125
Sessions where prices reflect judgment error	9	−0.007	−2.5	0.590

H_a: The percentage of shares converted in the all-or-none treatment of the market experiment is greater than the percentage of shares converted in the last eight trials of the individual experiment.

Using all individuals in all sessions, the hypothesis cannot be rejected at a significance level of 5%, but can be rejected at the 10% level. However, when the Wilcoxon test is performed using individuals from the sessions that were classified as rational, the p-value is 0.051. For the sessions classified as reflecting probability judgment errors, the null hypothesis cannot be rejected at standard significance levels. We do not observe significantly higher conversion rates in the market trials.

Even if individuals do not improve their conversion rates in the market experiment, it is still possible for the aggregate cohort conversion rate to improve. If certificates flow from the less rational to the more rational (where degrees of rationality are measured by conversion behavior), we would expect to see the rate of conversion improve with markets. However, a similar Wilcoxon signed-rank test performed on overall cohort rates shows that the overall conversion frequencies did not improve in the market trials. In our experiments, judgment errors often led to allocative inefficiency as well as biased prices.

Conversion rates in the partial-arbitrage treatment are more difficult to interpret. Even when all participants correctly understand the payoff probabilities, it is possible that some may prefer a certain 50 cents per share (obtainable by converting one-half of a subject's asset certificates) to a two-thirds chance of a $1 payoff per share. Further, risk-averse individuals who assess probabilities incorrectly will prefer to hedge. They can get a certain 50 cents per share rather than a 50% chance of a $1 payoff (based on the erroneous probabilities). Similar reasoning applies to the both-arbitrage treatment. In the partial-arbitrage treatment and the both-arbitrage treatment, conversion rates were generally higher than conversion rates in the individual trials. However, this does not necessarily indicate that probability judgment errors were reduced in these markets.

F. Judgment Errors and Event Efficiency

Our experiments also allow the study of event efficiency in markets where subjects make probability judgment errors. In the context of our market experiments, a losing card was revealed at the midpoint of each trial, but this event does not contain information relevant to pricing any of the assets. Following Fama's (1991) nomenclature, we call the lack of response to the revelation of the losing card as evidence that the markets are event-efficient.

Table XI presents the average of the prices both before and after we reveal the losing card. Formally:

Ho: The average price before the losing card was revealed equals the average price after the card was revealed.

Ha: The average price before the losing card was revealed is not equal to the average price after the card was revealed.

Paired t-test results are presented in Table XII. For each trial, the average transaction prices before and after a losing card is revealed are compared. The average difference between these prices is –1.39 cents, but significantly different from zero (p-value equals 0.001). Even though the entire sample is not event-efficient, this result is driven by sessions classified as PPME or mixed. When the paired t-tests are calculated for the subsample of trials in sessions that were earlier classified as rational, the null hypothesis cannot be rejected.

Further, Panel B of Table XII shows that a statistically significant price decline (about 2 cents) was observed in both the all-or-none convertible and the nonconvertible treatments. However, for the arbitrage treatments, the price response was not significant. In any case, the null hypothesis cannot be rejected either for the rational sessions or for the CPME sessions using a two-tailed test with standard significance levels.

Surprisingly, despite the presence of judgment errors, some markets were forward-looking. Even when our assets were priced incorrectly due to probability judgment error, markets could still be event-efficient.

Table XI

Average Prices Both before and after the Losing Card Is Revealed

No-conversion Treatment			All-or-none Conversion Treatment			Partial-conversion Treatment		
Session	Average Price Before Revealing the Losing Card	Average Price After Revealing the Losing Card	Session	Average Price Before Revealing the Losing Card	Average Price After Revealing the Losing Card	Session	Average Price Before Revealing the Losing Card	Average Price After Revealing the Losing Card
P1	50.38	50.15	P1	50.50	50.57	P1	52.12	52.69
P2	25.88	24.58	P2	26.88	24.75	P2	34.96	36.67
P3	41.63	36.19	P3	48.57	43.63	P3	64.34	62.55
P4	33.82	32.84	P4	60.03	62.04	P4	66.40	67.23
P5	48.38	50.46	P5	54.04	51.77	P5	59.62	61.44
P6	31.13	31.17	P6	39.63	40.13	P6	51.13	52.83

No-conversion Treatment			All-or-none Conversion Treatment			Both Treatment				
							Average Price Before Revealing the Losing Card		Average Price After Revealing the Losing Card	
Session	Average Price Before Revealing the Losing Card	Average Price After Revealing the Losing Card	Session	Average Price Before Revealing the Losing Card	Average Price After Revealing the Losing Card	Session	Conv	Nonconv	Conv	Nonconv
B1	33.81	29.48	B1	32.83	30.19	B1	41.35	41.61	44.92	40.50
B2	50.63	49.28	B2	53.36	49.52	B2	52.19	50.08	51.65	48.67
B3	35.34	34.72	B3	57.26	61.23	B3	65.17	45.50	64.71	39.00
B4	43.75	36.13	B4	48.88	40.38	B4	48.63	47.63	48.21	44.38
B5	30.02	30.46	B5	67.88	67.83	B5	72.13	35.50	72.92	32.38
B6	53.07	53.46	B6	61.37	63.55	B6	71.00	62.42	73.08	59.17

Table XII

Paired *t*-Tests to Compare the Average Price before the Losing Card Is Revealed to the Average Price after the Losing Card Is Revealed

Ho: The average price during a trial before the losing card was revealed equals the average price during a trial after the card was revealed.

Ha: The average price during a trial before the losing card was revealed does not equal the average price during a trial after the card was revealed.

	N	Mean Difference	*t*-Statistic	*p*-Value
Panel A: Testing Across Sessions				
All trials	168	−1.39	−3.39	0.001
P session trials	72	−1.14	−2.19	0.032
B session trials	96	−1.58	−2.60	0.011
Trials in rational sessions (P4, B3, B5)	44	−0.45	−0.49	0.626
Trials in PPME sessions (P1, P5, B2, B6)	56	−1.38	−2.34	0.023
Trials in CPME sessions (P2, P6, B1)	40	−0.20	−0.25	0.802
Trials in mixed error sessions (P3, B4)	28	−4.58	−4.90	<0.001
Panel B: Testing Across Treatments				
Nonconvertible treatment	48	−2.03	−2.41	0.020
All-or-none convertible treatment	48	−1.86	−2.80	0.007
Partial convertible treatment	24	0.38	0.42	0.173
Both treatment	48	−1.17	−1.38	0.676

IV. Conclusions

Our experiments use the Monty Hall probability judgment error as a tool to study how those errors aggregate into market prices and allocations. Our data offer the following general conclusions. In cohorts where all subjects exhibit individual judgment errors, market prices also reflect this error. Even after repetition to allow for learning, the error aggregates into market prices, as predicted by behavioral arguments. This result is robust to the inclusion of partial and complete arbitrage opportunities. Furthermore, these markets were not allocationally efficient because many subjects did not exercise their conversion rights optimally.

However, when markets are made up of both rational (no judgment error) and biased traders, we find the outcome is reversed. When at least two bias-free subjects are present, securities pricing does not reflect the cognitive bias seen at the individual level. Competition among two bias-free subjects is sufficient to drive prices to correct levels despite the presence of twice as many subjects who exhibit probability judgment errors.

We also find evidence that probability judgment errors can coexist with event study style market efficiency. Markets could be forward-looking, but still have prices consistent with behavioral biases. This result suggests that event efficiency and behavioral biases are not necessarily incompatible.

From our evidence two implications can be drawn. First, the conjecture that market trading alone always eliminates or substantially reduces judgment errors is unfounded. We show unambiguously that market errors can be linked to individual judgment errors.

The second implication concerns the role of "smart" traders in financial markets. We observe an error-free equilibrium when sufficient judgment error-free traders (smart traders) are present. This suggests that models that mix trader types must be concerned with the proportion of rational traders.

The result that individual behavior can be erratic, yet pricing can still be rational, has also been observed in another context. Bossaerts, Plott, and Zame (2002) conduct experiments that show that market prices can be consistent with the predictions of asset pricing models even though individuals regularly violate portfolio separation. Our results as well illustrate that building models based on the behavior followed by most people most of the time is problematic. Heterogeneity in behavior can be an important factor in asset pricing, and it may not require there to be many rational traders for prices to reflect no judgment errors.

REFERENCES

Anderson, Matthew, and Shyam Sunder, 1995, Professional traders as intuitive Bayesians, *Organizational Behavior and Human Decision Processes* 64, 185–202.

Arrow, Kenneth, 1982, Risk perceptions in psychology and economics, *Economic Inquiry* 20, 1–9.

Blume, Lawrence, and David Easley, 1982, Learning to be rational, *Journal of Economic Theory* 26, 340–351.

Bossaerts, Peter, 2002, *The Paradox of Asset Pricing* (Princeton University Press, Princeton and Oxford).

Bossaerts, Peter, Charles Plott, and William Zame, 2002, Prices and portfolio choices in financial markets: Theory and experimental evidence, Working paper, California Institute of Technology.

Camerer, Colin, 1987, Do biases in probability judgment matter in markets? Experimental evidence, *American Economic Review* 77, 981–997.

Camerer, Colin, 1995, Individual decision making, in John Kagel, and Alvin Roth, eds. *The Handbook of Experimental Economics* (Princeton University Press, Princeton, NJ).

Camerer, Colin, and Teck-Hua Ho, 1996, Experience weighted attraction learning in normal form games, *Econometrica* 67, 827–874.

Crack, Timothy, 2000, *Heard on the Street: Quantitative Questions for Wall Street Job Interviews* (Timothy Crack Publishing, Bloomington, IN).

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *Journal of Finance* 53, 1839–1886.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 2001, Mispricing, covariance risk and the cross section of security returns, *Journal of Finance* 56, 921–965.

De Bondt, Werner F. M., and Richard Thaler, 1985, Does the stock market overreact? *Journal of Finance* 40, 793–805.

DeLong, J. Bradford, Andrei Shleifer, Lawrence Summers, and Robert Waldman, 1990, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703–738.

Duh, Rong Ruey, and Shyam Sunder, 1986, Incentives, learning and processing of information in a market environment: An examination of the base-rate fallacy, in Shane Moriarty, ed. *Laboratory Market Research* (University of Oklahoma, Norman, OK).

Fama, Eugene, 1991, Efficient capital markets II, *Journal of Finance* 46, 1575–1617.

Fama, Eugene, 1998, Market efficiency, long-term returns, and behavioral finance, *Journal of Financial Economics* 49, 283–306.

Friedman, Daniel, 1998, Monty Hall's three doors: Construction and deconstruction of a choice anomaly, *American Economic Review* 88, 933–946.

Ganguly Ananda, John Kagel, and Donald Moser, 1994, The effects of biases in probability judgments on market prices, *Accounting, Organizations and Society* 19, 675–700.

Gervais, Simon, and Terrance Odean, 2001, Learning to be overconfident, *Review of Financial Studies* 14, 1–27.

Hirshleifer, David, 2001, Investor psychology and asset pricing, *Journal of Finance* 56, 1533–1597.

Rubinstein, Mark, 2001, Rational markets: Yes or no: The affirmative case, *Financial Analysts Journal* 57, 15–29.

Russell, Thomas, and Richard Thaler, 1985, The relevance of quasi rationality in competitive markets, *American Economic Review* 75, 1071–1082.

Shleifer, Andrei, and Robert Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52, 35–55.

Shleifer, Andrei, 2000, *Inefficient Markets* (Oxford University Press, Oxford).

Sunder, Shyam, 1995, Experimental asset markets: A survey, in John Kagel, and Alvin Roth, eds. *The Handbook of Experimental Economics* (Princeton University Press, Princeton, NJ).