Herding and Contrarianism: A Matter of Preference?

Chad Kendall*

July 27, 2017

Abstract

Herding and contrarian strategies in financial markets produce informational inefficiencies because investors ignore private information, instead following or bucking past trends. In a simple trading environment, I demonstrate theoretically that investors with prospect theory preferences ignore private information by following a strategy that *looks like* herding or contrarianism, but which is actually trend-independent. I confirm the theory's predictions in a laboratory experiment designed to rule out other sources of these behaviors, and find that approximately 70% of subjects exhibit herd-like behavior. Finally, I perform a calibration exercise using actual market data to demonstrate the applicability of the results to more general settings.

1 Introduction

Herding and momentum strategies, and their antithesis, contrarian strategies, have long been of interest to financial economists because of their implications for the efficiency of markets. What drives investors to use these strategies, and, in particular, to choose one over the other?[1] The most traditional explanation for herding (following the trend) is simply imitation, a form of 'herd mentality' (Mackay (1841)), but behavioral explanations such as extrapolative expectations (De Long et al. (1990)) and rational explanations based on information externalities (Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992))

*Department of Finance and Business Economics, Marshall School of Business, University of Southern California, 701 Exposition Blvd, Ste. 231 HOH-231, MC-1422, Los Angeles, CA 90089-1422 (e-mail chadkend@marshall.usc.edu).

[1]Momentum and contrarian strategies are frequently discussed by practitioners and a small academic literature suggests different types of individuals or firms pursue each (see Grinblatt, Titman, and Wermers (1995), Grinblatt and Keloharju (2000), Baltzer, Jank, and Smajlbegovic (2015), and Grinblatt et al. (2016)).

have also been suggested.[2] Contrarian strategies (going against the trend) can similarly be motivated through rational behavior given a suitable informational environment (Avery and Zemsky (1998), and Park and Sabourian (2011)). In all of these explanations, past observation of investor actions and/or prices is critical and, in fact, the strategies are *defined* in terms of changes in behavior induced by observing historical data. In this paper, I suggest instead, using both theory and a laboratory experiment, that what *appear* to be trend-dependent strategies may in fact not depend upon the observation of historical data at all. Instead, due to the nature of the relationship between past price trends and the subsequent expected returns, these strategies may instead be driven by *preferences* over future returns.

In the theoretical contribution, I introduce cumulative prospect theory (CPT) preferences - the most common descriptive model of choices involving risk - into a standard model of trading with asymmetric information, that of Glosten and Milgrom (1985). I show that an investor generically either buys or sells an asset at extreme prices, *independently* of her private information. This herd or contrarian-like behavior is not, however, driven by a change in prices: these investors would make the same decisions even without any knowledge of the history of events. Instead, decisions are driven by the fact that extreme prices imply highly skewed returns, which drive CPT investors' decisions (Barberis and Huang (2008)). Given the theoretical predictions, I then conduct a new laboratory experiment that controls for other possible explanations of behavior, demonstrating robust evidence of herding and contrarian-like behaviors even when traditional explanations cannot possibly be their cause. Thus, this paper suggests that CPT preferences likely contribute to the use of what appear to be trend-dependent strategies and also provide a unifying explanation for why some investors prefer to go with the trend and others against it.

In the model, investors arrive sequentially to a market, trading a single, binary-valued asset with a market maker who posts separate bid and ask prices. Each investor, after receiving a private, binary signal about the asset's fundamental value, may buy or sell a single unit of the asset (or abstain from trading). The standard result with expected utility investors (Avery and Zemsky (1998)) is that each investor trades according to her private information - buying with a favorable signal and selling otherwise. With CPT preferences, instead, an individual investor generically (i.e. for almost all combinations of preference

[2] A momentum strategy is one in which an investor simply trades with the past price trend, which may result from erroneous beliefs such as overextrapolation (De Long et al. (1990), Greenwood and Shleifer (2014), Barberis, Greenwood, Jin, and Shleifer (2015,2016)). Herding results in similar trade patterns but is generally defined in the context of asymmetric information environments: an investor herds if she initially trades in a way that reveals her private information, but switches her trade direction to follow the crowd after observing others' trades (Avery and Zemsky (1998), Park and Sabourian (2011)). These types of 'rational herding' models originated with papers by Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992). See Devenow and Welch (1996) and Hirshleifer and Teoh (2003) for surveys of herding in financial markets.

parameters) ignores her private information, trading in a single direction at extreme prices. The direction that she prefers to trade depends in a very simple way only upon the *difference* between the extent to which she weights probabilities and her utility curvature.[3]

The reference-dependent nature of CPT preferences implies preferences over skewness that drive trade decisions. At a high price, the expected return from buying the asset is *negatively skewed*: it is very likely to return a small positive amount (when the asset is valuable), but occasionally results in a large loss (when it is not). Selling the asset instead results in a *positively skewed* return. As has been recognized in the previous literature (Barberis and Huang (2008)), the overweighting of small probabilities in CPT generates a preference for positive skewness. When probability weighting dominates in an investor's preferences, we then observe contrarian-like behavior: selling at high prices. On the other hand, when utility curvature dominates in an investor's preferences, an investor exhibits a preference for negative skewness which leads to herd-like behavior: buying at high prices. In the model, the opposing effects of probability weighting and utility curvature take a particularly analytically tractable form, highlighting a feature of CPT that has received little emphasis within the literature that applies these preferences.[4]

To demonstrate that skewness preferences do in fact generate behavior that looks like herding and contrarianism, I conduct a laboratory experiment in which I directly control for other possible explanations (something which would be very difficult to do in naturally-occurring settings). To do so, I have subjects make decisions in an individual decision-making environment which rules out any social cause of behavior, including simply following others.[5] In a second treatment, I also rule out erroneous belief formation by providing subjects with the correct Bayesian beliefs given both publicly available information and their private signals. Evidence from both treatments provides strong evidence that skewness preferences generate both herding and contrarian-like behaviors. When Bayesian errors play no role, roughly three quarters of subjects make decisions that are better characterized by prospect theory than standard expected utility, and within these subjects, over 90% exhibit

[3]Other recent papers have considered non-expected utility preferences in very similar trading models. Qin (2015) shows that regret aversion generates behavior that looks like herding, but not contrarianism. Boortz (2016) builds on Ford (2013), to show that ambiguity can generate both behaviors, but herding behavior requires time-varying preferences. In contrast to these models, static CPT preferences generate both types of behavior.

[4]Barberis (2012), in his model of casino gambling, demonstrates the counteracting forces numerically. A literature on neuroeconomics (Glimcher and Fehr (2013)) has also recognized the offsetting effects of probability weighting and utility curvature.

[5]Past experiments have found both herding and contrarian-like behavior in the setup of this paper (Cipriani and Guarino (2005,2009) and Drehmann, Oechssler, and Roider (2005)), but with a design in which imitation and strategic ambiguity over past subjects' strategies may play a role. The individual decision-making environment rules out both these possible sources of behavior.

a preference for negative skewness, exhibiting behavior that is observationally equivalent to herding. Importantly, because of the offsetting effects of probability weighting and utility curvature, the prospect theory model has only a single degree of freedom, the same as I assume for expected utility (CRRA/CARA).[6]

Skewness in forward returns in the model and experiment are driven by the binary nature of the asset value. Binary asset values allow for analytic theoretical expressions and a very clean test for skewness preferences, but, of course, many asset values are not binary in real markets. To show that binary asset values are not necessary for the main conclusions of the paper, I conclude with a calibration exercise using actual market returns. I first show that market returns become more (less) skewed as prices rise (fall), a fact first documented by Chen, Hong and Stein (2011). Thus, the relationship between price trends and return skewness in actual markets is the same as with binary asset values, creating the necessary relationship for the mechanism of the model to take effect. Using the model parameters identified for the negative-skewness loving subjects in the experiment, I show numerically that these investors would be willing to trade against significantly strong private information in a herd-like manner, after periods of both rising and falling prices.

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

On the experimental side, this paper extends 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 within an expected

[6]In the model and fits of the experimental data, I also allow for loss aversion. I demonstrate that allowing for loss aversion provides a better fit to the experimental data, but also that CPT fits the data better than expected utility even when loss aversion is precluded.

utility framework. If a majority of subjects exhibit reference-dependence, as they appear to in the broad literature on estimating risk attitudes, it seems likely that reference-dependence would also play a role in determining behavior in games and markets.

The paper proceeds as follows. I describe the model in Section 2 and characterize its unique equilibrium in Section 3. In Section 4, I describe the experimental design, develop hypotheses, and provide evidence of skewness-loving preferences. In Section 5, I discuss the generalizability of both the theoretical and experimental results to other settings by performing a calibration exercise using market data.

2 Model

The model is a sequential trading model based on that of Glosten and Milgrom (1985). In each period $t = 1, 2, \ldots, T$, a single new investor arrives to the market to trade an asset of unknown value, $V \in \{0, 1\}$. I denote the initial prior that the asset is worth 1 by $p_1 \in (0, 1)$. Upon arrival, an investor may either buy or sell short a single unit, or not trade, $a_t \in \{buy, sell, NT\}$. After making her decision, the investor leaves the market. All trades are with a risk-neutral market maker who is assumed to face perfect competition, earning zero profits in expectation.[7] The market maker incorporates the information provided in the current order in setting prices. Specifically, he posts an ask price, A_t, at which he is willing to sell a unit of stock and a bid price, B_t, at which he is willing to buy a unit. When the asset value is realized at T, investors who purchased the asset at time t receive a payoff of $V - A_t$ and those who sold receive a payoff of $B_t - V$ (there is no discounting). All market participants observe the complete history of trades and prices, denoted $H_t = (a_1, a_2, \ldots, a_{t-1}) \cup (A_1, A_2, \ldots, A_{t-1}) \cup (B_1, B_2, \ldots, B_{t-1})$.

Investors are one of three types: risk-neutral, prospect theory, or uninformed investors. Uninformed investors, who arrive with probability $1 - \mu$, $\mu \in (0, 1)$, trade for exogenous reasons and are equally likely to buy or sell. Risk-neutral investors have standard risk-neutral expected utility preferences and arrive with probability, $\mu\gamma$, $\gamma \in (0, 1)$. Finally, prospect theory investors have the CPT preferences of Kahneman and Tversky (1992) (see Section 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

[7] This assumption is standard in the literature and follows Glosten and Milgrom (1985).

investors in the model primarily because we should expect heterogeneous preferences in any population.[8] In addition, including risk-neutral investors allows (partial) information to be revealed by every trade, ensuring dynamic price paths as in real markets, whereas prices generally stagnate in their absence. That said, risk-neutral investors are not necessary for the main conclusions of the model.

3 Theory

3.1 Solution Concept

Being a game of asymmetric information, the solution concept is Perfect Bayesian Equilibrium. An equilibrium consists of a specification of the strategies of the risk-neutral and 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 uniformed 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.

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 CPT investors.

The assumed zero-profit condition for the market maker results in him 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 at the start of the period, denoted 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,

[8] Bruhin, Fehr-Duda, and Epper (2010), for example, argue that both expected utility and prospect theory preferences should be included in applied theoretical work because they find a mixture of preferences among their experimental subjects.

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

An immediate consequence of Lemma 1 is that, 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* in which prices stagnate and subsequent trades reveal no new information never occurs. This fact implies that, by the law of large numbers, that the public beliefs and bid and ask prices converge to the true asset value in the limit as $T \to \infty$, as shown in Avery and Zemsky (1998). Informational efficiency is achieved in the limit.

3.3 Prospect Theory Investors

CPT differs from expected utility in that investors evaluate gains and losses relative to a reference point. Perhaps the simplest possible reference point is status quo wealth, which is the reference point I adopt. The behavioral asset pricing literature has tended to instead use the expected wealth from investing in a risk-free asset (see Barberis and Huang (2008), Barberis and Xiong(2009), and Li and Yang (2013)). In the absence of a risk-free asset, as is the case here, these different specifications are equivalent in the sense that the reference point is the amount an investor can attain without risk. Expectations-based reference points, such as those in Koszegi and Rabin (2006,2007) are another popular alternative. However, I show in Appendix E that they are generally inconsistent with the experimental evidence that follows.[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

[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.

[10]The 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 assume subjects use narrow framing, considering the experiment (and, in fact, each repetition of the game) in isolation.

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 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 CPT investor. Given a private belief, $b_t = Pr(V = 1|H_t, s_t)$, a prospect theory investor prefers buying to not trading if

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

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

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

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

The forms of equations (1) and (2) are sufficiently general that little can be said about the behavior of the investor without imposing additional structure. I proceed by using the functional forms for the value and decision weight functions provided in the original work of Kahneman and Tversky (1992), because they are tractable, parsimonious, and appear to fit decisions over binary gambles reasonably well.[12] Specifically, I assume

$$\begin{aligned} v^+(x) &= x^\alpha \\ v^-(y) &= -\lambda(-y)^\alpha \end{aligned}$$

and

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

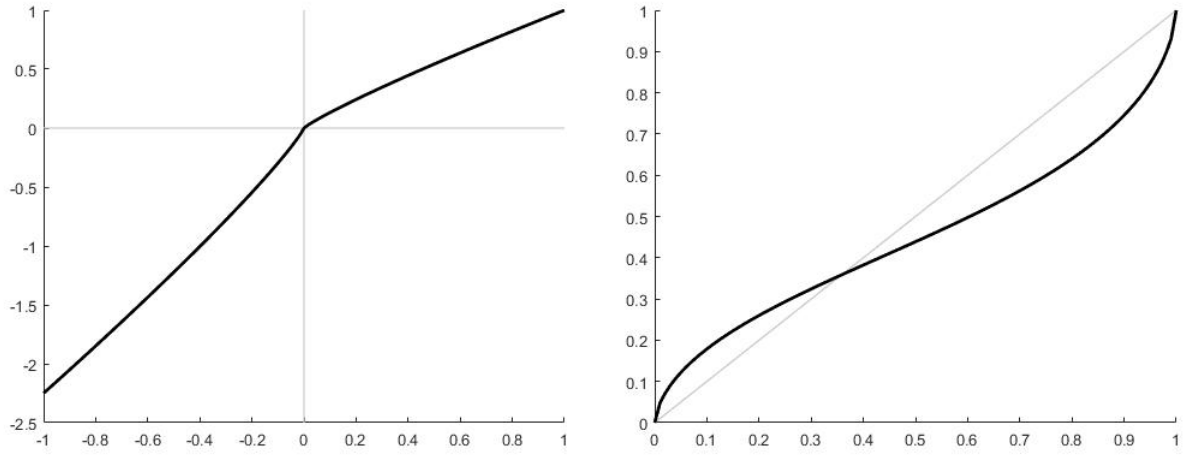
with $\alpha \in (0,1]$, $\lambda \geq 1$, and $\delta \in (0,1]$.[13] $\alpha \in (0,1)$ reflects the common experimental finding of risk-aversion over gains and risk-seeking over losses (an "S-shaped" value function).

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

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

[13] Kahneman and Tversky assume a slightly more general form allowing $w^+(r)$ and $w^-(r)$ to have different parameters, but their experimental estimates for the two parameters are quantitatively similar. I assume a common parameter, which results in a significant increase in tractability.

Figure 1: Examples of the Value and Probability Weighting Functions



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

$\lambda \geq 1$ 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 optimal decisions for a CPT investor with private belief, b_t:

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

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 explicitly evaluate the beliefs and prices.

Denoting the public belief at time t, $p_t = Pr(V = 1|H_t)$, 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:

$$\begin{array}{rcl} A_t & = & \frac{p_t Pr(a_t=buy|V=1)}{p_t Pr(a_t=buy|V=1)+(1-p_t)Pr(a_t=buy|V=0)} \\ B_t & = & \frac{p_t Pr(a_t=sell|V=1)}{p_t Pr(a_t=sell|V=1)+(1-p_t)Pr(a_t=sell|V=0)} \end{array} \tag{4}$$

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

$$\begin{array}{lcl} buy\ if & \left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} & \geq \lambda \left(\frac{1-q}{q}\right)^{\delta} \left(\frac{Pr(a_t=buy|V=1)}{Pr(a_t=buy|V=0)}\right)^{\alpha} \\ sell\ if & \left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} & \leq \frac{1}{\lambda} \left(\frac{1-q}{q}\right)^{\delta} \left(\frac{Pr(a_t=sell|V=1)}{Pr(a_t=sell|V=0)}\right)^{\alpha} \end{array} \tag{5}$$

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

Although the opposing effects of 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.

CPT investors, on the other hand, do have an incentive to trade. With the simplification, equations (5) become

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

We see that, unless the public belief is exactly $\frac{1}{2}$, either buying or selling is strictly preferable to not trading. Consider a public belief, $p_t > \frac{1}{2}$. As the decision weights become more distorted from linearity (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: she exhibits a *preference for positive skewness* which

is the consequence of prospect theory studied extensively in Barberis and Huang (2008).

Next, 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 small gain $(1 - p_t)$ that occurs with probability p_t if one buys is preferred to the large gain (p_t) that occurs with probability $1 - p_t$ if one sells, a simple consequence of risk-aversion (an investor with risk-neutral preferences would be indifferent). At the same time, the small probability of a large loss if one buys is preferred to the large probability of a small loss if one sells, due to risk-seeking. Both effects make buying more valuable than selling so that if $\delta > \alpha$, the investor always buys. The investor in this case exhibits a *preference for negative skewness*.[14]

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 abstain from trade. Perhaps surprisingly, however, loss aversion prevents trading only at intermediate public beliefs. Although the potential losses are larger at extreme public beliefs, one can take the side of the trade that either minimizes the probability ($\delta < \alpha$) or the size ($\delta > \alpha$) of a loss. At intermediate beliefs, on the other hand, the chance of a loss of medium size and probability can only be avoided by abstaining from trade.

3.4 Equilibrium

In this section, I describe the unique equilibrium of the trading game. The equilibrium strategies of the traders are a function of the public belief and their private signals, and follow

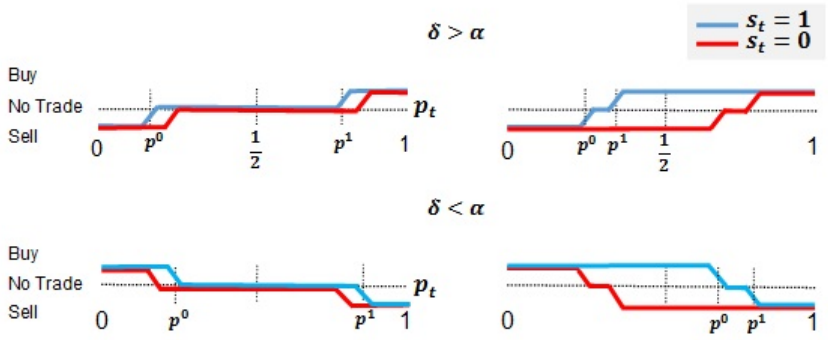
[14]Beyond the intuition, we can see that the preference to trade in a particular direction at an extreme price must be driven by the skewness of returns, and not by their mean or variance. For example, with a preference for negative skewness, a CPT investor switches from always selling to always buying as the price increases from 0 to 1. Over this range, the mean return is always positive if she trades with her private signal, so that, as in the case of a risk-neutral investor, an investor that cares only about the mean always trades with her private signal. The variance of the return increases from zero to its maximum value at $p = \frac{1}{2}$, and then decreases to zero again, so if the variance in returns were driving behavior, we'd expect behavior to be the same at both price extremes.

from 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. Note first the symmetry of the environment: an investor with a favorable signal at a public belief, p_t, is in a symmetric situation to an investor with an unfavorable signal at a public belief, $1 - p_t$. Therefore, behavior is symmetric around $p_t = \frac{1}{2}$, simplifying the description of an equilibrium.

As discussed in section 3.3, the behavior of CPT investors depends upon the difference in the prospect theory parameters, α and δ. When returns becomes sufficiently skewed, skewness preferences overcome private information, causing the investor to trade in the same direction for both realizations of her private signal. Thus, when the public belief is sufficiently large and $\delta > \alpha$, investors buy regardless of their private signal. For sufficiently small public beliefs, by symmetry, they always sell. Conversely, when $\delta < \alpha$, CPT investors sell regardless of their private signal when the public belief is sufficiently large, and buy when the public belief is sufficiently small.

For less extreme public beliefs, 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 CPT investor with a favorable signal transitions 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 back from not trading to trading, $p^1 \equiv (\underline{p}^1, \overline{p}^1)$, where $0 < \underline{p}^0, \overline{p}^0, \underline{p}^1, \overline{p}^1 < 1$. The other two transition regions are for a CPT 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: 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. *CPT investors' strategies are as follows:*
   (a) *If $\delta = \alpha$, there exist two cutoff values of loss aversion, $\overline{\lambda} > \underline{\lambda} > 1$ such that, at all $p_t \in (0,1)$, if $\lambda \leq \underline{\lambda}$, they buy with favorable signals and sell with unfavorable signals, and, if $\lambda \geq \overline{\lambda}$, they do not trade. If $\lambda \in (\underline{\lambda}, \overline{\lambda})$, they mix between buying and not trading with favorable signals 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 $\delta > \alpha$, CPT 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 $p_t \geq \overline{p}^1$. CPT 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 $\delta < \alpha$, CPT 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 $p_t \geq \overline{p}^1$. CPT 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 $\delta < \alpha$).*

3.5 Properties of Equilibrium

The fact that CPT investors buy or sell independently of their private signals as prices become extreme generates behavior that looks very much like herding or contrarian behavior. However, note that the history of prices is actually irrelevant for behavior: a CPT investor that faces an extreme price behaves identically even if she does not observe the past history. For this reason, I refrain from referring to CPT investor behavior as herding or contrarian.

Instead, I define two types of investors according to the skewness preference that drives their behavior.

Definition 1 (Skew-Loving Behavior):

1. *An informed investor exhibits negative skew-loving (NSL) behavior 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 exhibits positive skew-loving (PSL) behavior if, independently of her signal, she (i) buys when* $p_t < \frac{1}{2}$*, or (ii) sells when* $p_t > \frac{1}{2}$*.*

In addition to the definitions of positive and negative skew-loving, I use the term *unresponsive* to refer to an action that is independent of one's private information, encompassing both behaviors. Corollary 1 shows that either positive or negative skew-loving behavior occurs generically (unless $\delta = \alpha$) in the model. The corollary follows directly from the definitions in Definition 1 and Theorem 1.

Corollary 1 (Skew-Loving 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* $p_t < \underline{p}^0$ *and* $p_t > 1 - \underline{p}^0$*, CPT investors exhibit negative skew-loving behavior.*
2. *For all parameterizations with* $\delta < \alpha$*, there exists a public belief,* $\overline{p}^1 < 1$*, such that for all* $p_t > \overline{p}^1$ *and* $p_t < 1 - \overline{p}^1$*, CPT investors exhibit positive skew-loving behavior.*

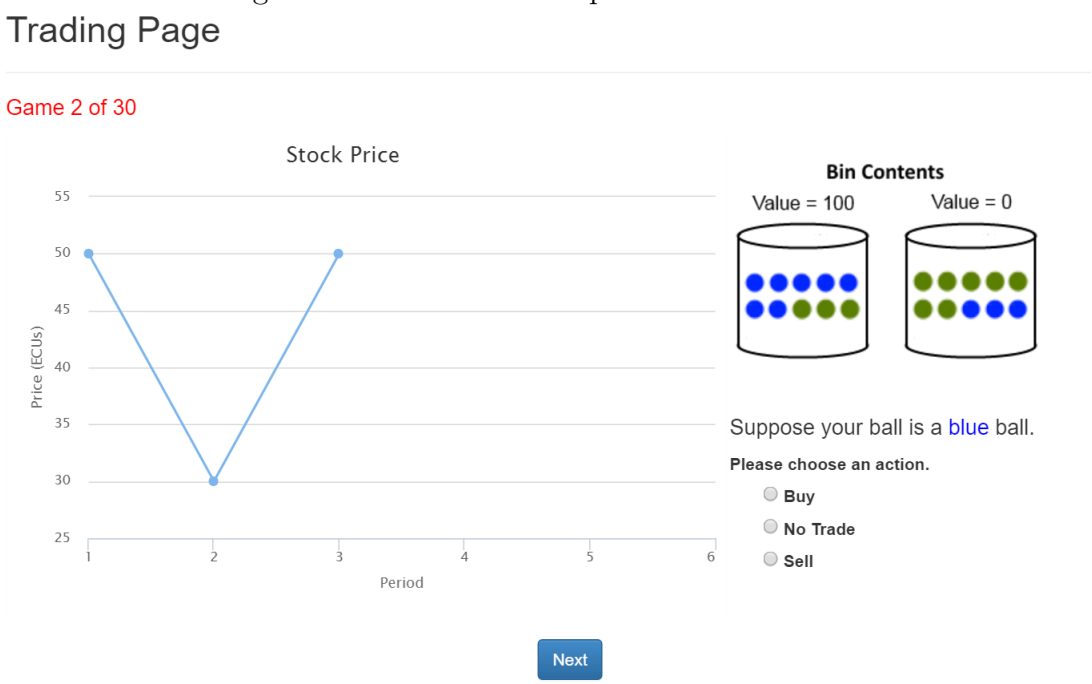
The definitions of positive and negative skew-loving are very similar to the typical definitions for herding and contrarian behavior (for example, see Avery and Zemsky (1998) and Park and Sabourian (2011)). The difference is that, herding and contrarianism are defined as a *change* in behavior after observing a series of trades. For example, an investors herds if she would sell at $p_t = \frac{1}{2}$, but buys at $p_t > \frac{1}{2}$ after others' trades increase prices. Herding and contrarianism cannot occur in a Glosten-Milgrom model (Avery and Zemsky (1998)) without introducing additional dimensions of uncertainty (Avery and Zemsky (1998)) or more than two states of the world (Park and Sabourian (2011)). Nevertheless, we see that CPT investors follow behavioral patterns that look very much like herding and contrarian behavior.

4 Experiment

4.1 Design

The goal of the experiment is to provide evidence of herding and contrarian-like behavior in a setting which controls for sources of these behaviors other than skew-loving preferences,

Figure 3: Screenshot of Experimental Interface



and, furthermore, to test the specific predictions of the model at an *individual* level (preferences, after all, being individual-specific). To achieve these goals, the implementation of the experiment differs from the model in one important way. Rather than have subjects arrive to the market and trade one at a time as in the model, I convert the strategic environment to an individual decision problem. Doing so eliminates the 'herd mentality' explanation for herding because there are no previous investors to imitate, and also also eliminates strategic ambiguity on the part of the subjects, which both Cipriani and Guarino (2005) and Drehmann, Oechssler, and Roider (2005) have shown can produce contrarian-like behavior. It also allows me to collect sufficient data to characterize individual behavior.

In the individual-decision version of the problem, asset values are represented by urns with 7 balls of one color and 3 of another ($q = 0.7$) as in Figure 3. A subject observes a history of prices determined by a sequence of random, public signals (urn draws). 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. After observing the price path, a subject is then asked if she would like to buy or sell the asset (or abstain) for each possible realization of her private signal (as in Cipriani and Guarino, 2009), allowing me to directly observe unresponsive behavior. Finally, the subject's private signal is drawn, her corresponding trade is executed, and she receives feedback on both her realized signal and the resulting payoff.

In the 'NO SOCIAL' treatment just described, where social factors are precluded, sub-

jects may still fail to form correct Bayesian posteriors given the public and private signals, which could induce herd or contrarian-like behavior. For example, if subjects overextrapolate from the price trend, they may trade in the same direction regardless of their private signal. To test whether or not belief errors are responsible for (or partially responsible for) unresponsive behavior, I conduct a second treatment (the 'NO INFERENCE' treatment) in which, in addition to shutting down social forces, I explicitly provide subjects with the posterior probability that the asset is valuable conditional on the price and their private signal.[15]

In both treatments, I follow the previous experimental literature (Cipriani and Guarino (2005) and Drehmann, Oechssler, and Roider (2005)) by having the market maker in the experiment (the experimentalist) post only a single price equal to the expected value of the asset, p_t, rather than separate bid and ask prices ($p_1 = \frac{1}{2}$). 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.[16] Importantly, however, the predictions of the model are unchanged when trades occur at a single price, except that the transition regions of Theorem 1 simplify to unique threshold prices (see Appendix B for details).

I collected the experimental data using undergraduate subjects 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. Each subject took part in 30 consecutive 'games' of a single treatment (NO SOCIAL or NO INFERENCE) in which they faced 30 historical price paths and made 60 trading decisions (one for each possible realization of their private signal in each game). Subjects were paid for their decisions in each game with average earnings of $17.13. The experiment typically finished in just over an hour.

4.2 Hypotheses

The first hypothesis is that we continue to observe herding and contrarian-like behaviors even in the absence of imitation motives or strategic ambiguity. Furthermore, if preferences, rather than mistaken beliefs, drive these behaviors, we expect to observe these skew-loving behaviors at least as as frequently in the NO INFERENCE treatment, where Bayesian errors

[15] Bisiere, 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 posterior in the NO INFERENCE treatment. See the instructions in Appendix F.

[16] Subjects are explicitly told that the price reflects the expected value of the asset given all public information (the public signals).

play no role, as in the NO SOCIAL treatment.

Hypothesis 1 (Treatment Effect):
A. Positive and negative-skew loving behavior is present in both treatments.
B. Providing subjects with the correct Bayesian beliefs does not reduce its frequency.

In addition to the presence of unresponsive behavior, Theorem 1 shows that prospect-theory preferences are expected to generate particular patterns in the data. Specifically, unresponsive behavior should become more frequent at more extreme prices, at the expense of risk-neutral behavior (trading in the direction of one's private signal as in Lemma 1) and abstaining from trade.

Hypothesis 2 (Aggregate Behavior):
A. The frequencies of no trade and risk-neutral behavior decrease as p_t increases or decreases from $p_t = \frac{1}{2}$.
B. The frequencies of unresponsive behavior increase as p_t increase or decreases from $p_t = \frac{1}{2}$.

4.3 Experimental Results

In Section 4.3.1, I compare behavior across treatments, showing that providing subjects with the correct Bayesian beliefs actually significantly *increases* the amount of herding-like behavior. I then take a more detailed look at aggregate behavior, providing evidence consistent with Hypothesis 2. In Section 4.3.2, I turn to individual behavior, showing that a majority of subjects exhibit skew-loving preferences, and that their decisions are better described 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 during which subjects are becoming familiar with the interface and environment.[17]

To test Hypothesis 1, I provide the percentage of each type of behavior in Table 1, by treatment. I discuss the 'Other' category, which consists primarily of trades for one signal realization but not the other, at the end of this section.

[17]Including this data does not affect any of the qualitative results, nor does restricting the analysis to only the second half of the data. Learning seems to play a very limited role: the results of classifying individual subjects in Section 4.3.2 are remarkably similar when using only the first or second half of the data. Therefore, for maximum statistical power, I drop only the first three trials.

Table 1: Aggregate Behavior by Treatment

Treatment	No Trade	Risk-Neutral	NSL ('Herding')	PSL ('Contrarian')	Other
NO SOCIAL	5.2	37.4	13.9	12.8	30.8
NO INFERENCE	5.5	27.5	34.9	10.6	21.6

Note: Percentages of each type of behavior in the NO SOCIAL treatment (no Bayesian beliefs) and the NO INFERENCE treatment (correct Bayesian beliefs provided to subjects). NSL and PSL refer to negative and positive skew-loving, respectively.

From Table 1 it is clear that the standard prediction of risk-neutral behavior does not provide an adequate description of the data in either treatment. Instead, unresponsive behavior is observed in both treatments, confirming Hypothesis 1, part A. Furthermore, providing subjects with the correct Bayesian posteriors does not reduce unresponsive behavior as we'd expect if Bayesian errors drive this behavior. In fact, the opposite occurs - providing subjects with the correct Bayesian beliefs significantly *increases* the frequency of negative skew-loving (herding-like) behavior, in contrast to what a model based on extrapolative expectations would predict.[18] Instead, these results are 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 overweighting their private signal or under-weighting the price (e.g. Goeree at. al. (2007) and Weizsacker (2010)). Logit regressions (with errors clustered by subject) of each type of behavior (versus not) on a treatment dummy confirm that negative skew-loving 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 1, part B, obtaining the first result.

Result 1 (Treatment Effect): *Providing subjects with the correct Bayesian posteriors does not reduce the frequency of unresponsive behavior. Instead, it leads to an increase in negative skew-loving (herding-like) behavior and a decrease in risk-neutral behavior.*

The fact that subjects observe both monotonic and non-monotonic price paths that lead to the same price allows me to test for the effect of beliefs in another way, by testing for path-dependence directly. If subjects are overextrapolating, we might expect herding-like behavior to be more frequent when the price path is monotonic. 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.[19]

[18] Bisiere, Decamps, and Lovo (2015) find a similar result when they compare their ME treatment (which requires Bayesian updating) to both their LE or SME treatments (which do not).

[19] Specifically, I run a logit regression of 'herding' versus no 'herding' on a dummy variable that indicates

Table 2: Aggregate Behavior by Price

Treatment	Normalized Price	No Trade	Risk-Neutral	NSL	PSL	Other
NO SOCIAL	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
NO INFERENCE	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

Note: Percentages of each type of behavior at a given price. NSL and PSL refer to negative and positive skew-loving, respectively.

To test for the particular patterns in the data predicted by Hypothesis 2, I now condition behavior on what I refer to as the 'normalized' price, a measure of the extremeness of the price that treats rising and falling prices symmetrically.[20] Table 2 breaks down the frequency of each type of behavior by this price.

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 NO SOCIAL treatment ($p = 0.047$), as hypothesized, but increases in the NO INFERENCE treatment ($p = 0.005$). One possible reason for the increase is that standard expected utility with a large degree of risk aversion can cause no trade at extreme prices, a possibility I consider when analyzing individual behavior in Section 4.3.2.

NSL (herding-like) behavior increases with the normalized price in both treatments ($p = 0.000$ in both). For PSL (contrarian-like) behavior, the corresponding regressions have positive, but insignificant coefficients, likely reflecting the fact that there is relatively little PSL behavior in either treatment and therefore low statistical power. Overall, the aggregate evidence is mostly supportive of Hypothesis 2, part B, with the exception of the increase in

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.

[20]Formally, the normalized price is defined as p_t if $p_t \geq \frac{1}{2}$ and $1 - p_t$ if $p_t < \frac{1}{2}$. Subjects do not always treat rising and falling prices symmetrically (see Section 4.3.2). Therefore, at the aggregate level, the results should be interpreted as average effects over rising and falling prices.

no trade behavior with the normalized price in the NO INFERENCE treatment.

Result 2 (Aggregate Behavior): *In the aggregate, risk-neutral behavior decreases with the normalized price, and negative skew-loving (herding-like) behavior increases, as predicted by the model. Positive skew-loving (contrarian-like) behavior increases, but not significantly so. No trade is predicted to decrease, but does so in the NO SOCIAL treatment only.*

A final aggregate prediction of the model is that we should observe partially unresponsive behavior in which a subject buys at a high (or low) price with a favorable signal but abstains with an unfavorable signal (or sells at a low (high) price with an unfavorable signal but abstains with a favorable signal). This behavior is 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.[21] Most of the behavior in the 'Other' category in Tables 1 and 2 consists of these behaviors.[22]

4.3.2 Individual Behavior

Given the preceding evidence that unresponsive behavior is driven by preferences, I now ask whether or not prospect theory preferences in particular are the best explanation. To do so, I assess the fit of the model's predictions on an individual basis by comparing four candidate models.

The first model is the main one of interest: CPT preferences. This model has two degrees of freedom: loss aversion (λ) and the difference between the utility curvature and probability weighting parameters ($\delta - \alpha$) (see Appendix B for details).[23] The second is CPT preferences, but without loss aversion ($\lambda = 1$). The third is expected utility. Because risk-neutrality is incapable of explaining either herding-like or contrarian-like behavior (as shown in Section 3.2), I consider general CRRA preferences (the results are almost identical with CARA preferences).[24] Finally, I consider a model which only requires that decisions respect

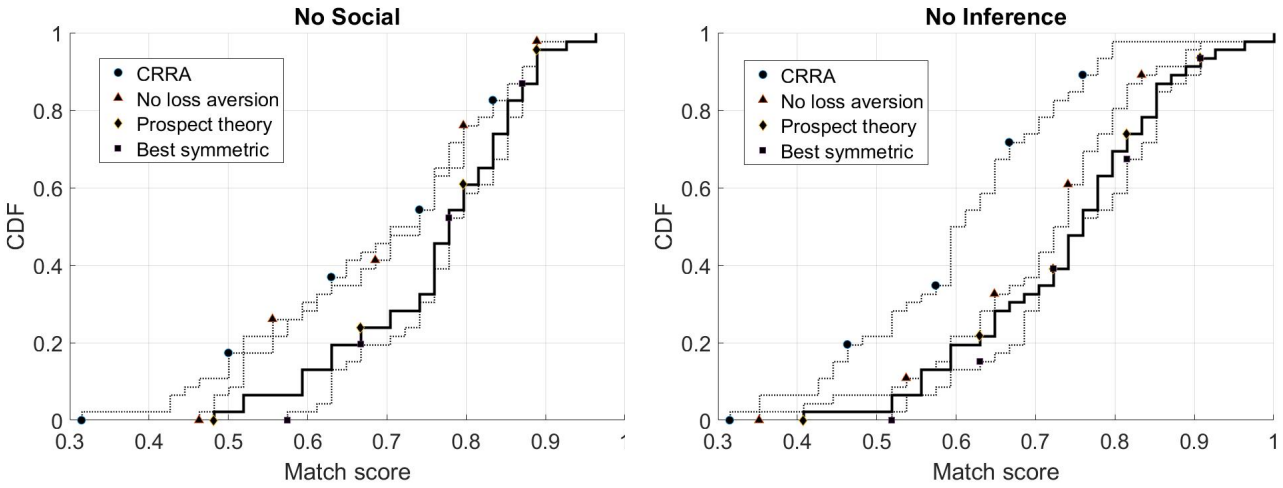
[21] Cipriani and Guarino (2009) also document partially unresponsive behavior in Table 1 of their paper.

[22] The remaining small fraction of behavior in the 'Other' category 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 NO SOCIAL and NO INFERENCE treatments, respectively.

[23] The theory states that an optimal CPT strategy is characterized by two thresholds, p^0 and p^1, at which an investor's behavior transitions. To generate the match score for CPT preferences, I compare the data to all possible trade patterns, subject to the restrictions on the thresholds that come out of the theory (see Appendix B for further details).

[24] Risk-aversion can not produce herding-like or contrarian-like behavior and risk-seeking cannot produce abstention (see Appendix D for formal propositions and proofs of these claims), which limits the ability of

Figure 4: Match Scores



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.

symmetry (the decision at a price p_t with a favorable signal must match that at a price $1 - p_t$ with an unfavorable signal). This model has 27 degrees of freedom, allowing different decisions for each normalized price and private signal. It provides an upper bound on the ability of any model that treats decisions and asset values symmetrically to fit the data (see Appendix D for further discussion).

To compare models, I employ the technique of Bisiere, Decamps, and Lovo (2015) to calculate a match score for each subject relative to each model of behavior. For a given model and a given (range of) preference parameters, I obtain predicted actions for each discrete price and for each private signal. With these predictions, I award 0.5 for each action that matches the prediction (and zero otherwise), and then divide by the maximum possible score so that the match score lies between 0 and 1. Specifically, for each model, I calculate the match score for each individual over the last 27 price paths (54 decisions) she faces. Figure 4 provides the empirical CDFs of the match scores in each treatment for each of the four models.

A particularly striking result in Figure 4 is that prospect theory explains choices in both treatments as well as the very lenient model that only imposes symmetry, even though it has an order of magnitude less degrees of freedom ($p = 0.765$ and $p = 0.269$ with a Kolmogorov-Smirnov test in the NO SOCIAL and NO INFERENCE treatments, respectively).[25] Furthermore, in the NO INFERENCE treatment in which there is no role for

expected utility to explain the data.

[25]The model that only imposes symmetry matches about 75% of a subject's decisions at the median, similar to the finding of Bisiere, Decamps, and Lovo (2015). Buying is more frequent in both treatments, perhaps because it is more familiar to subjects than selling short.

Table 3: Individual Types and Match Scores

Treatment	Models considered	Risk-Neutral	Risk-Averse	Risk-Seeking	NSL	PSL
NO SOCIAL	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
NO INFERENCE	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

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. NSL and PSL refer to negative and positive skew-loving, respectively.

Bayesian errors, even restricting prospect theory to not include loss aversion does almost as well as the full prospect theory model, and does significantly better than expected utility, both of which have a single degree of freedom ($p = 0.001$ in a Kolmogorov-Smirnov test). In the NO SOCIAL treatment, prospect theory without loss aversion and expected utility fit similarly well.[26] However, the full prospect theory model continues to do better than expected utility ($p = 0.057$) as it does in the NO INFERENCE treatment ($p = 0.000$).

Prospect theory provides an overall better fit of the data 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.[27]

Table 3 shows that the majority of subjects are better classified by prospect theory in both treatments. In the NO SOCIAL treatment where Bayesian errors play a role, a number of subjects are classified as risk-averse or risk-seeking, but in the NO INFERENCE treatment, these types almost disappear, being replaced by NSL (herding-like) types. In the NO INFERENCE treatment, which provides the cleanest test of the theory, 76% of subjects are better described by prospect theory, which is interestingly similar to the 80% of

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

[27] I break ties in favor of expected utility and don't include the symmetric model because it mechanically provides the best fit for every subject.

subjects that Bruhin, Fehr-Duda, and Epper (2010) report. 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-like (NSL) behavior is much more common than contrarian-like (PSL) behavior - approximately 90% exhibit herding-like behavior, providing the third result:[28]

Result 3 (Individual Behavior): *The majority of subjects in both treatments are better described by prospect theory than expected utility (CRRA) and, of these, negative skew-loving (herding-like) strategies are dominant. Furthermore, prospect theory fits the data as well as any model that imposes symmetry.*

The fact that negative skew-loving (herding-like) strategies are more popular means that utility curvature dominates probability weighting for most subjects ($\delta > \alpha$). In Kahenman and Tversky (1992) the median subject has the opposite preferences ($\delta < \alpha$). However, this finding is by no means unique. Across the ten studies summarized in Table A.3 of Glimcher and Fehr (2013) that use the same probability weighting function I use, four find $\delta > \alpha$ at the median.

The dominance of negative-skew loving behavior may also be surprising given the evidence of preferences for positively skewed assets found in real-world markets (Kumar (2009), Green and Hwang (2012), Eraker and Ready (2015)). However, this difference is perhaps easily reconciled by the fact that, in a controlled laboratory experiment, the participants are a random sample (of a particular population), whereas participants in real-world financial markets self-select into the market.

5 Discussion

Herding and contrarian strategies are of interest to financial economists because they do not reveal private information to the market, reducing the informational efficiency of prices. These strategies are defined by imitation (or anti-imitation), relying on the observation of past histories of actions and/or prices. In this paper, however, I have shown both theoretically and via a laboratory experiment that strategies which very much look like herding or contrarian strategies emerge from preferences over skewness, having little to do with past trends. In the experiment, the vast majority of subjects exhibit a preference for negative

[28] A likely explanation for the difference in behavior from previous experiments is the fact that the prior experiments implemented the actual strategic game. In the game, if some subjects understand the tendency of others to use herding-like strategies, then prices are on average too extreme, making contrarian-like behavior a best response.

skewness which generates herd-like strategies that are of particular concern due to their destabilizing effect on markets (De Long et al. (1990)).[29]

In the model, the relationship between price trends and return skewness is driven by the binary nature of the asset value.[30] However, increasing prices are associated with negatively skewed returns (and conversely) even in actual market data with unbounded asset prices (Harvey and Siddique (1999) and Chen, Hong, and Stein (2001)), providing the necessary relationship for the main mechanism of the model to hold.

To demonstrate that skewness causes CPT investors to trade against private information even when assets are non-binary, I conduct a numerical calibration exercise using market index data from the Center for Research in Security Prices (CRSP) for the period 1926 to 2015 (details in Appendix C). I first show a result similar to that of Chen, Hong, and Stein (2001): for six-month returns in the highest decile, the daily returns over the subsequent month are *negatively* skewed, but for past returns in the lowest decile, the subsequent returns are instead *positively* skewed. I then consider an investor with the preference parameters of the modal negative skew-loving subject in the experiment who has a private signal about the mean return.[31] I find that, after rising prices, this investor prefers to buy the asset even when her private information gives her an (annualized) expected return of -7.3% (with 95% confidence interval, -11.6% to -2.8%). After falling prices, the result is almost exactly reversed: the investor requires private information giving her at least an expected return of 7.1% (with 95% confidence interval, 0% to 15.7%) in order to prefer buying over selling. Therefore, preferences for skewness can cause CPT investors to exhibit herd-like behavior in real markets, forgoing economically significant returns.

Finally, a broader theme suggested by this paper is that prospect theory preferences can play a critical role in market and game experiments.[32] In interpreting experimental results, it is common to consider risk-aversion, but when decisions involve highly skewed payoffs, prospect theory may be a better candidate for describing behavior. An interesting avenue for future research is to study the extent to which subjects understand that other subjects'

[29]Herding and contrarian-like strategies have been observed in a similar experiment with financial market professionals (Drehmann, Oechssler, and Roider (2005)), suggesting that skewness preferences may play a role in their decisions as well.

[30]The results are therefore directly applicable to 'near binary' assets such as options or initial public offerings (Green and Hwang (2012)).

[31]With non-binary asset values, δ and α have independent effects on choices. Thus, I must impose one to specify the other. As the modal subject is negative skew-loving, I assume no probability weighting ($\delta = 1$) which then implies $\alpha = 0.5$ and $\lambda = 1$ for the modal subject.

[32]Examples of experimental environments in which subjects face decisions with skewed payoffs so that skewness 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)).

skewness preferences may drive their decisions.

References

[1] 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

[3] 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

[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., Greenwood, R., Jin, L., and Shleifer, A. (2015), "X-CAPM: An extrapolative capital asset pricing model", *Journal of Financial Economics*, 115, 1-24

[8] Barberis, N., Greenwood, R., Jin, L., and Shleifer, A. (2016), "Extrapolation and bubbles", mimeo, Yale University

[9] 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

[10] Barberis, N. and Xiong, W. (2012), "Realization utility", *Journal of Financial Economics*, 104, 251-271

[11] 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

[12] 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

[13] Boortz, C. (2016), "Irrational Exuberance and Herding in Financial Markets", SFB 649 Discussion Paper 2016-016

[14] 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

[15] Bond, P., Edmans, A., and Goldstein, I. (2012), "The real effects of financial markets", *Annual Review of Financial Economics*, 4, 339-360

[16] Bruhin, A, Fehr-Duda, H., and Epper, T. (2010), "Risk and rationality: uncovering heterogeneity in probability distortion", *Econometrica*, 78 (4), 1375-1412

[17] Charness and Levin (2009), "The origin of the winner's curse: A laboratory study", *American Economic Journal: Microeconomics*, 1, 1, 207-36

[18] Chen, J., Hong, H., and Stein, J. (2001), "Forecasting crashes: trading volume, past returns, and conditional skewness in stock prices", *Journal of Financial Economics*, 61, 345-381

[19] Cipriani, M. and Guarino, A. (2005), "Herd behavior in a laboratory financial market", *American Economic Review*, 95 (5), 1427-1443

[20] Cipriani, M. and Guarino, A. (2008), "Transaction costs and informational cascades in financial markets", *Journal of Economic Behavior & Organization*, 68, 581-592

[21] 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

[22] Cipriani, M. and Guarino, A. (2014), "Estimating a structural model of herd behavior in financial markets", *American Economic Review*, 104 (1), 224-251

[23] De Long, J., Shleifer, A, Summers, L, and Waldmann, R. (1990), "Positive feedback investment strategies and destabilizing rational speculation", *The Journal of Finance*, 45 (2), 379-395

[24] Devenow, A. and Welch, I. (1996), "Rational herding in financial economics", *European Economic Review*, 40, 603-615

[25] 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

[26] Ford, J., Kelsey, D., and Pang, W. (2013), "Information and ambiguity: herd and contrarian behaviour in financial markets", *Theory and Decision*, 75 (1), 1-15

[27] Eraker, B. and Ready, M. (2015), "Do investors overpay for stocks with lottery-like payoffs? An examination of the returns of OTC stocks", *Journal of Financial Economics*, 115 (3), 486-504

[28] Glimcher, P. and Fehr, E. (2013), "Neuroeconomics: decision making the brain", Academic Press, 2nd edition

[29] 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

[30] Green, T. and Hwang, B. (2012), "Initial public offerings as lotteries: skewness preference and first-day returns", *Management Science*, 58 (2), 432-444

[31] Greenwood, R. and Shleifer, A. (2014), "Expectations of returns and expected returns", *Review of Financial Studies*, 27, 714-746

[32] Goeree, J., Palfrey, T., Rogers, B., and McKelvey, R. (2007), "Self-correcting information cascades", *Review of Economic Studies*, 74, 733-762

[33] 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

[34] 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

[35] 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

[36] Harvey, C. and Siddique, A. (2000), "Conditional skewness in asset pricing tests", *The Journal of Finance*, LV (3), 1263-1296

[37] 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

[38] Ingersoll, J. and Jin, L. (2013), "Realization utility with reference-dependent preferences", *Review of Financial Studies*, 26 (3), 723-767

[39] Ingersoll, J. (2016), "Cumulative prospect theory, aggregation, and pricing", *Critical Finance Review*, 4, 1-55

[40] Kahneman, D. and Tversky, A. (1992), "Advances in prospect theory: cumulative representation of uncertainty", *Journal of Risk and Uncertainty*, 5, 297-323

[41] Koszegi, B. and Rabin, M. (2006), "A model of reference-dependent preferences", *The Quarterly Journal of Economics*, CXXI (4), 1133-1165

[42] Koszegi, B. and Rabin, M. (2007), "Reference-dependent risk attitudes", *American Economic Review*, 97 (4), 1047-1073

[43] Kumar, A. (2009), "Who gambles in the stock market?", *The Journal of Finance*, 64, 4, 1889-1933

[44] 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

[45] Li, Y. and Yang, L. (2013), "Prospect theory,the disposition effect, and asset prices", *Journal of Financial Economics*, 107, 715-739

[46] Mackay, C. (1841), "Memoirs of extraordinary popular delusions and the madness of crowds", (1 ed.). London: Richard Bentley

[47] 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.

[48] Palfrey, T. and Wang, S. (2012), "Speculative overpricing in asset markets with information flows", *Econometrica*, 80 (5), 1937-1976

[49] Park, A. and Sgroi, D. (2012), "Herding, contrarianism, and delay in financial market trading", *European Economic Review*, 56, 1020-1037

[50] Park, A. and Sabourian, H. (2011), "Herding and contrarian behavior in financial markets", *Econometrica*, 69 (4), 973-1026

[51] Qin, J. (2015), "A model of regret, investor behavior, and market turbulence", *Journal of Economic Theory*, 160, 150-174

[52] Rabin, M. (2000), "Risk-aversion and expected-utility theory: a calibration theorem", *Econometrica*, 68 (5), 1281-1292

[53] Roth, A. and Malouf, M. (1979), "Game theoretic models and the role of information in bargaining", *Psychological Review*, 86, 575-594

[54] 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, risk-neutral investors always trade according to their private information, independently of the equilibrium strategies of the CPT 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 CPT) sells at some p_t with positive probability in equilibrium, then a type 0 investor must also sell at p_t with probability one. For the CPT investor, this fact follows from 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 $B_t - b_t^1 \geq b_t^1 - A_t$ so that she weakly prefers selling to buying. Rearranging, $B_t + A_t \geq 2b_t^1 > 2b_t^0$ 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

$$\begin{array}{rcl} b_t^1 - A_t & > & 0 \\ \iff \frac{p_t q}{p_t q + (1-p_t)(1-q)} & > & \frac{p_t Pr(a_t = buy|V=1)}{p_t Pr(a_t = buy|V=1) + (1-p_t) Pr(a_t = buy|V=0)} \\ \iff \frac{q}{1-q} & > & \frac{Pr(a_t = buy|V=1)}{Pr(a_t = buy|V=0)} \end{array} \tag{7}$$

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

$$\begin{array}{rcl} Pr(a_t = buy|V = 1) & = & \frac{1-\mu}{2} + \mu\gamma\beta^{RN}|(V = 1) + \mu(1-\gamma)\beta^{PT}|(V = 1) \\ Pr(a_t = buy|V = 0) & = & \frac{1-\mu}{2} + \mu\gamma\beta^{RN}|(V = 0) + \mu(1-\gamma)\beta^{PT}|(V = 0) \end{array} \tag{8}$$

and $\beta^{RN}|(V = x)$ and $\beta^{PT}|(V = x)$, $x \in \{0, 1\}$ are the probabilities the market maker believes risk-neutral and CPT 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 CPT investor with $s_t = y$ buys. The right-hand side of equation (7) can be shown to be strictly increasing in $\beta^{1,PT}$ and strictly decreasing in $\beta^{0,PT}$ so that it 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 have[33]

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

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 $\mu < 1$. □

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 CPT investors).

Part 2 was proven in Lemma 1.

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

$$\begin{array}{rcl} buy\ if \quad 1 & \geq & \lambda\left(\frac{1-q}{q}\frac{Pr(a_t=buy|V=1)}{Pr(a_t=buy|V=0)}\right)^{\alpha} \\ sell\ if \quad 1 & \leq & \frac{1}{\lambda}\left(\frac{1-q}{q}\frac{Pr(a_t=sell|V=1)}{Pr(a_t=sell|V=0)}\right)^{\alpha} \end{array} \tag{9}$$

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

[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.

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

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

where, as in the proof of Lemma 1, $\beta^{PT}|(V = 1) = q\beta^{1,PT} + (1-q)\beta^{0,PT}$ and $\beta^{PT}|(V = 0) = (1-q)\beta^{1,PT} + q\beta^{0,PT}$. Now, as argued in the proof of Lemma 1, if a type 0 investor buys, then so must a type 1 investor. This fact implies $\beta^{1,PT} \geq \beta^{0,PT}$ which in turn implies 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 (9) 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

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

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

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

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

$$\begin{aligned} 1 &\geq \lambda \left(\tfrac{1-q}{q} \tfrac{\frac{1-\mu}{2} + \mu q}{\frac{1-\mu}{2} + \mu(1-q)} \right)^{\alpha} \\ \iff \lambda &\leq \left(\tfrac{q}{1-q} \tfrac{\frac{1-\mu}{2} + \mu(1-q)}{\frac{1-\mu}{2} + \mu q} \right)^{\alpha} \quad \equiv \underline{\lambda} \end{aligned}$$

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 (8), the two equations governing buy decisions are given by

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

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, $1 - \underline{p}^0$ and $\overline{p}^1$, respectively. These beliefs are given by the unique solutions to

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

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

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

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

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

where I have used $\eta^{0,PT}$ and $\eta^{1,PT}$ to denote the probabilities with which type 0 and type 1 CPT 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: $(\underline{p}^0, \overline{p}^0)$ for the type 1 investor and $(1-\overline{p}^1, 1-\underline{p}^1)$ for the type 0 investor, where $\overline{p}^0 < 1-\overline{p}^1$. 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 $(\underline{p}^0, \overline{p}^0)$), followed by the type 0 investor (over $(1-\overline{p}^1, 1-\underline{p}^1)$).

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,

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

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

B. Prospect Theory Parameters

With trades taking place at a single price, p_t, each period, the optimal strategy of a CPT investor with a favorable private signal (from (5)) becomes

$$\begin{array}{lll} buy\ if & \left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} & \geq \lambda\left(\frac{1-q}{q}\right)^{\delta} \\ sell\ if & \left(\frac{p_t}{1-p_t}\right)^{\delta-\alpha} & \leq \frac{1}{\lambda}\left(\frac{1-q}{q}\right)^{\delta} \end{array} \tag{10}$$

where t is the time of trade (after $t-1$ public signal draws). For an investor with an unfavorable signal, the ratio of $1-q$ to q is again inverted in each inequality. From (10), it is clear that one can define two threshold prices (that replace the transition regions that exist when there is asymmetric information) at which behavior transitions. The equilibrium is otherwise identical to that described in Theorem 1.

I now use (10) to establish that the model only has two degrees of freedom and then to find the set of pairs of threshold prices that it can support. I consider an investor with a favorable signal - the thresholds for an investor with an unfavorable signal follow by symmetry around $p_t = \frac{1}{2}$.

The price depends only on the difference in the number of favorable and unfavorable public signals. Denoting the difference, k, we have $p_t = \frac{q^k}{q^k+(1-q)^k}$. (10) can then be written

$$\begin{array}{lll} buy\ if & Q^{k(\delta-\alpha)+\delta} & \geq \lambda \\ sell\ if & Q^{k(\delta-\alpha)+\delta} & \leq \frac{1}{\lambda} \end{array}$$

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

$$\begin{array}{lll} buy\ if & Q^{k+\frac{1}{\delta'}} & \geq \lambda' \\ sell\ if & Q^{k+\frac{1}{\delta'}} & \leq \frac{1}{\lambda'} \end{array}$$

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 $\lambda = \lambda'^{\delta'}$. I make this assumption when performing the calibration exercise because, with non-binary asset values, both δ and α must be specified (not just their difference).

C. Calibration

Here, I use the parameters identified in the experiment to perform a calibration exercise using actual market data. Daily index return data comes from CRSP: the NYSE/AMEX/NASDAQ index (value-weighted, including distributions). I use the full history of available data: Jan, 1926 to Dec, 2015.[34] Following Chen, Hong, and Stein (2001), I use the prior six-month return as a measure of the current price level. Each month, I calculate the index return over the previous six months, sort the resulting returns, and take the lowest and highest deciles. For each month in each decile, I use the entire month's daily returns to create a conditional (daily) return distribution. In this way, I construct two conditional return distributions, one for which the previous six-month returns were abnormally large ('High Price') and one for which they were abnormally small ('Low price').[35]

Table A1 provides the mean, standard deviation, and skewness measures for each of the two conditional distributions.[36] It includes the corresponding 95% confidence intervals calculated using a percentile bootstrap. We see that the average return does not differ substantially at low and high prices, with significant overlap of the confidence intervals, but the standard deviation of returns is significantly lower at high prices. More importantly, after price increases, the index returns become more negatively skewed as in the case of a binary asset value, which provides the potential for unresponsive behavior.

Consider the investment decision of a negative-skew loving CPT investor with $\alpha = 0.5$ and $\delta = \lambda = 1$ who faces each of these conditional return distributions. The model's prediction is that as the price increases, the investor moves from selling even with a favorable signal to buying even with an unfavorable signal. In other words, at low prices the investor prefers to sell even when she could earn a positive return by buying and at high prices prefers to buy even when it results in a negative return. For simplicity, I model the investor's private information as information about the mean expected return and calculate the private signal that would make the investor just indifferent between buying and selling.[37] For a risk-neutral investor, the required signal for indifference is equal to the expected mean return of the asset, and the resulting expected return conditional on her private information is zero. Table A2

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

[35]For example, if the six-month return leading up to June 1, 1990 were in the lowest decile, the daily returns for June, 1990, would form part of the conditional distribution for the Low Price decile.

[36]The skewness is calculated with $SKEW = \frac{E(r-\overline{r})^3}{STD(r)^3}$ where $\overline{r}$ is the mean return and $STD(r)$ is its standard deviation.

[37]Because $\lambda = 1$ and the decision weight functions are the same for losses and gains, the utility from buying is equal to the opposite of the utility from selling. Therefore, indifference between buying and selling also implies indifference with not trading.

Table A1: Conditional Moments of the Market Index Daily Returns

Price Condition	Moment	Estimate	95% Confidence interval
Low Price	Mean	0.00041	(−0.00027,0.0011)
	Standard deviation	0.018	(0.017,0.019)
	Skew	0.65	(0.18,1.22)
High Price	Mean	0.00064	(0.00022,0.0011)
	Standard deviation	0.011	(0.010,0.012)
	Skew	−0.52	(−1.21,0.20)

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

Table A2: Required Returns for Prospect Theory Investor

Price Condition	Indifference Condition	Estimate	95% Confidence Interval
Low Price	Minimum return to buy	7.1%	(0.0%,15.7%)
High Price	Minimum return to sell	−7.3%	(−11.6%,−2.8%)

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

provides the corresponding annualized private return for the CPT investor along with its 95% confidence interval.

Table A2 shows that, at high prices, a prospect theory investor prefers to buy the asset any time her private information indicates an expected return of -7.4% or greater. At low prices, she prefers to sell the asset any time her private expected return is 6.8% or less.[38] We therefore obtain results similar to those with a binary asset value. Because the skewness of actual returns varies with the price level, and because skewness matters for an investor with prospect theory preferences, such an investor prefers to trade with the trend and against relatively strong private signals when facing actual market return distributions.

[38]The 95% confidence intervals do not overlap indicating the difference across high and low prices is significant.

D. Alternative Preference Models

D.1. Expected Utility

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

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

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

$$\begin{array}{lrcl} buy\ if & b_t u(1-A_t) + (1-b_t)u(-A_t) & \geq & u(0) \\ sell\ if & b_t u(B_t-1) + (1-b_t)u(B_t) & \geq & u(0) \end{array} \tag{13}$$

and otherwise abstain from trading.[39] As in the main model, it is possible to show that because of uninformed investors, we must have $b_t > A_t$ (favorable signal) or $b_t < B_t$ (unfavorable signal), in which case risk-neutral investors trade according to their private information as shown in Lemma 1. Consider an investor that is not risk-neutral then, and assume she has a favorable signal (symmetric arguments hold for unfavorable signals).

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

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

D.2. Symmetry

In Section 4.3.2, I consider the 'best symmetric' model, a model that only imposes symmetry around a price of one-half. Here, I discuss the generality of this model in more detail.

The model imposes only the requirement that if a subject buys, sells, or abstains at a price, p, when she has a particular private signal, then she must sell, buy, and abstain (respectively) at a price of $1-p$ with the opposing private signal. This requirement is very natural because of the symmetry in payoffs around a price of one-half. Buying at a price of p with a particular belief, b, provides a binary gamble which returns $1-p$ with probability b and $-p$ with probability $1-b$. Selling at a price of $1-p$ with a belief, b', provides a

[39] I have normalized initial wealth to zero without loss of generality.

binary gamble which returns $1-p$ with probability $1-b'$ and $-p$ with probability b'. Thus, if $b = 1-b'$, which is the case with Bayesian updating, the opposing actions at symmetric prices with opposing signals provide identical gambles.[40]

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

E. Expectations-Based Reference Point

Here, I show that the expectations-based reference point preferences of Koszegi and Rabin (2006,2007) are incapable of generating herding-like behavior in the laboratory environment, even when extended to allow for decision weights. Koszegi and Rabin (2007) consider multiple models of reference point formation. First, they consider the case in which expectations are taken as given, suggesting that this occurs when the outcome is realized shortly after a decision is made. For this case, they define an unacclimating personal equilibrium (UPE) and an associated refinement, preferred personal equilibrium (PPE), in which the decision-maker makes a choice she is willing to follow through on. Second, they consider the case in which outcomes are realized long after a decision is made, defining a choice-acclimating personal equilibrium (CPE) where the decision-maker's reference point has time to acclimate to the choice they made. Without taking a stance on which concept is appropriate in the context here, I show that herding is not possible in either a PPE or CPE.[41]

A necessary condition for a decision to be either a PPE or a CPE is that it must be optimal when the outcomes it induces become the reference point. Consider the decision to buy in the context of the model. Both a PPE and a CPE require $U(B|B) \geq U(NT|NT)$ and $U(B|B) \geq U(S|S)$ where $U(F|G) = \int\int u(w|r)dG(r)dF(w)$ is the decision-maker's expected utility. The expectation in this expression is over both the possible wealth levels (given by the distribution F) and the possible reference points (given by the distribution G) of a reference-dependent utility over wealth, $u(w|r)$. Koszegi and Rabin (2007) assume a reference-dependent utility function of the form, $u(w|r) = m(w) + \mu(m(w) - m(r))$. For tractability, and as assumed in many of the results of Koszegi and Rabin (2007), I assume $m(w) = w$. Note, however, that, under Koszegi and Rabin's assumptions on μ, the reference-

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

[41] The UPE concept, when also allowing for probability weighting, is more difficult to work with analytically. Herding-like behavior may be possible in a UPE under some conditions, but it nevertheless does not survive the PPE refinement that Koszegi and Rabin (2007) suggest.

dependent component of utility is S-shaped as in prospect theory. In addition, although not considered in their model, I allow for decision weights with the same functional form I assumed previously. Under these assumptions, we have

$$\begin{aligned}
U(NT|NT) &= 1 \\
U(B|B) &= \frac{b_t^\delta}{\left(b_t^\delta+(1-b_t)^\delta\right)^{\frac{1}{\delta}}}(1-p_t) - \lambda\frac{(1-b_t)^\delta}{\left(b_t^\delta+(1-b_t)^\delta\right)^{\frac{1}{\delta}}}p_t \\
&+ \frac{b_t^\delta(1-b_t)^\delta}{\left(b_t^\delta(1-b_t)^\delta+(1-b_t(1-b_t))^\delta\right)^{\frac{1}{\delta}}}\left(\mu(1)+\mu(-1)\right) \\
U(S|S) &= -\lambda\frac{b_t^\delta}{\left(b_t^\delta+(1-b_t)^\delta\right)^{\frac{1}{\delta}}}(1-p_t) + \frac{(1-b_t)^\delta}{\left(b_t^\delta+(1-b_t)^\delta\right)^{\frac{1}{\delta}}}p_t \\
&+ \frac{b_t^\delta(1-b_t)^\delta}{\left(b_t^\delta(1-b_t)^\delta+(1-b_t(1-b_t))^\delta\right)^{\frac{1}{\delta}}}\left(\mu(1)+\mu(-1)\right)
\end{aligned}$$

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

$$\begin{aligned}
& & U(B|B) &\geq U(S|S) \\
\iff & (b_t^0)^\delta(1-p_t) - \lambda(1-b_t^0)^\delta p_t &\geq & -\lambda(b_t^0)^\delta(1-p_t) + (1-b_t^0)^\delta p_t \\
\iff & (b_t^0)^\delta(1-p_t)(1+\lambda) &\geq & (1-b_t^0)^\delta p_t(1+\lambda) \\
\iff & \left(\frac{p_t}{1-p_t}\right)^{\delta-1}\left(\frac{1-q}{q}\right)^\delta &\geq & 1
\end{aligned} \tag{14}$$

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

F. Instructions

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

Instructions for Trading Experiment

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

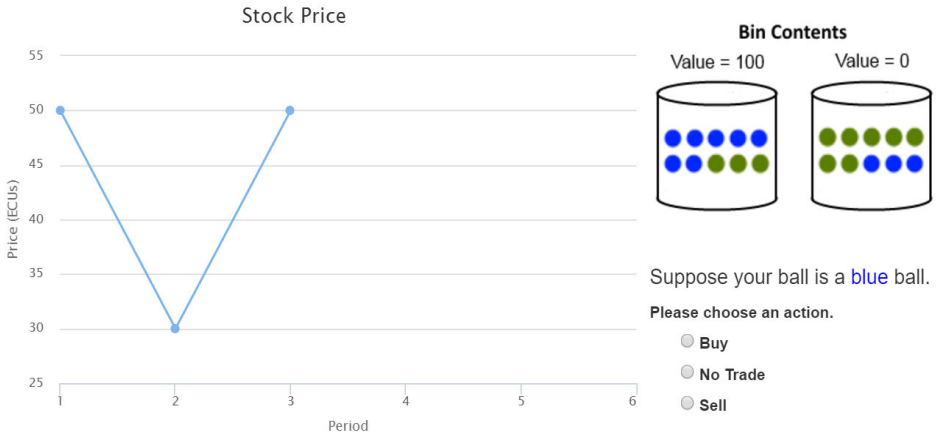
In each trading game, you will be given information about an artificial stock before having the opportunity to buy or sell it (or do neither). When the game is completed, a new game will begin. There will be 30 games in all and when you are done, you will be paid according to the trading decisions you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

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

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

Trading Page

Game 2 of 30



Next

Rules

Basics

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

Prior to each game, the computer will randomly choose whether the stock you can trade is worth 0 or 100 ECUs, with equal probability. Because the stock's value is randomly chosen each game, there is no dependence between its value in one game and its value in any other game.

You will be given 100 ECUs at the start of each game. These ECUs are **yours to keep** if you choose not to trade. If you choose to trade, however, you will have the opportunity to earn even more money.

The Stock's Price

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

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

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

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

Given the current price, you can choose to buy, sell, or not trade the stock. If you buy the stock, you will get 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.

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