Do Biases in Probability Judgment Matter in Markets? Experimental Evidence

By COLIN F. CAMERER*

Microeconomic theory typically concerns exchange between individuals or firms in a market setting. To make predictions precise, individuals are usually assumed to use the laws of probability in structuring and revising beliefs about uncertainties. Recent evidence, mostly gathered by psychologists, suggests probability theories might be inadequate *descriptive* models of *individual* choice. (See the books edited by Daniel Kahneman et al., 1982a, and by Hal Arkes and Kenneth Hammond, 1986.)

Of course, individual violations of normative theories of judgment or choice may be corrected by experience and incentives in markets, thus producing market outcomes which are consistent with the individual-rationality assumption even if that assumption is wrong for most agents. Whether judgment and choice violations matter in markets is a question that begs for empirical analysis.

In this paper I use experimental markets to address this issue (see also Rong Duh and Shyam Sunder, 1986; and Vernon Smith, 1982, for an overview). In these markets, traders are paid dividends for holding a one-period asset. The amount of the dividend depends upon which of two states occurred. Traders know the prior probabilities of the states, and a sample of likelihood information about which state occurred. The setting is designed so that prices and allocations will reveal whether traders use Bayes' rule to integrate the prior and the sample information, or whether they judge the likelihood of each state by the "representativeness" of the sample to the state (Amos Tversky and Kahneman, 1982b). (Several other non-Bayesian psychological theories can be tested, too.)

Evidence of judgment bias reported by psychologists poses an implicit challenge to economic theory based on rationality. Sometimes that challenge is made explicit, as when Kenneth Arrow suggested that use of the representativeness heuristic "typifies very precisely the excessive reaction to current information which seems to characterize all the securities and futures markets" (1982, p. 5). Others have warned that judgment biases will affect the judgments of well-trained experts who make societal decisions (about the risk of low-probability hazards, for instance, see Paul Slovic, Baruch Fischhoff, and Sarah Lichtenstein, 1976).

Assertions as bold as Arrow's are extremely rare, because the faith that individual irrationality will not affect markets is a strong part of the "oral tradition" in economics. This faith is often defended with Milton Friedman's (1953) famous claim that theories with false assumptions (such as strong assumptions of individual rationality) might still predict market behavior well (see Mark Blaug, 1980, pp. 104–14, for a cogent discussion). Besides that "*F*-twist," there is a standard list of arguments used to defend economic theories from the criticism that people are not rational. (Counterarguments are given in parentheses.)

1) In markets, agents have enough financial incentive, and experience, to avoid mistakes. (Incentives and experience were provided in David Grether's 1980 experiments on the representativeness heuristic. See

*Department of Decision Sciences, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104. I thank Mike Chernew, Marc Knez, Peter Knez, and Lisabeth Miller for research assistance. Helpful comments have been received from three anonymous referees, Greg Fischer, Len Green, David Grether, Dan Kahneman, John Kagel, Peter Knez, George Loewenstein, Charles Plott, Paul Slovic, Shyam Sunder, Richard Thaler, Keith Weigelt, and especially Howard Kunreuther. This research was funded by the Wharton Risk and Decision Processes Center, the Alfred Sloan Foundation grant no. 8551, and the National Science Foundation grant no. SES-8510758.

also Charles Plott and Louis Wilde, 1982, p. 97.)

2) Random mistakes of individuals will cancel out. (The biases found by psychologists are generally *systematic*—most people err in the same direction.)

3) Only a small number of rational agents are needed to make market outcomes rational, if those agents have access to enough capital or factors of production. (Institutional constraints may prevent those agents from making markets rational; see Thomas Russell and Richard Thaler, 1985.)

4) Agents who are less rational may learn *implicitly* from the actions of more rational agents. (This argument requires that "more rational" agents are identifiable, perhaps by their more vigorous trading.)

5) Agents who are less rational may learn *explicitly* from more rational agents by buying advice or information. (Institutional constraints, and the well-known problems of adverse selection and moral hazard, may limit the extent of information markets.)

6) Agents who are less rational may be driven from the market by bankruptcy, either by natural forces or at the hands of more rational competitors. (A new supply of agents who are less rational, or inexperienced, may be constantly entering the market.)

Most of these arguments, though not all of them, are put to the test in the market experiments described below. Subjects trade for up to 7 hours, observing nearly 100 realizations of the state variable, and every trade earns them a (small) dollar profit or loss (argument 1). The representativeness heuristic is systematic in direction (argument 2). Subjects trade with one another in a "double-oral" auction with no constraints on bidding or offering activity (argument 3), so they can learn implicitly from others' trading behavior (argument 4).

Many of the standard arguments are *not* tested in the experiments: There is no explicit market for advice (argument 5); subjects cannot sell short (argument 3); and bankruptcy is unlikely, though conceivable (argument 6). The first two arguments are being tested in further work. Even with these limits, the market experiments provide a greater combination of incentives, experience, and learning opportunity than in previous judgment experiments.

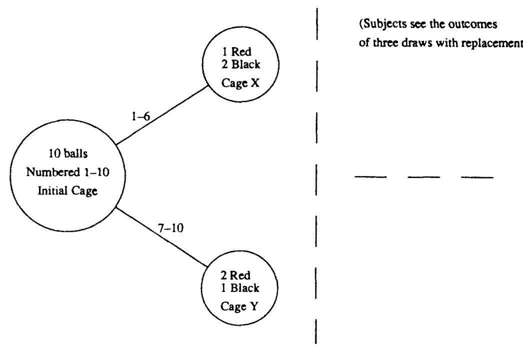

FIGURE 1

I. Experimental Design

In the experiments, each of 8 or 10 traders is endowed with two assets that live one period and pay a liquidating state-dependent dividend.

A. *State Probabilities*

The state is represented by which one of two bingo cages (X or Y) is chosen (Figure 1). A third bingo cage containing 10 balls is used to determine whether cage X or cage Y has been chosen. The X cage contains 1 red and 2 black balls. The Y cage contains 2 red balls and 1 black one. The prior probabilities of X and Y are .6 and .4.[1] Figure 1 is shown on a blackboard for all subjects to see, throughout the experiment.

After either X or Y is chosen (but *not* announced), a sample of three balls is drawn from the chosen cage, with replacement, and the sample is announced before trading begins. Since the cages X and Y contain different populations of balls, which are known to traders, they can use Bayes' rule to calculate $P(X/\text{sample})$ from the prior $P(X)$ and

[1]Unequal priors were chosen because priors of .5 and .5 might have made it too easy for subjects to intuit the Bayesian posteriors. Experiments with equal priors are a natural direction for future work.

TABLE 1—BAYESIAN EXPECTED DIVIDEND VALUES

				Bayesian Posterior $P(X/\text{sample})$			
				.923	.750	.429	.158
				Bayesian Expected Values[a]			
	No. of Traders	Dividend		No. of Reds			
Type	(Experiment No.)	X	Y	0	1	2	3
I	5 (1,3–5,11x–15xh) 4 (2,9r,10h)	500	200	477	425	329	247
II	5 (1,3–5,11x–15xh) 4 (2,9r,10h)	350	650	373	425	521	603
I	5 (6–7) 4 (8,12x)	525	225	502	450	354	272
II	5 (6–7) 4 (8,12x)	180	480	203	255	351	433

Note: All dividends were actually 80 francs higher, for both types of traders and in both states, in experienced subjects experiments 11x, 13x–15xh. (Therefore, all Bayesian expected values are 80 francs higher, too.) In all analyses prices are adjusted for this 80-franc difference.

[a]In francs.

the likelihood functions $P(\text{sample}/X)$ and $P(\text{sample}/Y)$ (which are determined by the cage contents). The top line of Table 1 gives the Bayesian posteriors for all three-ball samples. The possible samples are characterized by the number of reds only, since the order of draws should not matter and the data suggest the order did not matter to subjects. (In some experiments, like John Hey's 1982 experiments on price search, order does seem to matter. Subjects were paid in his 1987 experiments and order still mattered.)

B. *Market Procedure*

Subjects were undergraduate men, and some women, recruited from quantitative methods and economics classes at the Wharton School. These students have all taken statistics and economics courses. Experiments 1 to 10 used subjects who had not been in any previous market experiments. Five experiments used "experienced" subjects who had been in experiments 1 to 10; these experiments are numbered 11x to 15xh (the "x" reminds the reader that subjects were experienced). Experiments were conducted in one 3-hour session (experiments 1 and 2, 6, 9 and 10, 11x to 15xh) or two 2-hour sessions held on consecutive evenings (experiments 3 to 5, 7 and 8).

All trading and earnings are in terms of francs, which are converted to dollars at the end of the experiment at a rate of \$.001 dollars per franc (\$.0015 in experiment 1).[2] Traders are endowed with 10,000 francs and two certificates in each trading period, and 10,000 francs is subtracted from their total francs at the end of each period. In some experiments a known fixed cost (around 5,000 francs) was subtracted from their total earnings at the end of the experiment.

Traders voluntarily exchange assets in a "double-oral auction": Buyers shout out bids at which they will buy, sellers shout out offers at which they will sell. Bids must top outstanding bids and offers must undercut outstanding offers. A matching bid and offer is a trade, which erases all previous bids and offers. All bids, offers, and trades in a period are recorded by the experimenter on a transparency visible to subjects. (No history of previous periods of trading is posted.) Trading periods last 4 minutes in 10-subject experiments, 3 minutes in 8-subject experiments.

At the end of each trading period the state (X or Y) is announced and traders calculate

[2]In practice, using francs makes traders more precise in their trading than they would be with dollars, for example, traders routinely haggle over 5-franc differences between bids and offers, which represent half a penny. Francs may also alleviate competition among traders for relative status in dollar earnings, because traders' dollar conversion rates (while identical) are privately known.

their profits. Dollar profits are given by

$$(1)\ PROFITS = X\left[E_f - R_f + \sum_{i=1}^{x_s} 0_i - \sum_{j=1}^{x_b} B_j + D(S) \times (E_c - x_s + x_b) - F\right],$$

where X = dollar-per-franc conversion rate,
E_f = initial endowment in francs,
R_f = amount of francs repaid at period-end,
E_c = initial endowment in certificates,
x_s = number of certificates sold,
0_i = selling price of ith certificate sold,
x_b = number of certificates bought,
B_j = purchase price of jth certificate bought,
$D(S)$ = dividends per certificate in state S,
F = fixed cost per experiment in francs.

Traders may not sell short (that is, $E_c - x_s + x_b$ cannot be negative), and net francs on hand ($E_f + \Sigma 0_i - \Sigma B_j$) cannot be negative.

C. *Market Equilibrium*

Assuming risk neutrality, traders' reservation prices for assets are expected values. (If they are not risk neutral, their reservation prices are certainty equivalents.) Since each trader's endowment of francs is large enough to buy virtually the entire market supply of assets, and the supply is fixed (by the initial endowment, and the short-selling restriction), there is excess demand at any price less than the highest expected value. Thus, in competitive equilibrium, prices should be bid up to the largest expected value of any trader. One irrational trader who pays too much can therefore create a market price that is too high. The empirical question is whether such traders exist, and whether the experience and financial discipline of a market makes them more rational over the course of an experiment.

Of course, the double-oral auction is not Walrasian, so there is no theoretical assurance that competitive equilibrium will result. However, simple models of the double-oral auction as a dynamic game with incomplete information are beginning to establish the theoretical tendency of double-oral auctions to converge to competitive equilibrium (Daniel Friedman, 1984; Robert Wilson, 1985; see David Easley and John Ledyard, 1986). The empirical tendency to converge is well-established (for example, Smith, 1982), even in designs meant to inhibit convergence (Smith and Arlington Williams, in press).

D. *Competing Theories*

In each experiment, traders are randomly assigned to either of two "types," which differ in the dividends they receive in the two states X and Y (see Table 1). The dividends are chosen so that competing theories predict different patterns of prices and allocations (see Table 2). Each theory will now be described briefly.

Bayesian. If traders use Bayes' rule to calculate posterior probabilities given the sample data, prices should converge to the Bayesian expected values given in Table 2, assuming risk neutrality. (Tests and controls for risk neutrality are described below.) In the experiments described by the top panel of Table 2, for instance, type I traders should pay up to 477 if the sample is 0 reds, 425 if 1 red, 329 if 2 reds, and 247 if 3 reds. Type II traders should pay up to 373, 425, 521, and 603, respectively. Therefore, if the sample is 0 reds, then type I traders should buy from type II traders at a price of 477. If the sample is 2 or 3 reds, the type II traders should buy all the units, at prices of 521 or 603, respectively. If the sample is 1 red, then type I and type II traders both have a Bayesian expected value of 425 francs, so we expect half the units will be held by each of the two types of traders. (Trades might take place because of uncontrolled differences in risk tastes, but units are still equally as likely to end up in the hands of type I and type II traders.) In experiments 6 to 8 and $12x$, dividends were chosen so that the Bayesian expected values of the type I and type II

TABLE 2—PRICE AND ALLOCATION PREDICTIONS OF COMPETING THEORIES

	Predictions Expressed as: Price P (Type Holding Assets)			
		Number of Reds in Sample		
Theory	0 Reds	1 Red	2 Reds	3 Reds
Experiments 1–5, $9r$–$11x$, $13x$–$15xh$				
Bayesian	477 (I)	425 (I, II)	521 (II)	603 (II)
Exact Representativeness	477 (I)	$P > 425$ (I)	$P > 521$ (II)	603 (II)
Conservatism	$P < 477$ (I)	$P > 425$ (II)	$P < 521$ (II)	$P < 603$ (II)
Overreaction	$P > 477$ (I)	$P > 425$ (I)	$P > 521$ (II)	$P > 603$ (II)
Base-Rate Ignorance	467 (I)	450 (II)	550 (II)	617 (II)
Experiments 6–8, $12x$				
Bayesian	502 (I)	450 (I)	354 (I)	433 (II)
Exact Representativeness	502 (I)	$P > 450$ (I)	$P > 354$ (II)	433 (II)
Conservatism	$P < 502$ (I)	$P < 450$ (I)	$P > 354$ (I)	$P < 433$ (II)
Overreaction	$P > 502$ (I)	$P > 450$ (I)	$P > 354$ (II)	$P > 433$ (II)
Base-Rate Ignorance	492 (I)	425 (II)	380 (II)	447 (II)

traders were (nearly) equal when a 2-red sample was drawn.

Exact Representativeness. If subjects take the representativeness of the sample to the cage contents as a psychological index of the cage's likelihood, non-Bayesian expected values might result. Representativeness is a vague notion, but we can distinguish some precise variants of it. For instance, subjects might think $P(X/\text{sample}) = 1$, if the sample resembles the X-cage contents more closely than the Y-cage contents. Or they might think $P(X/\text{sample}) = 1$, if the sample exactly matches the X-cage contents. These extreme hypotheses are clearly ruled out by the data presented below.

More reasonably, subjects may be intuitively Bayesian for most samples, but overestimate a cage's likelihood when a sample resembles the cage *exactly*. This "exact representativeness" theory predicts that subjects will judge $P(X/1 \text{ red})$ to be greater than the Bayesian posterior .75 because a 1-red sample exactly matches the X-cage's contents. Similarly, $P(Y/2 \text{ red})$ will be judged to be greater than .57; other probabilities will be Bayesian. Of course, there are other possible interpretations but since they are either imprecise or clearly incorrect, only exact representativeness will be considered carefully.

Under exact representativeness, prices will be higher than Bayesian in 1- and 2-red periods (as shown in Table 2) and type I traders will hold units in 1-red periods. (Recall that the Bayesian theory predicts types I and II are equally likely to hold units in 1-red periods.)

Base-Rate Ignorance. If subjects judge $P(\text{state}/\text{sample})$ by the representativeness of samples to states, their judgments may ignore differences in the prior probabilities (or "base rates") of states (Tversky and Kahneman, 1982b). In our setting it is difficult to integrate this aspect of representativeness with other aspects, like the psychological power of exact representativeness, because the two aspects often work in opposite directions. In 1-red samples, for instance, exact representativeness predicts $P(X/1 \text{ red})$ will be overestimated, while ignorance of the higher base rate of X implies $P(X/1 \text{ red})$ will be underestimated. Since predictions of a theory that integrates representativeness with base-rate ignorance are ambiguous, I define base-rate ignorance as using Bayes' rule with erroneous priors $P(X) = P(Y) = .5$. Predictions of this theory are shown in Table 2.

Of course, ignoring base rates completely is rather implausible. For example, in an experiment with a prior probability of .001, it seems unlikely that subjects will act as if the prior is .5. If priors are simply underweighted, but not ignored, the data will show some statistical support for the complete base-rate ignorance theory. The theory should be considered an extreme benchmark that helps us judge whether priors are underweighted at all.

Conservatism. Subjects may be "conservative" in adjusting prior probabilities for sample evidence (for example, Ward Edwards, 1968).

Overreaction. Subjects may adjust prior probabilities *too much*, as if overreacting to sample evidence. The overreaction theory makes the same prediction as representativeness in 1- and 2-red periods, but it predicts bias in 0- and 3-red periods where representativeness does not. Note that the conservatism and overreaction theories make exactly opposite predictions. This implies quite a challenge for the Bayesian theory: Prices must be exactly at the Bayesian prediction, or insignificantly different from it, for both theories to be falsified.

II. Results

Fifteen experiments have been conducted—ten with inexperienced subjects, five with experienced subjects—excluding two inconclusive pilot experiments. For the sake of brevity, many details of the analyses are omitted and can be found in working papers available from the author.

There are two kinds of data which distinguish between theories: prices at which trades occurred, and the number of units of the asset that traders held at the end of trading periods.

A. *Trade Prices*

The mean prices across experiments 1 to 8 are summarized by a time-series of 90 percent confidence intervals, shown in Figure 2.[3] The upper (lower) solid line is the upper

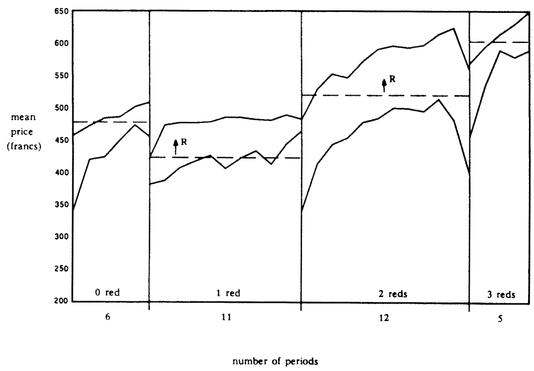

FIGURE 2

(lower) end of the confidence interval. Bayesian expected values are shown by dashed lines, and the direction of the exact representativeness prediction is shown by an arrow marked "*R*." Each of the four panels represents a different sample. From left to right, observations within a panel represent data from the first time that sample was drawn, the second time the same sample was drawn, and so forth.

Prices converge, from below, toward the Bayesian levels. These data clearly rule out many non-Bayesian theories of probability judgment (like the two extreme brands of representativeness mentioned above). However, prices do not converge exactly to the Bayesian expected values. There is some evidence of exact representativeness, because prices drift above the Bayesian expected values in 1- and 2-red periods. However, the confidence intervals are wide, and the degree of bias is rather small. Indeed, since prices should only converge to Bayesian predictions if the hypotheses of risk neutrality,

[3] Confidence intervals were constructed by first calculating mean prices in each period of each experiment, then separating the time-series of mean prices for each different sample. Data from experiments 6 to 8 were normalized so that the Bayesian predictions in those experiments were the same as in experiments 1 to 5. This yields groups of data such as 8 mean prices from the first 0-red period in each of the 8 experiments numbered 1 to 8. The mean of those means, and its standard error (the standard deviation divided by $8^{1/2}$) are used to calculate the 90 percent confidence interval. A second confidence interval was calculated using mean prices from the second 0-red period in each of the 8 experiments, and so on. Not all experiments have the same number of 0-red periods, so the number of observations in each confidence interval gradually decreases. The procedure was stopped just before there was only one experiment left with an Nth observation of a particular sample.

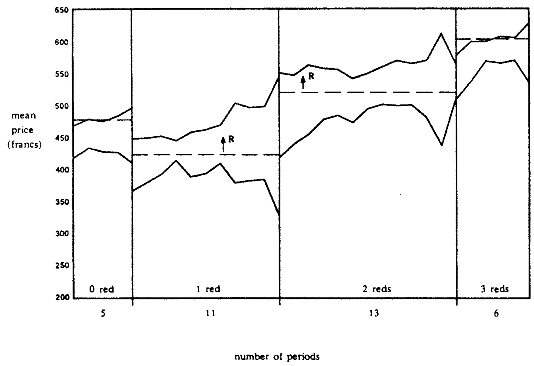

FIGURE 3

competitive equilibrium, and Bayesian updating are all true simultaneously, it is rather remarkable that prices converge as closely to the Bayesian predictions as they do.

Figure 3 shows confidence intervals from experiments with experienced subjects. Prices begin closer to the Bayesian expected value, and have less tendency to drift above it in 1- and 2-red periods. The confidence intervals are also wide, because they summarize a small number of experiments.[4]

We can define bias in prices as a deviation from the Bayesian prediction. If the Bayesian theory is true, biases will be around zero. To conduct statistical tests on price biases, the time-series of prices in each experiment must be independent. Since prices are typically autocorrelated, the equilibrium degree of bias is estimated from a simple partial adaptation model (a first-order autoregression),

$$(2) \quad P_t - P_{\text{Bayes}} = a + b\left(P_{t-1} - P_{\text{Bayes}}\right) + e_t,$$

where P_t is the tth observation of price and P_{Bayes} is the Bayesian prediction. This specification implies that the deviation from equilibrium is reduced by a fraction $1-b$ each trade. If b is close to 1, convergence is very slow; if b is close to 0, convergence is fast. While there is no theoretical rationale for (2), it works well empirically and there is no well-established theory of price convergence which suggests it is wrong.

Call the bias for the tth price B_t; it equals $P_t - P_{\text{Bayes}}$. If we define equilibrium as a bias that does not change each period, we impose $B_t = B_{t-1} = B$ on (2) and get

$$(3) \qquad B = a + bB + e_t.$$

Since $E(e_t) = 0$, a little algebra shows that we can estimate the degree of equilibrium bias B consistently by the estimator $B' = a'/(1-b')$, where a' and b' denote ordinary least squares estimators of a and b in (2). The standard error of B' can be calculated from a Taylor series approximation involving the variances of a' and b' and their covariance.[5]

Regressions were first run separately for each period, effectively allowing a and b to vary each period. The simple specification (2) fit fairly well: The convergence rate b was typically estimated precisely, and residuals were uncorrelated and roughly homoskedastic. An F-test (Jan Kmenta, 1971, p. 373) was used to test whether adjacent periods could be pooled at the 10 percent level. Periods were pooled, starting with the last period, until the F-test was violated.

The estimate B' resulting from the last group of poolable periods in each experiment are shown in Table 3. Also reported is the t-statistic testing the hypothesis that $B = 0$, which is simply B' divided by its (approximated) standard error. Sample sizes are shown in parentheses next to each experiment number. T-statistics marked with asterisks are unreliable because the assumption of normality of residuals was violated at the 1 percent level, by the studentized range

[4] Intervals flare out in Figure 3 when the number of different experiments used to construct them drops steeply and standard errors increase dramatically.

[5] I thank Dave Grether for correcting a mistake in earlier estimates of $V(B')$. The Taylor series approximation of $a'/(1-b')$ around its true value $a/(1-b)$ is $a/(1-b)+(a'-a)/(1-b)+a(b'-b)/(1-b)^2$, plus some higher-order terms. Using this expression to calculate (approximately) $V(a/(1-b))$, or $E[(a'/(1-b')-a/(1-b))^2]$ yields $V(a')/(1-b)^2 + a^2V(b')/(1-b)^4 + 2a\text{COV}(a',b')/(1-b)^3$. Evaluating this expression at a' and b' gives approximations of $V(B')$.

TABLE 3—ESTIMATES OF BIAS IN EQUILIBRIUM PRICES, AND TESTS OF THE BAYESIAN HYPOTHESIS AGAINST COMPETING HYPOTHESES

			0-Red Periods		
Experiment (n)	Bias B	t-Statistic	Significance Levels, Bayesian vs. Conservatism	Overreaction	Base-Rate Ignorance
Inexperienced subjects					
1 (24)	−28.31	2.31	.01	.99	.06
2 (46)	19.28	10.95	.999	.000	.000
3 (54)	−11.09	−3.67*	.000	.999	.999
4 (57)	4.97	4.56*	.999	.000	.000
5 (10)	−22.35	−6.57	.000	.000	.999
6 (29)	15.36	5.67*	.999	.000	.999
7 (70)	32.44	.65*	.76	.24	.57
8 (7)	−37.90	−2.20	.99	.01	.12
9r (9)	10.01	8.70	.000	.999	.999
10h (10)	−2.54	−1.16*	.12	.88	.999
mean	−2.95		.49	.51	.57
Experienced subjects					
11x (53)	−44.12	12.15	.000	.999	.999
12x (34)	15.17	1.78	.96	.04	.01
13x (8)	76.50	.49	.69	.31	.51
14x (18)	11.61	3.64	.999	.000	.999
15xh (16)	4.92	1.41	.92	.08	.35
mean		2.14		.71	.29.57

			1-Red Periods		
	Bayesian vs.		Exact Representativeness, Overreaction	Conservatism	Base-Rate Ignorance
1 (13)	5.00	2.63*	.005	.005	.999
2 (40)	56.34	7.94*	.000	.000	.000
3 (40)	1.18	.29*	.46	.46	.999
4 (25)	49.81	18.94	.000	.000	.000
5 (37)	31.19	10.23	.000	.000	.000
6 (28)	51.80	9.10	.000	.999	.999
7 (16)	23.12	4.65	.001	.999	.985
8 (57)	93.83	.18	.43	.57	.51
9r (8)	51.15	6.21	.000	.000	.000
10h (50)	54.63	14.12*	.000	.000	.000
mean	39.92		.09	.30	.45
11x (44)	−2.76	−2.08	.98	.98	.999
12x (7)	32.18	3.82	.001	.999	.999
13x (24)	.96	.49	.31	.31	.999
14x (33)	27.88	1.89*	.03	.03	.21
15xh (8)	29.77	3.49	.005	.005	.001
mean	17.61		.27	.47	.64

(continued)

test. Other diagnostic tests and estimates of b are reported in working papers.

Roughly speaking, biases are distributed around zero in 0-, 2-, and 3-red periods. Biases are positive in 1-red periods of every experiment except 11x, generally with large t-statistics. Biases are also positive in 2-red periods with experienced subjects, but not with inexperienced subjects.

The right-hand columns of Table 3 test the hypothesis that prices are Bayesian against each of the competing theories. The tests of the Bayesian theory against exact representativeness, conservatism, and overreaction are one-tailed t-tests of the null hypothesis $B = 0$ against one-sided alternative hypotheses (which vary depending upon the theory and the sample). Since the base-rate ignorance theory predicts a point estimate of the bias rather than a direction, the significance level of the Bayesian hypothesis against the base-rate ignorance alternative was esti-

TABLE 3—(CONTINUED)

	Bayesian vs.		2-Red Periods Exact Representativeness, Overreaction	Conservatism	Base-Rate Ignorance
1 (61)	−27.00	−4.84	.999	.000	.999
2 (24)	60.25	.11	.46	.54	.50
3 (22)	99.00	6.27	.000	.999	.000
4 (83)	77.39	7.74	.000	.999	.000
5 (52)	−53.31	6.55	.76	.24	.64
6 (77)	−1.74	−.02*	.51	.51	.52
7 (16)	98.24	18.16	.000	.000	.000
8 (16)	45.45	3.35	.001	.001	.000
9*r* (24)	−17.23	−4.55	.999	.000	.999
10*h* (27)	−7.57	−.95	.67	.33	.999
mean	27.35		.44	.36	.44
11*x* (18)	49.28	7.55	.000	.999	.000
12*x* (8)	20.80	12.16*	.000	.000	.000
13*x* (15)	17.22	14.35	.000	.999	.000
14*x* (11)	22.62	18.26	.000	.999	.000
15*xh* (17)	12.47	.80	.21	.79	.62
mean	24.48		.04	.78	.12

	Bayesian vs.		3-Red Periods Conservatism	Overreaction	Base-Rate Ignorance
1 (40)	2.47	.40	.65	.35	.04
2 (17)	−209.34	−3.51	.000	.999	.85
3 (41)	41.88	4.12*	.999	.000	.000
4 (48)	11.26	.64*	.74	.26	.41
5 (26)	−10.57	.70	.24	.76	.90
6 (32)	31.55	1.29*	.90	.10	.24
7 (22)	20.04	3.24*	.999	.001	.000
8 (28)	−26.46	−.79	.29	.71	.70
9*r* (29)	14.01	.29	.61	.39	.48
10*h* (7)	2.61	.02	.51	.49	.50
mean	−12.23		.59	.41	.41
(2 deleted)		9.64			
11*x* (37)	−22.51	−6.52	.000	.999	.999
12*x* (9)	24.98	.49	.69	.31	.39
13*x* (35)	11.65	.65	.74	.26	.40
14*x* (28)	114.21	.13	.55	.45	.50
15*xh* (9)	−38.00	−3.74	.000	.999	.999
mean	4.70		.40	.60	.66

Notes: * denotes studentized range of residuals greater than the 1 percent level for normality, so standard errors are unreliable. Biases are truncated in calculating means when the equilibrium price implied by the bias estimate is greater than the maximum dividend for the type of trader holding a majority of units (for example, 0-red period, experiment 7).

mated from likelihood ratios.[6] Significance levels were estimated by assuming the t-statistics were normally distributed (a reasonable approximation for most of the sample sizes in Table 3). Levels less than .001 or above .999 are reported as .000 or .999.

The significance levels of tests against most of the alternative theories are roughly 50 percent, suggesting departures from the Bayesian predictions are not systematic. However, the Bayesian theory can be strongly rejected against the alternative of exact representativeness in most 1-red periods and

[6] P(data/Bayesian) and P(data/Base-rate Ignorance) were calculated assuming the estimate B' was normally distributed with standard deviation $s(B')$. Assuming one of the two theories is true, and they are equally likely; Bayes' rule can then be used to calculate P(Bayesian/data).

many 2-red periods. Of course, the statistical significance of a bias is simply a measure of whether it could be due to chance. Whether the biases are economically significant is discussed in the conclusion.

Note that the graphs and the statistical tests seem to tell different stories because the confidence intervals are wide while the t-statistics are large. This simply means that biases are not random in each experiment (hence, the extreme significance levels in Table 3), but the degree of bias varies a lot across experiments (hence, the wide confidence intervals).

In most experiments subjects did not make probability calculations during the experiment (though they were given calculators to record profits). However, in experiment 1 two traders *did* write the correct likelihood ratios $P(X/\text{sample})/P(Y/\text{sample})$ on their profit sheets during the experiment; prices were quite close to Bayesian (for example, 1-red prices were only 5 francs too high). A small number of aggressive Bayesians apparently can make the market price Bayesian, but did not do so very often.

B. *Allocations of Assets*

For most samples, competing theories all predict the same type of trader will hold units. When the theories make the same prediction, they are extremely accurate. In 0-red and 3-red periods, for instance, virtually all of the units are held by the traders with the highest expected dividend type in every experiment.

The theories disagree about allocations in 1-red periods of some experiments and 2-red periods of other experiments. In these experiments, the average fraction of traders holding any units at the end of the period and the average fraction of units held were calculated for dividend types I and II. These data are shown in Table 4.

In the 1-red periods, the Bayesian theory predicts type I and type II traders are equally likely to hold units (since their expected values are equal, at 425). Exact representativeness predicts units will be held by type I's.

Across all experiments with inexperienced subjects, type I's hold 78 percent of the units. This fraction is quite stable across experiments, and is about the same in early periods (the first half of the periods) and late periods. With experienced subjects, about 90 percent of the units are held by type I's. Prices biases were apparently not due to simple one or two type I's buying units, because about 80 percent of the type I subjects held any units, compared to roughly 30 percent of the type II subjects. Significance tests using mean data from each experiment strongly reject the Bayesian theory against the alternative of exact representativeness.[7] Such cross-experiment tests are especially reliable because we can be confident that different experiments are genuinely independent because they contain different subjects.

The smaller amount of data from 2-red periods (the bottom panel of Table 4) are not very conclusive. The Bayesian theory predicts type I's will hold, exact representativeness predicts type II's, and holdings are about equal. This corroborates the finding from price data that exact representativeness has little effect in 2-red periods.

The results of Duh and Sunder (1986) are worth summarizing at this point. In their experiments, the two states (called R and W) are two bingo cages containing 16 red and 4 black balls (R) and 4 red and 16 black balls (W). The prior $P(R)$ varied from .65 to .85 across experiments, since their main concern was whether subjects ignored prior probabilities. One ball is drawn from whichever cage (state) is chosen (so there is no possibility of exact representativeness). They find that when an R is drawn, prices are close to Bayesian. When a W was drawn, the Bayesian theory predicted about as well as a base-rate ignorance theory (denoted $NBR2$) in which $P(R)$ and $P(W)$ are judged to be equal, and an extreme version of representativeness in which $P(W)$ is judged to be

[7] We can test the hypothesis that the average percentage holding of type I's was 50 percent by assuming the fractions across experiments 1 to 5, $9r$ and $10h$ are normally distributed (the t-statistic is 9.28). The more conservative binomial test of successes yields a significance level less than 1 percent. For experienced subjects these statistics are 10.33 and 6 percent.

TABLE 4—HOLDINGS OF UNITS AT PERIOD END, BY TRADER TYPE

Experiment (n = no. of periods)		Type I		Type II	
		1-Red Periods			
Theories predicting Each Type to Hold:		Bayesian, Exact Representativeness, Overreaction		Bayesian, Conservatism, Base-Rate Ignorance	
Inexperienced Subjects		Fraction Holding Any	Fraction Held	Fraction Holding Any	Fraction Held
1 ($n=5$)		.76	.85	.24	.15
2 ($n=7$)		.76	.82	.50	.18
3 ($n=9$)		.94	.75	.42	.25
4 ($n=12$)		.67	.73	.38	.27
5 ($n=10$)		.76	.64	.56	.36
9r ($n=8$)		.91	.85	.30	.15
10h ($n=7$)		.69	.85	.37	.15
Means	All Periods:	.787	.784	.396	.216
	Early Periods:	.82	.73	.53	.27
	Late Periods:	.76	.82	.31	.18
Experienced Subjects					
11x		.91	.93	.25	.07
13x		.95	.78	.56	.22
14x		.50	.95	.10	.05
15xh		.91	.97	.09	.03
Means	All Periods:	.818	.908	.250	.092
	Early Periods:	.88	.87	.29	.13
	Late Periods:	.76	.94	.21	.06
		2-Red Periods			
Theories Predicting Each Type to Hold		Bayesian Conservatism		Exact Representativeness, Overreaction, Base-Rate Ignorance	
6 ($n=11$)		.53	.36	.75	.64
7 ($n=11$)		.57	.60	.45	.40
8 ($n=8$)		.71	.49	.83	.51
12x ($n=13$)		.76	.40	.88	.60
Means	All Periods:	.643	.463	.728	.538
	Early Periods:	.64	.42	.77	.58
	Late Periods:	.66	.52	.67	.48

one (denoted $NBR1$). They do not estimate the degree of price bias parametrically, but it seems to be smaller in magnitude than the biases observed here. They conclude, "Although the Bayesian model performs best among the four models in its ability to predict transaction prices, the observed market behavior still deviates from the Bayesian prescription." I suspect the Bayesian model predicts better in their experiments than in mine because the exact representativeness in my experiments is a stronger psychological force than the base-rate ignorance in theirs.

C. *Further Controls for Risk and Incentives*

The analyses of prices and allocations lean heavily on the assumption that traders are risk neutral, so that they trade at expected values. If traders are risk seeking, prices will be above expected values. The higher prices observed in 1-red periods could therefore reflect risk seeking by Bayesian traders rather than judgment bias by risk-neutral traders.

This explanation is unlikely for several reasons. First, the Arrow-Pratt risk premium, which measures the approximate degree to which prices depart from expected values because of risk seeking, depends only on the variance of an asset's value (and possibly its mean) and the shape of traders' utility functions. The mean and variance of the value of units are identical for type I and type II traders in 1-red periods, so their risk premia should be equal (assuming no systematic differences in utility functions). Therefore, the Bayesian prediction that type I and type II traders hold equal amounts of units should be true even if traders are not risk neutral; but the equal holdings prediction is strongly rejected.

Second, most attempts at measuring risk tastes in experimental settings find evidence of risk aversion rather than risk seeking (for example, James Cox, Smith, and James Walker, 1985; Smith, Gerry Suchanek, and Williams, 1987). Third, the allocation data show that about 80 percent of the type I traders are holding units at the high prices in 1-red periods. It seems unlikely that almost every type I trader in every experiment would be risk seeking. Fourth, the data from all four samples can be used to estimate the degree of risk seeking implicit in prices, assuming a specific utility function. Adjusting the apparent price biases in 1-red periods for risk does reduce them by about two-thirds, but not quite to zero.[8]

More direct evidence of whether risk seeking can explain the biases comes from a control experiment (denoted $9r$) in which risk neutrality was induced by design (see Alvin Roth, 1983; Joyce Berg et al., 1986; though, see Cox et al., 1985). Traders accumulated earnings in francs but the francs were not converted into dollars at the end of the experiment. Instead, traders were paid \$15 plus a \$50 bonus if a uniformly distributed five-digit number between 0 and 50,000 was *below* their amount of earnings. Each franc they earned then raised their probability of winning the \$50 prize by 1/50,000; so francs were like units of probability. Since assets are lotteries over possible amounts of francs, and francs are probability units, assets are like compound lotteries. If traders satisfy the reduction of compound lotteries axiom in expected utility theory, they should regard a gamble with an expected franc value of G as identical to a certain payment of G francs, so they should act as if they are risk neutral toward francs. If biases observed in earlier experiments were due to risk seeking, those biases should disappear in experiment $9r$.

A second control experiment (denoted $10h$) used a "high-stakes" dollar-per-franc conversion rate of \$.005 rather than \$.001. Subjects in this experiment made about \$20 per *hour*. Experiment $15xh$ used the same level of high stakes with experienced subjects. If apparent price biases are due to insufficient incentive to think carefully about probabilities, biases should be smaller in experiments $10h$ and $15xh$.

Figure 4 shows the mean prices from the risk-control experiment $9r$ (thick line) and the high-stakes experiment $10h$ (thin line).[9] Compared to prices from inexperienced subjects shown in Figure 2, prices in these experiments are extremely close to the Bayesian expected values, except in 1-red periods. Price regression results and allocations (in Tables 3 and 4) suggest the exact representativeness bias in 1-red periods is highly significant. Therefore, biases in 1-red periods

[8]The value of the risk-seeking constant A was estimated in each experiment, assuming both constant absolute (CARS) and constant relative risk seeking (CRRS). The value of A was chosen to minimize the absolute deviations between observed price bias from Table 3 and the bias predicted by the Arrow-Pratt risk premium with parameter A, summed across the four possible samples. Weighting deviations by the number of trades in each sample minimized risk-adjusted biases better than not weighting them. The CARS and CRRS models fit almost identically. Using CARS, risk-adjusted biases in 1- and 2-red periods averaged 13.8 and -15.6 francs (experiments 1 to 8) and 4.5 and -6 francs (experiments $11x$ to $15xh$). In experiments $9r$ and $10h$ risk adjustment actually increased 1-red biases to 55.5 and 56.4 francs. Furthermore, in experiment $9r$, the estimated A was about equal in magnitude to A's in other experiments, though it should be zero in theory.

[9]The lines end abruptly because each experiment has a different number of periods of each sample. There are five 0-red periods in $10h$, for instance, and only three in $9r$. Also, the spike in the second 1-red period of experiment $9r$ was a short burst of irrational buying at very high prices, which defies explanation.

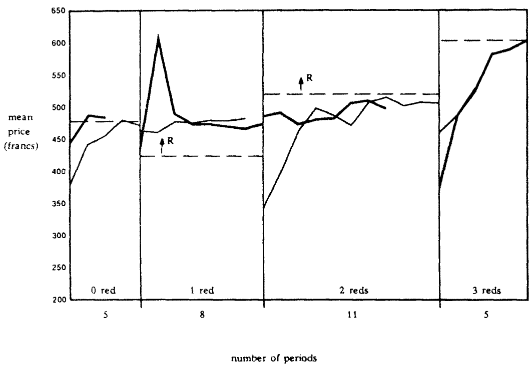

FIGURE 4

in other experiments are probably not due to risk seeking or insufficient motivation. At the same time, the control experiments give evidence *against* the exact representativeness prediction in 2-red periods.

D. *Individual Judgments and Market Prices*

The point of experiments like these is to compare behavior of individuals with behavior of markets in which individuals participate. So far there has been only an assumption individuals will err in using Bayes' rule, but no direct comparison between individuals and the market. However, we can make such a comparison because subjects did make individual probability judgments before trading began (except in experiments 1 and 2).

Judgments were rewarded with a quadratic scoring rule, with money incentives for accuracy.[10] The scoring rule is incentive compatible assuming risk neutrality (subjects should report their true subjective probabilities), but nonrisk neutrality will cause judgments to deviate from true beliefs. Subjects were given 10 to 20 three-ball samples from the bingo cages, with instant feedback about whether X or Y occurred. After completing the scoring-rule exercise, subjects were informed that they would ranked according to their earnings from the scoring-rule exercise, from 1 to N. They were told to predict their rank, choosing exactly one number between 1 and N, and they were paid \$5 if their rank was exactly correct.

We can compare the average scoring-rule judgment with a probability estimate imputed from the equilibrium price bias. For instance, in experiment 3 the estimated bias in 0-red periods was -11.09 francs (see Table 3). Since type I traders were holding in these periods, and their payoffs range from 200 ($P(X)=0$) to 500 ($P(X)=1$), the probability scale naturally corresponds to a 300-franc price scale from 200 to 500. A bias of -11.09 francs implies a probability judgment of $P(X/\text{0-red})$ that is $-11.09/300$, or $-.037$, different from the Bayesian posterior of .923. Probabilities were imputed from market prices for each sample and each experiment, using the estimated biases from Table 3.

Average individual probabilities from the scoring rule and probabilities imputed from market prices, averaged across experiments, are shown in Table 5. Both kinds of probabilities are close to Bayesian in 0- and 3-red samples. In 1- and 2-red samples, the individuals' probability estimates are closer to Bayesian than the market prices are,[11] but the gap is smaller with experienced subjects.

It seems that for exactly representative samples, markets are often *more* biased than

[10]Samples of three balls were drawn, exactly as in determining states, and subjects were asked to choose a two-digit "decision number" from 00 to 99. Define that number, divided by 100, as D. If event X occurred, subjects were paid $2D-D^2$ dollars. If event Y occurred, subjects were paid $1-D^2$ dollars. Subjects were shown a table of the possible numerical payoffs. If a subject's true subjective probability of X occurring was S, and she choose D, her expected payoff was $S(2D-D^2)+(1-S)(1-D^2)$. This payoff has a maximum at $D^*=S$, that is, subjects should truthfully choose their subjective probabilities as their decision numbers, except for risk aversion. If subjects are risk averse (risk seeking), their reported probabilities will be biased toward (away from) .5.

[11]The differences between averaged scoring-rule judgments and probabilities implicit in market prices are highly significant by parametric t-tests, or by nonparametric matched-pairs or rank-sum tests, except in 2-red periods with inexperienced subjects.

TABLE 5—AVERAGE PROBABILITY JUDGMENTS OF INDIVIDUALS AND PROBABILITIES IMPLICIT IN MARKET PRICES (EXPRESSED AS DEVIATIONS FROM THE BAYESIAN POSTERIOR $P(X/\text{SAMPLE})$)

Sample	0-Red	1-Red	2-Red	3-Red
Direction of Deviation Predicted by Exact Representativeness:	0	+	−	0
Inexperienced Subjects (8 Experiments)				
Individual Mean	−.009	−.030	−.031	−.022
(Standard Deviation)	(.044)	(.087)	(.033)	(.076)
Market Prices Mean	−.008	+.141	−.100	−.035
(Standard Deviation)	(.062)	(.081)	(.193)	(.073)
Experienced Subjects (5 Experiments)				
Individual Mean	+.037	−.026	−.043	−.084
(Standard Deviation)	(.022)	(.052)	(.061)	(.069)
Market Prices Mean	+.007	+.059	−.081	−.016
(Standard Deviation)	(.092)	(.049)	(.044)	(.117)

individuals are. One explanation is that market prices are determined by one or two highly biased traders, but almost all traders were holding units at the biased prices. Another possibility is that the market mechanism and the quadratic scoring rule simply elicit different probability judgments.

One consolation is that the biases shrink with experience. A closer look at individual data may suggest why. For market prices to be less biased than individuals, traders who are less biased must exert more influence on the market price. There is no external market to evaluate whether traders are unbiased and allocate more trading capital to them. Therefore, to exert more influence the traders who are less biased must realize they are less biased, and trade more aggressively.

Whether traders realize their relative ability at probability judgment can be measured by whether their predicted ranks in the scoring-rule exercise are correlated with their actual ranks. The two sets of ranks were somewhat correlated—averaging .49 for inexperienced subjects and .30 for experienced subjects[12]—so subjects do have some self-insight. However, predictions about relative ability are not highly correlated with the amount of arbitrage (defined as buying and selling in the same period). Those correlations averaged −.09 for inexperienced subjects, and .23 for experienced subjects. Furthermore, actual ranks and arbitrage were uncorrelated (.05 and −.12) with both inexperienced and experienced subjects. It seems that aggressive trading, as measured by arbitrage, is not something inexperienced subjects do only because they think they are better probability judges than others.

[12]These are high correlations considering that the range of the predicted rank variable was restricted by subjects' optimism about their ranks. Sixty-two of 74 inexperienced subjects (84 percent) thought they were in the top 50 percent in scoring-rule earnings, compared to 23 of 40 experienced subjects (58 percent). Apparently optimism is nearly erased after one experiment.

III. Conclusion and Future Research

In many experiments subjects do not follow the laws of probability, particularly Bayes' rule. However, subjects in these experiments are often unpaid and given little practice making judgments. In markets, traders often have incentives and experience, and people who are good at estimating probabilities can often exert more force on prices. Therefore, biases in *individual* judgments need not affect prices and allocations in *markets*.

Whether biases affect market outcomes is tested in a series of simple experimental markets. In the markets, traders exchange units of an asset that pays a state-dependent dividend. A random device yields sample evidence about which state has occurred. Traders' demand for assets depends upon their judgments about posterior state probability. If the market functions as if traders are Bayesians, a certain pattern of prices and allocations is predicted to occur. But if traders overestimate P(state/sample), relative to the Bayesian posterior, when the sample exactly matches the contents of a bingo cage that represents the state, then different prices and allocations will occur. This competing theory is called "exact representativeness." It is less useful than the Bayesian theory because it does not predict prices when samples do not exactly match states, but it does have some bite. Other non-Bayesian psychological theories can be defined too.

In eight experiments with inexperienced subjects, prices tend toward the Bayesian predictions, but there is some evidence of exact representativeness bias in prices and allocations. However, the degree of bias is small, and it is even smaller in experiments with experienced subjects. All other non-Bayesian theories can be rejected.[13] Furthermore, the Bayesian theory predicts prices remarkably well when the exact representativeness theory does not apply.

In most experiments, biases are statistically significant for only one of the two samples (the 1-red sample) in which exact representativeness predicts bias. Indeed, if the reader values the only experiment with controls for risk seeking ($9r$), exact representativeness predicts no better than chance: it predicts the significant bias in the 1-red period correctly, but it predicts the wrong sign on the significant bias in the 2-red period.

It is easy to imagine other market settings in which unbiased traders could correct market biases completely.[14] Some of these settings are the subject of ongoing research. However, if one pretends to not know the results, it is easy to imagine that biases could have been entirely eliminated in these experiments, too.

Whether the exact representativeness biases in 1-red periods are significant depends upon your yardstick of significance. By one overworked yardstick, the statistical test of whether they could be due to chance, the biases in 1-red periods are highly significant. The possibility of excess profits is an important yardstick in economics. There are apparently loss of profits to be earned from exploiting biased subjects, since they overpay by roughly $.20 per trade (a few dollars per experiment) in 1-red periods of the high-stakes experiments $10h$ and $15xh$. Excess profits are a lot smaller, only about $.03 per trade, in the other experiments. On the probability yardstick the biases are errors of about .10, which are large if your purpose is testing students' ability to make exact Bayesian calculations and small if your purpose is comparing these biases with errors found in other studies.[15]

Of course, if the stakes were large enough or (perhaps more importantly) traders had enough experience, the apparent biases might disappear entirely. Therefore, we should hesitate to generalize these results to the New York Stock Exchange (though some have tried[16]), but the results may generalize to settings in which stakes are relatively small and agents have little experience in a repeated situation. For instance, consumers might judge the quality of a new product by how much the product's packaging or advertising resembles that of well-known products. Financial journalists sometimes argue a depression is ahead because a pattern of economic indicators resembles a pattern from

[13] If subjects tended to ignore or underweight the unequal prior probabilities of the states, then 2-red biases would be larger than 1-red biases. Exactly the opposite is true. Notice also that overreaction predicts reasonably well in 1- and 2-red samples, when it overlaps with exact representativeness, but it predicts poorly in 0- and 3-red samples.

[14] For instance, if biases caused prices to be lower than expected values, then unbiased traders would pay higher prices than biased traders, effectively setting the market price, so prices might appear unbiased.

[15] For instance, in the well-known blue-green taxi problem (for example, Tversky and Kahneman, 1982b), the Bayesian posterior is around .4 but subjects often answer .80 because they ignore the low base rate of one type of taxi.

[16] Recall Arrow's (1982) suggestion cited above. Werner DeBondt and Richard Thaler (1985) also found empirical support for the representativeness prediction that investors do not expect regression in extreme earnings announcements.

before the Great Depression. (Whether such opinions affect market behavior is debatable.) The belief that the future is likely to be representative of the past could cause a failure to anticipate regression effects (Tversky and Kahneman, 1982b): Forgetting about regression, consumers may avoid all Hyatt hotels or DC-10's after an accident involving one of them; or studios might make movie sequels that are consistently unprofitable. The winner's curse in common-value auctions (see John Kagel and Dan Levin, 1986) might be caused by a heuristiclike representativeness. These conjectures, whether plausible or not, illustrate how representativeness bias akin to that observed in the experiments could affect economic outcomes in natural settings.

There are several directions for future experiments. Institutional extensions of these markets, like short selling or a parallel market for information about probabilities, might eliminate biases entirely. Experiments in which other judgment biases could affect markets might be interesting too (for example, myself, George Loewenstein, and Martin Weber, 1987). A program of empirical work, including both experiments and extending experimental results to natural settings, could establish what kinds of irrationality seem to persist under the incentives and learning opportunities present in natural markets. Such data might lead to economic theory that uses evidence of systematic irrationality to make better predictions.

REFERENCES

Arkes, Hal R. and Hammond, Kenneth R., *Judgment and Decision Making: An Interdisciplinary Reader*, Cambridge: Cambridge University Press, 1986.

Arrow, Kenneth, "Risk Perception in Psychology and Economics," *Economic Inquiry*, January 1982, *20*, 1–9.

Berg, Joyce E., Daley, Lane A., Dickhaut, John W. and O'Brien, John R., "Controlling Preferences for Lotteries on Units of Experimental Exchange," *Quarterly Journal of Economics*, May 1986, *101*, 281–306.

Blaug, Mark, *The Methodology of Economics: or How Economics Explain*, Cambridge: Cambridge University Press, 1980.

Camerer, Colin, Loewenstein, George and Weber, Martin, "The Curse of Knowledge in Economic Settings: An Experimental Analysis," Wharton Risk and Decision Processes Center, Working Paper No. 87-09-07, 1987.

Cox, James, Smith, Vernon and Walker, James, "Experimental Development of Sealed-bid Auction Theory: Calibrating Controls for Risk Aversion," *American Economic Review*, May 1985, *75*, 160–65.

DeBondt, Werner F. M. and Thaler, Richard, "Does the Stock Market Overreact?," *Journal of Finance*, July 1985, *40*, 793–805.

Duh, Rong Ruey and Sunder, Shyam, "Incentives, Learning and Processing of Information in a Market Environment: An Examination of the Base-Rate Fallacy," in S. Moriarty, ed., *Laboratory Market Research*, Norman, OK: Center for Economic and Management Research, University of Oklahoma, 1986.

Easley, David and Ledyard, John, "Theories of Price Formation and Exchange in Double Oral Auctions," Social Science Working Paper No. 611, California Institute of Technology, 1986.

Edwards, Ward, "Conservatism in Human Information Processing," in B. Kleinmuntz, ed., *Formal Representation of Human Judgment*, New York: Wiley & Sons, 1968.

Friedman, Daniel, "On the Efficiency of Experimental Double Auction Markets," *American Economic Review*, March 1984, *74*, 60–72.

Friedman, Milton, *Essays in Positive Economics*, Chicago: University of Chicago Press, 1953.

Grether, David M., "Bayes' Rule as a Descriptive Model: The Representativeness Heuristic," *Quarterly Journal of Economics*, November 1980, *95*, 537–57.

Hey, John D., "Search for Rules for Search," *Journal of Economic Behavior and Organization*, March 1982, *3*, 65–81.

______, "Still Searching," *Journal of Economic Behavior and Organization*, March 1987, *8*, 137–44.

Kagel, John H. and Levin, Dan, "The Winner's Curse and Public Information in Common Value Auctions," *American Economic Review*, December 1986, *76*, 894–920.

Kahneman, Daniel, Slovic, Paul and Tversky, Amos, *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press, 1982a.

Kmenta, Jan, *Elements of Econometrics*, New York: Macmillan, 1971.

Plott, Charles R. and Wilde, Louis L., "Professional Diagnosis vs. Self-diagnosis: An Experimental Examination of Some Special Features of Markets with Uncertainty," in V. Smith, ed., *Research in Experimental Economics*, Vol. 2, Greenwich: JAI Press, 1982.

Roth, Alvin, "Toward a Theory of Bargaining: An Experimental Study in Economics," *Science*, May 13, 1983, *220*, 687–91.

Russell, Thomas and Thaler, Richard, "The Relevance of Quasi Rationality in Competitive Markets," *American Economic Review*, December 1985, *75*, 1071–82.

Slovic, Paul, Fischhoff, Baruch and Lichtenstein, Sarah, "Cognitive Processes and Societal Risk Taking," in J. S. Carroll and J. W. Payne, eds., *Cognition and Social Behavior*, Hillsdale, NJ: Erlbaum, 1976.

Smith, Vernon L., "Microeconomic Systems as an Experimental Science," *American Economic Review*, December 1982, *72*, 923–55.

______ **, Suchanek, Gerry and Williams, Arlington,** "Bubbles, Crashes, and Endogeneous Expectations in Experimental Asset Markets," Department of Economics Working Paper No. 86-2, University of Arizona, 1987.

______ **and Williams, Arlington,** "The Boundaries of Competitive Price Theory: Convergence, Expectations, and Transaction Costs," in L. Green and J. Kagel, eds., *Advances in Behavioral Economics*, Vol. 2, Norwood, NJ: Ablex Publishing, in press.

Tversky, Amos and Kahneman, Daniel, (1982a) "Judgments of and by Representativeness," in D. Kahneman, P. Slovic, and A. Tversky, eds., *Judgment under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press.

______ **and** ______, (1982b), "The Evidential Impact of Base Rates," in D. Kahneman, P. Slovic, and A. Tversky, eds., *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press.

Wilson, Robert, "Incentive Efficiency of Double Auctions," *Econometrica*, September 1985, *53*, 1101–115.