Journal of Risk and Uncertainty, 20:3; 219–245, 2000

Do Asset Market Prices Reflect Traders' Judgment Biases?*

ANANDA R. GANGULY
University of Illinois at Urbana-Champaign

JOHN H. KAGEL
Ohio State University

DONALD V. MOSER
Katz Graduate School of Business, University of Pittsburgh, 256 Mervis Hall, Pittsburgh, PA 15260

Abstract

The existence of base rate fallacy (BRF) bias is explored employing: (i) a context treatment with a narrative story applied to asset markets and (ii) an isomorphic abstract setting using balls-and-bingo cages. Probability estimates reflect a BRF bias in both treatments, but is stronger with context. Prices track highest expected dividend values (HEDVs) with context, resulting in strongly biased prices relative to the Bayesian norm when biased traders have HEDVs. In the abstract treatment prices do *not* track HEDVs nearly as closely, resulting in prices closer to the BRF bias only when most traders hold biased beliefs.

Key words: asset markets, base rate fallacy, overreaction, Bayesian norm, experiment

JEL Classification: G12, D84

The effect of individual investors' decision-making on market outcomes has long been of interest to researchers in economics, finance and accounting. Research in psychology and a wide range of applied fields (e.g., accounting, marketing, medicine, and law) has identified a variety of systematic biases in individuals' judgments that could potentially affect market outcomes (see e.g., Kahneman, Slovic and Tversky, 1982; Plous, 1993; or Camerer, 1995, for reviews) The financial press regularly cautions investors against such "mental pitfalls" (e.g., Curran, 1989; Levinson, 1995). Despite the extensive academic literature on individual biases and the regular references to such biases in the financial press, the effects of such biases on market outcomes is still very much open to debate. It is difficult to obtain

*Earlier versions of this paper were presented at seminars at the University of Illinois at Urbana-Champaign, the University of Indiana, the University of Pittsburgh, and Washington University. We received helpful comments on earlier versions of this paper from Jake Birnberg, Jon Davis, Nick Dopoch, Harry Evans, Ron King, and Dan Stone.

unambiguous evidence regarding the effect of such biases on market outcomes using field data and there are important methodological differences between psychologists' studies and the 'methodology of the market.' The latter includes the fact that markets provide financial incentives to perform well and feedback regarding performance, both of which could potentially have a de-biasing effect and are typically absent in psychologists' studies. In addition, there is a strong presumption, in both the economics and finance literature, that the institutional forces underlying market processes can correct or overcome individual judgmental biases (see, for example, Camerer, 1987, 1995). These forces include biased traders being eliminated from markets because of heavy losses or bankruptcy or unbiased traders engaging in higher volumes of market activity, thereby reducing the influence of biased traders on market outcomes.

Over the past decade, researchers have begun to use laboratory asset markets to explore whether individual decision-making biases affect aggregate market outcomes. One area of relatively intense study concerns the effects of underweighting of base rates on asset market prices. Duh and Sunder (1986, 1993), Camerer (1987, 1990) and Anderson and Sunder (1995) presented subjects with tasks in which the probability of future high or low dividends of one-period securities could be estimated by Bayesian updating of prior probabilities with subsequent information. Observed market prices and allocations were compared with Bayesian predictions and several variations of the representativeness heuristic in which traders systematically underweight base rates in favor of current sample information about the likely outcome (Kahneman and Tversky, 1973; Bar-Hillel, 1980, 1990). The general conclusion from these studies is that prices converge close to the Bayesian prediction, but that the direction of any deviation is in favor of a representativeness heuristic (Camerer, 1995).

These experiments all implemented the methodology of the market and some of the institutional factors that presumably mitigate individual biases. However, they differed in one important respect from procedures that psychologists typically employ in identifying individual judgmental biases: The Bayesian judgment problem was operationalized using blandly labeled random devices as stimuli (balls-and-bingo cages) rather than natural stimuli related to the problem at hand (and which, presumably, would characterize information flows in field settings as well). There were other, potentially important limitations to these earlier studies as well. First, individual subject's probability estimates were not measured initially, before they traded in markets, so that the nearness to Bayesian prices reported may have been due, at least partially, to the fact that there never were any significant biases in traders' probability estimates to begin with. Second, biases in individual probability estimates were not monitored over time, leaving the mechanism underlying their elimination unclear: Did market feedback help traders learn to make less biased probability estimates over time or did the market process result in unbiased prices despite the presence of individual biases?[1] Finally, there was no distinction made between markets in which Bayesian traders had the highest expected payoffs compared to when biased traders had the highest expected payoffs. In a double-auction market with a fixed supply of securities and unlimited buying power of traders (the procedures employed in the asset markets), the competitive equilibrium model predicts that price will be driven up to the highest reservation value/expected dividend value. Therefore, competitive forces can be expected to produce biased (unbiased) prices in markets where the highest expected payoffs are biased (unbiased) unless traders with the highest expected payoffs are relatively inactive for one reason or another; e.g., biased traders are less secure in their beliefs (Camerer, 1987).

In the study that follows we address these issues in the following ways: First, we compare judgmental biases and market outcomes using a balls-and-bingo cages procedure versus a short vignette directly related to the problem at hand, with the vignette directly analogous to one psychologists have shown generates a strong representativeness heuristic. Second, in both treatments we first obtain subjects' estimates of expected outcomes prior to trading in the asset market. This two-step procedure, first soliciting probability estimates and then running the asset market, is repeated for all market periods. Third, we analyze the results separately for markets in which biased traders have the highest expected payoff compared to markets where unbiased traders have the highest expected payoffs. Finally, we conduct control sessions requiring subjects to compute trivial probability estimates involving extreme values identical to the Bayesian posterior probabilities employed in several of our treatments and run associated asset markets in the same fashion as in the more complicated main treatments. This is done to assure the reader (and ourselves) that, absent the need to update probabilities using Bayes rule, subjects can reliably report extreme probability outcomes and that asset prices converge to their (extreme) risk neutral expected values.

Our results may be summarized as follows: The vignette produced a large initial bias in probability estimates relative to the Bayesian norm, with only small changes in these estimates over time as a result of market feedback. Market prices closely track highest expected payoff values in this treatment. As a consequence, when biased traders have the highest expected payoffs, we observe strongly biased market prices with no tendency to converge to Bayesian expected values. But when Bayesian traders have the highest expected payoffs, market prices are reasonably close to the Bayesian norm even though only a minority of traders hold these beliefs. The more abstract balls-and-bingo cages procedures also produced relatively large initial biases in probability estimates, although not as large as when using natural stimuli. More importantly, in the abstract setting prices do *not* track highest expected dividend values nearly as closely as they do in the natural setting. As a result prices are only closer to the biased outcome when the overwhelming majority of traders have biased beliefs and highest expected dividend values. In contrast to the context sessions, when only a minority of traders have biased beliefs and highest expected dividend values, prices do not reflect these beliefs, being closer to the Bayesian norm. The mechanism underlying these different outcomes with respect to market prices is that the natural stimuli induce biased traders to have considerably more confidence in their beliefs than does the more abstract balls-and-bingo cages treatment. The control markets show that absent the difficulties of Bayesian calculations, subjects have no problem correctly estimating extreme probability outcomes and that asset prices converge quickly to their risk neutral expected values.

The paper proceeds as follows: Section I details our experimental design and procedures. Results are reported in Section II, with probability estimates reported first followed by results from the corresponding asset markets. A concluding section briefly summarizes our main results and discusses their potential broader implications.

1. Experimental design

Each experimental session had 16 market periods. There were two phases to each market period: a probability estimation phase (Phase I) and an asset market trading phase (Phase II). In the asset market subjects traded securities that had a one-period life and paid a liquidating outcome-dependent dividend of either 500 or 50 francs as a function of the realized state of nature.[2] These states of nature were characterized either in terms of Success or Failure of a fictional firm (we will refer to these procedures as the context treatment) or as Black or Red draws from one of two underlying bingo cages (we will refer to these procedures as the abstract treatment). As shown in Table 1, we employed two different context treatments: sessions in which Bayesian traders had the highest expected dividend value (Market type 1) and sessions in which Biased traders had the highest expected dividend value (Market type 2a). The abstract treatment only had sessions in which Biased traders had the highest expected dividend value, as the context sessions demonstrated that no price bias was to be expected when Bayesian traders had the highest expected dividend value. The control sessions employed an abstract setting replacing the Bayesian updating task with a trivial probability estimation task. In what follows we first describe the procedures used to characterize probabilities under the two treatments—context and abstract. We then describe the procedures

Table 1. Experimental design

	Context-based setting	Abstract setting
Unbiased traders with highest expected payoffs	Market Type 1 (2 sessions)	—
Biased traders with highest expected payoffs	Market Type 2a (2 sessions)	Market Type 2b (3 sessions)
Trivial probability task (as control)	—	Control sessions (2 sessions)

used to elicit subjects' probability estimates and the details of the trading procedures.

1.1. Basic procedures

The context treatment was designed to mimic the "cab problem" known to generate a significant base rate fallacy (BRF) bias in which respondents ignore or underweight base rates in favor of subsequent diagnostic information rather than using Bayes' rule (Kahneman and Tversky, 1973; Bar-Hillel, 1990).[3] Arrow (1982, p. 5) has suggested that such biases typify "very precisely the excessive reaction to current information which seems to characterize all the securities and futures markets." Although natural stimuli were used in the context treatment, there was no misrepresentation of actual probabilities or payoffs; i.e., good Bayesians had all the information necessary to make correct probability estimates and to compute expected payoff values.

Our context treatment dealt with buying and selling single-period securities of a company. Subjects were informed that the prior probability of success for companies whose assets would be traded was either 85% (Market Type 1—designed to result in Bayesian traders having the highest expected dividend value) or 15% (Market Type 2a—designed to result in BRF traders having the highest expected dividend value). In addition, in each period subjects were given a signal in the form of an analyst's prediction of either Success or Failure for the company. Subjects were informed that the analyst had an 80% accuracy rate in identifying firms that succeeded and firms that failed. They were also told that the analyst's accuracy rate was determined by testing him with a large sample of companies, half of which had succeeded and half of which had failed. This was done in order to prevent subjects from mistakenly concluding that the analyst's accuracy rate (80%) was more reliable for successful firms than for firms that failed because the sample included more successful firms.[4] Pilot tests of the context treatment showed that it produced biases similar in magnitude and frequency to those reported for the cab problem. Thus, our procedures insured a strong initial bias to begin with, which is what was intended.

Subjects were informed that they would be making probability estimates and trading securities for 16 companies consisting of 8 randomly selected companies from those the analyst said would succeed and 8 randomly selected companies from those the analyst said would fail. This was operationalized in the form of a bag containing 16 envelopes, one for each market period, placed in the front of the room. On the outside of each envelope was the analyst's prediction, Success or Failure. On the inside was the actual outcome for the company in that period.

Using an equal number of Success and Failure signals gives traders an equal opportunity to learn in both signal cases. It was explicitly announced that this did *not* mean that the sample of 16 firms was drawn from a population containing half successful firms and half firms which failed. Further, the prior probability of success and the analyst's accuracy rate were written on the blackboard as the instructions were read and remained in full view throughout the session so that traders had ready access to the information necessary to produce unbiased estimates. Although we used natural stimuli to frame the problem, the base rate probabilities, the analyst's predictions, and actual outcomes were generated using balls-and-bingo cage procedures identical to those described below for the abstract treatment.

In the abstract treatment the probability estimation task was statistically identical to the task in Market Type 2a except that the 'story' underlying the description of probabilities was replaced by a more abstract balls-and-bingo cages representation (these sessions are referred to as Market Type 2b; see Table 1). At the beginning of each abstract-market session, a demonstration was conducted in which a chip was first drawn from an urn containing 100 chips (15 black and 85 red). This chip was shown to all subjects, and the experimenter wrote its color on a slip of paper, and then sealed the paper in an unmarked envelope. The chip drawn was then returned to the urn. This first draw was analogous to the outcome (success or failure of the company) in the context treatment. A second chip was then drawn from either Urn A or Urn B, depending on the color of the first chip drawn. If the first draw was a red chip, the second draw was made from Urn A containing 80 red and 20 black chips. If the first draw was a black chip, the second draw was made from Urn B containing 80 black and 20 red chips. The experimenter wrote the color of the second chip drawn (either "Red" or "Black") on the envelope containing the slip of paper indicating the color of the first draw. The second chip was then returned to its urn. This second draw was analogous to the analyst's prediction in the context treatment.

After demonstrating this procedure several times, the experimenter explained that a computer program had been used to repeat this balls-and-bingo cages procedure 500 times, thereby generating 500 envelopes with the signal ("Red" or "Black") written on the outside of each envelope and the outcome sealed inside the envelope. Subjects were informed that each of these 500 envelopes was now in one of two boxes sitting before them (all Red-signal envelopes were put in one box, and all Black-signal envelopes were put in the other box). Then one of the subjects randomly drew 8 envelopes from each box which comprised the sample for that session.

In the control markets the probability estimation task was trivial. At the beginning of each session, two urns were shown to subjects: Urn A consisted of 24 black chips and 1 red chip, and Urn B consisted of 24 red chips and 1 black chip, with the contents of these urns publicly announced and posted. At the start of each period a coin was tossed to determine if Urn A or Urn B was selected for that period, subject to the constraint that there would be a total of 8 occurrences of Urn A and 8 occurrences of Urn B. Whether Urn A or Urn B was selected was publicly announced and written on the chalkboard. A single chip was drawn (with replacement) from the urn in question following the close of the asset market.

1.2. Phase I: probability elicitation procedures

In both context and abstract sessions, in each market period one of the 16 envelopes was randomly drawn and the signal on the outside of the envelope was announced. Subjects then estimated the probability that the outcome was Success (Black) or Failure (Red). They responded on either a Success (Black) scale or a Failure (Red) scale specifying a probability from .5 to 1.0 on the appropriate scale. We explained that probabilities from 0 to .5 were crossed off both scales since such values indicated that subjects should be using the other scale.

Although the probability calculations were trivial in the control sessions, we still required subjects to convert the raw chip frequencies—24 out of 25 or 1 out of 25 —into probability estimates. The purpose of making subjects go through this simple conversion was to provide them with the opportunity to exhibit probability biases, such as either ignoring or over-weighting extremely low probability events, as for example predicted by the probability weighting function in prospect theory (Kahneman and Tversky, 1979).

The payment scheme for probability estimates was the same for all markets. Subjects were informed that one market period would be randomly selected at the end of the session and their probability estimates for that period would be compared to the answer given by a statistician. Use of the terms "Bayesian posterior" or "correct answer" was avoided so as not to suggest to subjects that a particular calculation was expected.[5] If their probability estimate was the same as the Bayesian posterior, they received 2000 francs. For every 1% of absolute deviation from the Bayesian posterior, their payment was reduced by 20 francs.[6]

1.3. Phase II: market procedures

After making their probability estimates subjects bought and sold securities in a continuous double-auction market. In each market period, each trader was endowed with two securities that had a one-period life and paid a liquidating outcome-dependent dividend. The liquidating dividend could have one of two values: 500 francs for a Success (Black) outcome or 50 francs for a Failure (Red) outcome. Each trader was also endowed with 11,000 francs in each trading period, with the 11,000 francs subtracted from his or her total francs at the end of each period. The 11,000 franc endowment ensured that any trader could purchase all the securities in the market at their maximum possible dividend value of 500 francs per security.[7]

Subjects were seated at terminals of a computer network and voluntarily traded securities using the MUDA double auction asset market program (Plott, 1991).[8] Each trader could both buy and sell securities. Details of past bids, offers and transactions could be obtained at any point by pressing a key on the computer terminal. Each trading period lasted 4 minutes. At the end of each trading period the actual outcome for that period was announced.

Traders' dollar profits were determined as follows:

$$\text{Profits} = X\left[E_f - R_f + \Sigma S_i - \Sigma B_j + D(O)(E_c - x_s + x_b)\right] \tag{1}$$

where X = dollar-per-franc conversion rate, E_f = initial endowment in francs, R_f = amount of francs subtracted (repaid) at period-end, S_i = selling price of ith security sold, B_j = purchase price of jth security bought, $D(O)$ = dividends per security given outcome O, E_c = initial endowment in securities, x_s = number of securities sold, x_b = number of securities bought.

The computer program prevented traders from selling short (that is, $E_c - x_s + x_b$ could not be negative), and net francs on hand ($E_f + \Sigma S_i - \Sigma B_j$) were not allowed to go below zero. During each trading period, the program recorded all purchases and sales of securities and computed ending balances in francs and securities.

One hundred and eight subjects (mostly graduate business students at the University of Pittsburgh) participated in the study on a voluntary basis. Nine market sessions were conducted, with an average of 12 subjects per market session. As shown in Table 1, there were two sessions each of market Types 1, 2a, and the control markets, and three sessions of Market Type 2b.[9]

2. Predictions

Table 2 reports predictions for Market Types 1 (top panel) and 2 (bottom panel). Columns 1 and 2 show the prior probabilities of success (black) and failure (red) along with the underlying dividend values for the two states of nature. Columns 3 and 4 show the posterior probabilities and expected dividend values following Success (Black) and Failure (Red) signals for Bayesian traders. The last two columns show the corresponding BRF predictions. The latter are based on the extreme assumption that subjects focus solely on the signal information, totally ignoring the prior probabilities of the two outcomes.

In Market Type 1, the strongest contrast between Bayesian and BRF predictions is in the Failure signal case with success probabilities of .59 for Bayesians versus .20 BRF types. This results in an expected difference in dividend payoffs of 176 francs (316 versus 140). In contrast, the probabilities of success are quite similar for the Success signal case (.96 versus .80) and expected payoff differences are much smaller (482 versus 410). As such we would anticipate the Failure signal cases to more easily distinguish between the two hypotheses.[10] Note that in Market Type 1, Bayesian traders have the higher expected dividend values for both Success and Failure signal cases. As such it is not necessary that all traders, or even a majority of traders, hold Bayesian beliefs for prices to reflect Bayesian expected dividend values. Rather, with traders having heterogeneous beliefs it is only necessary that some traders be Bayesian, since with unlimited buying power and a fixed supply of securities there will be excess demand at any price lower than the Bayesian expected payoff. This will put pressure on prices to increase toward the Bayesian expected payoff.

Table 2. Model predictions for Market Types 1, 2a, and 2b: dividend values and probabilities of success

	Outcomes			Signal[a]			Signal[a]	
	(1) Success or Black	(2) Failure or Red	Bayesian predictions	(3) Success or Black	(4) Failure or Red	BRF predictions	(5) Success or Black	(6) Failure or Red
Market Type 1								
Prior probability of success	.85	.15	Bayesian posterior probability of success	.96	.59	BRF-biased predicted probability of success	.80	.20
Dividend value	500	50	Bayesian expected payoff	482	316	BRF-biased expected payoff	410	140
Market Types 2a and 2b								
Prior probability of success or Black	.15	.85	Bayesian posterior probability of S or B	.41	.04	BRF-biased predicted probability of S or B	.80	.20
Dividend value	500	50	Bayesian expected payoff	235	68	BRF-biased expected payoff	410	140

[a] In Market Types 1 and 2a, the signal was the analyst's prediction. In Market Type 2b, the signal was the color of the chip drawn from Urn A or Urn B, as recorded on the covers of the envelopes.

In Market Type 2, the strongest contrast between Bayesian and BRF predictions is in the Success (Black) signal cases with success (Black) probabilities of .41 for Bayesians versus .80 for BRF types. This results in an expected difference in dividend payoffs of 185 francs (235 versus 410). In contrast, the probabilities of success are quite similar for the Failure (Red) signal case (.04 versus .20) with expected payoff differences much smaller as well (68 versus 140). As such we would anticipate that the Success (Black) signals would more easily distinguish between the two hypotheses for Market Type 2. Further, in this market it is the BRF traders who have the highest expected dividend values. As such, even if there is only a minority of BRF types, we would expect dividend values to reflect BRF beliefs. Nevertheless, Bayesian beliefs could emerge in these markets if BRF traders are substantially less active in the market than Bayesian traders, and/or BRF traders have "weak" beliefs to begin with, and/or losses and repeated failures of beliefs to materialize disabuse BRF traders of their incorrect beliefs over time.

In the control markets, given the chip frequencies announced and the dividend values (500 francs for a black chip; 50 francs for a red chip), traders could compute the expected payoffs for the two urns: 482 francs for Urn A and 68 francs for Urn B. Assuming no bias in probability estimates and risk neutrality, asset prices would be expected to converge to one of these two values depending on which urn was chosen.

3. Results

3.1. Probability estimates

In the judgment and decision-making literature on the BRF the medians and modes (but not the means) of subjects' probability estimates are typically reported (e.g., Bar-Hillel, 1980, 1990). Therefore, for comparative purposes, we will report the medians and modes as well as the means.

The probability estimates for the context markets are reported in Figures 1 and 2. The data for the 16 market periods are separated into the 8 Success signal cases and 8 Failure signal cases, and reported in the order of occurrence.[11] For each occurrence the figures report the mean probability estimate with bars indicating plus and minus one standard deviation of the estimates (not standard deviations of the mean, but of the raw data). For each occurrence a Wilcoxon Matched-Pairs Signed-Rank test (Conover, 1971, pp. 206–215) was performed comparing the absolute deviation of each trader's probability estimate from the Bayesian model with the absolute deviation from the BRF model. The model with significantly smaller absolute deviations is considered to be the better predictor for that occurrence.

As shown in Figure 1, traders' median and modal probability estimates for the first Success signal period of Market Type 1 were .80, identical to the BRF prediction. The mean probability estimate of .77 was closer to the BRF prediction

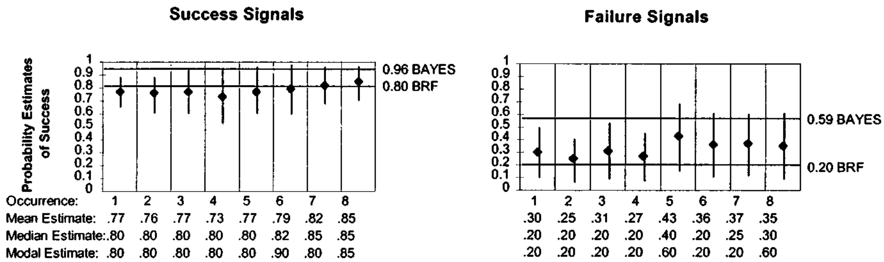


Figure 1. Market Type 1: context treatment. Traders' probability estimates plus or minus one standard deviation. Note: "Bayes" represents the Bayesian probability estimate. "BRF" represents the Base Rate Fallacy probability estimate. A diamond represents the mean probability estimate in a period.

of .80 than to the Bayesian prediction of .96, and the Wilcoxon test indicates that the BRF model was a significantly better predictor of probability estimates ($|z| = 4.16$, $p < .01$). Similarly, for the first Failure signal period of Market Type 1, the median and modal probability estimates of success were .20, identical to the BRF prediction. The mean probability estimate of .30 was closer to the BRF prediction of .20 than to the Bayesian prediction of .59, and the Wilcoxon test indicates that the BRF model was a significantly better predictor of probability estimates ($|z| = 2.73, p < .01$).

The initial probability estimates collected in Market Type 2a also reflected a very strong BRF bias. As shown in Figure 2, the median and modal probability estimates for the first Success signal were .80, identical to the BRF prediction. The mean probability estimate of .76 was closer to the BRF prediction of .80 than to the Bayesian prediction of .41, and the Wilcoxon test indicates that the BRF model was a significantly better predictor of probability estimates ($|z| = 3.57, p < .01$). For the first Failure signal of Market Type 2a, the median probability estimate of .19 and the mean of .24 were closer to the BRF prediction of .20 than to the Bayesian prediction of .04, with the mode (.10) closer to the Bayesian prediction.

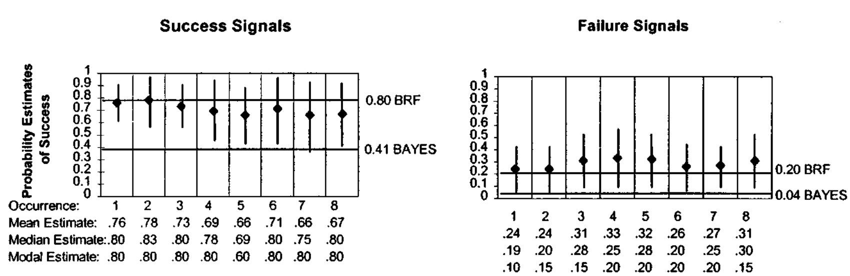


Figure 2. Market Type 2a: context treatment. Traders' probability estimates plus or minus one standard deviation. Note: "Bayes" represents the Bayesian probability estimate. "BRF" represents the Base Rate Fallacy probability estimate. A diamond represents the mean probability estimate in a period.

The Wilcoxon test indicates that the BRF was a significantly better predictor of probability estimates ($|z| = 3.69, p < .01$).

The probability estimate data for the abstract sessions (Market Type 2b) are reported in Figure 3. In the first occurrence of the Black signal, the modal probability estimate is .35, the median is .49, and the mean is .50, all closer to the Bayesian posterior of .41 than to the BRF prediction of .80.[12] This is in marked contrast to the first occurrence of the Success signal in the context treatment which was closer to the BRF prediction. In addition, the Wilcoxon test indicates that the Bayesian model was a significantly better predictor of probability estimates in this period ($|z| = 2.28, p < .05$). Thus, the abstract treatment did not produce an initial BRF bias in the Black signal case. In contrast, the first occurrence in the Red signal case did produce a strong BRF bias. The median (.21), the mode (.32), and the mean (.24) probability estimates were all closer to the BRF prediction of .20 than to the Bayesian prediction of .04. In addition, the Wilcoxon test indicates that the BRF model was a significantly better predictor of probability estimates ($|z| = 4.49, p < .01$). In summary, the abstract treatment produced an individual probability-estimate bias as strong as that in the context treatment in the Red (Failure) signal case, but not in the Black (Success) signal case.

Examination of Figure 1 suggests some reduction in the BRF bias with experience in Market Type 1, particularly for the Failure signal cases, with the mean probability estimate for the last Failure occurrence (.35) no longer being significantly closer to BRF prediction than the Bayesian ($|z| = 1.13, p > .10$). As a more formal test of whether traders' probability estimates became more Bayesian over time, the signed deviation of the traders' mean probability estimate from the Bayesian posterior (Probdev_t) in each market period was regressed against the occurrence number indicated in Figure 1. Separate regressions were run for the Success and Failure signal cases. The results, which are reported in Table 3, are consistent with the results based on examination of the first and last periods in Figure 1: For Market Type 1 in both signal cases, traders' mean probability estimates became more Bayesian over time, starting below the Bayesian prediction

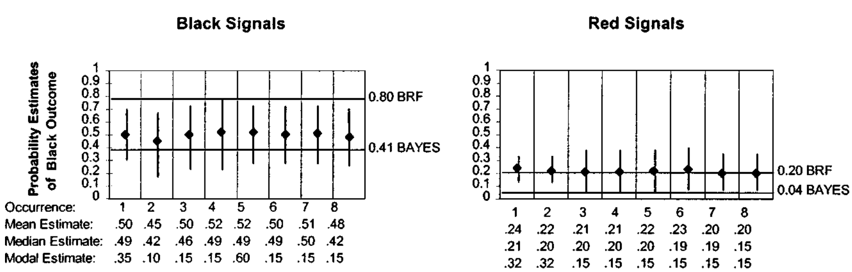


Figure 3. Market Type 2b: abstract treatment. Traders' probability estimates plus or minus one standard deviation. Note: "Bayes" represents the Bayesian probability estimate. "BRF" represents the Base Rate Fallacy probability estimate. A diamond represents the mean probability estimate in a period.

Table 3. Changes in probability judgments over time (Market Types 1, 2a, and 2b)

	a'	b'	$\|t\text{-stat}\|$ for b'	p	R^2
Market Type 1					
Success signal	−.2293	.0114	2.586	.04	.53
Failure signal	−.3276	.0146	1.958	.10	.39
Market Type 2a					
Success signal	.3657	−.0154	3.571	.01	.68
Failure signal	.2190	.0056	1.046	.34	.15
Market Type 2b					
Black signal	.0767	.0019	.485	.65	.04
Red signal	.1918	−.0034	2.017	.09	.40

Signed mean $\text{Probdev}_t = a + bt + \epsilon_t$, where t is the occurrence number.

(negative intercept a'), and moving upwards (positive slope b'). The improvement was significant in the Success signal cases ($|t| = 2.59, p < .05$) and marginally significant in the Failure signal cases ($|t| = 1.96, p = .10$).

Examination of Figure 2 shows that in Market Type 2a there was some reduction in the BRF bias over time in the Success signal case, but no change or even a possible increase in bias over time in the Failure signal case. The regression results reported in Table 3 lead to the same conclusions. In the Success signal cases, the mean probability estimates started above the Bayesian prediction (positive intercept a') and moved significantly downward (negative slope b', $|t| = 3.57, p = .01$) towards the Bayesian posterior, indicating improvement over time. In the Failure signal case, the mean probability estimates started above the Bayesian prediction (positive intercept), but the slope coefficient was not significant ($|t| = 1.05, p > .10$), indicating no significant change in probability estimates over time.

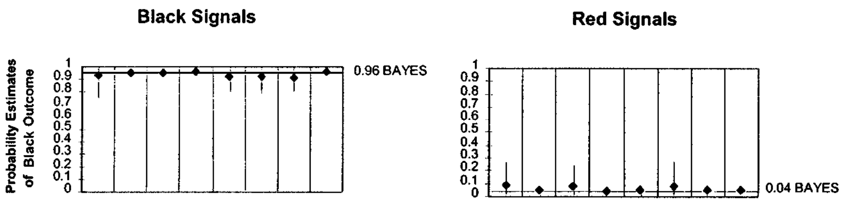
Figure 3 suggests no significant change in probability estimates over time in both Black and Red signal cases in Market Type 2b. The regression results in Table 3 lead to much the same conclusions. In the Black signal cases, the slope coefficient was nowhere close to being significant ($|t| = .49, p > .10$). In the Red signal cases, the mean probability estimates started above the Bayesian prediction, and there was a marginally significant reduction in the slope coefficient ($|t| = 2.02, p < .10$), possibly indicating a small reduction in bias, but still a bias that clearly favored the BRF prediction.

In summary, as expected the BRF prediction is stronger in the context treatment than in the abstract treatment for both signal cases and closer to the BRF to begin with for all but the Black abstract treatment. Further, there is very little movement in the probability predictions over time with the possible exception of the Failure signal case in Market Type 1.

Probability estimates for the control markets are reported in Figure 4. The distribution of probability estimates here is clearly unimodal, with the mode in each period being the correct probability estimate. t-tests for differences between traders' probability estimates and the correct probability indicate that estimates were not significantly different from the correct probability (all ps $> .10$) except for the first occurrence of a Red signal. Thus, the failure of the probability estimates to converge to the extreme outcomes required by the Bayesian prediction for the Success signal case in Market 1 and the Failure/Red signals in Markets 2a and b cannot be attributed to any general tendency on the part of traders to ignore very small probability outcomes as prospect theory suggests (Kahneman and Tversky, 1979). Rather the failure to converge to the Bayesian price prediction results from a clear inability to compute the correct posterior probabilities.

3.2. Asset prices

Since equilibrium is expected to attain only after traders have gained some experience with the market, only the last few trades in each period or the last few periods of each treatment type are typically examined in experimental market studies. Consistent with this approach, our statistical analyses will focus on prices for the last three occurrences of each signal. Figures 5–8 report the mean price for each occurrence, with the bars indicating plus and minus one standard deviation of the prices (not standard deviation of the mean, but of the raw data). For each of the last three occurrences of each signal type, a Wilcoxon Matched-Pairs Signed-Rank test was performed comparing the absolute deviation of each trading price from the Bayesian prediction to the absolute deviations from the Biased prediction (see Table 4). The model with significantly smaller absolute deviations from actual prices for an occurrence is considered to be the better predictor for that occurrence.



Black Signals								
Occurrence:	1	2	3	4	5	6	7	8
Mean Estimate:	93	95	95	96	92	92	91	96
Median Estimate:	96	96	96	96	96	96	96	96
Modal Estimate:	96	96	96	96	96	96	96	96

Red Signals								
Occurrence:	1	2	3	4	5	6	7	8
Mean Estimate:	.09	.05	.08	.04	.05	.08	.05	.05
Median Estimate:	.04	.04	.04	.04	.04	.04	.04	.04
Modal Estimate:	.04	.04	.04	.04	.04	.04	.04	.04

Figure 4. Control markets: probability estimates plus or minus one standard deviation. Note: "Bayes" represents the Bayesian probability estimate. A diamond represents the mean probability estimate in a period. The line representing a standard deviation above and below the mean has been limited to the maximum (1.0) or minimum (zero) possible.

Table 4. Transaction prices (Market Types 1, 2a, and 2b)[a]

	Signal					
	Success or Black			Failure or Red		
Occurrence number	6	7	8	6	7	8
Market Type 1						
Mean Price	448.5	440.5	437.5	361.2	331.1	295.2
Std Dev	45.2	47.1	48.8	89.6	107.9	84.5
Bayes Prediction	482	482	482	316	316	316
Biased Prediction	410	410	410	140	140	140
Better Predictor	n-sig[b]	n-sig	n-sig	Bayes	Bayes	Bayes
| *z*-stat |	.86	.49	.42	5.73	4.73	3.63
p	.39	.63	.68	.0000	.0000	.0003
Market Type 2ab						
Mean Price	398.6	406.9	407.9	266.4	307.8	310.5
Std Dev	39.0	45.7	40.2	32.2	34.9	21.1
Bayes Prediction	235	235	235	68	68	68
Biased Prediction	410	410	410	140	140	140
Better Predictor	BRF	BRF	BRF	BRF	BRF	BRF
| *z*-stat |	6.22	6.42	6.54	6.68	6.74	6.96
p	.0000	.0000	.0000	.0000	.0000	.0000
Market Type 2b						
Mean Price	290.3	280.7	295.5	215.9	199.6	192.9
Std Dev	121.2	106	110.5	70.7	65.2	68.2
Bayes Prediction	235	235	235	68	68	68
Biased Prediction	410	410	410	140	140	140
Better Predictor	n-sig	Bayes	n-sig	BRF	BRF	BRF
| *z*-stat |	.79	2.06	.57	6.26	6.90	5.65
p	.43	.03	.57	.0000	.0000	.0000

[a] Performance of models in last 3 occurrences of each signal case: Descriptive statistics and Wilcoxon Matched-Pairs Signed-Rank test between absolute deviations of each transaction price from the model predictions. The Wilcoxon Matched-Pairs Signed-Rank tests compare (a) the absolute deviation of each transaction price from Bayes Prediction with (b) the absolute deviation of the same transaction price from the Biased prediction. This is done for each of the last three occurrences of each signal. The number of observations included in each period consists of the total number of transactions in that period.

[b] n-sig = non-significant difference in predictive ability of the models.

3.2.1. Market Type 1. In Market Type 1, Bayesian traders had the highest expected payoffs so that the competitive equilibrium prediction is that prices will bid up to the Bayesian expected payoff level in both Success and Failure signal cases, as long as some small percentage of traders hold Bayesian beliefs. As Figure 5 shows, mean prices in the last three Failure periods were much closer to the Bayesian prediction of 316 than to the Biased prediction of 140. The Wilcoxon test statistics

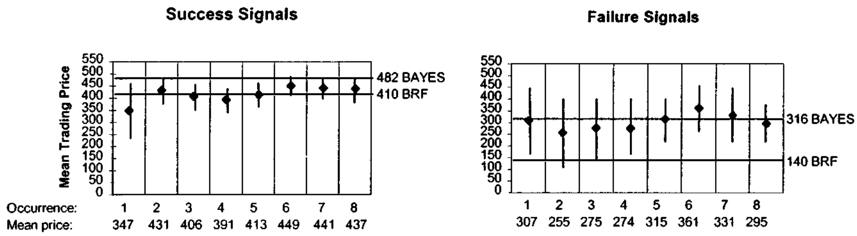


Figure 5. Market Type 1: context treatment, Bayesian expected payoff highest traders' mean trading prices plus or minus one standard deviation. Note: "Bayes" represents the Bayesian price prediction. "BRF" represents the Base Rate Fallacy price prediction. A diamond represents the mean trading price in a period.

confirm the Bayesian model is the better predictor of prices in each of these last three periods (all ps $< .001$). In contrast, prices were in between the Bayesian and BRF price predictions for the last three Success signal cases. Consistent with this observation, the Wilcoxon test indicates that the two models are not significantly different in their predictive accuracy (all ps $> .39$). Part of the reason for the failure to distinguish between Bayesian and BRF price predictions in the Success signal cases is the small spread between the price predictions. A second element, no doubt, has to do with the fact that there was little scope for overly optimistic beliefs relative to the Bayesian outcome in the Success signal case, with much more room for such over-optimism in the Failure signal case. This is supported by the fact that traders with expected dividend values higher than the Bayesian price prediction averaged 5% in the Success signal cases and 15% in the Failure signal cases, with mean deviations for these traders above the Bayesian price prediction of 15.8 and 69.1 francs, respectively (averages taken over the last three occurrences in each case). Thus, there was considerably more pressure for prices to rise above the Bayesian price prediction in the Failure compared to the Success signal case.

For prices to be driven by beliefs, we would expect end of period asset holdings to be positively correlated with traders' expected dividend values based on their stated beliefs. Correlations of this sort were run in each market period. These correlations are positive and significantly different from zero for both Success and Failure signal cases in virtually all periods, confirming that in Market Type 1 trading activity was driven by agents' stated beliefs.

Mean prices had converged to, or were slightly above, the Bayesian price prediction in the last three Failure signal cases. Although the majority of traders clearly did not hold Bayesian beliefs in these three periods (recall Figure 1), the strength of traders' BRF beliefs had weakened substantially, to the point that the Wilcoxan matched-pairs signed rank test no longer showed beliefs to be significantly closer to the BRF norm. However, even in the first several Failure signal cases, when traders' beliefs were clearly closer to the BRF prediction, average market prices were closer to the Bayesian prediction, as the significant minority of traders with Bayesian beliefs drove prices up.[13] These results are entirely consistent with the competitive equilibrium prediction that prices will be driven up to Bayesian levels provided Bayesians have the highest expected dividend value and there is a sizable minority holding these beliefs. We turn next to Market Type 2 in which traders with BRF beliefs have the highest expected dividend values.

3.2.2. Market Type 2a. Recall that in Market Type 2a traders' probability estimates were strongly biased in both signal cases, and these biased traders had the highest expected payoffs in the market. Corresponding to these biased probability estimates, Figure 6 shows that prices in Market Type 2a were either bid up to the BRF price prediction (Success signals) or beyond (Failure signals). Wilcoxon matched-pairs signed-rank tests confirm the statistical significance of these deviations from Bayesian price predictions as the BRF model clearly provides a better price predictor for the last three occurrences of both signal cases (all ps < .001). Further, the price biases reported are considerably greater than any reported in the literature to date, and were even well above the BRF prediction in the Failure signal cases, with *no* indication of learning in favor of the Bayesian outcome in either case.

The driving force behind prices being bid up well above the BRF prediction in the Failure signal cases was the large number of traders with expected dividend values higher than the BRF prediction: 46% of all traders in the Failure signal cases, with mean deviations above the BRF norm for these traders of 85.2 francs (measured over the last three Failure periods). Similar factors underlie prices being driven up to the BRF prediction in the last three Success signal cases, even though mean and median expectations were below the BRF prediction (25% of all traders had expectations above the BRF prediction, with mean deviations above the BRF norm of 61.3 francs).

Within period correlations of final asset holdings and expected dividend values were positive and statistically significant in virtually all periods in both signal cases. Thus, there was no weakening of BRF beliefs in the face of losses in these markets, virtually the only mechanism whereby prices could converge to the Bayesian norm.

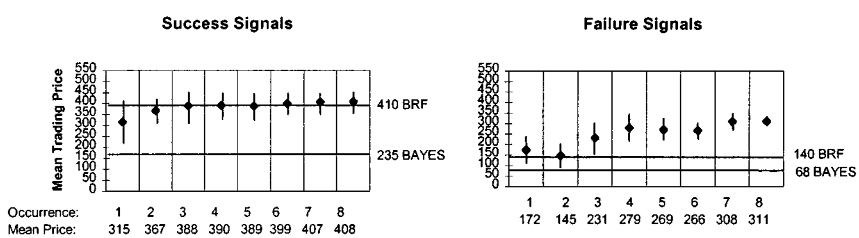


Figure 6. Market Type 2a: context treatment, biased expected payoff highest traders' mean trading prices plus or minus one standard deviation. Note: "Bayes" represents the Bayesian price prediction. "BRF" represents the Base Rate Fallacy price prediction. A diamond represents the mean trading price in a period.

Finally, traders who consistently succumbed to the BRF bias earned considerably less money than traders who were consistently closer to the Bayesian norm: The top three most consistent BRF traders (closer to the BRF norm for 16, 16, and 14 periods, respectively) earned an average trading profit of 114.4 francs (11¢) per period compared to an average of 2,145.3 francs ($2.15) per period for the three most consistent Bayesian traders (11, 10, and 10 periods, respectively). This compares with average earnings of 218.8 francs (22¢) per period for subjects who did no trading at all, collecting the average realized dividends on their initial asset endowment instead. Thus, although earnings for BRF traders were typically not negative, the opportunity costs were substantial. We now turn to the abstract markets (Market Type 2b) in which traders with BRF beliefs have the highest expected dividend values.

3.2.3. Market Type 2b. Probability estimates in Market Type 2b for the Red signal cases were strongly biased towards the BRF; in the last three Red signal cases only 5% of all traders had expected dividend values at or below the Bayesian prediction and only 27% had expected dividend values closer to the Bayesian than the BRF prediction. Given this overwhelming bias, and the fact that biased traders had the highest expected dividend values, it is hardly surprising that asset prices were closer to the BRF prediction in the last three market periods (all ps $< .01$; see Table 4). In fact, mean prices were above the BRF prediction, although not as much above it as in the Failure signal case in Market Type 2a. As in Market Type 2a, these deviations beyond the BRF norm were no doubt driven by traders holding overly optimistic estimates relative to the BRF norm (33% of all traders, with mean expected dividends for these traders 73.7 francs greater than the BRF prediction). However, unlike the corresponding context sessions, correlations between expected dividend values and final asset holding in these Red signal markets were not positive. Rather, the correlation coefficients were close to zero in all periods and not significantly different from zero at the 5% level in any of the Red signal markets. The absence of a positive correlation between expected dividends and final asset holdings suggests that traders were less secure in their beliefs under the abstract (balls-and-bingo cages) procedures. Nevertheless, prices continued to reflect the BRF bias given the overwhelming bias in traders' beliefs and the fact that biased traders had the highest expected dividend values.

In contrast to the Red signal case, in the Black signal case, the probability estimates were significantly closer to the Bayesian price prediction than the BRF prediction; over the last three Black signal cases 41% of all traders had expected dividend values at or below the Bayesian price prediction and 69.5% had beliefs closer to the Bayesian than the BRF prediction. Figure 7 shows that prices in this case were closer to the Bayesian prediction than to the BRF prediction (the Wilcoxon tests reported in Table 4 show prices significantly closer to the Bayesian prediction in the second-to-last occurrence, with neither model clearly outperforming the other in the third-to-last and last occurrences). However, there was a significant minority of traders with beliefs closer to the BRF bias, and a few traders

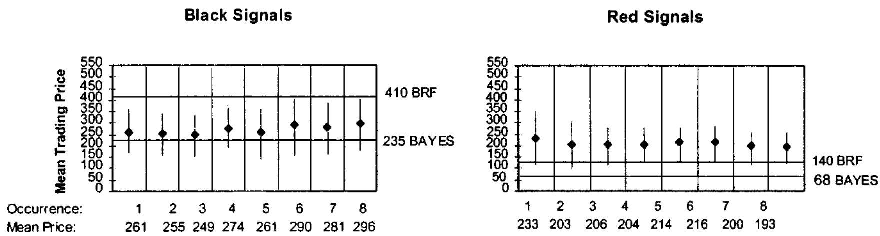


Figure 7. Market Type 2b: abstract treatment, biased expected payoff highest traders' mean trading prices plus or minus one standard deviation. Note: "Bayes" represents the Bayesian price prediction. "BRF" represents the Base Rate Fallacy price prediction. A diamond represents the mean trading price in a period.

with beliefs more optimistic than the BRF bias (8% of all traders), with this minority having the highest expected dividend values. Why didn't this minority drive prices significantly closer to the BRF prediction as the minority of traders with Bayesian beliefs was able to drive prices up to the Bayesian price prediction in the Market Type 1 Failure signal cases? The explanation appears to be that BRF traders had much weaker beliefs in this abstract treatment than in either of the context treatments, as evidenced by *negative* correlations between expected dividend values and final asset holdings in each of the last three Black signal cases (significantly below zero in two of these last three occurrences). Further, these beliefs appear to never have been strongly held as there was only a modest positive correlation between expected payoffs and asset holdings in the first Black signal case (not significantly different from zero; $p > .25$) and a modest negative correlation (again not significantly different from zero; $p > .25$) in the second Black signal case, followed by relatively large (in absolute value) negative correlations in all subsequent periods (significantly different from zero at the 5% level in four of the six periods). In contrast, the signs of the correlations in the corresponding context treatment (Success signals in Market Type 2a) were positive in seven of the eight periods, with significant positive correlations ($p < .05$) in five of the seven periods (the one negative correlation was not significantly different from zero, $p > .25$).

As in Market Type 2a, BRF traders paid a financial price for sticking to their beliefs compared to their Bayesian counterparts. Overall, the three traders with the most consistent BRF beliefs (closer to BRF beliefs in all 16 periods) had average profits of 435.6 francs (44¢) per period compared to the three most consistent Bayesian traders who earned an average profit of 788.9 francs (79¢) per period. This compares to average earnings of 190.6 francs per period (19¢) that would have been earned with no trading at all. The superior performance of these BRF traders relative to the no trade benchmark and compared to the results reported in the corresponding context treatment (Market Type 2b) can be accounted for by the fact that (i) they did not act as strongly on their beliefs as in the context sessions and (ii) prices were closer to the Bayesian price predictions.

3.2.4. Summary. Prices tend to reflect the beliefs of traders with the highest expected dividend values. This is clearest in the context treatments where prices were closer to the highest expected dividend values whether or not these traders were Bayesians or BRF types and whether or not these traders were in the majority (as in Market Type 2a) or the minority (as in Market Type 1). Given the structure of our markets, it is clear that traders with the highest expected dividend values will drive prices to these values provided there are enough traders with these beliefs and traders do not waiver in their beliefs. The wide dispersion in expected dividend values and the positive correlations between final asset holdings and expected dividend values in the context treatments indicate that both of these conditions were satisfied. The net result is that when BRF traders had the highest expected dividend values, asset prices were strongly biased, being close to, or even above, the BRF prediction.

In contrast, in the abstract treatments, we fail to observe positive correlations between expected dividend values and final asset holdings (with significant negative correlations in the Black signal case). The net result is that prices were closer to the Bayesian prediction in the Black signal case, even though there were a significant minority of traders whose beliefs were closer to the BRF bias and these traders had the highest expected dividend values. The negative correlations indicate that these traders did not trade consistently with their expressed beliefs. In contrast, in the Red signal cases, BRF prices emerge despite the weakness of these beliefs because of the sheer number of traders holding these, or more optimistic, beliefs.

One striking (unexplained) feature of the data is that prices are considerably higher than the BRF prediction, which in turn is higher than the Bayesian prediction, in the Failure/Red cases for Market Types 2a and b. As already noted, the cause of these outcomes appears to be the same in both cases: Mean beliefs were squarely at or slightly above the BRF prediction, there was virtually no support for the Bayesian prediction, there were relatively large numbers of traders with overly optimistic beliefs relative to the BRF norm, and BRF traders had the highest expected dividend values.

However, this still leaves unexplained why some traders held beliefs so far above the BRF norm in the Red/Failure cases for Markets 2a and b. Two possibilities immediately suggest themselves. First is a kind of gambler's fallacy, whereby some traders think that after a streak of Black/Success outcomes in markets with Red/Failure signals, that a Red/Failure outcome is due in these markets. This explanation is consistent with the increase in average prices over time in Market Type 2a. However, it is inconsistent with the stable average prices in Market Type 2b. Second, perhaps some of our subjects were confused by the fact that since we were sampling equally from Black/Red (Success/Failure) signal outcomes the posterior probabilities of the two events were equal, thereby generating overly optimistic beliefs for the Red/Failure signal case. However, this explanation is suspect because there were only six out of sixty traders in the Red/Failure signal

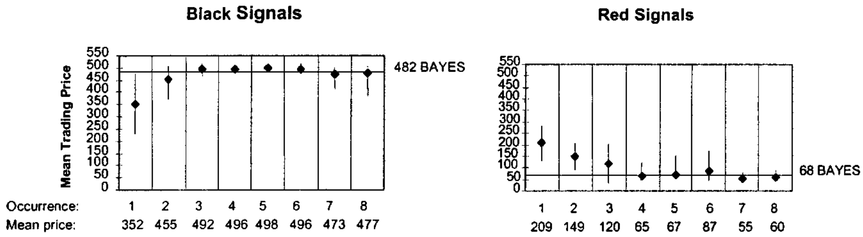


Figure 8. Control markets: mean trading prices plus or minus one standard deviation. Note: "Bayes" represents the Bayesian price prediction. A diamond represents the mean trading price in a period. The line representing a standard deviation above and below the mean has been limited to the maximum (500) or minimum (50) possible.

markets who had mean beliefs between .4 and .6. Further, for all such traders the variance around these beliefs was quite high, averaging .235 per subject indicating that they did not hold such beliefs on a consistent basis, and only two of these six traders were consistently confused in the sense of having mean beliefs between .4 and .6 in the Black/Success markets as well. Thus, we have no obvious explanation as to why some traders held beliefs above the BRF norm in these two cases. But the fact is that they did, with the net effect being that their mistaken beliefs drove market prices since they held the highest expected dividend values.

3.2.5. Control Markets. Figure 8 reports price data for the control markets. With the exception of the first couple of market periods, average market prices were very close to their predicted levels. Although t-statistics comparing mean prices to the competitive equilibrium prediction for each of the last three occurrences of each signal type show that we can reject a null hypothesis of no difference at better than the 10% level in three of the six cases, (i) the differences in these three cases averaged 11.7 francs, which is not economically meaningful and (ii) pooling the data from the last three occurrences of each signal case, mean prices averaged 481 and 68, statistically indistinguishable from the competitive equilibrium prediction of 482 and 68. Thus, the results for the control markets provide evidence that the competitive equilibrium model of price formation, assuming risk neutrality, is a good predictor of price in the asset markets employed in this study, even though the dividend probabilities, and the corresponding expected dividend values, are quite extreme.

Also note that in the control markets there were minimal differences in probability estimates across traders so that it is unlikely that differences in expected dividend values drove the trading activity. Correlations between expected dividend values and final asset holdings confirm this as they were close to zero in absolute value, and not significantly different from zero, in any market period. Nevertheless we observe a significant amount of asset trading, with prices close to the risk neutral expected dividend value. Further, this occurs in markets where all traders had common dividend values. Since it is clearly not differences in beliefs that drive these trades, we are left with the presumption that small differences in traders risk tolerances underlie the trading. If we assume that traders have risk tolerances that range from risk averse to risk neutral to perhaps slightly risk loving, it would be the risk neutral and risk loving types who would be willing to pay the highest prices in both signal cases. Further, as we have seen, these asset markets do not require a majority of traders with these preferences, but only a strong minority, for something approximating the risk neutral competitive equilibrium price to emerge.

4. Discussion

We have looked at the effect of an individual judgmental bias, neglect of base rates, on prices in asset markets. In contrast to traditional psychological studies that have identified this base rate neglect, we have employed the 'methodology of the market,' examining behavior with financial incentives and feedback from previous choices. In contrast with earlier experiments investigating the market impact of base rate neglect, we have compared behavior using natural stimuli to characterize probabilities (a methodology commonly employed by psychologists) versus blandly labeled random devices (a methodology preferred by economists). We have also distinguished between cases where Bayesians or BRF types had the highest expected dividend values, as under the asset market rules employed one would expect traders with the highest expected dividend values to dictate market prices. As such it is considerably more challenging for Bayesian prices to emerge when BRF types have the highest expected dividend values. Finally, we have monitored traders' beliefs throughout to help determine the mechanism underlying the prices that emerge in the market, and have conducted control sessions with equally extreme expected dividend values but a trivial probability estimation task.

Important differences in outcomes reported occurred as a function of whether natural stimuli (context) are used to characterize probabilities versus the abstract stimuli economists' prefer. With natural stimuli there were strong BRF biases in probability estimates in all treatments to begin with, and these biases remained relatively strong throughout. In these markets, when Bayesian traders had the highest expected dividend values (Market Type 1), and there was a significant spread between Bayesian and BRF prices, mean prices closely approximated the Bayesian price prediction. However, when BRF traders had the highest expected dividend values, and there was a significant spread in price predictions, mean prices closely approximated the BRF prediction. This resulted in substantially larger price biases than any reported in the literature to date. Further, agents traded in accordance with their beliefs as evidenced by the strong positive correlations between final asset holdings and expected dividend values reported throughout."

In contrast, with abstract stimuli traders' probability estimates were less biased.[14] Further, in markets with a significant spread between BRF and Bayesian prices, prices were close to the Bayesian prediction, although biased in favor of the BRF. This corresponds to the outcome reported in earlier studies of the impact of base rate neglect on asset market prices (see Camerer, 1995, for a review of earlier results). Why didn't biased traders, although in the minority, still strongly bias market prices in this treatment as Bayesian traders did when they were in the minority and had the highest expected dividend values in the context treatment (Market Type 1)? Apparently biased traders in the abstract treatment did not have strongly held beliefs as evidenced by significant negative correlations between final asset holdings and expected dividend values. Further, even with traders having weakly held beliefs in the abstract treatment, when they were overwhelmingly biased in favor of the BRF, and these traders had the highest expected dividend values (the Red signal outcomes in Market Type 2b), asset market prices were at or above the BRF prediction, and significantly above the Bayesian price prediction. Finally, the control markets indicate that none of the outcomes reported here can be attributed to the extreme expected dividend values employed in these markets or to any strange elements in our procedures or subject population, as prices converged rapidly to expected dividend values. This clearly suggests that the deviations from 'rational' prices reported in the markets requiring Bayesian updating resulted from the complicated Bayesian updating task that traders faced.

Camerer (1995), in reviewing the differences in methodology between economists and psychologists in studying base rate neglect, suggests that one reason economists prefer abstract stimuli is that they are inclined to believe that reasoning about bingo cages and corresponding natural stimuli are similar, so that they serve as good substitutes for each other. The results reported here are consistent with a growing body of results reported in the psychology literature demonstrating that this is clearly *not* the case.[15] This raises the really interesting question of whether or not 'experts'—professional traders in security markets—behave in the same way. Unfortunately, we know of no direct evidence on this point.

The one element missing from our laboratory markets that might be expected to have an important impact on market outcomes was the inability of traders to sell short. When BRF traders have the highest expected dividend values, short sales by Bayesian traders, although in the minority, might lessen the upward pressure on prices. However, experiments with short selling in this and a similar setting (asset market bubbles) have found little effect (Camerer, 1990, King et al., 1993). Nevertheless, the introduction of short sales would no doubt result in increased losses by biased traders as they could buy more and suffer larger losses on average than in the absence of short sales. This might eventually correct biased traders base rate neglect through the shock of larger losses and/or bankruptcy and elimination of biased traders.[16] The latter mechanism presumes, however, that these bankrupt traders will not be replaced by equal numbers of biased and unbiased traders, or that the adverse selection that biased traders face will eventually reduce their numbers to the point where they have no significant effect on market prices. The impact of these longer term effects is, of course, an empirical question that goes well beyond the scope of the present study. Our results do, however, suggest that markets can be wildly out of equilibrium for rather extended periods of time as a result of base rate neglect and are entirely consistent with the belief, held by many who study securities and futures markets, that they often overreact to current information (e.g., De Bondt and Thaler, 1986; Chopra, Lakonishok and Ritter, 1992; Barberis, Schleifer, and Vishney, 1998).

Acknowledgments

This research has been partially supported by an Arthur Andersen Foundation doctoral dissertation grant to Ananda Ganguly, the National Science Foundation through a grant to John Kagel and the Katz Graduate School of Business through a research award to Don Moser.

Notes

1. Camerer (1987, 1990) did measure traders' probability-estimate biases, but found only very small pre-trading biases relative to the biases reported in most previous BRF Studies. However, the measurement was made only at the beginning of a market session and not at the beginning of each period, so it was not possible to test for possible improvement in probability estimates over time.
2. Francs served as the experimental currency with francs converted to U.S. dollars at a rate of 1000 francs = $1.00.
3. One version of the "cab problem" is as follows:

   Two cab companies operate in the same city, the Blue and Green (according to the color of the cab they run). Eighty-five percent of the cabs in the city are Blue, and 15 percent are Green.

   A cab was involved in a hit-and-run accident at night in which a pedestrian was run over. An eyewitness identified the cab as a Green cab. The court tested the witness's ability to distinguish between Blue and Green cabs under nighttime visibility conditions. It found that the witness was correct 80 percent of the time but confused it with the other color 20 percent of the time.

   What is the probability that the hit-and-run cab was Green?

   The median and modal response for this problem is typically .80, while the Bayesian posterior is .41 (Bar-Hillel, 1990). Thus, subjects appear to underweight or even ignore the prior probability and anchor on the witness's 80% accuracy rate.
4. Our instructions also incorporated other features identified in the literature as ways to ensure that Bayesian updating is the normatively appropriate way for subjects to respond to the task. In particular, the criticism that in some word problems the Bayesian likelihood ratio may not be independent of either base rates or prior probabilities (Birnbaum, 1983) was addressed by stating directly that the analyst "based his judgment entirely on the output of a computerized analysis package that exclusively uses the accounts and other company-specific data of each company as input and produces a measure of the project's (and therefore the company's) success potential." In addition, the issue sometimes raised in word problems that base rates and prior probabilities are not necessarily the same thing (base rates help people set prior probabilities but need not be identical to prior probabilities; Koehler, 1996; Cohen, 1981) was avoided by stating the base rates in terms of prior probabilities as follows: "If you had no access to more specific information about the

company, you would have estimated the chance of success for each project to be 15% (85%), which is the normal chance of success for similar projects and similar companies." This wording also insures that the prior information is "causally relevant" to the task at hand (Cohen, 1981).

5. Grether (1978, 1980) criticizes payment procedures of the sort employed here on the grounds that with an incentive to behave as experts, subjects may or may not interpret this as an incentive to give the right answer. Note, however, that (i) the same procedures were used in the control market and, as will be shown, resulted in no systematic distortion for the probability estimates and (ii) this concern applies only to payment for subjects' probability estimates and not to their actions in the asset market.
6. We did not use a more complicated 'proper' scoring rule, such as the quadratic scoring rule explicitly designed to elicit incentive compatible probability estimates since explaining such a rule is time consuming and the experiment was already rather long. Our payment scheme introduces no obvious biases into subjects' responses and appears not to have resulted in any bias as evidenced by the results from the control sessions.
7. In the first session of each of Market Types 1 and 2a, the initial endowment was 2,300 francs per period, which did not permit any single trader to buy all the securities available for sale. Although results indicated no significant wealth constraint, the initial endowment was raised to 11,000 francs in subsequent sessions. In these initial sessions the amount of the dividend also depended on one of two randomly assigned trader types. Dropping the different trade types simplified procedures with no apparent effect on behavior.
8. The first session of each of market type was conducted manually. We observe no material differences within treatments between manual and computerized sessions.
9. We conducted a third session of Market Type 2b because, although the probability estimates were very similar for the first two sessions, prices were consistently lower in the first session than in the second session. In the third session probability estimates were again similar to those in the first two sessions, and prices were very similar to those in the second session. We include all three sessions in the results reported for Market Type 2b. Excluding the first session (the potential outlier) does not affect the conclusions reported in the paper.
10. We intentionally selected prior probabilities that were relatively extreme (i.e., 85% chance of success and 15% chance of failure) to increase the spread between the Bayesian and BRF price predictions, and thereby provide a clearer test of which model better predicts prices and probabilities. Nevertheless, the nature of experimental task is such that if the spread for one signal type is large, the spread for the other signal type must be small.
11. The data from the different market sessions within each market type have been combined by signal and by occurrence. For example, data from the first occurrence of the success signal in session 1 of Market Type 1 have been combined with the data from the first occurrence of the success signal in session 2 of Market Type 1.
12. There are actually three modes for the first occurrence of a black signal in Market Type 2b, two of which (.35 and .60) are closer to the Bayesian posterior, and one (.65) which is closer to the BRF prediction.
13. The Wilcoxon test shows prices to be significantly closer to the Bayesian prediction in the first Failure signal case ($p < .05$) even though beliefs were significantly closer to the BRF prediction.
14. It has been suggested that since in the abstract treatment subjects saw a demonstration of how the envelopes were constructed, and in the context treatment they were only told how the envelopes were constructed, this difference in treatments might be an additional factor promoting forecasts closer to the true conditional probabilities in the abstract case. We do not necessarily dispute this claim. However, (i) it hardly seems possible to do a comparable demonstration in the context treatment and (ii) *if* this claim is correct, then our abstract treatment might better be viewed as "abstract plus," with the only implication being that we may not yet fully understand the full set of factors underlying the differences between treatments reported.
15. For entree into this literature see the fascinating paper by Windschitl and Weber (1998) and the many references cited therein.

16. However, Shefrin and Statman (1994) construct a model in which "noisetraders" employing non-Bayesian decision rules similar to the ones our traders employ survive in the long run.

References

Anderson, Matthew, and Shyam Sunder. (1995). "Professional Traders as Intuitive Bayesian," *Organizational Behavior and Human Decision Processes* 64, 185–202.

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

Barberis, Nicholas, Andrei Schleifer, and Robert Vishney. (1998). "A Model of Investor Sentiment," *Journal of Financial Economics* 49, 307–343.

Bar-Hillel, Maya. (1980). "The Base-Rate Fallacy in Probability Judgments," *Acta Psychologica* 44, 211–233.

Bar-Hillel, Maya. (1990). "Back to Base Rates." In Robin M. Hogarth (ed.), *Insights in Decision Making: A Tribute to Hillel J. Einhorn*, pp. 200–216. Chicago: University of Chicago Press.

Birnbaum, Michael. (1983). "Base Rates in Bayesian Inference: Signal Detection Analysis of the Cab Problem," *American Journal of Psychology* 96, 85–94.

Camerer, Colin. (1987). "Do Biases in Probability Judgment Matter in Markets? Experimental Evidence," *American Economic Review* 77, 981–997.

Camerer, Colin. (1990). "Do Markets Correct Biases in Probability Judgment? Evidence from Market Experiments." In L. Green and John H. Kagel (eds.), *Advances in Behavioral Economics*, pp. 126–172. Norwood, NJ: Ablex.

Camerer, Colin. (1995). "Individual Decision Making." In John H. Kagel and Alvin E. Roth (eds.), *Handbook of Experimental Economics*, pp. 587–703. Princeton, NJ: Princeton University Press.

Chopra, Navin, Josef Lakonishok, and Jay R. Ritter. (1992). "Do Stocks Overreact?" *Journal of Financial Economics* 31, 235–268.

Cohen, L. Jonathan. (1981). "Can Human Irrationality Be Experimentally Demonstrated?" *Behavioral and Brain Sciences* 4, 317–370.

Conover, W. J. (1971). *Practical Nonparametric Statistics*. New York: Wiley and Sons.

Curran, J. J. (1989). "Why Investors Misjudge the Odds," *Fortune*: *1989 Investor's Guide* 85–97.

De Bondt, Werner F. M., and Richard Thaler. (1986). "Does the Stock Market Overreact?" *The Journal of Finance* XL, 793–807.

Duh, Rong Ruey, and Shyam S. Sunder. (1986). "Incentives, Learning and Processing of Information in a Market Environment: An Examination of the Base-Rate Fallacy." In S. Moriarty (ed.), *Laboratory Market Research*, pp. 50–79. Norman, OK: University of Oklahoma, Center for Economic and Management Research.

Duh, Rong Ruey, and Shyam S. Sunder. (1993). "Economic Agent as an Intuitive Bayesian: Experimental Evidence," *Cuadernos Economicos* 54, 101–127.

Grether, David M. (1978). "Recent Psychological Studies of Behavior Under Uncertainty," *American Economic Review* 68, 70–77.

Grether, David M. (1980). "Bayes' Rule as a Descriptive Model: The Representativeness Heuristic," *Quarterly Journal of Economics* XCV, 537–557.

Kahneman, Daniel, and Amos Tversky. (1973). "On the Psychology of Prediction," *Psychological Review* 237–251.

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

Kahneman, Daniel, Paul Slovic, and Amos Tversky (eds.). (1982). *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

King, Ronald R., Vernon L. Smith, Arlington W. Williams, and Mark Van Boening. (1993). "The Robustness of Bubbles and Crashes in Experimental Stock Markets." In R. H. Day and P. Chen (eds.), *Nonlinear Dynamics and Evolutionary Economics*, pp. 183–200. Oxford: Oxford University Press.

Koehler, Jonathan J. (1996). "The Base Rate Fallacy Reconsidered: Descriptive, Normative, and Methodological Challenge," *Behavioral and Brain Sciences* 19, 1–53.

Levinson, M. (1995). "Dismal Science Grabs a Couch," *Newsweek* April 10, 41–42.

Plott, Charles R. (1991). "Social Science Working Paper 783, California Institute of Technology.

Plous, Scott. (1993). *The Psychology of Judgment and Decision Making*. New York: McGraw-Hill, Inc.

Shefrin, Hersh, and Meir Statman. (1994). "Behavioral Capital Asset Pricing Theory," *Journal of Financial and Quantitative Analysis* 27, 323–349.

Windschitl, Paul D. and Elke U. Weber. (1998). "The Interpretation of 'Likely' Depends on the Context, But '70%' is 70%—Right? The Associative and Rule-Based Processing of Uncertainty Information," Working Paper, Ohio State University.