Herding and Contrarianism: A Matter of Preference?

Chad Kendall*

June 1, 2018

Abstract

Herding and contrarian strategies produce informational inefficiencies when investors ignore private information, instead following or bucking past trends. In a simple market model, I show theoretically that investors with prospect theory preferences generically follow strategies that are observationally equivalent to herding or contrarianism, but which are actually trend-independent. I confirm the theory’s predictions in a laboratory experiment designed to rule out other sources of these behaviors, and find that approximately half of subjects exhibit herding-like behavior. Finally, I simulate decisions of the modal subject when facing actual market returns to demonstrate that this behavior extends to more general settings.

1 Introduction

Herding and momentum strategies, and their antithesis, contrarian strategies, interest financial economists because of their implications for the informational efficiency of market prices, which are in turn important due to potential effects on the real economy.1 What drives investors to use these strategies, and, in particular, to choose one over the other?2 The most basic explanation for herding/momentum (following the trend) is simply imitation, a form

---

1Firm prices are important both for the allocation of capital and as signals to managers about their underlying investment decisions. See Bond, Edmans, and Goldstein (2012) for a recent review of the literature on the real effects of secondary financial markets.

2By a ‘momentum strategy’, I’m referring to the strategy of buying if prices have risen and selling if they’ve fallen, sometimes referred to as positive feedback trading (De Long (1990)), and not to be confused with the cross-sectional strategy of shorting losers and going long winners (which goes by the same name). Momentum and contrarian strategies are frequently discussed by practitioners and a small academic literature suggests different types of individuals or firms pursue each (see Grinblatt, Titman, and Wermers (1995), Grinblatt and Keloharju (2000), Baltzer, Jank, and Smajlbegovic (2015), and Grinblatt et al. (2016)).
of ‘herd mentality’ (Mackay (1841)), but behavioral explanations such as extrapolative expectations (De Long et al. (1990), Greenwood and Shleifer (2014), Barberis, Greenwood, Jin, and Shleifer (2015,2016)) and rational explanations based on information externalities (Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992)) have also been theorized.\(^3\) Similarly, contrarian strategies (going against the trend) have rational explanations given a suitable informational environment (Avery and Zemsky (1998), and Park and Sabourian (2011)). In all of these explanations, the observation (or inference) of previous investor actions and/or prices is critical and, in fact, the strategies are typically defined in terms of behaviors that depend upon past trends (Avery and Zemsky (1998)). In this paper, I suggest instead that what appear to be strategies that depend on past trends may in fact actually be independent of historical data. Rather, due to the nature of the relationship between past price trends and subsequent expected returns, these strategies may instead simply be driven by preferences over future returns.

I make my case for a preference-based explanation using both theory and a laboratory experiment. In the theoretical contribution, I revisit a standard model of trading with asymmetric information, that of Glosten and Milgrom (1985). In this model, herding and contrarianism are impossible with risk-neutral investors (Avery and Zemsky (1998)), but both have nevertheless been frequently observed in past experiments (Cipriani and Guarino (2005,2009) and Drehmann, Oechssler, and Roider (2005)). To resolve this apparent contradiction, I consider the behavior of an investor with cumulative prospect theory (CPT) preferences (Kahneman and Tversky (1992)). I show that such an investor generically (for almost all preference parameters) either buys or sells the tradable asset at extreme prices, independently of her private information. This herding-like and contrarian-like behavior is not, however, driven by a change in prices: these investors make the same decisions regardless of the history of events.\(^4\) Instead, decisions are driven by the fact that extreme prices imply highly skewed returns, which drive CPT investors’ decisions (Barberis and Huang (2008)).

Given the theoretical predictions, I conduct a new laboratory experiment that, unlike the seminal experiments of Cipriani and Guarino (2005) and Drehmann, Oechssler, and Roider (2005), controls for other possible explanations of behavior. I provide robust evidence of herding and contrarian-like behaviors even when traditional explanations cannot possibly be

\(^3\)Momentum and herding produce similar trade patterns, but herding is often defined in the context of asymmetric information environments: an investor herds if she initially trades in a way that reveals her private information, but switches her trade direction to follow the crowd after observing others’ trades (Avery and Zemsky (1998), Park and Sabourian (2011)). See Devenow and Welch (1996) and Hirshleifer and Teoh (2003) for surveys of herding in financial markets.

\(^4\)Precisely because this behavior does not depend on historical events, I refrain from calling it herding or contrarian behavior, instead using the terms ‘herding-like’ and ‘contrarian-like’. See Section 2.6 for further discussion.
their cause. Thus, I suggest that CPT preferences likely contribute to the use of what appear to be trend-dependent strategies, and furthermore provide a unifying explanation for both trend-following (herding-like) and trend-bucking (contrarian-like) behaviors.

In the theoretical model, investors arrive sequentially to a market, trading a single, binary-valued asset with a market maker who posts separate bid and ask prices. Each investor, after receiving a private, binary signal about the asset’s fundamental value, may buy or sell a single unit of the asset (or abstain from trading). The standard result with risk-neutral, expected utility investors (Avery and Zemsky (1998)) is that each investor trades according to her private information - buying with a favorable signal and selling otherwise. With CPT preferences, I show that an investor ignores her private information at extreme prices, instead trading in a single direction. The direction that she prefers to trade depends in a very simple way upon the difference between the extent to which she weights probabilities and her utility curvature.

The reference-dependent nature of CPT preferences implies preferences over skewness that drive trading behavior. At a high price, the expected return from buying the asset is negatively skewed: it is very likely to return a small positive amount (when the asset is valuable), but occasionally results in a large loss (when it is not). Selling the asset instead results in a positively skewed return. As has been recognized (Barberis and Huang (2008)), the overweighting of small probabilities in CPT generates a preference for positive skewness. Thus, when probability weighting dominates utility curvature, we observe contrarian-like behavior: selling at high prices (regardless of what led to the high price). On the other hand, when utility curvature dominates, an investor exhibits a preference for negative skewness, leading to herding-like behavior: buying at high prices. The opposing effects of probability weighting and utility curvature have received relatively little attention in the literature, but here are clearly illustrated through the resulting simple analytic expressions that govern behavior.5

Guided by the theoretical predictions, I conduct a laboratory experiment which serves two purposes. First, it demonstrates the extent to which preferences are responsible for herding and contrarian-like behaviors (something which would be very difficult to determine in naturally-occurring settings). In both of the two treatments, I have subjects make decisions in an individual decision-making environment that rules out any social cause of behavior, including simply following others. In one of the treatments, I also provide subjects with the correct Bayesian beliefs corresponding to their information sets (including both public

5Barberis (2012), in his model of casino gambling, demonstrates the counteracting forces numerically. A literature on neuroeconomics (Glimcher and Fehr (2013)) has also recognized the offsetting effects of probability weighting and utility curvature.
and private information). The difference across treatments allows me to assess the extent to which belief-based errors, such as overextrapolation, drive behavior. In the aggregate, I find strong evidence of herding and contrarian-like behaviors in both treatments, suggesting that preferences are overwhelmingly responsible for behavior.

The second purpose of the experiment is to assess the type of preferences which best describe individual behavior. Across treatments, approximately two thirds of subjects make decisions that are better characterized by prospect theory than expected utility. Among these, three quarters exhibit behavior that is observationally equivalent to herding, with the remainder exhibiting contrarian behavior. Among the third of subjects whose behavior is consistent with expected utility, almost all display risk-neutral behavior (which is technically consistent with either theory), but several abstain from trading at extreme prices. The latter behavior cannot be reconciled with prospect theory preferences (but is consistent with a high level of risk-aversion), demonstrating that the theory is falsifiable.

Lastly, taking the estimated preferences of the modal subject in the experiment (who exhibits herding-like behavior), I ask how an investor with these preferences would behave when facing actual market returns. Because market returns become more (less) skewed as prices rise (fall), a fact first documented by Chen, Hong and Stein (2011), if preferences over skewness drive behavior as in the theory and experiment, we should expect this investor to be willing to purchase an asset after rising prices even with negative private information, and to sell with positive information after falling prices. A simulation exercise confirms this conjecture, demonstrating that herding-like behavior can be driven by preferences in environments more general than the simple, binary asset value setup of the theory and experiment.

The seminal paper on herding in financial markets is that of Avery and Zemsky (1998) who show that herding and contrarian behavior are impossible under standard preference assumptions (unless additional sources of uncertainty are added to the model). More recent theoretical papers have studied whether or not non-expected utility preferences can generate these behaviors. Qin (2015) shows that regret aversion generates herding-like, but not contrarian-like, behavior. Boortz (2016) builds on Ford (2013), to show that ambiguity can generate both behaviors, but only with preferences that vary with the state. In contrast to these models, I show that static CPT preferences generate both types of behavior, consistent with previous experimental evidence.

The experiment contributes to the experimental literature on herding in financial markets (see Cipriani and Guarino (2005), Drehmann, Oechssler, and Roider (2005), Cipriani and Guarino (2008), Cipriani and Guarino (2009), Park and Sgroi (2012), and Bisiere, Decamps, and Lovo (2015)). I follow the methodology of Bisiere, Decamps, and Lovo (2015)) in con-
verting the game-theoretic model to an individual decision-making task in order to rule out social causes of herding and contrarianism (including imitation and strategic ambiguity). An important difference is that I allow for both prospect theory and expected utility preferences when attempting to explain the experimental data, whereas they consider expected utility only.

This paper also contributes to the growing literature that applies CPT preferences to understanding investor behavior in financial markets. Several papers study the disposition effect, the tendency to sell recent winners but hang onto recent losers (Barberis and Xiong (2009), Barberis and Xiong (2012), Ingersoll and Jin (2013), Li and Yang (2012), Meng and Weng (2016)). Barberis and Huang (2008) study the pricing of securities when investors have CPT preferences, and Barberis, Huang, and Thaler (2006) use loss aversion to explain stock market non-participation. Levy, De Giorgi, and Hens (2012) and Ingersoll (2016) study CAPM with prospect theory. Although not directly related to financial markets, Barberis (2012) is a closely related paper that shows how CPT preferences can explain the popularity of casino gambling.

2 Theory

2.1 Model

The model is a sequential trading model based on that of Glosten and Milgrom (1985). In each period $t = 1, 2, \ldots, T$, a single new investor arrives to the market to trade an asset of unknown value, $V \in \{0, 1\}$. I denote the initial prior that the asset is worth 1 by $p_1 \in (0, 1)$. Upon arrival, an investor may either buy or sell short a single unit, or not trade, $a_t \in \{\text{buy, sell, NT}\}$. After making her decision, the investor leaves the market. All trades are with a risk-neutral market maker who is assumed to face perfect competition, earning zero profits in expectation.\footnote{This assumption is standard in the literature and follows Glosten and Milgrom (1985).} The market maker incorporates the information provided in the current order in setting prices. Specifically, he posts an ask price, $A_t$, at which he is willing to sell a unit of stock and a bid price, $B_t$, at which he is willing to buy a unit. When the asset value is realized at $T$, investors who purchased the asset at time $t$ receive a payoff of $V - A_t$ and those who sold receive a payoff of $B_t - V$ (there is no discounting).

All market participants observe the complete history of trades and prices, denoted $H_t = (a_1, a_2, \ldots, a_{t-1}) \cup (A_1, A_2, \ldots, A_{t-1}) \cup (B_1, B_2, \ldots, B_{t-1})$.

Investors are one of three types: risk-neutral, prospect theory, or uninformed investors. Uninformed investors, who arrive with probability $1 - \mu$, $\mu \in (0, 1)$, trade for exogenous
reasons and are equally likely to buy or sell. Risk-neutral investors have standard risk-neutral expected utility preferences and arrive with probability, $\mu \gamma$, $\gamma \in (0, 1)$. Finally, prospect theory investors have the CPT preferences of Kahneman and Tversky (1992) (see Section 2.4) and arrive with the remaining probability, $\mu (1 - \gamma)$. The risk-neutral and CPT investors receive private information upon arrival to the market: a binary signal, $s_t \in \{0, 1\}$, which has the correct realization with probability $q = Pr(s_t = 1|V = 1) = Pr(s_t = 0|V = 0) \in (\frac{1}{2}, 1)$. All signals are independent conditional on $V$. I refer to $s_t = 1$ as a favorable signal, and $s_t = 0$ as unfavorable.

Although I focus on the behavior of the CPT investors, I include risk-neutral investors in the model primarily because we should expect heterogeneous preferences in any population. In addition, including risk-neutral investors allows (partial) information to be revealed by every trade, ensuring dynamic price paths as in real markets, whereas, as I show, prices generally stagnate in their absence. Risk-neutral investors, however, are not necessary for the main conclusions of the model.

### 2.2 Solution Concept

Being a game of asymmetric information, the solution concept is Perfect Bayesian Equilibrium. An equilibrium consists of a specification of the strategies of the risk-neutral and CPT investors, along with the bid and ask prices of the market maker, which depend upon his beliefs about investors’ strategies. As usual, these beliefs, which are pinned down at every history due to the presence of the uniformed investors, must be correct in equilibrium. Strategies are functions of the complete history of prices and trades, as well as one’s private signal, to an action: buy, sell, or not trade. As these details are standard, I omit formal definitions.

### 2.3 Risk-Neutral and Uninformed Investors

The roles of the risk-neutral and uninformed investors, as well as the market maker, are standard. I describe them first before discussing the more novel behavior of the CPT investors.

In each period, the market maker posts separate bid and ask prices given by $B_t = Pr(V = 1|H_t, a_t = sell)$ and $A_t = Pr(V = 1|H_t, a_t = buy)$, respectively, a consequence of the assumption of perfect competition. Intuitively, the ask price exceeds the public belief, denoted $p_t = Pr(V = 1|H_t)$, because a buy decision reflects favorable private information, $s_t = 1$.

---

7Bruhin, Fehr-Duda, and Epper (2010), for example, argue that both expected utility and prospect theory preferences should be included in applied theoretical work because they find a mixture of these preferences among their experimental subjects.
in equilibrium. Similarly, the public belief exceeds the bid price, resulting in the standard bid-ask spread, $A_t - B_t > 0$. Importantly, uninformed investors allow the adverse selection problem between informed investors and the uninformed market maker to be overcome. Due to their presence, the bid and ask prices do not fully reflect the private information of informed investors, who are then able to make profitable trades. The market maker loses money to informed investors, but recoups it from uninformed investors. This intuition is formalized in Lemma 1, which characterizes the behavior of the risk-neutral investors, showing that the standard result of Glosten and Milgrom (1985) continues to hold even in the presence of prospect theory investors. All proofs are provided in Appendix A.

**Lemma 1:** In any equilibrium, for all $p_t \in (0, 1)$, risk-neutral investors always trade: those with favorable signals ($s_t = 1$) buy and those with unfavorable signals ($s_t = 0$) sell.

An immediate consequence of Lemma 1 is that, if risk-neutral investors arrive with positive probability, some information is partially revealed in every period: an information cascade in which prices stagnate and subsequent trades reveal no new information never occurs. Thus, by the law of large numbers, public beliefs and bid and ask prices converge to the true asset value in the limit as $T \to \infty$, achieving informational efficiency. In Section 2.6, I contrast this result with the case in which we only have CPT investors.

### 2.4 Prospect Theory Investors

CPT differs from expected utility in that investors evaluate gains and losses relative to a reference point. Perhaps the simplest possible reference point is status quo wealth, which is the reference point I adopt. The behavioral asset pricing literature has tended to instead use the expected wealth from investing in a risk-free asset (see Barberis and Huang (2008), Barberis and Xiong(2009), and Li and Yang (2013)). In the absence of a risk-free asset, as is the case here, these different specifications are equivalent in the sense that the reference point is the amount an investor can attain without risk. Expectations-based reference points, such as those in Koszegi and Rabin (2006, 2007) are another popular alternative. However, I show in Appendix D that they are generally inconsistent with the experimental evidence
CPT specifies value functions, $v^+()$ and $v^-(())$, and decision weight functions, $w^+()$ and $w^-(())$, over gains and losses, respectively. The decision weight functions apply to capacities, a generalization of probabilities, but for binary outcomes result in simple non-linear transformations of the objective probabilities. The utility a CPT investor derives from a binary lottery, $\mathcal{L}$, which returns a gain of $x$ with probability $r$ and a loss of $y$ with probability $1-r$ is given by $U(\mathcal{L}) = w^+(r)v^+(x) + w^-(1-r)v^-(y)$.

Given this utility function, a CPT investor with a private belief, $b_t = Pr(V = 1|H_t, s_t)$, prefers buying to not trading if

$$w^+(b_t)v^+(1-A_t) + w^-(1-b_t)v^-(-A_t) \geq 0$$  \hspace{1cm} (1)$$

where the utility of not trading results in no gain or loss and is normalized to zero. Similarly, she prefers selling to not trading if

$$w^+(1-b_t)v^+(B_t) + w^-(b_t)v^-(B_t - 1) \geq 0$$  \hspace{1cm} (2)$$

If neither equation (1) nor equation (2) is satisfied, then a CPT investor abstains from trading.

The forms of equations (1) and (2) are sufficiently general that little can be said about the behavior of the investor without imposing additional structure. I proceed using the functional forms for the value and decision weight functions provided in the original work of Kahneman and Tversky (1992), because they are tractable, parsimonious, and appear to fit decisions over binary gambles reasonably well. Specifically, I assume

---

8I’m implicitly assuming investors evaluate their gains or losses when the asset value is realized, either by closing their position so that the gains or losses are realized (corresponding to the realization utility of Shefrin and Statman (1985)) or by evaluating their gains or losses on paper. In the experiment, this assumption is satisfied. Barberis and Xiong (2009) discuss the difference between paper gains and losses and realization utility, showing that the distinction can be important in a model in which investors make multiple trading decisions.

9The issue of narrow or broad framing (Barberis, Huang, and Thaler (2006)) does not play a role in the model because only a single asset is available. With multiple assets or other sources of background risk, it becomes important to distinguish between gains and losses on one’s overall portfolio and narrow framing in which each asset is evaluated individually. In applying the model to the experimental results, I assume subjects use narrow framing, considering the experiment (and, in fact, each repetition of the game) in isolation.

10See Kahneman and Tversky (1992) for the more general formulation for any number of outcomes, as well as an axiomatic foundation for the preferences.

11Other functional forms, especially for the decision weighting function, have appeared in the literature. See Bruhin, Fehr-Duda, and Epper (2010) and the references therein.
Figure 1: Examples of the Value and Probability Weighting Functions

Note: Value function (left graph) and probability weighting function (right graph) for the case of $\alpha = 0.88$, $\lambda = 2.25$, and $\delta = 0.65$ (taken from the median estimates in Kahneman and Tversky (1992) and averaging the probability weighting parameters they separately estimate for gains and losses).

\[
v^+(x) = x^\alpha \\
v^-(y) = -\lambda(-y)^\alpha
\]

and

\[
w^+(r) = w^-(r) = \frac{r^\delta}{(r^\delta + (1-r)^\delta)^{\frac{1}{\delta}}}
\]

with $\alpha \in (0,1]$, $\lambda \geq 1$, and $\delta \in (0,1]$.\(^\text{12}\) $\alpha \in (0,1)$ reflects the common experimental finding of risk-aversion over gains and risk-seeking over losses (an “S-shaped” value function). $\lambda \geq 1$ reflects loss-aversion: losses are weighted more heavily than gains. Finally, $\delta \in (0,1]$ implies that low-probability event are overweighted. Figure 1 illustrates examples of each function.

Substituting the functional forms into equations (1) and (2) results in the following optimal decisions for a CPT investor:

\[
\text{buy if } \left( \frac{b_t}{1-b_t} \right)^\delta \geq \lambda \left( \frac{A_t}{1-A_t} \right)^\alpha \\
\text{sell if } \left( \frac{b_t}{1-b_t} \right)^\delta \leq \frac{1}{\lambda} \left( \frac{B_t}{1-B_t} \right)^\alpha
\]  
\(^{12}\)Kahneman and Tversky assume a slightly more general form allowing $w^+(r)$ and $w^-(r)$ to have different parameters, but their experimental estimates for the two parameters are quantitatively similar. I assume a common parameter, which results in a significant increase in tractability.
Risk-neutral investors are a special case of CPT investors with $\alpha = \delta = \lambda = 1$. Under this parameterization, equations (3) state that an investor buys when her belief exceeds the bid price and sells when her belief is below the ask price as in Lemma 1. More generally, we must explicitly evaluate the beliefs and prices. An investor with a favorable signal, $s_t = 1$, has a private belief conditional on the history and her private signal (denoted $b^1_t$) given by Bayes’ rule:

$$b^1_t = \frac{p_t q}{p_t q + (1 - p_t)(1 - q)}$$

Similarly, an investor with an unfavorable signal, $s_t = 0$, has private belief (denoted $b^0_t$):

$$b^0_t = \frac{p_t(1 - q)}{p_t(1 - q) + (1 - p_t)q}$$

The bid and ask prices can also be written as functions of the public belief:

$$A_t = \frac{p_t Pr(a_t = \text{buy}|V = 1)}{p_t Pr(a_t = \text{buy}|V = 1) + (1 - p_t)Pr(a_t = \text{buy}|V = 0)}$$

$$B_t = \frac{p_t Pr(a_t = \text{sell}|V = 1)}{p_t Pr(a_t = \text{sell}|V = 1) + (1 - p_t)Pr(a_t = \text{sell}|V = 0)}$$

(4)

where the conditional probabilities of observing a purchase or a sale depend upon the equilibrium strategies of the investors.

After substituting the expressions for her private belief and the bid and ask prices, the optimal decision of a CPT investor with a favorable signal becomes

$$\text{buy if } \left(\frac{p_t}{1 - p_t}\right)^{\delta - \alpha} \geq \lambda \left(\frac{1 - q}{q}\right)^{\delta} \left(\frac{Pr(a_t = \text{buy}|V = 1)}{Pr(a_t = \text{buy}|V = 0)}\right)^{\alpha}$$

$$\text{sell if } \left(\frac{p_t}{1 - p_t}\right)^{\delta - \alpha} \leq \frac{1}{\lambda} \left(\frac{1 - q}{q}\right)^{\delta} \left(\frac{Pr(a_t = \text{sell}|V = 1)}{Pr(a_t = \text{sell}|V = 0)}\right)^{\alpha}$$

(5)

The corresponding equations for an investor with an unfavorable signal are identical except that the ratio of $1 - q$ to $q$ on the right-hand side is inverted in each.

Although the opposing effects of $\alpha$ and $\delta$ have received relatively little attention in applications of prospect theory (with the exception of Barberis (2012)), they are immediately clear in (5). To understand the intuition, consider a simplified example. Remove all private information so that the bid and ask prices collapse to the public belief, $p_t$. In this case, risk-neutral investors have no incentive to trade given that their private beliefs correspond to that of the public belief (equal to price): the gambles corresponding to a purchase or a sale have zero expected value. With this simplification, equations (3) for a CPT investor become
\[
\text{buy if } \left( \frac{p_t}{1-p_t} \right)^{\delta-\alpha} \geq \lambda \\
\text{sell if } \left( \frac{p_t}{1-p_t} \right)^{\delta-\alpha} \leq \frac{1}{\lambda}
\]

(6)

We see that, unless the public belief is exactly \( \frac{1}{2} \), either buying or selling is strictly preferable to not trading.

Consider a public belief, \( p_t > \frac{1}{2} \). As the decision weights become more distorted from linearity (\( \delta \) decreases from one), the propensity to buy decreases and the propensity to sell increases. Intuitively, a decrease in \( \delta \) increases the weight assigned to the small probability, \( 1-p_t \), of a loss and reduces the weight assigned to the larger probability, \( p_t \), of a gain, thereby making buying less attractive. Conversely, it increases the utility from selling because the small probability is instead associated with a gain. Therefore, for \( p_t \) sufficiently large and \( \delta < \alpha \), the investor strictly prefers to sell the stock: she exhibits a \textit{preference for positive skewness} which is the consequence of prospect theory studied extensively in Barberis and Huang (2008).

Conversely, consider an increase in the curvature of the value function (decrease in \( \alpha \) from one). Mathematically, we see that we get exactly the opposite effect from that due to an increase in the distortion of probabilities. Intuitively, as the curvature increases, the small gain \((1-p_t)\) that occurs with probability \( p_t \) if one buys is preferred to the large gain \((p_t)\) that occurs with probability \( 1-p_t \) if one sells, a simple consequence of risk-aversion (an investor with risk-neutral preferences would be indifferent). At the same time, the small probability of a large loss if one buys is preferred to the large probability of a small loss if one sells, due to risk-seeking. Both effects make buying preferable to selling so that for \( p_t \) sufficiently large and \( \delta > \alpha \), the investor always strictly prefers to buy: she exhibits a \textit{preference for negative skewness}.\(^{13}\)

Finally, consider the role of loss aversion. An increase in \( \lambda \) reduces the range of public beliefs at which an investor is willing to trade, because it simultaneously makes each inequality in (6) more difficult to satisfy. The intuition here is simple: an increase in loss-aversion increases the dis-utility of losses, which makes one more likely to abstain from taking on a position in the asset that can result in a loss. Perhaps surprisingly, however, loss aversion

\(^{13}\)Beyond the intuition, we can see that the preference to trade in a particular direction at an extreme price must be driven by the skewness of returns, and not by their mean or variance. For example, with a preference for negative skewness, a CPT investor switches from always selling to always buying as the price increases from 0 to 1. Over this range, the mean return is always positive if she trades with her private signal, so that, as in the case of a risk-neutral investor, an investor that cares only about the mean always trades with her private signal. The variance of the return increases from zero to its maximum value at \( p = \frac{1}{2} \), and then decreases to zero again, so that if the variance in returns were driving behavior, we’d expect behavior to be the same at both price extremes.
prevents trading only at intermediate public beliefs. Although the potential losses are larger at extreme public beliefs, one can take the side of the trade that either minimizes the probability ($\delta < \alpha$) or the size ($\delta > \alpha$) of a loss. At intermediate beliefs, on the other hand, the chance of a loss of medium size and probability can only be avoided by abstaining from trade.

This simple example captures the countervailing forces of distortions due to decision weights and utility curvature. The intuition carries over to the full equilibrium characterization I pursue in the following section. Importantly, the intuition also suggests that, if we believe individuals simultaneously exhibit both utility curvature and probability distortions, then we cannot necessarily consider only one aspect of prospect theory in isolation, because its effect is likely to be diminished by the countervailing force of the other aspect.

2.5 Equilibrium

Equilibrium strategies of the investors are a function of the public belief and their private signals. They are given by Lemma 1 for risk-neutral investors, and follow from equations (5) for CPT investors with a favorable signal and the corresponding equations for those with unfavorable signals. Note first the symmetry of the environment: an investor with a favorable signal at a public belief, $p_t$, is in a symmetric situation to an investor with an unfavorable signal at a public belief, $1 - p_t$. Therefore, equilibrium behavior is symmetric around $p_t = \frac{1}{2}$, simplifying its description.

As discussed in the previous section, absent private information, skewness preferences can cause CPT investors to strictly prefer to trade in a particular direction. As prices become extreme, because the skewness in returns is unbounded, eventually skewness preferences dominate private information, causing the investor to trade in the same direction for both realizations of her private signal. Thus, when the public belief is sufficiently large and $\delta > \alpha$, investors buy regardless of their private signal, and, for sufficiently small public beliefs, they always sell. Conversely, when $\delta < \alpha$, CPT investors sell regardless of their private signal when the public belief is sufficiently large, and buy when the public belief is sufficiently small.

For less extreme public beliefs, CPT investors may either trade according to their private information or abstain from trading. The public beliefs at which behavior transitions depend upon an investor’s private signal, so that an equilibrium is characterized by four transition regions. I denote the transition region in which a CPT investor with a favorable signal transitions from trading to not trading (as the public belief increases), $p^0 \equiv (p^0_0, p^0_1)$, and that in which she transitions back from not trading to trading, $p^1 \equiv (p^1_0, p^1_1)$, where $0 < p^0_0, p^0_1, p^1_0, p^1_1 < 1$. The other two transition regions, those for a CPT investor with an
unfavorable signal, are located at $1 - p^0$ and $1 - p^1$ (see Figure 2).

In each of the transition regions, there exists an equilibrium in which the investor mixes between trading and not trading due to strategic interaction with the market maker through the bid and ask prices that depend upon the investor’s strategy (each of the equations in (5) holds with equality in this equilibrium). This equilibrium is not unique, however. In a transition region in which investors with both favorable and unfavorable signals are trading with positive probability, there exist pure strategy equilibria as well. For example, at high public beliefs where the investor with a favorable signal is buying, the investor with an unfavorable signal may buy, abstain, or mix. When she buys, she pools with the investor with the favorable signal, thus revealing less information to the market maker, lowering the ask price, and making buying more profitable. If she instead abstains, a favorable signal is (partially) revealed to the market maker, raising the ask price, and making buying less profitable. Figure 2 provides a general illustration of the equilibria. Transition regions with background shading indicate a multiplicity of equilibrium strategies.

Figure 2: Prospect Theory Investor Behavior in Equilibrium

![Figure 2: Prospect Theory Investor Behavior in Equilibrium](image)

Note: Behavior of prospect theory investors in equilibrium. The upper two plots correspond to $\delta > \alpha$, and the bottom two plots to $\delta < \alpha$. The left two plots illustrate a parameterization for which investors do not trade with either signal over some intermediate range of public beliefs. The right two plots illustrate a second parameterization in which investors instead trade according to private information over this range. Transition regions with background shading indicate a multiplicity of equilibrium strategies. Within a transition region, the distance between No Trade and Buy or Sell reflects the mixing probability in the mixed strategy equilibrium (a higher probability is associated with the closer action).

In Figure 2, the upper two plots correspond to $\delta > \alpha$ and the lower two to $\delta < \alpha$. Within each of these two cases, I illustrate the two possible relationships between the locations of the transition regions. For the plots on the left of Figure 2, the parameters are such that the two transition regions lie on opposite sides of $p = \frac{1}{2}$. In this case, neither type of investor
trades over some range of intermediate beliefs. The plots on the right of Figure 2 illustrate a second parameterization in which the transition regions lie on the same side of $p = \frac{1}{2}$. In this case, for intermediate public beliefs, a separating equilibrium exists in which CPT investors’ trades reveal their private information. Theorem 1 is the main theorem of the paper formalizing the illustration of Figure 2.

**Theorem 1:** In equilibrium:

1. The market maker posts bid and ask prices (given by (4)) where the conditional buy and sell probabilities are determined by the equilibrium strategies of informed investors that follow.
2. For all $p_t \in (0, 1)$, risk-neutral investors buy with favorable signals and sell with unfavorable signals.
3. CPT investors’ strategies are as follows:
   (a) If $\delta = \alpha$, there exist two cutoff values of loss aversion, $\bar{\lambda} > \lambda > 1$ such that, at all $p_t \in (0, 1)$, if $\lambda \leq \bar{\lambda}$, they buy with favorable signals and sell with unfavorable signals, and, if $\lambda \geq \bar{\lambda}$, they do not trade. If $\lambda \in (\lambda, \bar{\lambda})$, they mix between buying and not trading with favorable signals and between selling and not trading with unfavorable signals.
   (b) If $\delta \neq \alpha$, strategies are characterized by four transition regions in public beliefs, $p^0 \equiv (\underline{p}^0, \bar{p}^0), p^1 \equiv (\underline{p}^1, \bar{p}^1)$, and their symmetric counterparts.
   i. If $\delta > \alpha$, CPT investors with favorable signals sell for $p_t \leq \underline{p}^0$, don’t trade for $p_t \in [\underline{p}^0, \underline{p}^1]$, and buy for $p_t \geq \bar{p}^1$. CPT investors with unfavorable signals sell for $p_t \leq 1 - \bar{p}^1$, don’t trade for $p_t \in [1 - \bar{p}^1, 1 - \bar{p}^0]$, and buy for $p_t \geq 1 - \underline{p}^0$.
   ii. If $\delta < \alpha$, CPT investors with favorable signals buy for $p_t \leq \underline{p}^0$, don’t trade for $p_t \in [\underline{p}^0, \underline{p}^1]$, and sell for $p_t \geq \bar{p}^1$. CPT investors with unfavorable signals buy for $p_t \leq 1 - \bar{p}^1$, don’t trade for $p_t \in [1 - \bar{p}^1, 1 - \bar{p}^0]$, and sell for $p_t \geq 1 - \underline{p}^0$.
   iii. Within each transition region, there’s an equilibrium strategy in which the CPT investor mixes between abstaining and trading. For transition regions in which CPT investors with favorable and unfavorable signals trade in the same direction, there also exist pure strategy equilibria.
   iv. The transition regions do not overlap: $\bar{p}^0 < \underline{p}^1$ and $\bar{p}^1 < 1 - \underline{p}^0$ if $\delta > \alpha$ ($1 - \underline{p}^1 < \bar{p}^0$ if $\delta < \alpha$), implying $\bar{p}^0 < \frac{1}{2}$ if $\delta > \alpha$ ($1 - \underline{p}^1 < \frac{1}{2}$ if $\delta < \alpha$).
2.6 Equilibrium Properties

The fact that CPT investors may buy or sell independently of their private signals as prices become extreme generates behavior that looks very much like herding or contrarian behavior. However, there is a subtle difference. Herding and contrarian behavior, at least in the context of informational models, are typically defined as behaviors that depend on the history of past actions (Avery and Zemsky (1998) and Park and Sabourian (2011)).\textsuperscript{14} For example, an investor is said to ‘herd’ if she would sell at some initial price (based upon her private information), but buys after others’ trades increase prices. A CPT investor that faces an extreme price instead exhibits behavior that is independent of the past history. If she buys at a particular price regardless of her signal, she does so whether the price rose from \( p_1 = \frac{1}{2} \) or fell from a more extreme price. Of course, extreme prices are more likely to occur after a sequence of purchases which is why her behavior will tend to look like herding. For these reasons, I refrain from referring to CPT investor behavior as herding or contrarian, instead I define ‘herding-like’ and ‘contrarian-like’ behavior:

**Definition 1:**

1. An informed investor exhibits herding-like (H-L) behavior if, independently of her signal, she (i) buys when \( p_t > \frac{1}{2} \), and (ii) sells when \( p_t < \frac{1}{2} \).
2. An informed investor exhibits contrarian-like (C-L) behavior if, independently of her signal, she (i) buys when \( p_t < \frac{1}{2} \), and (ii) sells when \( p_t > \frac{1}{2} \).

In addition to the previous definitions, I use the term unresponsive to encompass both: behavior is unresponsive if it is either herding-like or contrarian-like.

Corollary 1, which follows directly from Definition 1 and Theorem 1, shows that herding-like and contrarian-like behaviors occur generically (unless \( \delta = \alpha \)) in the model.

**Corollary 1:** In any equilibrium:

1. For all parameterizations with \( \delta > \alpha \), there exists a public belief, \( p^0 < 1 \), such that for all \( p_t < p^0 \) and \( p_t > 1 - p^0 \), CPT investors exhibit herding-like behavior.
2. For all parameterizations with \( \delta < \alpha \), there exists a public belief, \( p^1 < 1 \), such that for all \( p_t > p^1 \) and \( p_t < 1 - p^1 \), CPT investors exhibit contrarian-like behavior.

The fact that behaviors that look like herding and contrarianism arise is in stark contrast to the standard result that these behaviors are impossible in a Glosten-Milgrom model (Avery and Zemsky (1998)) with risk-neutral investors. These behaviors imply that, if there are no

\textsuperscript{14} An older definition of herding is ‘acting the same as everyone else’. CPT investors may exhibit herding in this sense, provided they have identical preferences.
risk-neutral investors, information cascades will develop. In the simplest case, if the prior is intermediate \((p_1 = \frac{1}{2})\) and the CPT investors are loss averse, no trade will ever occur - a form of non-participation. On the other hand, if the investors are not sufficiently loss averse, they initially trade according to private information until prices become extreme, at which point their behavior becomes unresponsive. At this point, an information cascade occurs - no further information is ever revealed and prices stagnate. Thus, herding and contrarian-like behaviors can be detrimental to the informational efficiency of markets and can lead to very misleading prices (i.e. if initial signals are incorrect).

3 Experiment

3.1 Design

The goal of the experiment is to test for evidence of herding and contrarian-like behavior in a setting that controls for sources of these behaviors other than preferences, and, furthermore, to test the specific predictions of the model at an individual level (preferences, after all, being individual-specific). To achieve these goals, the implementation of the experiment differs from the model in one important way. Rather than have subjects arrive to the market and trade one at a time, I implement an individual decision problem version of the model. Doing so eliminates the ‘herd mentality’ explanation for herding (because there are no previous investors to imitate), but also strategic ambiguity on the part of the subjects, which both Cipriani and Guarino (2005) and Drehmann, Oechssler, and Roider (2005) have shown can produce contrarian-like behavior.

In the individual-decision version of the problem, asset values are represented by urns with 7 balls of one color and 3 of another \((q = 0.7)\) as in Figure 3. Starting from a uniform prior, a subject observes a history of prices determined by a sequence of random, public signals (one to five urn draws). Although historical prices should be irrelevant if behavior is solely driven by preferences, I didn’t want to preclude other possibilities by eliminating the historical prices altogether. Instead, I have subjects face a variety of price paths (monotonic and non-monotonic) that lead to the same price so that I can explicitly test for history dependence. After observing each price path, a subject is asked if she would like to buy or sell the asset (or abstain) for each possible realization of her private signal, allowing me to directly observe unresponsive behavior (as in Cipriani and Guarino (2009)). After a subject makes her decision, her private signal is drawn, her corresponding trade is executed, and she receives feedback about both her realized signal and the resulting payoff.

In the ‘NO SOCIAL’ treatment just described, social factors are precluded, but subjects
may fail to form correct Bayesian posteriors given the public and private signals, which could also potentially lead to unresponsive behavior. For example, if subjects overextrapolate from the price trend, we would expect their beliefs to be too extreme, causing herding-like behavior. To test whether or not belief errors are responsible for (or partially responsible for) unresponsive behavior, I conduct a second treatment (the ‘NO INFERENCE’ treatment) in which, in addition to shutting down social forces, I explicitly provide subjects with the posterior probability that the asset is valuable conditional on the price and their private signal.\(^\text{15}\)

In both treatments, I follow the previous experimental literature (Cipriani and Guarino (2005) and Drehmann, Oechssler, and Roider (2005)) by having the market maker in the experiment (the experimentalist) post only a single price equal to the expected value of the asset, \(p_t\), rather than separate bid and ask prices. This procedure both simplifies the problem for subjects, and, more importantly, strengthens incentives to trade according to private information by making the difference between private beliefs and prices larger.\(^\text{16}\)

\(^{15}\)Bisiere, Decamps, and Lovo (2015) developed this approach. However, their main treatments (LE and ME) confound framing effects (lotteries vs. trading environment) with the provision of the correct Bayesian beliefs. They conduct another treatment (SME) that keeps the framing consistent with their ME treatment, but do not statistically compare the behavior across these two treatments. I keep the framing across treatments identical - the only difference is that subjects are given an additional statement of the correct posterior in the NO INFERENCE treatment. See the instructions in Appendix E.

\(^{16}\)Subjects are explicitly told that the price reflects the expected value of the asset given all public information (the public signals).
Importantly, the predictions of the model are unchanged when trades occur at a single price, except that the transition regions of Theorem 1 become degenerate threshold prices, guaranteeing unique predictions (see Appendix B for details).

I recruited undergraduate subjects at the University of California, Santa Barbara over the month of August, 2016. They participated in a computerized implementation of the experiment (using oTree (Chen, Schonger, and Wickens (2016)). I conducted three sessions of each treatment for a total of 46 subjects in each. Each subject took part in 30 consecutive ‘games’ of a single treatment (NO SOCIAL or NO INFERENCE) in which they faced 30 historical price paths and made 60 trading decisions (one for each possible realization of their private signal in each game). Subjects were paid for their decisions in each game with average earnings of $17.13. The experiment typically finished in just over an hour.

3.2 Experimental Results

I organize the results around the two main questions the experiment is designed to answer. In Section 3.2.1, I ask to what extent preferences in general are likely to be responsible for behavior, relative to other explanations. In Sections 3.2.2 and 3.2.3, I assess how well CPT preferences in particular fit the data, first by looking for the patterns they predict, and then by modeling individual behavior.

3.2.1 The Role of Preferences

In the data analysis, I decided ex ante to drop the first 3 games during which subjects are becoming familiar with the interface and environment.\(^{17}\) I begin by categorizing behavior by type in Table 1. ‘Abstain’ refers to not trading for either signal realization. ‘Other’ consists primarily of partially unresponsive behavior in which a subject trades with one signal but abstains with the other. This behavior is predicted for investors with CPT preferences: we expect transitions from trading to abstention to occur at different threshold prices for different private signal realizations.\(^{18}\)

The results of Table 1 strongly confirm that social factors, which are absent in both treatments, can not be the sole cause of unresponsive behavior: more than 25% of behavior

\(^{17}\)Including this data does not affect any of the qualitative results, nor does restricting the analysis to only the second half of the data. Learning seems to play a very limited role: the results of classifying individual subjects in Section 3.2.2 are remarkably similar when using only the first or second half of the data.

\(^{18}\)Cipriani and Guarino (2009) also document partially unresponsive behavior (Table 1). A small fraction of behavior in the ‘Other’ category is difficult to reconcile with any theory: trading against both signals, always buying or always selling at \(p_t = 0.5\), buying with an unfavorable signal but abstaining with a favorable signal, or selling with a favorable signal but abstaining with an unfavorable signal. This behavior makes up only 5.2% and 7.0% of behavior in the NO SOCIAL and NO INFERENCE treatments, respectively.
Table 1: Aggregate Behavior by Treatment

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Abstain</th>
<th>Risk-Neutral</th>
<th>Herding-Like</th>
<th>Contrarian-Like</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO SOCIAL</td>
<td>5.2</td>
<td>37.4</td>
<td>13.9</td>
<td>12.8</td>
<td>30.8</td>
</tr>
<tr>
<td>NO INFERENCE</td>
<td>5.5</td>
<td>27.5</td>
<td>34.9</td>
<td>10.6</td>
<td>21.6</td>
</tr>
</tbody>
</table>

Note: Percentages of each type of behavior in the NO SOCIAL treatment (no Bayesian beliefs) and the NO INFERENCE treatment (correct Bayesian beliefs provided to subjects).

is herding or contrarian-like in each treatment and less than 40% is consistent with the predictions of risk-neutral subjects. Furthermore, providing subjects with the correct Bayesian posteriors does not reduce unresponsive behavior as we’d expect if Bayesian errors drive this behavior. In fact, the opposite occurs - with correct Bayesian beliefs, we observe significantly more herding-like behavior. Logit regressions (with errors clustered by subject) of each type of behavior (versus not) on a treatment dummy confirm this result: herding-like behavior significantly increases ($p = 0.000$), while risk-neutral behavior ($p = 0.041$) and Other behavior ($p = 0.011$) significantly decrease in the NO INFERENCE treatment.\(^\text{19}\)

The fact that we observe less herding-like behavior when belief errors may play a role is contrary to what we’d expect under extrapolative expectations. Instead, belief errors appear to work against a preference-based tendency to herd, consistent with subjects having a belief that is too conservative (too close to one-half) when observing a private signal opposite to the price trend, perhaps due to either overweighting their private signal or under-weighting the price (see also Goeree et. al. (2007) and Weizsacker (2010)).

Result 1: Unresponsive behavior is common when social factors are absent. Bayesian errors are also not the cause of this behavior: providing the correct Bayesian posteriors leads to an increase in herding-like behavior.

Historical price paths could affect behavior even with correct Bayesian beliefs if subjects simply ignore their beliefs. To test whether or not herding and contrarian-like behavior is path dependent, I ask whether or not it depends upon (i) a monotonic price path (no contradictory public signals), or (ii) a price path that is either monotonic or ends in a ‘streak’ (two or more public signals in the same direction). To do so, I first define a ‘normalized’ price, a measure of the extremeness of the price which treats rising and falling prices symmetrically.\(^\text{20}\)

For each treatment and each normalized price at which unresponsive behavior is possible, I

---

\(^{19}\)Bisiere, Decamps, and Lovo (2015) find a similar result when they compare their ME treatment (which requires Bayesian updating) to both their LE or SME treatments (which do not).

\(^{20}\)Formally, the normalized price is defined as $p_t$ if $p_t \geq \frac{1}{2}$ and $1 - p_t$ if $p_t < \frac{1}{2}$. Subjects do not always treat rising and falling prices symmetrically (see Section 3.2.2). Therefore, the results should be interpreted as average effects over rising and falling prices.
run a logit regression of a dummy variable indicating herding or contrarian-like behavior on a dummy variable indicating either a monotonic price path or a streak.

For herding-like behavior, I find no statistical evidence of path-dependence \((p > 0.370)\) in any of the 12 regressions (minimum of 276 observations in each regression). For contrarian-like behavior, I find some evidence of path-dependence, but evidence that is inconclusive. The most robust evidence occurs at a normalized price of 0.93 (three more public signals in one direction than the other). At this price, contrarian-like behavior is significantly less likely after a monotonic price path or a streak in both treatments \((p < 0.043)\), suggesting that contrarians hesitate to trade against a trend. Consistent with this hypothesis, contrarian-like behavior is also less likely at a normalized price of 0.7 after a streak \((p = 0.008\) in NO SOCIAL, \(p = 0.075\) in NO INFERENCE). On the other hand, contrarian-like behavior is more likely at a normalized price of 0.85 after a streak in NO INFERENCE \((p = 0.027)\) and after a monotonic price path in NO SOCIAL \((p = 0.086)\). Overall, the evidence suggests that contrarians may prefer to observe contradictory evidence before acting contrarian. However, the fact that no clear picture emerges also leaves open the possibility that the findings may be artifacts of the fact that contrarian-like behavior is relatively infrequent.

**Result 2:** Herding-like behavior is independent of historical price paths. Contrarian-like behavior is often more likely when historical price paths contain contradictory information, but not consistently so.

### 3.2.2 Trading Patterns

The evidence from the previous section demonstrates that social factors and belief errors are not necessary for herding and contrarian-like behavior, indicating a role for preferences, but which preferences in particular? The results of Table 1 suggest that expected utility preferences cannot fully explain behavior. In particular, risk-aversion can not produce herding or contrarian-like behaviors, and risk-seeking cannot produce abstention, behaviors which together make up a sizable fraction of overall behavior (I give formal proofs of these assertions in Appendix C). CPT preferences, on the other hand, can generate these behaviors and in fact predict particular characteristic patterns in the data.

At the aggregate level, the main prediction of the model is that unresponsive behavior becomes more frequent as prices become more extreme, while risk-neutral behavior and abstention become less frequent. To test this prediction, Table 2 breaks down the frequency of each type of behavior by the normalized price.

The frequency of risk-neutral behavior almost perfectly monotonically decreases with the normalized price in both treatments, while the frequency of herding-like behavior almost
perfectly monotonically increases. Both of these results are exactly as predicted by the model and are confirmed by logit regressions of the type of behavior (versus not) on the normalized price ($p = 0.000$ in all four regressions). For the frequency of abstention, the evidence is mixed because it decreases in the NO SOCIAL treatment ($p = 0.047$), in accordance with the model, but increases in the NO INFERENCE treatment ($p = 0.005$). One possible reason for the increase is that standard expected utility with a large degree of risk aversion can cause abstention at extreme prices, a possibility I consider when analyzing individual behavior in the following section. Finally, contrarian-like behavior increases in both treatments (positive coefficients in the regressions), but insignificantly so, perhaps reflecting the fact that the behavior is not frequent enough to provide sufficient statistical power.

**Result 3:** In the aggregate, risk-neutral behavior decreases with the normalized price, and herding-like behavior increases, as predicted by the model. Contrarian-like behavior increases, but not significantly so. Abstention is predicted to decrease, but does so in the NO SOCIAL treatment only.

At the individual level, the model predicts that an individual makes herding-like or contrarian-like decisions, but not both. To test this prediction, Figure 4 plots the fraction of contrarian-like behavior versus the fraction of herding-like behavior for each subject. If the theory were perfect, we’d expect contrarian-like individuals to lie on the y-axis and herding-like individuals on the x-axis. Although noisy, there is a clear negative correlation - indicating that most subjects have a clear predisposition for either herding-like or contrarian-like behavior. This predisposition allows me to classify subjects according to their preferences.
Figure 4: Herding and Contrarian-like Behavior By Subject

Note: Each point represents a single subject and indicates the fraction of their decisions that are herding-like or contrarian-like. Most subjects have a clear predisposition for one behavior or the other, allowing them to be classified accordingly.

(herding-like, contrarian-like, or expected utility) as illustrated in the figure and described in detail in the following section.

Result 4: The majority of subjects have a predisposition towards either herding or contrarian-like behavior, consistent with the model.

3.2.3 Preference Models

To classify subjects, I consider four candidate preference models. The first model is CPT preferences. It has two degrees of freedom: loss aversion ($\lambda$) and the difference between the utility curvature and probability weighting parameters ($\delta - \alpha$) (see Appendix B for details). The second is CPT preferences, but without loss aversion ($\lambda = 1$), which gives it a single degree of freedom making it directly comparable to the third model, expected utility. Because risk-neutrality is incapable of explaining either herding-like or contrarian-like behavior, I consider general CRRA preferences (the results are almost identical with CARA preferences). Finally, I consider a model which only requires that decisions respect symmetry (the decision at a price $p_t$ with a favorable signal must match that at a price $1 - p_t$ with an unfavorable signal). This model has 27 degrees of freedom, allowing different decisions for each normalized price and private signal. It provides an upper bound on the

21The theory states that an optimal CPT strategy is characterized by two thresholds, $p^0$ and $p^1$, at which an investor’s behavior transitions. For CPT preferences, I compare the data to all possible trade patterns, subject to the restrictions on the thresholds that come out of the theory (see Appendix B for further details).
Figure 5: Individual Match Scores

Note: Empirical CDFs of the match scores of individuals for expected utility, prospect theory without loss aversion, prospect theory with loss aversion, and a model that imposes only symmetry.

ability of any model that treats decisions and asset values symmetrically to fit the data (see Appendix C for further discussion).

To compare models, I employ the technique of Bisiere, Decamps, and Lovo (2015) to calculate a match score for each subject relative to each model of behavior. For a given model and a given (range of) preference parameters, I obtain predicted decisions for each discrete price and for each private signal. I award 0.5 for each of a subject’s 54 decisions that matches the prediction (and zero otherwise), and then divide by the maximum possible score so that the match score lies between 0 and 1. Figure 5 provides the empirical CDFs of the match scores in each treatment for each of the four models.

A particularly striking result in Figure 5 is that prospect theory explains choices in both treatments as well as the very lenient model that only imposes symmetry, even though it has an order of magnitude less degrees of freedom ($p = 0.765$ and $p = 0.269$ with a Kolmogorov-Smirnov test in the NO SOCIAL and NO INFERENCE treatments, respectively). Furthermore, in the NO INFERENCE treatment in which there is no role for

---

22 It would be tempting to structurally estimate subjects’ preference parameters. However, doing so is problematic because, under a standard random utility specification, $\alpha$ and $\delta$ would appear to be separately identifiable, contradicting the theoretical result that only the difference affects behavior. This problem of ‘incredible structural inference’ is nicely illustrated by Thomas Rosenberg (2005) with the example of a research assistant that repeatedly measures the length and width of a table, but reports only the table’s area. Absent measurement error, it is clear that one cannot recover the length and width from the area. However, with assumptions on distributions of the measurement errors in each dimension, one can miraculously recover the two dimensions. Clearly, the resulting estimates would depend highly on assumptions made about the error distributions (see also Wilcox (2008)). The match score approach, although less technically sophisticated, has the distinct advantage that an error distribution does not need to be specified.

23 The model that only imposes symmetry matches about 75% of a subject’s decisions at the median, similar to the finding of Bisiere, Decamps, and Lovo (2015). Buying is more frequent in both treatments,
Table 3: Individual Types and Match Scores

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Models considered</th>
<th>Risk-Neutral</th>
<th>Risk-Averse</th>
<th>Risk-Seeking</th>
<th>H-L</th>
<th>C-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO SOCIAL</td>
<td>Expected utility</td>
<td>25</td>
<td>15</td>
<td>6</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Both (no loss aversion)</td>
<td>13</td>
<td>12</td>
<td>2</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Both (with loss aversion)</td>
<td>12</td>
<td>8</td>
<td>2</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>NO INFRINGEMENT</td>
<td>Expected utility</td>
<td>37</td>
<td>8</td>
<td>1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Both (no loss aversion)</td>
<td>8</td>
<td>4</td>
<td>0</td>
<td>32</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Both (with loss aversion)</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>32</td>
<td>3</td>
</tr>
</tbody>
</table>

Note: Number of subjects best matched to each model of behavior under different combinations of models: expected utility (EU), EU and prospect theory without loss aversion, and EU and prospect theory. H-L and C-L are shorthand for herding-like and contrarian-like, respectively.

Bayesian errors, even restricting prospect theory to not include loss aversion does almost as well as the full prospect theory model, and does significantly better than expected utility, both of which have a single degree of freedom \((p = 0.001)\) in a Kolmogorov-Smirnov test. In the NO SOCIAL treatment, prospect theory without loss aversion and expected utility fit similarly well. However, the full prospect theory model continues to do better than expected utility \((p = 0.057)\) as it does in the NO INFRINGEMENT treatment \((p = 0.000)\).

Prospect theory provides an overall better fit of the data when I impose a single model for all subjects, but we can also allow the model to vary on an individual basis. Table 3 provides the number of subjects for which each model provides the best match. The first row considers expected utility only for comparison purposes. The next two include both expected utility and prospect theory, with and without loss aversion.

Table 3 shows that the majority of subjects are better classified by prospect theory in both treatments. Within the set of subjects for which prospect theory is the best match, approximately three-quarters exhibit herding-like behavior. Importantly, however, the other quarter of subjects are classified as contrarian-like, which is consistent with prospect theory but inconsistent with the other non-expected utility theory that has been studied in this perhaps because it is more familiar to subjects than selling short.

\(^{24}\)In the NO SOCIAL treatment, partial herding and contrarian behavior are relatively frequent (much of the ‘Other’ category in Tables 1 and 2). This behavior cannot be justified by either risk-aversion or prospect theory without loss aversion, which leads to the relatively poor fit of both models.

\(^{25}\)I break ties in favor of expected utility and don’t include the symmetric model because it mechanically provides the best fit for every subject.
environment, regret aversion (Qin (2015)). Among the subjects for which expected utility is a better match, twenty subjects are classified as risk-neutral which is technically consistent with both prospect theory and expected utility. However, thirteen subjects are classified as risk-seeking or risk-averse, exhibiting behavior that is inconsistent with prospect theory. The majority of these subjects abstain as prices approach zero or one, consistent with a high level of standard risk-aversion.

In the NO INFERENCE treatment, which provides the cleanest measurement of preferences, 76% of subjects are better described by prospect theory, which is interestingly similar to the 80% of subjects that Bruhin, Fehr-Duda, and Epper (2010) report. Returning to Figure 4, it is these subjects that have a clear disposition for either herding or contrarian-like behavior, with many never making a single decision that contradicts their ‘type’. The few subjects that do tend to make both types of decisions are generally classified (poorly) as expected utility types.

**Result 5:** The majority of subjects in both treatments are better described by prospect theory than expected utility (CRRA) and, of these, herding-like strategies are dominant. Furthermore, prospect theory fits the data as well as any model that imposes symmetry.

The finding that herding-like strategies are most popular implies that utility curvature dominates probability weighting for most subjects ($\delta > \alpha$). Kahneman and Tversky (1992) find the opposite, but they report only the parameters for the median subject which masks any underlying heterogeneity. In addition, across the ten preference measurement studies summarized in Table A.3 of Glimcher and Fehr (2013) that use the same probability weighting function I use, four also find $\delta > \alpha$ at the median. Thus, although Kahneman and Tversky’s original study suggested probability weighting is more significant than utility curvature, the literature overall reports mixed results, many of which are consistent with those reported here.

4 An Extension to Actual Market Returns

In the model, the relationship between price trends and return skewness is driven by the binary nature of the asset value, which limits the generalizability of the results. However, previous research has noted that rising prices are associated with negatively skewed returns (and conversely) in actual market data (Harvey and Siddique (1999) and Chen, Hong, and Stein (2001)). Here, I replicate this finding and then show what it implies for the behavior

\[26\text{ The results are, however, directly applicable to ‘near binary’ assets such as options or initial public offerings (Green and Hwang (2012)).}\]
Table 4: Conditional Moments of the Market Index Daily Returns

<table>
<thead>
<tr>
<th>Price Condition</th>
<th>Moment</th>
<th>Estimate</th>
<th>95% Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Price</td>
<td>Mean</td>
<td>0.00041</td>
<td>(−0.00027, 0.0011)</td>
</tr>
<tr>
<td></td>
<td>Standard deviation</td>
<td>0.018</td>
<td>(0.017, 0.019)</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>0.65</td>
<td>(0.18, 1.22)</td>
</tr>
<tr>
<td>High Price</td>
<td>Mean</td>
<td>0.00064</td>
<td>(0.00022, 0.0011)</td>
</tr>
<tr>
<td></td>
<td>Standard deviation</td>
<td>0.011</td>
<td>(0.010, 0.012)</td>
</tr>
<tr>
<td></td>
<td>Skew</td>
<td>−0.52</td>
<td>(−1.21, 0.20)</td>
</tr>
</tbody>
</table>

Note: Mean, standard deviation, and skewness of the daily returns from two constructed conditional distributions. The High Price distribution comes from days in the month following a six-month return in the highest decile. The Low Price distribution comes from days in the month following a six-month return in the lowest decile.

The results are very similar if October 1987 is excluded, ensuring that they are not driven by the market crash during this period.

Skewness is calculated with $SKEW = \frac{E(r - \bar{r})^3}{STD(r)^3}$, where $\bar{r}$ is the mean return and $STD(r)$ is its standard deviation.

If skewness is in fact responsible for the herding-like behavior of the modal subject in the experiment, the results in Table 4 suggest that we should expect such individuals to buy after rising prices and sell after falling prices when faced with actual market returns. I confirm this hypothesis via numerical simulation. In particular, I consider the trading decision of a CPT investor with $\alpha = 0.5$ and $\delta = \lambda = 1$ who faces each of the return distributions in Table 4. For simplicity, I model the investor’s private information as information about the mean expected...
Table 5: Required Annualized Returns for Prospect Theory Investor

<table>
<thead>
<tr>
<th>Price Condition</th>
<th>Indifference Condition</th>
<th>Estimate</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Price</td>
<td>Minimum return to buy</td>
<td>7.1%</td>
<td>(0.0%,15.7%)</td>
</tr>
<tr>
<td>High Price</td>
<td>Minimum return to sell</td>
<td>−7.3%</td>
<td>(−11.6%,−2.8%)</td>
</tr>
</tbody>
</table>

Note: Each return is the annualized return (with private information) that makes an investor with prospect theory preferences \( \alpha = 0.5, \lambda = \delta = 1 \) indifferent between buying and selling. It is calculated assuming the investor faces the distribution of historic daily returns conditional on the return over the previous six months being in the upper or lower decile of such returns.

return and calculate the private signal that would make her indifferent between buying and selling in each case.\(^{30}\) For a risk-neutral investor, the required signal for indifference is equal to the expected mean return of the asset, and the resulting required expected return conditional on her private information is zero. Table 5 provides the corresponding required private returns for the CPT investor along with their 95% confidence intervals.

Table 5 shows that, at low prices, a prospect theory investor prefers to sell the asset if her private information indicates an expected return of 7.1% or less. At high prices, she instead prefers to buy the asset if her private expected return is −7.3% or greater.\(^{31}\) Thus, skewness in returns is sufficient to cause individuals with prospect theory preferences to trade against economically significant private information even in more general environments than those considered in the theory and experiment.

5 Discussion

Herding and contrarian strategies are of interest to financial economists because they don’t reveal private information to the market, reducing the informational efficiency of prices. These strategies are defined by imitation (or anti-imitation), relying on the observation of past histories of actions and/or prices. In this paper, however, I have shown both theoretically and via a laboratory experiment that strategies which very much look like herding or contrarian strategies emerge from preferences. The strategies therefore depend only upon

\(^{30}\)Because \( \lambda = 1 \) and the decision weight functions are the same for losses and gains, the utility from buying is equal to the opposite of the utility from selling. Therefore, indifference between buying and selling also implies indifference with not trading.

\(^{31}\)The 95% confidence intervals do not overlap indicating the difference across high and low prices is significant.
past price trends to the extent to which these trends determine future returns. In the experiment, I control for social and belief-based causes of behavior, yet observe frequent herding and contrarian-like behavior, consistent with a preference-based explanation. Among the two-thirds of subjects whose behavior is more consistent with prospect theory than expected utility, the majority of subjects exhibit a preference for negative skewness which generates herding-like strategies that are of particular concern due to their destabilizing effect on markets (De Long et al. (1990)).

The finding that herding-like strategies are more frequent than contrarian-like strategies is opposite to the findings of the previous experimental papers of Cipriani and Guarino (2005,2009) and Drehmann, Oechssler, and Roider (2005). However, a notable difference is that, in these experiments, subjects trade one at a time so that beliefs about the strategies of previous subjects are important. As Cipriani and Guarino (2005) and Drehmann, Oechssler, and Roider (2005) both show, if subjects believe previous trades do not always correspond to private information, then prices are perceived to be too extreme, potentially making it optimal to act in a contrarian manner (these beliefs could be due to believing other subjects make mistakes or believing that they are acting in a herding-like manner). Confirming their intuition, when I remove the need to form beliefs about others’ strategies, contrarian behavior is reduced (but does not disappear entirely). The contribution of Drehmann, Oechssler, and Roider (2005) is interesting for another reason - they run their experiment with financial market professionals rather than undergraduate subjects, demonstrating that herding and contrarianism are not restricted to undergraduates.

Interpreted as a preference for negative skewness, the dominance of herding-like behavior may be surprising because thinking about skewness may bring to mind participation in activities with positive skewness (gambling, lottery tickets, etc.). But, a negatively-skewed activity provides frequent positive feedback (small, positive rewards interspersed with larger losses) which may tend to reinforce its usage. With such feedback, it may be difficult to recognize that one is in fact, on average, facing a loss, and even if recognized, it may be difficult to forgo the frequent rewards.

Evidence from real-world markets has tended to provide evidence of preferences for posi-

---

32 I'm not arguing that preferences are the only source of herding and contrarian behavior. In particular, in the laboratory setting, the data generating process is known, but it is entirely plausible that in actual markets where it is not, overextrapolative beliefs may play a role (Greenwood and Shleifer (2014)).

33 Trades in financial markets that result in negatively-skewed returns include shorting volatility and selling options, among others. In fact, exchange-traded products have been created to directly cater to people’s desire to trade inverse volatility (for example, VelocityShares Daily Inverse VIX Short-Term ETN and ProShares Short VIX Short-Term Futures ETF). Examples of negatively-skewed, non-financial activities with active participation include adrenaline sports such as skydiving or free climbing, as well as criminal activity (thanks to my colleagues, Larry Harris and Fernando Zapatero, for these examples).
tive skewness (Kumar (2009), Green and Hwang (2012), Eraker and Ready (2015)). However, this evidence reflects selection into these markets, as documented by Kumar (2009). In experiments, instead, the findings represent a random sample of (a particular) population. In line with my finding, Huber, Palan, and Zeisberger (2017) document that positively-skewed assets are more likely to be perceived as risky (having a higher probability of a loss) and thus trade at lower prices than negatively-skewed assets. Given the inherent heterogeneity of preferences, we should not be surprised to find that a single type of skewness is not preferred by everyone.

A broader theme suggested by this paper is that prospect theory preferences can play a critical role in market and game experiments. In strategic environments, risk-neutral, expected utility preferences are often assumed, citing the Rabin critique (Rabin (2000)). But, Rabin’s critique rules out risk aversion only within an expected utility framework. Given that many subjects exhibit reference-dependence in the broad literature on estimating risk attitudes, it seems likely that these types of preferences also play a role in determining behavior in games and markets. An interesting avenue for future research is to study not only how subjects’ own reference-dependence affects their behavior, but also whether or not they form beliefs about, and respond to, other subjects’ reference-dependent behavior.

References


34 Examples of experimental environments in which subjects face decisions with skewed payoffs so that preferences for skewness may be important include: herding in the absence of prices (e.g. Goeree et al. (2007)), overpricing and bubbles (e.g. Palfrey and Wang (2012)), markets with private information (e.g. Brocas et al. (2014)), and common-value auctions (e.g. Charness and Levin (2009)).


32


Appendices

A. Omitted Proofs

Proof of Lemma 1:
For convenience, I refer to investors with favorable signals \( s_t = 1 \) as type 1, and investors with unfavorable signals \( s_t = 0 \) as type 0. I provide the proof that a type 1 risk-neutral investor buys. The proof that a type 0 risk-neutral investor sells is symmetric.
I first claim that if a type 1 informed investor (risk-neutral or CPT) sells at some \( p_t \) with positive probability in equilibrium, then a type 0 investor must also sell at \( p_t \) with probability one. For the CPT investor, this fact follows from inspecting equations (5) and their counterparts for a type 0 investor. For risk-neutral investors, if a type 1 investor sells with positive probability, then it must be the case that \( B_t - b^1_t \geq b^1_t - A_t \) so that she weakly prefers selling to buying. Rearranging, \( B_t + A_t \geq 2b^1_t > 2b^0_t \) so that the type 0 investor must strictly prefer selling. Similarly, if a type 1 investor weakly prefers selling to not trading, the type 0 investor must strictly prefer the same.
Given that type 0 investors must sell if type 1 investors sell, it follows that a sell trade either reveals no information or negative information. Therefore, the bid price must be weakly less than the public belief, \( B_t \leq p_t \) which implies that a type 1 risk-neutral investor will never sell because her expected profit is negative: \( B_t - b^1_t < B_t - p_t \leq 0 \) (using \( b^1_t > p_t \)). I now show that she also never abstains.
Using the formula for the ask price, (4), a type 1 risk-neutral investor prefers buying over abstaining if
\[
\left( \frac{p_t}{p_tq+(1-p_t)(1-q)} \right) > \frac{q}{1-q} > \frac{p_tPr(a_t=buy|V=1)}{p_tPr(a_t=buy|V=1)+(1-p_t)Pr(a_t=buy|V=0)} \tag{7}
\]
where \( Pr(a_t = buy|V = 1) \) and \( Pr(a_t = buy|V = 0) \) depend upon the equilibrium strategies of the informed investors,
\[
Pr(a_t = buy|V = 1) = \frac{1-\mu + \mu\gamma\beta^{RN}(V = 1) + \mu(1-\gamma)\beta^{PT}(V = 1)}{2} \tag{8}
Pr(a_t = buy|V = 0) = \frac{1-\mu + \mu\gamma\beta^{RN}(V = 0) + \mu(1-\gamma)\beta^{PT}(V = 0)}{2}
\]
and $\beta^{RN}(V = x)$ and $\beta^{PT}(V = x)$, $x \in \{0, 1\}$ are the probabilities of observing buy orders from risk-neutral and CPT investors, conditional on $V = x$, respectively. $\beta^{PT}(V = 1) = q\beta^{1, PT} + (1 - q)\beta^{0, PT}$ and $\beta^{PT}(V = 0) = (1 - q)\beta^{1, PT} + q\beta^{0, PT}$ where $\beta^{0, PT}$ is the probability that the market maker believes a CPT investor with $s_t = 1$ buys. The right-hand side of equation (7) can be shown to be strictly increasing in $\beta^{1, PT}$ and strictly decreasing in $\beta^{0, PT}$ so that it can be bounded above by the case of $\beta^{1, PT} = 1$ and $\beta^{0, PT} = 0$.

Furthermore, if a type 1 risk-neutral investor were to abstain, then a type 0 risk-neutral investor must also abstain (this fact is established in the same manner as the fact that, if a type 1 risk-neutral investor sells, then a type 0 risk-neutral investor must also sell), so that $\beta^{RN}(V = 1) = \beta^{RN}(V = 0) = 0$. Thus, if a type 1 risk-neutral investor were to abstain, we must have

$$\frac{q}{1 - q} \leq \frac{1 - \mu}{2} \cdot (1 - \gamma)q + \mu(1 - \gamma)(1 - q)$$

But, this inequality never holds for $\mu < 1$. Conversely, a type 1 risk-neutral investor strictly prefers buying to abstaining (so that $\beta^{RN}(V = 1) = q$ and $\beta^{RN}(V = 0) = 1 - q$) because a type 0 risk-neutral investor can never buy, just as a type 1 risk-neutral investor can never sell) if

$$\frac{q}{1 - q} > \frac{1 - \mu}{2} \cdot \gamma q + \mu(1 - \gamma)(1 - q)$$

This inequality holds for all $\mu < 1$. □

**Proof of Theorem 1:**

Part 1 follows directly from the assumption that the market maker faces perfect competition and Bayes’ rule.

Part 2 is proven in Lemma 1.

Part 3a. Given $\alpha = \delta$, the optimality conditions for a type 1 CPT investor, (5), become

\[
\begin{align*}
\text{buy if } 1 & \geq \lambda \frac{1 - q \cdot Pr(a_t = \text{buy}|V = 1)}{q \cdot Pr(a_t = \text{buy}|V = 0)}^\alpha \\
\text{sell if } 1 & \leq \frac{1}{\lambda} \frac{1 - q \cdot Pr(a_t = \text{sell}|V = 1)}{q \cdot Pr(a_t = \text{sell}|V = 0)}^\alpha
\end{align*}
\]

(9)

From (9), and the corresponding equations for a type 0 investor (in which the ratios of $q$ and $1 - q$ are inverted), we see that whether or not an investor trades is independent of the current public belief.

I first show that a type 0 investor can never buy with positive probability. Using the equilibrium strategies of the risk-neutral investor, we can write the probabilities of observing a buy conditional on $V = 1$ and $V = 0$ as

\[
\begin{align*}
Pr(a_t = \text{buy}|V = 1) & = \frac{1 - \mu}{2} + \mu \gamma q + \mu(1 - \gamma)\beta^{PT}(V = 1) \\
Pr(a_t = \text{buy}|V = 0) & = \frac{1 - \mu}{2} + \mu \gamma(1 - q) + \mu(1 - \gamma)\beta^{PT}(V = 0)
\end{align*}
\]

where, as in the proof of Lemma 1, $\beta^{PT}(V = 1) = q\beta^{1, PT} + (1 - q)\beta^{0, PT}$ and $\beta^{PT}(V = 0) = (1 - q)\beta^{1, PT} + q\beta^{0, PT}$.

Now, as argued in the proof of Lemma 1, if a type 0 investor buys, then so must a type 1 investor. This fact implies $\beta^{1, PT} \geq \beta^{0, PT}$ which in turn implies that the ratio $\frac{Pr(a_t = \text{buy}|V = 1)}{Pr(a_t = \text{buy}|V = 0)}$ is bounded below by $\beta^{1, PT} = \beta^{0, PT}$, because this ratio is increasing in $\beta^{1, PT}$ and decreasing in $\beta^{0, PT}$. Therefore, for a type 0 investor to buy, we must have
\[ 1 \geq \lambda \left( \frac{q}{1-q} \frac{1-q}{\frac{1-q}{1-q} + \mu p(1-q) + \mu(1-q)q \beta^{1,PT}} \right)^\alpha. \] But, the ratio inside the the parentheses is strictly greater than one for any \( \beta^{0,PT} \), so does not hold for any \( \lambda \geq 1 \).

Given that a type 0 investor never buys and, as can be shown similarly, a type 1 investor never sells, we are left to determine the conditions under which investors trade according to their private information, and when they do not trade. Substituting the probabilities of observing a buy into the first equation of (9) (using the fact that a type 0 investor never buys \( (\beta^{0,PT} = 0) \)), we have that a type 1 investor buys if

\[
\text{buy if } 1 \geq \lambda \left( \frac{1-q}{q} \frac{1-q}{\frac{1-q}{1-q} + \mu p(1-q) + \mu(1-q)q \beta^{1,PT}} \right)^\alpha.
\]

For \( \lambda \) sufficiently large, the investor will not buy. Setting \( \beta^{1,PT} = 0 \), we can find the cutoff value of \( \lambda \)

\[
1 \leq \lambda \left( \frac{1-q}{q} \frac{1-q}{\frac{1-q}{1-q} + \mu p(1-q) + \mu(1-q)q} \right)^\alpha \quad \Leftrightarrow \quad \lambda \geq \left( \frac{1-q}{q} \frac{1-q}{\frac{1-q}{1-q} + \mu p(1-q) + \mu(1-q)q} \right)^\alpha \quad \equiv \quad \bar{\lambda}
\]

For \( \lambda \) sufficiently small, the investor will buy with probability one. Setting \( \beta^{1,PT} = 1 \), we can find the cutoff value of \( \lambda \)

\[
1 \geq \lambda \left( \frac{1-q}{q} \frac{1-q}{\frac{1-q}{1-q} + \mu p(1-q) + \mu(1-q)q} \right)^\alpha \quad \Leftrightarrow \quad \lambda \leq \left( \frac{1-q}{q} \frac{1-q}{\frac{1-q}{1-q} + \mu p(1-q) + \mu(1-q)q} \right)^\alpha \quad \equiv \quad \lambda
\]

Simple algebra shows that \( \bar{\lambda} > \lambda > 1 \) for all parameterizations. Finally, for intermediate values of \( \lambda \), the investor mixes between buying and not trading such that the ask price makes him indifferent between the two. The mixing probability, \( \beta^{1,PT} \), satisfies

\[
1 = \lambda \left( \frac{1-q}{q} \frac{1-q}{\frac{1-q}{1-q} + \mu p(1-q) + \mu(1-q)q} \right)^\alpha
\]

which has a unique solution for \( \beta^{1,PT} \). The conditions for a type 0 investor to always sell, not trade, and mix between not trading and selling are easily shown to be identical to those for a type 1 investor.

Part3b. Consider \( \delta > \alpha \). I first evaluate the decisions to buy for both types. Beginning with equation (5) and the formulae for the probability of observing a buy (8), the two equations governing buy decisions are given by

\[
s_t = 1: \left( \frac{p_t}{1-p_t} \right)^{\delta - \alpha} \geq \lambda \left( \frac{1-q}{q} \frac{1-q}{\frac{1-q}{1-q} + \mu p(1-q) + \mu(1-q)q \beta^{1,PT}} \right)^\alpha
\]

\[
s_t = 0: \left( \frac{p_t}{1-p_t} \right)^{\delta - \alpha} \geq \lambda \left( \frac{1-q}{q} \frac{1-q}{\frac{1-q}{1-q} + \mu p(1-q) + \mu(1-q)q \beta^{1,PT}} \right)^\alpha
\]

Because the left-hand side of each equation is an unbounded function of the public belief and the right-hand side of each is independent of the public belief, we immediately see that both types must buy for sufficiently high public beliefs. Furthermore, because the ask price always exceeds the bid price, from (3), we can see that each type of CPT investor must not trade for some range of public beliefs.\(^{35}\) It remains to characterize behavior in the transition

\(^{35}\)The ask price always exceeds the bid price because risk-neutral investors always trade according to their
region from not trading to buying. From the proof of Lemma 1, we know the type 1 investor must buy with probability one if the type 0 investor buys with positive probability, so that the transition regions of each type cannot overlap. Thus, in the transition region of the type 1 investor, the type 0 investor is buying with probability zero such that the type 1 investor buys if
\[
\left( \frac{p_t}{1 - p_t} \right)^{\delta - \alpha} \geq \lambda \left( \frac{1 - q}{q} \right)^{\delta} \left( \frac{1 - \mu}{2} + \mu \gamma q + \mu (1 - \gamma) \beta_{1, PT} \right)^{\alpha}
\]
(11)
The right-hand side of (11) is monotonically increasing in \( \beta_{1, PT} \), which ensures a unique fixed point between the market maker’s belief and the action of the type 1, CPT investor. In particular, as the public belief increases, \( \beta_{1, PT} \) must increase such that the equation holds with equality: the type 1 investor mixes with \( \beta_{1, PT} \in (0, 1) \). Intuitively, the more that the type 1 investor buys, the more information revealed by her purchase, the higher the ask price, and the less profitable it is to buy. The lower and upper bounds of the transition regions, \( p^1 \) and \( p^1 \), satisfy
\[
\left( \frac{p^1}{1 - p^1} \right)^{\delta - \alpha} = \lambda \left( \frac{1 - q}{q} \right)^{\delta} \left( \frac{1 - \mu}{2} + \mu \gamma q \right)^{\alpha}
\]
(12)
In the transition region of the type 0 investor, the type 1 investor buys with probability one so that the type 0 investor buys if
\[
\left( \frac{p_t}{1 - p_t} \right)^{\delta - \alpha} \geq \lambda \left( \frac{q}{1 - q} \right)^{\delta} \left( \frac{1 - \mu}{2} + \mu \gamma q + \mu (1 - \gamma) \beta_{0, PT} \right)^{\alpha}
\]
(13)
Contrary to the case of the type 1 investor, the right-hand side of (13) is monotonically decreasing in \( \beta_{0, PT} \). Thus, for a given public belief, multiple fixed points between the market maker’s belief and the type 0 CPT investor’s action may exist. In particular, in the transition region, there exists an equilibrium in which the type 0 mixes with \( \beta_{0, PT} \in (0, 1) \) such that (13) holds with equality. The mixing probability, \( \beta_{0, PT} \), decreases as the public belief increases. But, if (and only if) this mixed strategy equilibrium exists, we also have equilibria with \( \beta_{0, PT} = 0 \) (the right-hand side is larger than when mixing, so not buying is a best response) and \( \beta_{0, PT} = 1 \) (the right-hand side is smaller, so buying is a best response). Intuitively, if the type 0 investor buy, she pools with the type 1 investor which decreases the ask price because no information is revealed, making buying more profitable. The lower and upper bounds of type 0’s transition region, denoted \( 1 - \bar{p}^0 \) and \( 1 - \bar{p}^0 \) are given by
\[
\left( \frac{1 - \bar{p}^0}{\bar{p}^0} \right)^{\delta - \alpha} = \lambda \left( \frac{q}{1 - q} \right)^{\delta} \left( \frac{1 - \mu}{2} + \mu \gamma q + \mu (1 - \gamma) \right)^{\alpha}
\]
private information and CPT investors’ trades either reveal their private information or no information. We cannot have only a type 0 investor buying, for example, because the proof of Lemma 1 shows that the type 1 investor buys whenever the type 0 investor does (in which case a buy reveals no information).
\[ \left( \frac{1 - p^0}{p^0} \right)^{\delta - \alpha} = \lambda \left( \frac{q}{1 - q} \right)^{\delta} \left( \frac{1 - \frac{\mu}{2} + \mu q}{1 - \frac{\mu}{2} + \mu (1 - q)} \right)^{\alpha} \]  

Inspecting (12) and (14), we can see that the transition region for type 0 investors always occurs at public beliefs greater than one-half (the right-hand side of the equation governing \( 1 - p^0 \) is always greater than one so that we must have \( 1 - p^0 > \frac{1}{2} \)), but the transition region for type 1 investors can occur at lower public beliefs.

For the decision to sell, the two equations of interest are

\[ s_t = 1 : \left( \frac{p_t}{1 - p_t} \right)^{\delta - \alpha} \leq \frac{1}{\lambda} \left( \frac{1 - q}{q} \right)^{\delta} \left( \frac{1 - \frac{\mu}{2} + \mu q (1 - q) + \mu (1 - \gamma)}{1 - \frac{\mu}{2} + \mu q (1 - q) + \mu (1 - \gamma)} \right)^{\alpha} \]

\[ s_t = 0 : \left( \frac{p_t}{1 - p_t} \right)^{\delta - \alpha} \leq \frac{1}{\lambda} \left( \frac{1 - q}{q} \right)^{\delta} \left( \frac{1 - \frac{\mu}{2} + \mu q (1 - q) + \mu (1 - \gamma)}{1 - \frac{\mu}{2} + \mu q (1 - q) + \mu (1 - \gamma)} \right)^{\alpha} \]

where \( I \) denote the probabilities with which type 0 and type 1 CPT investors sell, \( \eta^{0,PT} \) and \( \eta^{1,PT} \), respectively. Note the symmetry between the sell decision of the type 1 investor and the buy decision of the type 0 investor, and that between the sell decision of the type 0 investor and the buy decision of the type 1 investor. The problems are in fact identical to no trade in the same region that type 1 investors transit from sell to buy, except that the transition regions occur over symmetric intervals: \( (p^0, 1 - p^0) \) for the type 1 investor and \( (1 - p^1, 1 - p^1) \) for the type 0 investor, where \( p^0 < 1 - p^1 \). At sufficiently low public beliefs, both types of investors sell. As the public belief increases, the type 1 investor first transitions to not selling (over \( (p^0, 1 - p^0) \)), followed by the type 0 investor (over \( (1 - p^1, 1 - p^1) \)).

For \( \delta < \alpha \), the only difference is a relabeling of the transition regions. Type 0 investors now transition from buy to no trade in the same region that type 1 investors transition from sell to no trade in the \( \delta > \alpha \) case, and similarly for the other transitions, as illustrated in Figure 2. This duality is easily verified by comparing the inequalities that govern each transition.

For part iv), \( 1 - p^0 > p^1 \) for \( \delta > \alpha \) follows from the statement in the proof of Lemma 1 that if the type 0 investor buys with positive probability, the type 1 investor must buy with probability one. Similarly, \( 1 - p^1 < p^0 \) for \( \delta < \alpha \). Lastly, we must show \( p^0 < p^1 \). Using (12) and (14), this inequality is equivalent to

\[ \frac{1}{\lambda} \left( \frac{1 - q}{q} \right)^{\delta} \left( \frac{1 - \frac{\mu}{2} + \mu q (1 - q) + \mu (1 - \gamma)}{1 - \frac{\mu}{2} + \mu q (1 - q) + \mu (1 - \gamma)} \right)^{\alpha} \leq \lambda \left( \frac{1 - q}{q} \right)^{\delta} \left( \frac{1 - \frac{\mu}{2} + \mu q (1 - q) + \mu (1 - \gamma)}{1 - \frac{\mu}{2} + \mu q (1 - q) + \mu (1 - \gamma)} \right)^{\alpha} \]

which is more easily satisfied for higher \( \lambda \), so take \( \lambda = 1 \). Then, the inequality becomes

\[ \frac{1}{2} \mu q (1 - q) + \mu (1 - \gamma) < \frac{1}{2} \mu q (1 - q) + \mu (1 - \gamma) \]

\[ \iff 2 \left( \frac{1}{2} \mu q (1 - q) + \mu (1 - \gamma) \right) < 0 \]

which holds for all parameterizations. □

**B. Prospect Theory Parameters**

When trades take place at a single price, \( p_t \), the optimal strategy of a CPT investor with a favorable private signal (from (5)) becomes
\begin{align*}
\text{buy if } & \left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} \geq \lambda \left(\frac{1-q}{q}\right)^{\delta} \\
\text{sell if } & \left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} \leq \frac{1}{\lambda} \left(\frac{1-q}{q}\right)^{\delta}
\end{align*}

(15)

where \( t \) is the time of trade (after \( t - 1 \) public signal draws). For an investor with an unfavorable signal, the ratio of \( 1 - q \) to \( q \) is again inverted in each inequality. From (15), it is clear that one can define two threshold prices (that replace the transition regions that exist when there is asymmetric information) at which behavior transitions. The equilibrium is otherwise identical to that described in Theorem 1.

I now use (15) to establish that the model only has two degrees of freedom and then to find the set of pairs of threshold prices that it can support. I consider an investor with a favorable signal - the thresholds for an investor with an unfavorable signal follow by symmetry around \( p_t = \frac{1}{2} \).

The price depends only on the difference in the number of favorable and unfavorable public signals. Denoting the difference, \( k \), we have \( p_t = \frac{q^k}{q^k + (1-q)^k} \). (15) can then be written

\begin{align*}
\text{buy if } & \left(\frac{q}{1-q}\right)^k \geq \lambda \\
\text{sell if } & \left(\frac{q}{1-q}\right)^k \leq \frac{1}{\lambda}
\end{align*}

If we define \( \lambda' \equiv \lambda^{\frac{1}{1-q}} \) and \( \delta' \equiv \frac{\delta-\alpha}{\delta} \), we can rewrite these conditions as

\begin{align*}
\text{buy if } & \left(\frac{q}{1-q}\right)^{k+\frac{1}{\delta'}} \geq \lambda' \\
\text{sell if } & \left(\frac{q}{1-q}\right)^{k+\frac{1}{\delta'}} \leq \frac{1}{\lambda'}
\end{align*}

which establishes that the model only has two degrees of freedom, \( \lambda' \) and \( \delta' \). One can solve for these parameters in terms of the two threshold differences in public signals, \( k^0 \) and \( k^1 \), at which behavior is observed to transition, resulting in

\[ \lambda' = \sqrt{\left(\frac{q}{1-q}\right)^{k^1-k^0}} \]

(16)

and

\[ \delta' = \frac{-2}{k^0 + k^1} \]

(17)

\( k_0 \) and \( k_1 \) must satisfy several restrictions. First, from the theory, we require \( p^0 < p^1 \) which implies \( k^0 < k^1 \). Second, when \( \delta > \alpha \) such that herding occurs, we must have \( p^1 < 1 - p_0 \), which implies \( k^1 < -k^0 \). However, given upper bounds on \( \alpha \) and \( \delta \), we must have \( \delta' < 1 \) which, from (17), implies the tighter restriction, \( k^1 + k^0 < -2 \). When \( \alpha > \delta \) such that contrarianism occurs, we must have \( 1 - p^1 < p_0 \), and therefore \( -k^1 < k^0 \). The restrictions on \( \alpha \) and \( \delta \) make \( \delta' \) negative in this case so that (17) leads to the same constraint. I impose
these restrictions when generating the set of possible trade patterns that prospect theory is capable of explaining.

Focusing on the more common herding types, if we assume no probability weighting, \( \delta = 1 \), we can relate \( \lambda' \) and \( \delta' \) to the primitives of the model. Specifically, \( \alpha = 1 - \delta' \) and \( \lambda = \lambda' \delta' \). I make this assumption when performing the calibration exercise because, with non-binary asset values, both \( \delta \) and \( \alpha \) must be specified (not just their difference).

C. Alternative Preference Models

C.1. Expected Utility

Theorem B1 demonstrates that expected utility can not generate abstention and unresponsive behavior within the same subject, limiting its ability to explain the experimental findings.

**Theorem B1:** If CPT investors are replaced with investors with standard expected utility preferences:
(i) if risk-averse, herding and contrarian-like behaviors do not occur.
(ii) if risk-seeking, abstention does not occur.

**Proof of Theorem B1:**
Under expected utility, an investor with continuous utility function \( u(x) \) and private belief, \( b_t \), will

\[
\text{buy if } b_t u(1 - A_t) + (1 - b_t) u(-A_t) \geq u(0) \\
\text{sell if } b_t u(B_t - 1) + (1 - b_t) u(B_t) \geq u(0)
\]

and otherwise abstain from trading.\(^{36}\) As in the main model, it is possible to show that because of uninformed investors, we must have \( b_t > A_t \) (favorable signal) or \( b_t < B_t \) (unfavorable signal), in which case risk-neutral investors trade according to their private information as shown in Lemma 1. Consider an investor that is not risk-neutral then, and assume she has a favorable signal (symmetric arguments hold for unfavorable signals).

(i) If risk-averse, then not trading is always preferable to selling, so herding and contrarian-like behaviors are not possible: \( b_t u(B_t - 1) + (1 - b_t) u(B_t) < u(B_t(B_t - 1) + (1 - b_t)(B_t)) = u(B_t - b_t) < u(0) \). The first inequality holds because utility is strictly concave and the second because \( b_t > A_t > B_t \) with a favorable signal.

(ii) If risk-seeking, then buying is always preferable to not trading so abstention is not possible: \( b_t u(1 - A_t) + (1 - b_t) u(-A_t) > u(b_t(1 - A_t) + (1 - b_t)(-A_t)) = u(b_t - A_t) > u(0) \). The first inequality holds because utility is strictly convex and the second because \( b_t > A_t \) with a favorable signal. \( \blacksquare \)

\(^{36}\) I have normalized initial wealth to zero without loss of generality.
C.2. Symmetry

In Section 3.2.2, I consider the ‘best symmetric’ model, a model that only imposes symmetry around a price of one-half. Here, I discuss the generality of this model in more detail.

The model imposes only the requirement that if a subject buys, sells, or abstains at a price, $p$, when she has a particular private signal, then she must sell, buy, and abstain (respectively) at a price of $1 - p$ with the opposing private signal. This requirement is very natural because of the symmetry in payoffs around a price of one-half. Buying at a price of $p$ with a particular belief, $b$, provides a binary gamble which returns $1 - p$ with probability $b$ and $-p$ with probability $1 - b$. Selling at a price of $1 - p$ with a belief, $b'$, provides a binary gamble which returns $1 - p$ with probability $1 - b'$ and $-p$ with probability $b'$. Thus, if $b = 1 - b'$, which is the case with Bayesian updating, the opposing actions at symmetric prices with opposing signals provide identical gambles.\(^{37}\)

Any model that respects Bayesian updating and has choices that depend only on monetary payoffs must therefore respect the symmetry requirement. In particular, all utility-based models, including those with probability weighting and reference-dependence fall into this category. Even typical models of non-Bayesian updating, such as conservativeness (over-weighting the prior) or overweighting one’s private signal would respect symmetry. To capture asymmetric behavior, a model necessarily has to introduce some asymmetric primitive. The asymmetry could act through utility, such a pure preference for buying over selling, or through beliefs, such as updating differently for favorable versus unfavorable signals. However, constructing such a model would be a post hoc exercise designed to fit the data: I’m not aware of any off-the-shelf model that microfounds such asymmetries.

D. Expectations-Based Reference Points

Here, I show that the expectations-based reference point preferences of Koszegi and Rabin (2006,2007) are incapable of generating herding-like behavior in the laboratory environment, even when extended to allow for decision weights. Koszegi and Rabin (2007) consider multiple models of reference point formation. First, they consider the case in which expectations are taken as given, suggesting that this occurs when the outcome is realized shortly after a decision is made. For this case, they define an unacclimating personal equilibrium (UPE) and an associated refinement, preferred personal equilibrium (PPE), in which the decision-maker makes a choice she is willing to follow through on. Second, they consider the case in which outcomes are realized long after a decision is made, defining a choice-acclimating personal equilibrium (CPE) where the decision-maker’s reference point has time to acclimate to the choice they made. Without taking a stance on which concept is appropriate in the context here, I show that herding is not possible in either a PPE or CPE.\(^{38}\)

\(^{37}\)With a favorable signal at $p$ and unfavorable signal at $1 - p$, $b = \frac{pq}{pq + (1-p)(1-q)} = 1 - \frac{(1-p)(1-q)}{pq + (1-p)(1-q)} = 1 - b'$, and similarly with the opposite signals.

\(^{38}\)The UPE concept, when also allowing for probability weighting, is more difficult to work with analytically. Herding-like behavior may be possible in a UPE under some conditions, but it nevertheless does not survive the PPE refinement that Koszegi and Rabin (2007) suggest.
A necessary condition for a decision to be either a PPE or a CPE is that it must be optimal when the outcomes it induces become the reference point. Consider the decision to buy in the context of the model. Both a PPE and a CPE require $U(B|B) \geq U(NT|NT)$ and $U(B|B) \geq U(S|S)$ where $U(F|G) = \int \int u(w|r)dG(r)dF(w)$ is the decision-maker’s expected utility. The expectation in this expression is over both the possible wealth levels (given by the distribution $F$) and the possible reference points (given by the distribution $G$) of a reference-dependent utility over wealth, $u(w|r)$. Koszegi and Rabin (2007) assume a reference-dependent utility function of the form, $u(w|r) = m(w) + \mu(m(w) - m(r))$. For tractability, and as assumed in many of the results of Koszegi and Rabin (2007), I assume $m(w) = w$. Note, however, that, under Koszegi and Rabin’s assumptions on $\mu$, the reference-dependent component of utility is S-shaped as in prospect theory. In addition, although not considered in their model, I allow for decision weights with the same functional form I assumed previously. Under these assumptions, we have

$$
U(NT|NT) = 1, \\
U(B|B) = \frac{b_1^\delta}{(b_1^\delta + (1 - b_1^\delta)\delta)}(1 - p_t) - \lambda \frac{(1 - b_1^\delta)\delta}{(b_1^\delta + (1 - b_1^\delta)\delta)^{\frac{1}{\delta}}} p_t \\
+ \frac{b_1^\delta}{b_1^\delta + (1 - b_1^\delta)\delta}(1 - p_t) + \frac{1 - b_1^\delta}{b_1^\delta + (1 - b_1^\delta)\delta}\lambda p_t \\
U(S|S) = -\lambda \frac{b_1^\delta}{(b_1^\delta + (1 - b_1^\delta)\delta)^{\frac{1}{\delta}}} (1 - p_t) + \frac{1 - b_1^\delta}{b_1^\delta + (1 - b_1^\delta)\delta}\lambda p_t \\
+ \frac{b_1^\delta}{(b_1^\delta + (1 - b_1^\delta)\delta)^{\frac{1}{\delta}}} (1 - p_t) \lambda \frac{b_1^\delta}{b_1^\delta + (1 - b_1^\delta)\delta}\lambda p_t
$$

In each of the latter two expressions, the first two terms correspond to the direct utility from wealth and the third term corresponds to the additional utility relative to the reference point. Importantly, in both expressions, utility relative to the reference point only derives from the cases in which the reference outcome and the actual outcome differ (e.g. for the buy case, the reference is $2 - p_t$ and the outcome is $1 - p_t$, or vice versa). Thus, the difference is always equal to one so that the contribution to utility from reference-dependence is the same for both a buy and a sell and drops out in the comparison of the two. The S-shaped utility function therefore plays no role, which is the reason herding-like behavior is not possible. To see this, consider an investor with an unfavorable signal. For buying to be optimal, we require

$$
\Leftrightarrow (b_1^\delta)^{\delta}(1 - p_t) - \lambda(1 - b_1^\delta)^{\delta}p_t \geq -\lambda(b_1^\delta)^{\delta}(1 - p_t) + (1 - b_1^\delta)^{\delta}p_t \\
\Leftrightarrow (b_1^\delta)^{\delta}(1 - p_t)(1 + \lambda) \geq (1 - b_1^\delta)^{\delta}p_t(1 + \lambda) \tag{19}
$$

(19) requires $p_t < \frac{1}{2}$ to hold given that $\delta < 1$, so that an investor with an unfavorable signal can only buy at low prices, implying contrarian-like behavior occurs, but herding-like behavior does not. Therefore, the expectations-based reference point of Koszegi and Rabin (2007) is incapable of explaining the most frequent type of behavior observed in the experiment.
E. Instructions

I have included the instructions for the NO INFERENCE treatment below. The instructions for the NO SOCIAL treatment are identical except that the second paragraph under 'A Valuable Clue' is removed.
Instructions for Trading Experiment

You are about to participate in an experiment in the economics of decision-making. In the experiment you will make decisions in several repetitions of a simulated trading game. If you follow these instructions carefully and make good decisions, you can earn a CONSIDERABLE AMOUNT OF MONEY!

In each trading game, you will be given information about an artificial stock before having the opportunity to buy or sell it (or do neither). When the game is completed, a new game will begin. There will be 30 games in all and when you are done, you will be paid according to the trading decisions you make in each game.

A screenshot of the Trading Page you will use to trade the stock is shown below. The rules of the game shown below it describe the game and the interface in detail. These rules always appear under the trading interface so that you can refer to them whenever you need to.

Please read the rules carefully and then press the Next button. You will be asked two questions to make sure you understand the rules before the games begin.
Rules

Basics

The currency for the trading game is *experimental currency units (ECUs)* which will be converted to dollars and paid to you as a bonus at the completion of all games. The conversion rate is 100 ECUs = $0.40.

Prior to each game, the computer will randomly choose whether the stock you can trade is worth 0 or 100 ECUs, with equal probability. Because the stock's value is randomly chosen each game, there is no dependence between its value in one game and its value in any other game.
You will be given 100 ECUs at the start of each game. These ECUs are **yours to keep** if you choose not to trade. If you choose to trade, however, you will have the opportunity to earn even more money.

**The Stock's Price**

In each game, you can trade at most one unit of the stock. As shown on the Trading Page, the stock's value is represented by one of two bins of balls. If the stock is worth 100 ECUs, there are 7 blue balls and 3 green balls in the bin. If the stock is worth 0 ECUs, there are 7 green balls and 3 blue balls.

The price of the stock you can trade at is set by the computer in each game. The initial price is set to 50 ECUs reflecting the fact that there is a 50% chance the stock is worth 0 and a 50% chance it is worth 100 ECUs.

Before your turn to trade the stock, the computer will randomly draw between 1 and 5 balls from the bin that corresponds to the stock (**with replacement**). After each ball is drawn, the computer will update the price of the stock, keeping it equal to the **expected value** of the stock given the information that the balls reveal (but no other information). The expected value is how much the stock is worth on average - sometimes it is worth 0 and sometimes it is worth 100 ECUs - but on average it is worth the price the computer sets given the information revealed by the balls. The stock's price can therefore provide you with information about whether the stock's value is 0 or 100 ECUs.

The price graph on the Trading Page shows the history of prices as each ball is drawn by the computer. When it is your turn, you can trade the stock at the most recent price. To see this price (or a past price) exactly, on the actual Trading Page you can hover over the point on the graph. (In the example in the screenshot, the current price is 50 ECUs.)

Given the current price, you can choose to buy, sell, or not trade the stock. If you buy the stock, you will get its value minus the price you pay for it. If you sell it, you will get its price, but will have to pay back its value. If we call the value of the stock, $V$, (either 0 or 100 ECUs), and the price of the stock, $P$, your trading profit is

- $V - P$, if you buy the stock
- $P - V$, if you sell the stock

Note importantly, that your trading profit can be positive or negative! For example, if you buy the stock at a price of 75 ECUs and it turns out to be worth 100 ECUs, you **GAIN** 25 ECUs. But, if you sell the stock for 75 ECUs and it turns out to be worth 100 ECUs, you **LOSE** 25 ECUs. Conversely, if the stock turns out to be worth 0 ECUs, you would lose 75 ECUs if you bought the stock and gain 75 ECUs if you sold it.

**A Valuable Clue**

To help you determine whether the stock is worth 0 or 100 ECUs, you will be shown a valuable clue: the color of one ball from the bin corresponding to the stock. **Only you** get to see this ball, so it does not affect the price the computer sets - you can trade at the most recent price as described above. After seeing the valuable clue, you may choose to buy, sell, or not trade the stock.
To help you interpret your valuable clue, an expert in probabilities will give you the true probability that the stock is worth 100 ECUs. This expert knows the price and your valuable clue, but not the stock value itself. This probability will be shown above where you are asked what trade you would like to make (not shown in the screenshot example).

Your Trading Decisions

In each game, you will actually make two decisions. You will be asked whether you would like to buy, sell, or not trade the stock before you see your valuable clue. You will first be asked how you’d like to trade if your valuable clue is a blue ball (as shown in the screenshot). After choosing buy, sell, or not trade, you will hit the Next button to confirm your trade. You will then be asked for how you’d like to trade if your valuable clue is a green ball. Then, the computer will draw a ball from the appropriate bin, show it to you, and your trading decision for that color will take place. In this way, it as if you first saw the ball and then made your trading decision.

Your Payment

After all 30 games have been completed, your ECUs in each game will be converted at a rate of 100 ECUs = $0.40. For example, if you have 150 ECUs after trading in a game, they are worth 150/100*$0.40=$0.60.