Herding and Contrarianism: A Matter of Preference

Chad Kendall^{*}

November 25, 2016

Abstract

Herding and contrarianism in financial markets produce informational inefficiencies when investors ignore their private information, instead following or bucking recent trends. I theoretically establish a preference-based link between the two behaviors: investors with prospect theory preferences follow one of the two strategies generically, depending only upon the relative strengths of their utility curvature and non-linearity in decision weights. The third component of prospect theory, loss aversion, further exacerbates informational efficiencies, causing traders to abstain. A laboratory experiment provides strong evidence in support of the model's theoretical predictions and shows that herding is by far more common than contrarianism.

1 Introduction

Informationally efficient financial markets in which prices reflect fundamental asset values are important for the real economy, signaling efficient investments and allocating capital efficiently (Bond, P., Edmans, A., and Goldstein, I. (2012)). To achieve informational efficiency, markets must aggregate diverse private information, requiring individual investors to follow trading strategies that are responsive to this information. If investors instead herd (unconditionally buy as prices rise), act contrarian (unconditionally sell as prices rise), or simply abstain from trading, informational inefficiencies arise. Since the seminal works of Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992), researchers have sought to understand these informationally inefficient strategies theoretically, empirically, and through

^{*}Department of Finance and Business Economics, Marshall School of Business, University of Southern California, 3670 Trousdale Parkway, Ste. 308, BRI-308, MC-0804, Los Angeles, CA 90089-0804 (e-mail chadkend@marshall.usc.edu). I would like to thank Cary Frydman and Ryan Oprea for valuable discussions as well as participants in seminars at USC Marshall brown bag and the Toulouse School of Economics.

laboratory experiments.¹ In this paper, I develop a model that ties all three sources of inefficiency to a single underlying mechanism - preferences.²

I consider the standard sequential trading model of Glosten and Milgrom (1985). 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 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 expected utility investors (Avery and Zemsky (1998)) is that each investor trades according to her private information: herding, contrarianism, and abstention are impossible because prices both aggregate information and impact payoffs (which is the main difference from models such as Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992) in which there are no market prices).³ Contrary to this result, however, experimental tests of the Glosten and Milgrom (1985) model (Cipriani and Guarino (2005,2009) and Drehmann, Oechssler, and Roider (2005)) find that these behaviors are common. I show that introducing investors with cumulative prospect theory (CPT) preferences provides a unifying explanation.

With CPT preferences, an individual investor generically (i.e. for almost all combinations of preference parameters) herds or acts contrarian as prices become more extreme. Furthermore, at intermediate prices the investor may abstain from trading. Thus, introducing standard non-expected utility preferences into one of the simplest trading models of private information simultaneously generates all three forms of inefficient trade behavior.

To understand this result, consider the three main differences between CPT and expected utility. First, in prospect theory, investors derive utility from gains and losses relative to a reference point rather than from final wealth. Investors' value functions are 'S-shaped' being risk-averse over gains and risk-seeking over losses. Second, investors apply non-linear decision weights to probabilities so that small probability events are over-weighted. Third, investors are loss-averse, the dis-utility from a loss is greater than the utility from a gain of equal magnitude. As shown by Kahneman and Tversky (1992), all three departures from expected utility are required to explain the observed patterns of choices over lotteries in laboratory experiments.

¹Devenow and Welch (1996) provide a survey of the older literature. Hirschleifer and Teoh (2003) is a more recent survey. See also Avery and Zemsky (1998) and Park and Sabourian (2011) for theoretical papers, Cipriani and Guarino (2014) for an empirical test of herding, and Cipriani and Guarino (2005,2009) and Drehmann, Oechssler, and Roider (2005) for laboratory experiments.

 $^{^{2}}$ Qin (2015) also uses preferences, specifically regret aversion, to generate herding in a setup very similar to mine. His model does not generate contrarian behavior, however.

³Avery and Zemsky (1998) produce herding and contrarianism in the presence of prices by adding additional sources of uncertainty. Park and Sabourian (2011) show that moving beyond two states of the world also allows herding and contrarianism to occur.

As has been noted previously (Barberis (2012)), the S-shaped value function of prospect theory gives investors a preference for negatively-skewed assets. Intuitively, risk-aversion over gains and risk-seeking over losses together induce a preference for purchasing an asset with zero expected value that has a high chance of a small gain, but a small chance of a large loss. With binary asset values, a preference for negatively-skewed assets translates to a propensity to buy assets that have increased in value. On the other hand, as noted by Barberis (2012) and Barberis and Huang (2008), non-linear decision weights induce a preference for positively-skewed assets. Because small probabilities are over-weighted, investors desire 'lottery-like' assets that have a small chance of a large payoff even if they provide zero expected value. The decision weights therefore induce a propensity to buy assets that have fallen in value. If either of these preferences for skewness are strong enough, investors can trade against private information, herding or acting contrarian depending upon which preference dominates.

Although the opposing effects of utility curvature and decision weights have been noted in previous applications of prospect theory, the model here is sufficiently tractable to provide analytic solutions for the first time.⁴ I find that, with the functional forms assumed by Kahneman and Tversky (1992), the utility curvature and decision weights exactly offset each other. Therefore, what drives the behavior of a particular individual to either herd or act contrarian is simply the difference between the two parameters in her utility function.

The third and final departure from expected utility, loss aversion, causes investors to not trade the asset at all. Intuitively, loss aversion causes an extreme form of risk aversion such that simply holding one's endowment is preferable to either buying or selling the asset.

Previous experimental tests of the Glosten and Milgrom (1985) model (Cipriani and Guarino (2005,2009) and Drehmann, Oechssler, and Roider (2005)) document *aggregate* behavior consistent with the model's predictions. However, the model predicts that behavior is specific to the *individual*. Furthermore, behavior in previous experiments could be due to strategic uncertainty or Bayesian errors, in addition to preferences. For these reasons, I design an experiment to (i) collect enough data to be able to characterize individual behavior, and (ii) strip out other potential sources of herding and contrarian behavior. To rule out these other explanations, I have subjects make decisions in an individual decision-making environment in which strategic uncertainty cannot be the cause of behavior. In a second treatment, I'm also able to rule out Bayesian errors, providing the cleanest test of the theory. Evidence from both treatments provides strong evidence for the model's predictions and, when Bayesian

 $^{^{4}}$ Barberis (2012), in his model of casino gambling, studies the opposing effects through numerical simulation and provides the basic intuition as to why utility curvature and decision weighting oppose each other.

errors play no role, roughly three quarters of subjects are better characterized by prospect theory than standard expected utility. Within these subjects, over 90% prefer herding to contrarianism.

Within the context of the model, herding is also a form of a momentum strategy in that traders chase past price trends. In this sense, the model suggests that preferences drive investors' decisions to follow momentum versus contrarian strategies. 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), Brozynski et al. (2003), Baltzer, Jank, and Smajlbegovic (2015), and Grinblatt et al. (2016)).

This paper contributes to the growing literature that applies prospect theory preferences to understanding 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 prospect theory 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 prospect theory can explain the popularity of casino gambling.

On the experimental side, this paper 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)), but more broadly makes the point that the preferences that subjects bring into the lab may be important determinants of behavior in the tasks they are asked to perform. Risk-neutral preferences are often assumed, citing the Rabin critique (Rabin (2000)), but this critique only rules out risk aversion in an expected utility framework. If, as in the broad literature on estimating risk attitudes, a majority of subjects exhibit reference-dependence, it seems likely that they would continue to do so when participating in games or markets.

The paper proceeds as follows. Section 2 describes the model. Section 3 provides an equilibrium characterization.studies the model theoretically and generates testable predictions. Section 4 describes the experimental design, hypotheses, and results. Section 5 discusses the generalizability of both the theoretical and experimental results.

2 Model

The model is a sequential trading model based on that of Glosten and Milgrom (1985). In each period t = 1, 2, ..., 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 $p_0 \in (0, 1)$. Upon arrival, an investor may either buy or sell short a single unit, or not trade. I denote the trade decision $a_t \in \{buy, sell, NT\}$, where NT stands for no trade. 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. 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 and 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 3.3) and arrive with the remaining probability, $\mu(1 - \gamma)$. Only the risk-neutral and prospect theory investors are informed, receiving private information upon arrival to the market. They receive a private, 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 prospect theory investors, I include risk-neutral investors in the model for several reasons. First, if one takes the stance that prospect theory is truly capturing preferences and not irrationality, then including both is logical because both have been observed in previous measurement experiments.⁵ On the other hand, if one feels that prospect theory preferences are irrational, then one can think of the risk-neutral investors as being more sophisticated than the prospect theory investors.⁶ Finally, we'll see that including risk-neutral investors allows (partial) information to be revealed by every

⁵Bruhin, Fehr-Duda, and Epper (2010) find that approximately 20% of their subjects are risk-neutral expected utility maximizers, but the remainder are better described by prospect theory preferences. They argue for this reason that both types should be included in applied theoretical work.

⁶Brozynski et al. (2003) find some evidence that less experienced fund managers are more likely to follow momentum and contrarian strategies than those with more experience. An interpretation in light of the model is that fund managers learn away their prospect theory behavior.

trade, ensuring dynamic price paths, as in real markets, whereas prices can stagnate in their absence.

3 Theory

3.1 Preliminaries

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 prospect theory investors, along with the bid and ask prices of the market maker, which depend upon his beliefs about these strategies. As usual, these beliefs, which are pinned down at every history due to the presence of the noise traders, 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.

Different definitions of herding and contrarian behavior have been used in the literature. I use definitions very similar to those in Park and Sabourian (2011) and Avery and Zemsky (1998), but that allow for departures from risk-neutrality.⁷ I state the definitions in terms of the public belief that the asset is valuable, $p_t = Pr(V = 1|H_t)$, which serves as a sufficient statistic for public information revealed prior to time t.

Definition 1 (Herding and Contrarianism): Given the public belief that the asset is valuable, p_t :

- 1. An informed investor herds if, independently of her signal, she (i) buys when $p_t > \frac{1}{2}$, or (ii) sells when $p_t < \frac{1}{2}$.
- 2. An informed investor trades in a contrarian manner if, independently of her signal, she (i) buys when $p_t < \frac{1}{2}$, or (ii) sells when $p_t > \frac{1}{2}$.

Informally, herding refers to ignoring one's private information and instead going with the consensus, and contrarianism refers to the opposite. A related concept is that of an information cascade, which occurs when no further information is revealed to the market and prices stagnate. I do not provide a formal definition here because, as we'll see in the following section, due to the presence of risk-neutral traders, information cascades do not

⁷I state the definition in terms of signals and actions, rather than expected values, because expected values do not directly map to actions when investors are not risk-neutral. A subtle difference arises when preferences, rather than information externalities, induce herd or contrarian behavior in that particular price paths are not necessary to induce such behavior. For this reason, I do not state the definitions in terms of investors' actions changing from what they would have done at t = 0, as in some previous definitions.

occur. I also use the term *unresponsive* to refer to an action that is independent of one's private information, encompassing both herding and contrarian behavior.

3.2 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 prospect theory investors.

Due to the assumption of perfect competition, the market maker earns zero profits in expectation. This zero-profit condition results in the market market posting 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. Intuitively, the ask price exceeds the public belief, p_t , because a buy decision reflects favorable private information, $s_t = 1$, 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 (Risk-neutral Investors): In any equilibrium, for all $p_t \in (0, 1)$, riskneutral 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, because risk-neutral investors arrive with positive probability and trade according to their private information, information is partially revealed in every period: an information cascade never occurs. This fact implies that, by the law of large numbers, the public beliefs and bid and ask prices converge to the true asset value in the limit as $t, T \to \infty$, as shown in Avery and Zemsky (1998).

3.3 Prospect Theory Investors

CPT differs from expected utility in that investors evaluate gains and losses relative to a reference point. The behavioral finance literature has tended to use the expected wealth from investing in a risk-free asset as the reference point (see Barberis and Huang (2008),

Barberis and Xiong(2009), and Li and Yang (2013)), with the interpretation that this is the amount of wealth an investor could have had without risk. Here, because there is no risk-free asset, I equivalently adopt the status quo as the reference point, which is also risk-free.^{8,9,10}

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 prospect theory 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 then given by $U(\mathcal{L}) = w^+(r)v^+(x) + w^-(1-r)v^-(y)$.¹¹

Given this utility function, I now derive the two main equations that characterize the behavior of a prospect theory investor. Given a private belief, $b_t = Pr(V = 1|H_t, s_t)$, a prospect theory investor prefers buying to not trading if

$$w^{+}(b_{t})v^{+}(1-A_{t}) + w^{-}(1-b_{t})v^{-}(-A_{t}) \ge 0$$
(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

$$w^{+}(1-b_{t})v^{+}(B_{t}) + w^{-}(b_{t})v^{-}(B_{t}-1) \ge 0$$
(2)

If neither equation (1) nor equation (2) is satisfied, then a prospect theory refrains from trading, preferring to keep her endowment.

The forms of equations (1) and (2) are sufficiently general that little can be said about

⁸An alternative for the reference point is expectations, as in Koszegi and Rabin (2006,2007). In their model, reference points can be stochastic and depend upon 'recent' beliefs. The assumption of status quo corresponds to 'surprise' in their model: one does not expect the availability of a trade. If, instead, one assumes expectations adapt to the decision made, what they call 'choice-acclimated expectations', I can show that traders either trade according to their private information or abstain, which is counterfactual to the experimental evidence I provide. Other interpretations of 'recent' beliefs are possible, but given that the status quo assumption is so successful in explaining the data, I do not pursue these possibilities here.

 $^{^{9}}$ I'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 the distinction can be important in a model in which investors make multiple trading decisions.

¹⁰The issue of narrow or broad framing (Barberis, Huang, and Thaler (2006)) is not important in the model given that only one 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'm assuming subjects use narrow framing, considering the experiment (and, in fact, each game) in isolation.

¹¹See Kahneman and Tversky(1992) for the more general formulation for any number of outcomes, as well as an axiomatic foundation for the preferences.

Figure 1: Examples of the Value and Probability Weighting Functions



Note: The figure illustrates the 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).

the behavior of the investor without imposing additional structure. I proceed by using the functional forms for the value and decision weight functions provided in the original work of Kahneman and Tversky (1992), because these are tractable, parsimonious, and appear to fit decisions over binary gambles reasonably well.¹² Specifically, I assume

$$v^+(x) = x^{\alpha}$$

 $v^-(y) = -\lambda(-y)^{\alpha}$

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 $\delta \in (0, 1]$.¹³ $\alpha \in (0, 1)$ 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, $\delta \in (0, 1)$ matches the experimental finding that subjects overweight low-probability events. Figure 1 illustrates examples of each function.

Substituting the functional forms into equations (1) and (2) results in the following op-

¹²Other functional forms, especially for the probability weighting function, have appeared in the literature. See Bruhin, Fehr-Duda, and Epper (2010) and the references therein.

¹³Kahneman 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 for a significant increase in tractability.

timal decisions given private belief b_t :

$$\begin{aligned} buy \ if \quad \left(\frac{b_t}{1-b_t}\right)^{\delta} &\geq \lambda \left(\frac{A_t}{1-A_t}\right)^{\alpha} \\ sell \ if \quad \left(\frac{b_t}{1-b_t}\right)^{\delta} &\leq \frac{1}{\lambda} \left(\frac{B_t}{1-B_t}\right)^{\alpha} \end{aligned} \tag{3}$$

Risk-neutral investors are a special case of prospect theory 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 need to consider how beliefs and prices are formed.

An investor with a favorable signal, $s_t = 1$, has a private belief conditional on the history and her private signal (denoted b_t^1) given by Bayes' rule:

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

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

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

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

$$A_{t} = \frac{p_{t}Pr(a_{t}=buy|V=1)}{p_{t}Pr(a_{t}=buy|V=1)+(1-p_{t})Pr(a_{t}=buy|V=0)}$$

$$B_{t} = \frac{p_{t}Pr(a_{t}=sell|V=1)}{p_{t}Pr(a_{t}=sell|V=1)+(1-p_{t})Pr(a_{t}=sell|V=0)}$$
(4)

Substituting the equations for her private belief and the bid and ask prices, for a trader with a favorable signal, (3) becomes

$$buy if \left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} \geq \lambda \left(\frac{1-q}{q}\right)^{\delta} \left(\frac{Pr(a_t=buy|V=1)}{Pr(a_t=buy|V=0)}\right)^{\alpha} \\ sell if \left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} \leq \frac{1}{\lambda} \left(\frac{1-q}{q}\right)^{\delta} \left(\frac{Pr(a_t=sell|V=1)}{Pr(a_t=sell|V=0)}\right)^{\alpha}$$
(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 α and δ 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. Set $\lambda = 1$ and 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. Prospect theory investors, on the other hand, do have an incentive to trade. With the simplification, equations (5) become

$$\begin{aligned} buy \ if \quad \left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} &\geq 1 \\ sell \ if \quad \left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} &\leq 1 \end{aligned}$$
(6)

so 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 (δ decreases), the propensity to buy decreases and the propensity to sell increases. Intuitively, an increase in the distortion 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 in which case the small probability is associated with a gain. In fact, for $\delta < \alpha$, the investor strictly prefers to sell the stock. This example represents a *preference for positive skewness* which is a consequence of prospect theory studied extensively in Barberis and Huang (2008).

Now consider an increase in the curvature of the value function (decrease in α). It is clear mathematically that we get exactly the opposite effect from an increase in the distortion of probabilities due to decision weights. Intuitively, as the curvature increases, the large probability of a small gain $(1 - p_t)$ if one buys becomes more valuable than the small probability of a large gain (p_t) if one sells, a simple consequence of risk-aversion. At the same time, the small probability of a large loss if one buys becomes more valuable than the large probability of a small loss if one sells due to risk-seeking. Both effects make buying more valuable than selling so that if $\delta > \alpha$, the investor always buys. The investor in this case exhibits a *preference for negative skewness*.

This simple example captures the countervailing forces of distortions due to decision weights and utility curvature. These intuitions carry over to the full equilibrium characterization I pursue in the following section. Importantly, the example 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 any effect is likely to be diminished by the countervailing force of the other aspect.

Finally, consider the role of loss aversion. An increase in λ reduces the range of public beliefs at which an investor is willing to trade, because it simultaneously makes each inequality in (5) more difficult to satisfy. The intuition here is simple: an increase in loss-aversion increases the dis-utility of losses, leaving the utility of gains unchanged. Because taking on either a long or short position in the stock can result in a loss, this change makes one more likely to stick with one's endowment. Perhaps surprisingly, however, loss aversion prevents trading only at intermediate public beliefs. Although the potential losses are larger at extreme public beliefs, one can always take the side of the trade that either minimizes the probability ($\delta < \alpha$) or the size ($\delta > \alpha$) of the 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.

3.4 Equilibrium

In this section, I derive the equilibrium implied by Lemma 1 for risk-neutral investors, the equations in (5) for prospect theory investors with a favorable signal, and the corresponding equations for those with unfavorable signals. I describe the equilibrium in terms of the strategies of the traders as a function of the public belief and their private signal. Note first the symmetry of the environment: an investor with a favorable signal at a public belief, p_t , is in the symmetric situation to an investor with an unfavorable signal at a public belief, $1 - p_t$. Therefore, behavior is symmetric around $p_t = \frac{1}{2}$, simplifying the description of an equilibrium.

As in section 3.3, the behavior of prospect theory investors depends upon the difference in the prospect theory parameters, α and δ . When the preference for negative or positive skewness is strong enough, it can overcome private information, causing herding or contrarian behavior. When the public belief is sufficiently large and $\delta > \alpha$, investors herd, buying regardless of their private signal. For sufficiently small public beliefs, by symmetry, they herd sell. Conversely, when the public belief is sufficiently large and $\delta < \alpha$, prospect theory investors act in a contrarian manner, selling regardless of their private signal. For sufficiently small public beliefs, they make contrarian purchases.

For less extreme public beliefs, prospect theory investors may either trade according to their private information or abstain from trading (if loss averse, $\lambda > 1$). As one may expect, 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 prospect theory investor with a favorable signal transitions back from trading to not trading (as the public belief increases), $p^0 \equiv (\underline{p}^0, \overline{p}^0)$, and that in which she transitions from not trading to trading, $p^1 \equiv (\underline{p}^1, \overline{p}^1)$, where $0 < \underline{p}^0, \underline{p}^1, \underline{p}^1 < 1$. The other two transition regions are for a prospect theory investor with an unfavorable signal, and are symmetric around $p_t = \frac{1}{2}$ (i.e. the transition regions are at $1 - p^0$ and $1 - p^1$; see Figure 2). In each of these transition regions, the investor mixes between trading and not trading due to strategic interaction with the market maker through the bid and ask prices which depend upon the investor's strategy (each of the equations in (5) holds with equality for a transition region). Figure 2 provides a general illustration of the unique equilibrium for the four possible cases.

Figure 2: Prospect Theory Investor Behavior in Equilibrium



Note: The figure illustrates the 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.

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. Were it not for the risk-neutral investors, no trade would take place and the public belief would remain unchanged in an information cascade. The plots on the right of Figure 2 illustrate a second possible 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 prospect theory investors' trades reveal their private information. Theorem 1 is the main theorem of the paper formalizing the illustration of Figure 2.

Theorem 1 (Equilibrium): In the unique equilibrium:

- 1. The market maker posts unique 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. Prospect theory investors' strategies are as follows:
 - (a) If δ = α, there exist two cutoff values of loss aversion, λ̄ > λ̄ > 1 such that, at all pt ∈ (0,1), if λ ≤ λ̄, they buy with favorable signals and sell with unfavorable signals, and, ifλ ≥ λ̄, they do not trade. If λ ∈ (λ̄, λ̄), they mix between buying and not trading with favorable signals and and between selling and not trading with unfavorable signals.
 - (b) Otherwise, strategies are characterized by four transition regions in public beliefs, $p^0 \equiv (\underline{p}^0, \overline{p}^0), p^1 \equiv (\underline{p}^1, \overline{p}^1)$, and their symmetric counterparts.
 - i. If $\delta > \alpha$, prospect theory investors with favorable signals sell for $p_t \leq \underline{p}^0$, don't trade for $p_t \in [\overline{p}^0, \underline{p}^1]$, and buy for $p_t \geq \overline{p}^1$. Prospect theory investors with unfavorable signals sell for $p_t \leq 1 \overline{p}^1$, don't trade for $p_t \in [1 \underline{p}^1, 1 \overline{p}^0]$, and buy for $p_t \geq 1 \underline{p}^0$.
 - ii. If $\delta < \alpha$, prospect theory investors with favorable signals buy for $p_t \leq \underline{p}^0$, don't trade for $p_t \in [\overline{p}^0, \underline{p}^1]$, and sell for $p_t \geq \overline{p}^1$. Prospect theory investors with unfavorable signals buy for $p_t \leq 1 \overline{p}^1$, don't trade for $p_t \in [1 \underline{p}^1, 1 \overline{p}^0]$, and sell for $p_t \geq 1 \underline{p}^0$.
 - iii. Within the transition regions, the investors mix such that they are indifferent between trading in the direction of the adjacent region and not trading.
 - iv. The transition regions do not overlap: $\overline{p}^0 < \underline{p}^1$ and $\overline{p}^1 < 1 \overline{p}^0$ if $\delta > \alpha$ $(1 - \underline{p}^1 < \underline{p}^0$ if $\delta < \alpha)$, implying $\overline{p}^0 < \frac{1}{2}$ if $\delta > \alpha$ $(1 - \underline{p}^1 < \frac{1}{2}$ if $\delta < \alpha)$.

Corollary 1 follows from the definitions of herd and contrarian behavior given in Definition 1 and Theorem 1, highlighting the regions at which herding and contrarian behavior occur.

Corollary 1 (Herding and Contrarian Behavior): In the unique equilibrium:

- 1. For all parameterizations with $\delta > \alpha$, there exists a public belief, $\underline{p}^0 < 1$, such that for all $p_t < \underline{p}^0$ prospect theory investors herd sell and for all $p_t > 1 - \underline{p}^0$, prospect theory investors herd buy.
- 2. For all parameterizations with $\delta < \alpha$, there exists a public belief, $\overline{p}^1 < 1$, such that for all $p_t > \overline{p}^1$ prospect theory investors contrarian sell and for all $p_t < 1 - \overline{p}^1$, prospect theory investors contrarian buy.

4 Experiment

4.1 Design

In order to simplify the environment and provide a clean test of prospect theory behavior as the source of herding and contrarianism, the experiment differs from the model in two ways.

First, consistent with the previous experimental literature, the market maker in the experiment (the experimentalist) posts 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 a subject's incentives, making the difference between her private belief and the price at which she can trade larger.

Second, rather than have subjects arrive to the market and trade one at a time as in the model, I convert the problem to an individual decision problem. Doing so allows me to avoid taking a stand *ex ante* on the distribution of preferences in the population, which is necessary to update the price. It also eliminates strategic uncertainty on the part of the subjects, removing one potential explanation for any observed deviations from risk-neutral behavior, which is the usual standard.¹⁴

In the individual-decision version of the problem, subjects observe a past history of prices which is determined simply by a sequence of random, public signals, rather than by a sequence of trades. I vary the number of signals between one and five to create many different prices, as well as price paths that that are both monotonic and non-monotonic. From a purely theoretical perspective, if subjects do not have to form beliefs about how previous subjects behaved, the past history of prices becomes irrelevant - only the price at which a subject can trade matters. Nevertheless, I include a history of prices to keep the problem subjects face

¹⁴Both Cipriani and Guarino (2005) and Drehmann, Oechssler, and Roider (2005) put forth a model in which subjects believe that previous subjects may have made mistakes, showing that it can explain contrarian, but not herding behavior.

more similar to the multiple-player version of the game, and to allow for the possibility that the sequence of prices matters behaviorally due to some form of Bayesian error.¹⁵

Next, to rule out Bayesian errors as a possibility, and provide an even cleaner test of the theory, I conduct a second treatment (the BELIEFS treatment) in which I explicitly provide subjects with the probability that the asset is valuable conditional on the price and their private signal. If Bayesian errors account for (or partially account for) unresponsive behavior, then we should observe a decrease in this behavior relative to the previous treatment (the MAIN treatment) in which subjects have to use Bayesian updating to form their private beliefs.¹⁶

Subjects in the experiment take part in 30 consecutive 'games' in which they are faced with 30 historical price paths. For each price path, I elicit their strategy for both signal realizations, allowing me to directly observe unresponsive behavior (as in Cipriani and Guarino, 2009). Each subject therefore makes 60 decisions, which provides adequate data with which to characterize their individual behavior. I set $p_0 = \frac{1}{2}$ and q = 0.7 in the experiment.

Data was collected from undergraduates at the University of California, Santa Barbara over the month of August, 2016. I conducted three sessions of each treatment for a total of 46 subjects in each. Average earnings were \$17.13 and the experiment typically finished in just over an hour.

4.2 Hypotheses

With the model and experimental design in mind, I construct several hypotheses about behavior. Consider the optimal strategy of a prospect theory investor who faces a single price (conditional on a series of public signals). For a prospect-theory investor with a favorable private signal, the optimal trading strategy in (5) becomes

$$\begin{aligned} buy \ if \quad \left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} &\geq \lambda \left(\frac{1-q}{q}\right)^{\delta} \\ sell \ if \quad \left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} &\leq \frac{1}{\lambda} \left(\frac{1-q}{q}\right)^{\delta} \end{aligned}$$
(7)

¹⁵I'm unaware of any belief formation theory that can simultaneously explain herding, contrarian, and abstention behavior. Nevertheless, given that departures from Bayesianism are well known, it seems plausible that theories such as conservatism (too much weight on the prior) or representativeness (too little weight on the prior) could play a role.

¹⁶Bisiere, Decamps, and Lovo (2015) use a similar 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 belief in the BELIEFS treatment. See the instructions in Appendix D.

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.

In the absence of the bid-ask spread, the transition regions of Theorem 1 become simple threshold prices, so that mixing is no longer part of an optimum strategy. With this exception, the unique equilibrium of the model (up to indifference at the threshold prices) is as in Theorem 1 which, together with Corollary 1, leads to predictions at the individual level. In the predictions, risk-neutral behavior refers to trading in the direction of one's private signal (as in Lemma 1).

Hypothesis 1 (Individual Behavior):

A. An individual may herd or act contrarian, but not both.

B. If an individual herds or acts contrarian at a price, p_t , then she also does so at all prices greater than p_t if $p_t > \frac{1}{2}$, or less than p_t if $p_t < \frac{1}{2}$.

C. If an individual doesn't trade or trades in a risk-neutral manner at a price, p_t , then she also does so at all prices closer to $\frac{1}{2}$ than p_t (including $p_t = \frac{1}{2}$).

D. Behavior is symmetric around $p_t = \frac{1}{2}$: the individual's decision at $p_t > \frac{1}{2}$ with a given signal realization is identical to her decision when trading at $1 - p_t$ with the opposite realization.

Aggregating across a distribution of preference parameters in the population (which may include risk-neutral types), results in the predictions of Hypothesis 2.

Hypothesis 2 (Aggregate Behavior):

A. The frequency of no trade and risk-neutral behavior decreases as p_t increases or decreases from $p_t = \frac{1}{2}$. B. The frequency of herding and contrarianism increase as p_t increases or decreases from $p_t = \frac{1}{2}$.

Finally, if herding and contrarianism are solely driven by preferences, we expect this behavior to be at least as frequent in the BELIEFS treatment, where Bayesian errors play no role, as in the MAIN treatment.

Hypothesis 3 (Treatment Effects): Providing subjects with the correct Bayesian beliefs does not reduce the frequency of herding or contrarian behavior.

Treatment	No Trade	Risk-Neutral	Herding	Contrarian	Other
MAIN BELIEFS	$5.2 \\ 5.5$	$37.4 \\ 27.5$	$13.9 \\ 34.9$	12.8 10.6	$30.8 \\ 21.6$

Table 1: Aggregate Behavior by Treatment

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

4.3 Experimental Results

Section 4.3.1 compares behavior across treatments, showing that providing subjects with the correct Bayesian beliefs actually significantly *increases* the amount of herding. It then takes a more detailed look at aggregate behavior, providing evidence consistent with Hypothesis 2. Section 4.3.2 characterizes behavior on an individual basis, showing that the majority of subjects are best characterized by prospect theory rather than standard expected utility preferences.

4.3.1 Aggregate Behavior

In the data analysis, I decided *ex ante* to drop the first 3 games a subject participates in because they are becoming familiar with the interface and environment during these games.¹⁷ I therefore have 27*2=54 decisions for each subject for a total of 54*46=2484 observations in each treatment.

Hypothesis 3 states that providing subjects with the correct Bayesian beliefs given the price and their private signal does not reduce the incidence of unresponsive behavior. To test this hypothesis, Table 1 provides percentages of each type of behavior in each treatment. The 'Other' category consists of cases in which the subject trades for one realization of her private signal but not the other, as well as trading contrary to both her signals.

From Table 1, it is clear that providing subjects with correct Bayesian beliefs does not reduce unresponsive behavior as we'd expect if Bayesian errors drive this behavior. In fact, the opposite occurs - providing subjects with the correct Bayesian beliefs significantly *increases* the frequency of herding behavior.¹⁸ Logit regressions (with errors clustered by subject) of each type of behavior (versus not) on a treatment dummy confirm that herding behavior significantly increases (p = 0.000), at the expense of risk-neutral behavior (p = 0.041) and other behavior (p = 0.011). We therefore confirm Hypothesis 3, obtaining the first result.

¹⁷Including this data does not affect any of the qualitative results.

¹⁸Bisiere, 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).

Result 1 (Treatment Effects): Providing subjects with the correct Bayesian beliefs does not reduce the frequency of unresponsive behavior. Instead, it leads to an increase in herding behavior and a decrease in risk-neutral behavior.

The higher frequency of herding in the BELIEFS treatment means Bayesian errors work against the expression of preferences through herding behavior. This result is consistent with subjects having a belief too close to one-half when observing a private signal opposite to the price trend, perhaps due to either over-weighting their private signal or under-weighting the price (e.g. Goeree at. al. (2007) and Weizsacker (2010)).¹⁹

I designed the experiment with both monotonic and non-monotonic price paths that lead to the same price to test for a specific type of Bayesian error. If errors are due to confirmation bias (e.g. Rabin and Schrag (1999)), we'd expect a contradictory private signal to be more likely to be ignored when all public signals indicate the same asset value, which could lead to more frequent herding. However, I find no statistical difference across the cases in which all public signals indicate the same asset value relative to cases in which there is at least one contradictory public signal, ruling out this form of Bayesian error.²⁰

To provide a more detailed look at the data, I now condition behavior on price, which takes on only nine discrete values in the experiment.²¹ To do so, I assume behavior is symmetric as conjectured in Hypothesis 1, part D, allowing me to treat each price, p_t , and the corresponding price, $1 - p_t$, symmetrically in the analysis.²² Specifically, I define a normalized price equal to p_t if $p_t \geq \frac{1}{2}$ and $1 - p_t$ if $p_t < \frac{1}{2}$ and perform the analysis using this normalized price. Table 2 summarizes the frequency of each type of behavior.

The frequency of risk-neutral behavior almost perfectly monotonically decreases with the normalized price in both treatments, consistent with Hypothesis 2, part A. A logit regression of risk-neutral behavior (versus not) on the normalized price provides a significant result in both treatments (p = 0.000 in both). For the frequency of no trade, the evidence is mixed because it decreases in the MAIN treatment (p = 0.047), as hypothesized, but increases in

¹⁹However, if subjects did either consistently, we'd expect to see a decrease in contrarian behavior as well, because their beliefs should be too extreme in the case of a signal confirming the price trend, but instead the frequency of contrarian behavior is similar across treatments.

 $^{^{20}}$ Specifically, I run a logit regression of herding versus no herding on a dummy variable that indicates all public signals agree separately for each treatment and for each of the three normalized prices (see below) that can be reached with both monotonic and non-monotonic price paths (0.7, 0.84, and 0.93). Most of the coefficients are small in magnitude and all are insignificant even though I have a minimum of 276 observations in each regression.

²¹The maximum price is for a sequence of four favorable signals, and conversely the minimum price is for four unfavorable signals.

²²This assumption does not strictly hold because some subjects act asymmetrically, which I explore more fully in Section 4.3.2. Therefore, at the aggregate level, the results should be interpreted as average effects over rising and falling prices.

Treatment	Normalized Price	No Trade	Risk-Neutral	Herding	Contrarian	Other
MAIN	0.50	12.3	63.0	0.0	0.0	24.6
	0.70	4.9	43.8	10.0	13.0	28.3
	0.84	3.8	32.3	13.9	14.4	35.6
	0.93	4.0	27.1	23.1	16.3	29.4
	0.97	4.4	23.9	21.7	14.1	35.9
BELIEFS	0.50	2.1	61.6	0.0	0.0	36.2
	0.70	2.1	39.7	18.5	10.3	29.3
	0.84	5.7	19.6	44.9	12.5	17.3
	0.93	8.0	10.1	55.4	12.7	13.8
	0.97	15.2	12.0	50.0	14.1	8.7

Table 2: Aggregate Behavior by Price

Note: Percentages of each type of behavior at a given price.

the BELIEFS treatment (p = 0.005). One possible reason for the increase is that standard expected utility with a large degree of risk aversion can cause no trade at extreme prices, a possibility I consider in greater detail when looking at individual behavior in Section 4.3.2.

Importantly, both herding and contrarian behavior increase with the normalized price in both treatments, confirming Hypothesis 2, part B (all four logit regressions have p = 0.000). Prospect theory generates not only both behaviors but also predicts this particular pattern due to the increase in skewness as prices become more extreme. Overall, with the exception of the increase in no trade behavior with the normalized price in the BELIEFS treatment, aggregate behavior provides fairly strong support for the theory.

Result 2 (Aggregate Behavior): In the aggregate, risk-neutral behavior decreases with the normalized price, and herding and contrarian behavior both increase, as predicted by the model. No trade is predicted to decrease, but does so in the MAIN treatment only.

A final aggregate prediction of the model is that we should observe both partial herding and partial contrarian behavior. Partial herding behavior occurs when a subject buys at a high price with a favorable signal but abstains with an unfavorable signal (or sells at a low price with an unfavorable signal but abstains with a favorable signal). Partial contrarian behavior is the opposite - a subject buys with a favorable signal at a low price but abstains with an unfavorable signal (or sells at a high price with an unfavorable signal but abstains with a favorable signal). Both types of behavior are predicted by the theory from the fact that the transitions from trade to no trade occur at different threshold prices for different private signal realizations.²³ Most of the behavior in the 'Other' category in Tables 1 and 2 consists of these behaviors. A small fraction though is behavior which is difficult to reconcile with any theory: trading contrary to 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 irrational behavior makes up only 5.2% and 7.0% of behavior in the MAIN and BELIEFS treatments, respectively.

4.3.2 Individual Behavior

Given that the aggregate data strongly suggests prospect theory preferences generate the observed behavior in the experiment, we'd expect it to vary by individual. In this section, I identify the model of preferences that best describes each particular individual's behavior.

Hypothesis 1 is a very strict interpretation of the model and is unlikely to hold exactly for any one individual given the large number of decisions each makes. In order to test the hypothesis less strictly, I identify individuals which have a general tendency to herd or act contrarian. To do so, I use the technique of Bisiere, Decamps, and Lovo (2015) to calculate a match score for each subject relative to each potential model of behavior. In this technique, I award 0.5 for each action that matches the model's prediction at each price (and zero otherwise) and then divide by the maximum possible score so that the match score lies between 0 and 1.

For prospect theory preferences, the theory states that an optimal strategy is characterized by two thresholds, p^0 and p^1 , at which an investor with a favorable signal transitions to not trading and then again to trading as the public belief increases. Given the discrete nature of the prices in the experiment, all that matters observationally is which contiguous prices these two thresholds fall between: only a finite number of trade patterns are possible. To generate the match score for prospect theory, 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).

I also consider expected utility as a candidate model, with risk-neutrality being the standard assumption. Importantly, risk-aversion can not produce herding or contrarian behavior and risk-seeking cannot produce abstention (see Appendix C for formal propositions and proofs of these claims). Therefore, risk preferences cannot simultaneously explain all three behaviors. At very high levels of risk-aversion (assuming CRRA preferences), risk-aversion can produce abstention at extreme prices.²⁴ Similarly, at very high levels of risk-

 $^{^{23}\}mathrm{Cipriani}$ and Guarino (2009) also document partial herding and partial contrarian behavior in Table 1 of their paper.

²⁴From numerical simulations, I find that the highest value of the CRRA coefficient that can produce

seeking, risk-seeking can produce contrarianism.²⁵ Although the required levels of riskaversion and risk-seeking are much higher than is typically measured for laboratory subjects, I nevertheless include them.

One (fair) criticism of prospect theory is that it has many degrees of freedom, relative to expected utility. Here, I chose the reference point *ex ante* as the status quo, thus eliminating one potential degree of freedom. At face value, the three parameters of the prospect theory model provide three degrees of freedom, but in fact they only provide two. Loss aversion provides the first. Once it is fixed, due to the fact that probability weighting and utility curvature work against each other, only one other degree of freedom remains (see Appendix B). To put prospect theory and expected utility on a level playing field, I also consider a model without loss aversion $(\lambda = 1)$. This model is characterized by a single transition, going from selling to buying in the herding case and conversely for the contrarian case. It does not allow for no trade with either signal.

Finally, I consider a model in which the only constraint is that decisions be symmetric (the decision at a price p_t with a favorable signal must match that at a price $1 - p_t$ with an unfavorable signal). The fit of this model provides an upper bound on the ability of any model that treats decisions and asset values symmetrically (i.e. does not favor buying over selling, for example) to fit the data.²⁶

For each model, I calculate the match score for each individual's behavior over the last 27 price paths they face (54 decisions for each subject: one for each realization of signal at 27 prices) Figure 3 provides the empirical CDFs of the match scores in each treatment for (i) expected utility, (ii) prospect theory without loss aversion, (iii) full prospect theory, and (iv) the model that only imposes symmetry.

A particularly striking result in Figure 3 is that prospect theory does almost as good of a job of explaining choices in both treatments as the simplistic model which only imposes symmetry, although it has an order of magnitude less degrees of freedom.²⁷ A Kolmogorov-Smirnov test does not reject the null of these two CDFs being drawn from the same distribution (p = 0.765 and p = 0.269 in the MAIN and BELIEFS treatments, respectively). Furthermore, in the BELIEFS treatment in which there is no role for Bayesian errors, even

abstention at a price greater than or equal to 0.84, 0.93, and 0.97 is 0.01,0.17, and 0.27, respectively.

²⁵Contrarianism at a price greater than or equal to 0.84, 0.93, and 0.97 requires a CRRA coefficient greater than 4.92, 4.56, and 4.43, respectively.

²⁶This model has 27 degrees of freedom. At $p_t = 0.5$, any of buy, sell, or not trade with a favorable signal are possible, with symmetry then determining the action with an unfavorable signal. At the four prices greater than one half, there are three possible actions multiplied by two possible signals, with actions then pinned down by symmetry for all four prices less than one half.

²⁷The 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, perhaps because it is more familiar to subjects than selling short.



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

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 MAIN treatment, prospect theory without loss aversion and expected utility fit similarly well.²⁸ However, the full prospect theory model continues to do better than expected utility (p = 0.057) as it does in the BELIEFS treatment (p = 0.000).

Prospect theory provides a better fit than expected utility 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.²⁹

Table 3 shows that the majority of subjects are better classified by prospect theory in both treatments. In the MAIN treatment where Bayesian errors play a role, a number of subjects are classified as risk-averse or risk-seeking, but in the BELIEFS treatment where subjects express their preferences directly, these types of subjects are reduced substantially, being replaced by herding types. In this treatment, 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.

 $^{^{28}}$ In the MAIN 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.

 $^{^{29}\}mathrm{I}$ break ties in favor of expected utility.

		v .	L			
Treatment	Models considered	Risk- Neutral	Risk- Averse	Risk- Seeking	Herding	Contrarian
MAIN	Expected utility	25	15	6	N/A	N/A
	Both (no loss aversion)	13	12	2	13	6
	Both (with loss aversion)	12	8	2	14	10
BELIEFS	Expected utility	37	8	1	N/A	N/A
	Both (no loss aversion)	8	4	0	32	2
	Both (with loss aversion)	8	3	0	32	3

Table 3: Individual Types and Match Scores

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.

Within the set of subjects for which prospect theory is the best match, in contrast to the previous experimental papers of Cipriani and Guarino (2005,2009) and Drehmann, Oechssler, and Roider (2005), herding behavior is much more common than contrarian behavior - approximately 90% are herding types. One likely explanation for this is the change from a sequential game to an individual-decision problem.³⁰ In the sequential game used in the previous papers, if subjects understand other subjects may have traded against their signals (either because of prospect theory preferences or mistakes), then subjects may best respond by acting contrarian. In the individual-decision problem this motive is removed by design.

The fact that herding is a more popular strategy means that utility curvature dominates probability weighting for most subjects ($\delta > \alpha$). Because the model only has two degrees of freedom, to identify the other parameters, we must constrain the third. If we set $\delta = 1$, we can identify the set of values of α and λ consistent with the observed behavior (see Appendix B for details).³¹ Among the herding subjects, the maximum value of α consistent with behavior ranges from 0.33 to 0.75. The corresponding levels of λ range from 1 to 2.33. In both treatments, we observe a negative correlation between the two parameters (significantly so in the BELIEFS treatment where more herding subjects exist), but this correlation is mainly driven by the two most common, and closely related, modes of behavior. In both of these, herding is observed at prices at which the difference in favorable and unfavorable signals is

 $^{^{30}}$ Kendall (2016) also finds herding to be much more frequent than contrarian behavior in a similar environment to the one here in which strategic ambiguity plays a limited role.

³¹We cannot directly estimate the parameters even after constraining one of them because they are not point-identified.

two or more. In the first, a subject trades according to her signal at more intermediate prices. In the second, she instead either abstains completely (at $p_t = \frac{1}{2}$) or only trades for one signal realization (at the remaining two prices). Result 3 summarizes the individual results.

Result 3 (Individual Behavior): The majority of subjects in both treatments are better described by prospect theory than expected utility and, of these, herding is the more popular strategy. Furthermore, prospect theory fits the data as well as any model that imposes symmetry.

Surprisingly little is known in the literature about the joint distribution of the prospect theory parameters, particularly at the individual level. Bruhin, Fehr-Duda, and Epper (2010) identify representative prospect theory types and find that, with one exception, utility curvature is fairly linear but decision weights are quite distorted (although they use a twoparameter probability weighting model), which would suggest contrarian behavior. The exception is Chinese subjects whose utility curvature over gains is quite pronounced, which would lead to herding behavior. More individual data is needed to determine whether or not the joint distributions of Bruhin, Fehr-Duda, and Epper (2010) or the joint distribution suggested by the data here are representative of the population as a whole.

5 Discussion

An obvious limitation of the theory is that the model is restricted to binary asset values. This restriction allows me to derive simple, analytic results directly applicable to 'near binary' assets such as, for example, options or initial public offerings (Green and Hwang (2012)). But, it leads to the question as to whether or not similar results would attain more generally. As emphasized, the results are driven by a preference for (negative or positive) skewness - increasing prices lead to herding or contrarianism because the asset values become more skewed. This relationship has in fact been documented in the cross-sectional returns of firms: Chen, Hong, and Stein (2001) find that stocks that have recently risen in value are more negatively skewed, just as in the case of binary asset values. So, to the extent that a preference for skewness carries over to non-binary asset values, it seems reasonable that herding or contrarianism could occur more generally. Unfortunately, applying CPT to general distributions is analytically considerably more difficult.

On the experimental side, one can question whether the results are likely to be similar with financial professionals. Normally we don't have direct evidence, but in a very nice contribution, Drehmann, Oechssler, and Roider (2005) ran experiments with such professionals and found similar levels of non-risk-neutral behavior to those I report. I suspect, therefore, that prospect theory preferences may play a role in professional decisions as well as those of undergraduate subjects.

In addition to explaining herding and contrarian behavior within one particular environment, a broader theme suggested by this paper is that prospect theory preferences can play a critical role in markets and games. In interpreting the results of experiments, it is common to address the question of whether or not risk-aversion may explain results, but, when decisions involve highly skewed payoffs, prospect theory is likely to play an even more important role. To cite just a few, environments in which prospect theory preferences 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)). In all of these environments, subjects face decisions that result in binary gambles with skewed payoffs. Until we find a robust way to manipulate preferences, we should consider the way in which subjects' innate preferences affect the decisions they make.³² Future experimental work will hopefully continue the process of identifying the joint distribution of prospect theory parameters and its impact in applications.

References

- Avery, C. and Zemsky, P. (1998), "Multidimensional uncertainty and herd behavior in financial markets", *American Economic Review*, 88 (4), 724-748
- [2] Baltzer, M, Jank, S., and Smajlbegovic, E. (2015), "Who trades on momentum?", mimeo, SSRN, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2517462
- Banerjee, A. (1992), "A simple model of herd behavior", The Quarterly Journal of Economics, 107 (3), 797-817
- [4] Barberis, N. (2012), "A model of casino gambling", Management Science, 58 (1), 35-51
- [5] Barberis, N. and Huang, M. (2008), "Stocks as lotteries: the implications of probability weighting for security prices", *American Economic Review*, 98 (5), 2066-2100

 $^{^{32}}$ The practice of inducing risk-neutrality by paying in lottery tickets (Roth and Malouf (1979)) does not necessarily suffice: subjects may simply try to maximize the number of lottery tickets they obtain, leading to preferences for skewness over lottery tickets. Kendall (2016) provides evidence consistent with this hypothesis.

- [6] Barberis, N., Huang, M, and Thaler, R. (2006), "Individual preferences, monetary gambles, and stock market participation: a case for narrow framing", *American Economic Review*, 96 (4), 1069-1090
- [7] Barberis, N. and Xiong, W. (2009), "What drives the disposition effect? An analysis of a long-standing preference-based explanation", *The Journal of Finance*, LXIV (2), 751-784
- [8] Barberis, N. and Xiong, W. (2012), "Realization utility", Journal of Financial Economics, 104, 251-271
- [9] Bisiere, C., Decamps, J., and Lovo, S. (2015), "Risk attitude, beliefs updating, and the information content of trades: an experiment", *Management Science*, 61 (6), 1378-1397
- [10] Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992), "A theory of fads, fashion, custom, and cultural change as informational cascades", *The Journal of Political Economy*, 100 (5), 992-1026
- [11] Brocas, I., Carrillo, J., Wang, S., and Camerer, C. (2014), "Imperfect choice or imperfect attention? Understanding strategic thinking in private information games", *Review of Economic Studies*, 81, 944-970
- [12] Bond, P., Edmans, A., and Goldstein, I. (2012), "The real effects of financial markets", Annual Review of Financial Economics, 4, 339-360
- [13] Brozynski, T., Menkhoff L., and Schmidt, U. (2003), "The use of momentum, contrarian, and buy-&-hold strategies: survey evidence from fund managers", mimeo, University of Hannover
- [14] Bruhin, A, Fehr-Duda, H., and Epper, T. (2010), "Risk and rationality: uncovering heterogeneity in probability distortion", *Econometrica*, 78 (4), 1375-1412
- [15] Charness and Levin (2009), "The origin of the winner's curse: A laboratory study", American Economic Journal: Microeconomics, 1, 1, 207-36
- [16] Cipriani, M. and Guarino, A. (2005), "Herd behavior in a laboratory financial market", American Economic Review, 95 (5), 1427-1443
- [17] Cipriani, M. and Guarino, A. (2008), "Transaction costs and informational cascades in financial markets", Journal of Economic Behavior & Organization, 68, 581-592

- [18] Cipriani, M. and Guarino, A. (2009), "Herd behavior in financial markets: an experiment with financial market professionals", *Journal of the European Economic Association*, 7 (1), 206-233
- [19] Cipriani, M. and Guarino, A. (2009), "Estimating a structural model of herd behavior in financial markets", American Economic Review, 104 (1), 224-251
- [20] Devenow, A. and Welch, I. (1996), "Rational herding in financial economics", European Economic Review, 40, 603-615
- [21] Drehmann, M., Oechssler, J. and Roider, A. (2005), "Herding and Contrarian Behavior in Financial Markets: An Internet Experiment", *American Economic Review*, 95 (5), 1403-1426
- [22] Glosten, L. and Milgrom, P. (1985), "Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders", *Journal of Financial Economics*, 14, 71-100
- [23] Green, T. and Hwang, B. (2012), "Initial public offerings as lotteries: skewness preference and first-day returns", *Management Science*, 58 (2), 432-444
- [24] Goeree, J., Palfrey, T., Rogers, B., and McKelvey, R. (2007), "Self-correcting information cascades", *Review of Economic Studies*, 74, 733-762
- [25] Grinblatt, M., Jostova, G., Petrasek, L., and Philipov, A. (2016), "Style and skill: hedge funds, mutual funds, and momentum", mimeo, SSRN, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2712050
- [26] Grinblatt, M., Titman, S, and Wermers, R. (1995), "Momentum investment strategies, portfolio performance, and herding: a study of mutual fund behavior", American Economic Review, 85 (5), 1088-1105
- [27] Grinblatt, M., and Keloharju, M. (2000), "The investment behavior and performance of various investor types: a study of Finland's unique data set", *Journal of Financial Economics*, 55 (1), 43-67
- [28] Hirshleifer, D. and Teoh, S.H. (2003), "Herd behavior and cascading in capital markets: a review and synthesis", *European Financial Management*, 9 (1), 25-66
- [29] Ingersoll, J. and Jin, L. (2013), "Realization utility with reference-dependent preferences", *Review of Financial Studies*, 26 (3), 723-767

- [30] Ingersoll, J. (2016), "Cumulative prospect theory, aggregation, and pricing", Critical Finance Review, 4, 1-55
- [31] Kahneman, D. and Tversky, A. (1992), "Advances in prospect theory: cumulative representation of uncertainty", *Journal of Risk and Uncertainty*, 5, 297-323
- [32] Kendall, C. (2016), "Rational and heuristic-driven trading panics in an experimental asset market", mimeo, University of Southern California
- [33] Koszegi, B. and Rabin, M. (2006), "A model of reference-dependent preferences", The Quarterly Journal of Economics, CXXI (4), 1133-1165
- [34] Koszegi, B. and Rabin, M. (2007), "Reference-dependent risk attitudes", American Economic Review, 97 (4), 1047-1073
- [35] Levy, H., De Giorgi, E, and Hens, T. (2012), "Two paradigms and Nobel prizes in economics: a contradiction or coexistence?", *European Financial Management*, 18 (2), 163-182
- [36] Li, Y. and Yang, L. (2013), "Prospect theory, the disposition effect, and asset prices", Journal of Financial Economics, 107, 715-739
- [37] Meng, L. and Weng, X. (2016), "Can prospect theory explain the disposition effect? A new perspective on reference points", mimeo, SSRN, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1851883.
- [38] Palfrey, T. and Wang, S. (2012), "Speculative overpricing in asset markets with information flows", *Econometrica*, 80 (5), 1937-1976
- [39] Park, A. and Sgroi, D. (2012), "Herding, contrarianism, and delay in financial market trading", *European Economic Review*, 56, 1020-1037
- [40] Park, A. and Sabourian, H. (2011), "Herding and contrarian behavior in financial markets", *Econometrica*, 69 (4), 973-1026
- [41] Qin, J. (2015), "A model of regret, investor behavior, and market turbulence", Journal of Economic Theory, 160, 150-174
- [42] Rabin, M. (2000), "Risk-aversion and expected-utility theory: a calibration theorem", *Econometrica*, 68 (5), 1281-1292
- [43] Rabin, M. and Schrag, J. (1999), "First impressions matter: a model of confirmatory bias", The Quarterly Journal of Economics, 114 (1), 37-82

- [44] Roth, A. and Malouf, M. (1979), "Game theoretic models and the role of information in bargaining", *Psychological Review*, 86, 575-594
- [45] Weizsacker, G. (2010), "Do we follow others when we should? A simple test of rational expectations", American Economic Review, 100 (5), 2340-2360

Appendices

A. Omitted Proofs

Proof of Lemma 1:

For convenience, I refer to traders with favorable signals $(s_t = 1)$ as type 1, and traders with unfavorable signals $(s_t = 1)$ as type 0. The claim is that, in any equilibrium, riskneutral investors always trade according to their private information, independently of the equilibrium strategies of the prospect theory investors. Here I provide the proof that a risk-neutral, type 1, investor buys. The proof that a type 0 investor sells is similar.

I first show that if a type 1 informed investor (risk-neutral or prospect theory) 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 prospect theory investor, this fact follows from equations (5) and their counterparts for a type 0 investor. For a risk-neutral investor, if a type 1 investor sells with positive probability, then it must be that case that $B_t - b_t^1 \ge b_t^1 - A_t$ so that she weakly prefers selling to buying. Rearranging, $B_t + A_t \ge 2b_t^1 > 2b_t^0$ so that the type 0 investor must strictly prefer selling. Similarly, if the 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 and therefore the bid price must be weakly less than the public belief, $B_t \leq p_t$. In this case, a risk-neutral, type 1 investor will never sell because her expected profit is negative, $B_t - b_t^1 < 0$, given that $b_t^1 > p_t$. It remains to be shown that she always has an expected profit from buying at the ask price so that she doesn't abstain from trading.

Using the formula for the ask price, (4), the investor buys if

where $Pr(a_t = buy|V = 1)$ and $Pr(a_t = buy|V = 0)$ depend upon the equilibrium strategies of the informed traders,

$$Pr(a_t = buy|V = 1) = \frac{1-\mu}{2} + \mu\gamma\beta^{RN}|(V = 1) + \mu(1-\gamma)\beta^{PT}|(V = 1)$$

$$Pr(a_t = buy|V = 0) = \frac{1-\mu}{2} + \mu\gamma\beta^{RN}|(V = 0) + \mu(1-\gamma)\beta^{PT}|(V = 0)$$
(9)

and $\beta^{RN}|(V = x)$ and $\beta^{PT}|(V = x)$, $x \in \{0, 1\}$ are the probabilities the market maker believes risk-neutral and prospect theory investors buy 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 $\beta^{y,PT}$ is the equilibrium probability that the market maker believes a prospect theory investor with $s_t = y$ buys. The right-hand side of equation (8) can be shown to be strictly increasing in $\beta^{1,PT}$ and strictly decreasing in $\beta^{0,PT}$ so that it it can be bounded above by the case of $\beta^{1,PT} = 1$ and $\beta^{0,PT} = 0$.

Using this upper bound, if a risk-neutral, type 1 investor were to not buy, we must have³³

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

But, this inequality never holds given $\mu < 1$. Therefore, we cannot have an equilibrium in which the risk-neutral, type 1 investor does not trade. To show that she profits from buying in expectation, we must have

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

again using the upper bound and the fact that a risk-neutral type 0 investor can never buy (just as the type 1 investor can never sell). This inequality holds for all $\mu < 1.\square$

Proof of Theorem 1:

Part 1 follows directly from the assumption that the market maker faces perfect competition and Bayes' rule. The fact that the bid and ask prices are unique follows from uniqueness of the equilibrium strategies of the informed investors (see Lemma 1 for risk-neutral investors and below for prospect theory investors).

Part 2 was proven in Lemma 1.

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

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

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

$$(10)$$

From (10), 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

$$\begin{aligned} Pr(a_t = buy | V = 1) &= \frac{1-\mu}{2} + \mu\gamma q + \mu(1-\gamma)\beta^{PT} | (V = 1) \\ Pr(a_t = buy | V = 0) &= \frac{1-\mu}{2} + \mu\gamma(1-q) + \mu(1-\gamma)\beta^{PT} | (V = 0) \end{aligned}$$

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

 $^{^{33}}$ If a risk-neutral, type 1 investor doesn't buy, then a risk-neutral type 0 investor must also not buy. This fact is established in the same manner as the fact that, if a risk-neutral, type 1 investor sells, then a risk-neutral, type 0 investor must also sell.

 $\begin{array}{l} 0)=(1-q)\beta^{1,PT}+q\beta^{0,PT}. \mbox{ Now, as argued in the proof of Lemma 1, if a type 0 investor buys, then so must a type 1 investor. This fact implies <math display="inline">\beta^{1,PT}\geq\beta^{0,PT}$ which in turn implies a lower bound for the ratio $\frac{Pr(a_t=buy|V=1)}{Pr(a_t=buy|V=0)}.$ Because this ratio is increasing in $\beta^{1,PT}$ and decreasing in $\beta^{0,PT}$, it is bounded below by $\beta^{1,PT}=\beta^{0,PT}$ under the constraint $\beta^{1,PT}\geq\beta^{0,PT}.$ Therefore, for a type 0 investor to buy, we must have $1\geq\lambda\left(\frac{q}{1-q}\frac{\frac{1-\mu}{2}+\mu\gamma q+\mu(1-\gamma)\beta_0^{PT}}{\frac{1-\mu}{2}+\mu\gamma(1-q)+\mu(1-\gamma)\beta_0^{PT}}\right)^{\alpha}.$ But, the ratio inside the the parentheses is strictly greater than one for any $\beta^{0,PT}$, so this inequality 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. Due to the symmetry of the problems, it suffices to consider when a type 1 investor buys. That is, when the first equation in (10) holds. Substituting for the probabilities of observing a buy and using the fact that a type 0 investor never buys ($\beta^{0,PT} = 0$), we obtain

$$buy \ if 1 \ge \lambda \left(\frac{1-q}{q} \frac{\frac{1-\mu}{2} + \mu\gamma q + \mu(1-\gamma)q\beta^{1,PT}}{\frac{1-\mu}{2} + \mu\gamma(1-q) + \mu(1-\gamma)(1-q)\beta^{1,PT}} \right)^{\alpha}$$

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

$$1 \leq \lambda \left(\frac{1-q}{q} \frac{\frac{1-\mu}{2} + \mu\gamma q}{\frac{1-\mu}{2} + \mu\gamma(1-q)} \right)^{\alpha} \\ \iff \lambda \geq \left(\frac{q}{1-q} \frac{\frac{1-\mu}{2} + \mu\gamma(1-q)}{\frac{1-\mu}{2} + \mu\gamma q} \right)^{\alpha} \equiv \overline{\lambda}$$

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

$$1 \geq \lambda \left(\frac{1-q}{q} \frac{\frac{1-\mu}{1-\mu} + \mu q}{\frac{1-\mu}{2} + \mu(1-q)} \right)^{\alpha} \\ \iff \lambda \leq \left(\frac{q}{1-q} \frac{\frac{1-\mu}{2} + \mu(1-q)}{\frac{1-\mu}{2} + \mu q} \right)^{\alpha} \equiv \underline{\lambda}$$

Simple algebra shows that $\overline{\lambda} > \underline{\lambda} > 1$ for all parameterizations. Finally, for intermediate values of λ , 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

$$1 = \lambda \left(\frac{1-q}{q} \frac{\frac{1-\mu}{2} + \mu\gamma q + \mu(1-\gamma)q\beta^{1,PT}}{\frac{1-\mu}{2} + \mu\gamma(1-q) + \mu(1-\gamma)(1-q)\beta^{1,PT}} \right)^{\alpha}$$

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 identical.

Part3b. Consider the case of $\delta > \alpha$. I first consider the decision to buy or not for both types. Beginning again with equation (5) and the formulas for the probability of observing a buy (9), the two equations governing buy decisions are given by

$$\begin{split} s_t &= 1: \quad \left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} &\geq \quad \lambda \left(\frac{1-q}{q}\right)^{\delta} \left(\frac{\frac{1-\mu}{2} + \mu\gamma q + \mu(1-\gamma)\left(q\beta^{1,PT} + (1-q)\beta^{0,PT}\right)}{\frac{1-\mu}{2} + \mu\gamma(1-q) + \mu(1-\gamma)\left((1-q)\beta^{1,PT} + q\beta^{0,PT}\right)}\right)^{\alpha} \\ s_t &= 0: \quad \left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} &\geq \quad \lambda \left(\frac{q}{1-q}\right)^{\delta} \left(\frac{\frac{1-\mu}{2} + \mu\gamma q + \mu(1-\gamma)\left(q\beta^{1,PT} + (1-q)\beta^{0,PT}\right)}{\frac{1-\mu}{2} + \mu\gamma(1-q) + \mu(1-\gamma)\left((1-q)\beta^{1,PT} + q\beta^{0,PT}\right)}\right)^{\alpha} \end{split}$$

Because the public belief enters the inequalities only on the left-hand side, we can see that for a sufficiently large public belief, both types of investors buy. Denote the upper threshold public beliefs at which the the type 0 and type 1 investors buy with probability one, $1 - p^0$ and \overline{p}^1 , respectively. These beliefs are given by the unique solutions to

$$s_t = 1: \quad \left(\frac{\overline{p}^1}{1-\overline{p}^1}\right)^{\delta-\alpha} = \lambda \left(\frac{1-q}{q}\right)^{\delta} \left(\frac{\frac{1-\mu}{2}+\mu q}{\frac{1-\mu}{2}+\mu(1-q)}\right)^{\alpha}$$
$$s_t = 0: \quad \left(\frac{1-\underline{p}^0}{\underline{p}^0}\right)^{\delta-\alpha} = \lambda \left(\frac{q}{1-q}\right)^{\delta} \left(\frac{\frac{1-\mu}{2}+\mu\gamma q+\mu(1-\gamma)}{\frac{1-\mu}{2}+\mu\gamma(1-q)+\mu(1-\gamma)}\right)^{\alpha}$$

where I have substituted $\beta^{1,PT} = 1$ and $\beta^{0,PT} = 0$ into the first equation and $\beta^{1,PT} = \beta^{0,PT} = 1$ into the second, using the fact that the type 1 investor's upper threshold belief must be less than the belief at which the type 0 investor begins to buy with positive probability, $\overline{p}^1 < 1 - \overline{p}^0$. This fact 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. As the public belief decreases from \overline{p}_1 , a transition region exists in which the type 1 investor mixes between buying and not trading according to $\beta^{1,PT}$ which uniquely solves

$$\left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} = \lambda \left(\frac{1-q}{q}\right)^{\delta} \left(\frac{\frac{1-\mu}{2} + \mu\gamma q + \mu(1-\gamma)q\beta^{1,PT}}{\frac{1-\mu}{2} + \mu\gamma(1-q) + \mu(1-\gamma)(1-q)\beta^{1,PT}}\right)^{\alpha}$$

Similarly, as the public belief decreases from $1 - \underline{p}^0$, the type 0 investor mixes between buying and not trading according to $\beta^{0,PT}$ which solves

$$\left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} = \lambda \left(\frac{q}{1-q}\right)^{\delta} \left(\frac{\frac{1-\mu}{2} + \mu\gamma q + \mu(1-\gamma)\left(q + (1-q)\beta^{0,PT}\right)}{\frac{1-\mu}{2} + \mu\gamma(1-q) + \mu(1-\gamma)\left((1-q) + q\beta^{0,PT}\right)}\right)^{\alpha}$$

These transition regions ends at lower threshold prices (at which point each type of investor ceases to buy with positive probability) given by the unique solutions to

$$s_t = 1: \quad \left(\frac{\underline{p}^1}{1-\underline{p}^1}\right)^{\delta-\alpha} = \lambda \left(\frac{1-q}{q}\right)^{\delta} \left(\frac{\frac{1-\mu}{2}+\mu\gamma q}{\frac{1-\mu}{2}+\mu\gamma(1-q)}\right)^{\alpha}$$
$$s_t = 0: \quad \left(\frac{1-\overline{p}^0}{\overline{p}^0}\right)^{\delta-\alpha} = \lambda \left(\frac{q}{1-q}\right)^{\delta} \left(\frac{\frac{1-\mu}{2}+\mu q}{\frac{1-\mu}{2}+\mu(1-q)}\right)^{\alpha}$$

Turning to the sell decisions for each type, the two equations of interest are

$$s_{t} = 1: \quad \left(\frac{p_{t}}{1-p_{t}}\right)^{\delta-\alpha} \leq \frac{1}{\lambda} \left(\frac{1-q}{q}\right)^{\delta} \left(\frac{\frac{1-\mu}{2}+\mu\gamma(1-q)+\mu(1-\gamma)\left(q\eta^{1,PT}+(1-q)\eta^{0,PT}\right)}{\frac{1-\mu}{2}+\mu\gamma q+\mu(1-\gamma)((1-q)\eta^{1,PT}+q\eta^{0,PT})}\right)^{\alpha}$$

$$s_{t} = 0: \quad \left(\frac{p_{t}}{1-p_{t}}\right)^{\delta-\alpha} \leq \frac{1}{\lambda} \left(\frac{q}{1-q}\right)^{\delta} \left(\frac{\frac{1-\mu}{2}+\mu\gamma(1-q)+\mu(1-\gamma)\left(q\eta^{1,PT}+(1-q)\eta^{0,PT}\right)}{\frac{1-\mu}{2}+\mu\gamma q+\mu(1-\gamma)((1-q)\eta^{1,PT}+q\eta^{0,PT})}\right)^{\alpha}$$

where I have used $\eta^{0,PT}$ and $\eta^{1,PT}$ to denote the probabilities with which type 0 and type 1 prospect theory investor sells, respectively. Now, 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 if one replaces p_t with $1 - p_t$. It thus follows that selling behavior is identical to buying behavior except that the transition regions occur over symmetric intervals: (p^0, \bar{p}^0) for the type 1 investor and $(1 - \bar{p}^1, 1 - p^1)$ for the type 0 investor, where $\bar{p}^0 < 1 - \bar{p}^1$. 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 $(\underline{p}^0, \bar{p}^0)$), followed by the type 0 investor (over $(1 - \bar{p}^1, 1 - p^1)$).

Now consider the case of $\delta < \alpha$. The only difference from the case of $\delta > \alpha$ 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. This completes the proof of parts i) to iii).

For part iv), $1 - \overline{p}^0 > \overline{p}^1$ 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, $1 - \underline{p}^1 < \underline{p}^0$ for $\delta < \alpha$. Lastly, we must show $\overline{p}^0 < \underline{p}^1$. From the formulas for these threshold public beliefs, this inequality is equivalent to

$$\frac{1}{\lambda} \left(\frac{1-q}{q}\right)^{\delta} \left(\frac{\frac{1-\mu}{2} + \mu(1-q)}{\frac{1-\mu}{2} + \mu q}\right)^{\alpha} < \lambda \left(\frac{1-q}{q}\right)^{\delta} \left(\frac{\frac{1-\mu}{2} + \mu\gamma q}{\frac{1-\mu}{2} + \mu\gamma(1-q)}\right)^{\alpha}$$

which is more easily met as λ increases, so take $\lambda = 1$. In this case,

$$\begin{array}{rcl} & \frac{\frac{1-\mu}{2}+\mu(1-q)}{\frac{1-\mu}{2}+\mu q} & < & \frac{\frac{1-\mu}{2}+\mu q q}{\frac{1-\mu}{2}+\mu q(1-q)} \\ \mu(1-q)\frac{1-\mu}{2}+\frac{1-\mu}{2}\mu \gamma(1-q)+\mu^2\gamma(1-q)^2 & < & \mu\gamma q\frac{1-\mu}{2}+\frac{1-\mu}{2}\mu q+\mu^2\gamma q^2 \\ \mu(1-2q)\frac{1-\mu}{2}(1+\gamma)+\mu^2\gamma\left((1-q)^2-q^2\right) & < & 0 \end{array}$$

which is true given $0 < \mu < 1$ and $q > \frac{1}{2}$.

B. Prospect Theory Parameters

I first show that the model only has two degrees of freedom and then find the supportable set of pairs of threshold prices at which behavior transitions. 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 $p_t = \frac{q^k}{q^k + (1-q)^k}$. (7), can then be written

buy if
$$Q^{k(\delta-\alpha)+\delta} \ge \lambda$$

sell if $Q^{k(\delta-\alpha)+\delta} \le \frac{1}{2}$

where $Q \equiv \frac{q}{1-q}$. If we define $\lambda' \equiv \lambda^{\frac{1}{\delta-\alpha}}$ and $\delta' \equiv \frac{\delta-\alpha}{\delta}$, we can rewrite these conditions as

buy if
$$Q^{k+\frac{1}{\delta'}} \ge \lambda'$$

sell if $Q^{k+\frac{1}{\delta'}} \le \frac{1}{\lambda'}$

which establishes that the model only has two degrees of freedom, λ' and δ' . 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

$$\lambda' = \sqrt{Q^{k^1 - k^0}} \tag{11}$$

and

$$\delta' = \frac{-2}{k^0 + k^1} \tag{12}$$

 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 restrictions on α and δ , we must have $\delta' < 1$ which, from (12), 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 α and δ make δ' negative in this case so that (12) leads to the same constraint. I impose these restrictions when generating the set of possible trade patterns prospect theory is capable of explaining.

Focusing on the more common herding types, if we assume no probability weighting, $\delta = 1$, we can relate λ' and δ' to the primitives of the model. Specifically, $\alpha = 1 - \delta'$ and $\lambda = {\lambda'}^{\delta'}$. The parameters are not point-identified so I report the pair of parameters associated with the minimum utility curvature consistent with observed behavior.

C. Expected Utility Preferences

Theorem B1 (Expected Utility Investors): If prospect theory investors are replaced

with investors with standard expected utility preferences:

(i) if risk-averse, herding and contrarian behavior 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

and otherwise abstain from trading.³⁴ 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 behavior 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 last 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 last because $b_t > A_t$ with a favorable signal. \Box

C. Instructions

I have included the instructions for the BELIEFS treatment below. The instructions for the MAIN treatment are identical except that the second paragraph under 'A Valuable Clue' is removed.

³⁴I have normalized initial wealth to zero without loss of generality.

Instructions for Trading Experiment

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.

https://marshall-fbe-lab.herokuapp.com/p/vzrlk7tr/seqtrade/Instructions/2/

Instructions for Trading Experiment



Trading Page

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.

https://marshall-fbe-lab.herokuapp.com/p/vzrlk7tr/seqtrade/Instructions/2/

8/12/2016

8/12/2016

Instructions for Trading Experiment

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 it's 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.

https://marshall-fbe-lab.herokuapp.com/p/vzrlk7tr/seqtrade/Instructions/2/

8/12/2016

Instructions for Trading Experiment

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.

Next

https://marshall-fbe-lab.herokuapp.com/p/vzrlk7tr/seqtrade/Instructions/2/