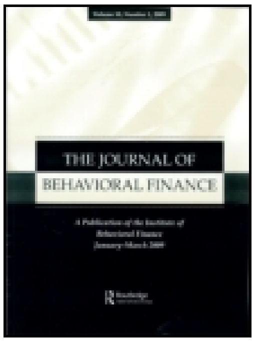

Journal of Behavioral Finance

ISSN: 1542-7560 (Print) 1542-7579 (Online) Journal homepage: http://www.tandfonline.com/loi/hbhf20

Risk Preferences, Investor Sentiment and Lottery Stocks: A Stochastic Dominance Approach

Wai Mun Fong

THE JOURNAL OF BEHAVIORAL FINANCE, 14: 42–52, 2013

ISSN: 1542-7560 / 1542-7579 online
DOI: 10.1080/15427560.2013.759579

Risk Preferences, Investor Sentiment and Lottery Stocks: A Stochastic Dominance Approach

Wai Mun Fong
National University of Singapore

Lottery stocks are a puzzle: individual investors value these stocks highly despite their low average returns and high volatility (Kumar [2009]). I argue that individuals are attracted to lottery stocks because they are risk-seeking and sentiment-prone. Because risk preferences are not directly observable, I use a model-free approach based on stochastic dominance to infer aggregate risk preferences. I also use a direct measure of individual investor sentiment, the bull-bear spread, to test whether sentiment affects the returns of lottery stocks. My results show that lottery stock investors are indeed risk seekers. Sentiment also plays an important role in explaining the demand and returns of lottery versus nonlottery stocks. The lottery stock puzzle can thus be understood only by incorporating unusual risk preferences and the propensity for individual investors to trade on sentiment.

Keywords: Lottery stocks, Risk-seeking, Risk-aversion, Investor sentiment, Stochastic dominance

INTRODUCTION

Stocks with lottery-like payoffs ("lottery stocks") attract individual investors. Why they do so is puzzling since on average, lottery stocks underperform nonlottery stocks. For example, Kumar [2009] find that from 1980–2005, the average return spread between lottery and nonlottery stocks is -7.96% a year. The poor performance of lottery stocks is especially puzzling because lottery stocks are stocks with low prices, high idiosyncratic volatility, and high idiosyncratic skewness. Low price stocks are mostly small firms with higher distress risk. Moreover, standard finance theory argues that idiosyncratic volatility should not be priced. If individual investors are not well diversified, as is typically the case, then a stock's expected return should be positively, not negatively, related to its idiosyncratic risk.

I argue that individuals are attracted to lottery stocks because they are risk seeking and sentiment prone. My empirical analysis is partly motivated by the recent theoretical work of Barberis and Xiong [2010] who propose a realization utility model to explain why investors seek volatile stocks that are known to have low average returns. Under this approach, individuals think of their investments not in terms of overall portfolio returns but rather as a series of investing episodes, each defined by gains or losses. Risk-seeking behavior results because by buying highly volatile stocks, investors stand a chance to realize huge gains. While this upside potential comes with high downside risks, investors can manage disutility of losses by creating a mental "loss account" in which loss realizations are deferred to a discounted future. The model also predicts that risk appetite may decrease if investors' sensitivity to realized losses exceeds their sensitivity to realized gains although it does not identify what actually drives changes in investors' risk appetites.

My analysis is also motivated by noise trader theory (Delong, Shleifer, Summers, and Waldmann [1990], Shleifer and Summers [1990]). Being relatively uninformed, individual investors are the prime examples of noise traders who are prone to trade on sentiment, or beliefs which are not justified by fundamentals. Not all stocks are equally affected by investor sentiment, however. Baker and Wurgler [2006] show that stocks of small firms, young firms, unprofitable firms, distressed firms, firms with extreme growth, and firms with high idiosyncratic volatility are more sensitive to sentiment than other stocks. They argue that this is because these stocks are hard to value, costly to trade, and difficult to sell short (D'Avolio [2002]). This makes them difficult to arbitrage and

Address correspondence to Wai Mun Fong, Department of Finance, NUS Business School, Mochtar Riady Building, National University of Singapore, Kent Ridge Crescent, Singapore 119245. E-mail: bizfwm@nus.edu.sg

thus prone to mispricing. Since lottery stocks embody many of the same characteristics, their returns are likely to be sentiment driven. In short, lottery stocks provide an interesting laboratory-type setting to study the role of risk preferences and investor sentiment on asset pricing.

Since risk preferences are not directly observable, I use a preference-based econometric technique to infer investors' risk preferences from market data. Specifically, I use stochastic dominance tests to infer whether lottery stock investors are risk averse or risk seeking. The basic inputs for my tests are the empirical cumulative distribution functions (cdfs) of stock returns. Stochastic dominance comparisons of empirical cdfs for lottery and nonlottery stocks allow me to back out as it were, investors' aggregate preferences for these stocks for broad classes of risk-averse as well as risk-seeking investors. No assumptions about the parametric form of return distributions are needed, though the approach does assume investors' preferences for one asset over another can be revealed by stochastic dominance comparisons of their return distributions. Border [1992] provides rigorous theoretical justification for this revealed preference approach. To account for sampling errors in the empirical cdfs, I use a general test for stochastic dominance developed by Linton, Maasoumi and Whang [2005] to conduct formal statistical inference.

Investor sentiment is also not directly observable, but studies have used various sentiment proxies. Most studies use individual proxies, such as the closed-end fund discount (Lee, Shleifer, and Thaler [1991]), share turnover (Baker and Stein [2004]), and mutual funds flow (Frazzini and Lamont [2008]), while others adopt a broader approach by combining several proxies into a composite sentiment index. This approach is taken by Baker and Wurgler [2006], who construct a sentiment index based on six underlying sentiment proxies: NYSE shares turnover, dividend premium, closed-end fund discount, number of initial public offerings (IPOs), first-day returns of IPOs, and equity share in new issues. They show that hard-to-value stocks such as those mentioned earlier significantly affected by sentiment.

While broad sentiment measures may capture overall market sentiment, they may not be a good proxy for *individual* investors' sentiment, which is my focus. I therefore take a different approach by using sentiment measures that are directly related to individuals. My main sentiment measure is the bull-bear spread, defined as the percentage of bullish investors minus the percentage of bearish investors which is obtained from the American Association of Individual Investors (AAII). This measure has several advantages compared with other sentiment indicators used in the academic literature: it is well known in financial circles, easily available to the public, and, most importantly, expresses the optimism or pessimism of retail investors, who are main clienteles of lottery stocks.

The main findings of this study are easy to summarize. Consistent with the predictions of the realization model, I find that risk-averters generally avoid lottery stocks even in periods when investor sentiment is highly positive. On the other hand, risk seekers are strongly attracted to lottery stocks when investor sentiment is high but reverse their preference when sentiment wanes. My results thus support a sentiment-based model of lottery stocks in which sentiment affects the price of such stocks by reinforcing the disposition to seek risk during good times and dampening it during bad times. My results also throw light on the lottery puzzle. Although investors know that lottery stocks generally underperform nonlottery stocks, their risk-seeking disposition explains why they are willing to risk a large chance of negative returns for a small chance of a huge reward. In other words, lottery stocks attract precisely the "right" type of investors with the "right" risk attitude. These gambling-motivated investors may also believe that they can profit handsomely from buying lottery stocks when other investors are also doing the same. My results indicate that such correlated trades are more likely to occur when investor sentiment is high. For the sample period in this study, these investors have not been disappointed, as lottery stocks outperformed nonlottery stocks by 22% on an annualized basis in high sentiment periods. More importantly perhaps is that lottery stocks provided much better odds of realizing huge returns compared to nonlottery stocks during high sentiment periods. In short, the lottery stock phenomenon is better understood if our models incorporate both unusual risk preferences and the propensity for individual investors to trade on sentiment.

My study is related to three areas of the behavioral finance literature which may be captioned by the three "Cs": clientele, correlated trades, and costly arbitrage. The clientele effect is the idea that different types of stocks attract different groups of investors. As Kumar [2009] has shown, individual investors are the main clienteles of lottery stocks. The literature on correlated trades, for example, Barber, Odean, and Zhu [2009, forthcoming] show that individual investors tend to buy or sell similar stocks at the same time. Moreover, this herding effect persists over months and it affects returns. While herding behavior may be driven by investor sentiment, these papers do not directly relate returns to sentiment as we do. Nonetheless, my findings suggest that sentiment drives affects lottery stock returns by coordinating the buy/sell decisions of individual risk-seeking investors. Finally, as mentioned earlier, my study draws directly from the noise trader literature, which argues that investor sentiment can affect asset pricing if there are noise traders who "trade on noise as if it were information" (Black [1986], pp. 529–530) and if there are limits to arbitrage by better informed traders. Lottery stocks, by virtue of their investor clientele and their defining characteristics (especially high idiosyncratic risk) epitomize both of these conditions.

The rest of this article is organized as follows. The second section describes my data sources and variables construction. The third section presents a descriptive summary of the sample and the returns of lottery and nonlottery stocks. I introduced the stochastic dominance approach and the econometric test for stochastic dominance in the fourth section. Results of the tests are discussed in the fifth section. The final section concludes with a summary of the key results.

DATA AND VARIABLES CONSTRUCTION

Measuring Sentiment

My primary data source for investor sentiment is the American Association for Individual Investors (AAII). Since July 1987, the AAII has conducted a weekly sentiment survey by polling a random sample of its members. Respondents are asked whether they are bullish, neutral or bearish about the stock market over the coming six months. Only members of the AAII are eligible to participate in these surveys once during each survey period. The average response rate typically varies between 50% and 70%. Since AAII members are individuals, the survey response can be regarded as a measure of individual investor sentiment. Using the AAII survey results, I compute the percentage of bullish investors minus the percentage of bearish investors each week and use this proxy for individual investor sentiment. In financial circles, this measure is popularly known as the bull-bear spread.[1]

Several studies have also used AAII sentiment indicators. Fisher and Statman [2000] examine whether the percentage of bullish investor predicts future returns on the S&P 500 index and small firms. Brown [1999] studies the relationship between individual investor sentiment and the volatility of closed-end investment funds. Bange [2000] investigates whether individual investors act on sentiment by changing their portfolio allocations. She finds that changes in the percentage of bullish and bearish investors predict individual investors' equity allocations over the next one to two months. Her finding is important as it shows that individual investors' portfolio decisions are consistent with their beliefs about market movements.[2]

Stock Data and Variable Construction

Data on stock prices and characteristic come from various sources. From the Center for Research on Security Prices (CRSP) database, I obtain stock prices and daily returns of common stocks (share codes 10 or 11) that trade on NYSE, AMEX, and NASDAQ. From COMPUSTAT, I obtain data to construct book-to-market values of firms according to the methodology used by Fama and French [1992, 1993]. From the website of Professor Kenneth French, I obtain data for equity risk factors related to the market, firm size, value/growth, and momentum. Finally, from Thomson Reuter's 13F database (available quarterly from 1980), I construct for each firm an aggregate measure of institutional ownership.

Following Kumar [2009], I define lottery stocks as those with low stock price, high idiosyncratic volatility, and high idiosyncratic skewness. Stocks with low prices appeal to gambling-motivated investors who are looking for cheap bets. Stocks with high idiosyncratic volatility are attractive to speculators because such stocks are perceived as more likely to produce extreme returns than more stable stocks. Finally, stocks with high idiosyncratic skewness are lottery-like because they offer a small chance of very large gains.

Idiosyncratic volatility

Following Kumar [2009], I compute the idiosyncratic volatility of a stock as the standard deviation of the residuals obtained by fitting a four-factor model that comprises three risk factors from Fama and French [1992, 1993] and a momentum factor. I denote this model as the FF4 model. The regression specification is as follows:

$$R_{i,t} = \alpha_i + \beta_{1,t} R_{M,t} + \beta_{2,i} SMB_t + \beta_{3,i} HML_t + \beta_{i,4} UMD + \varepsilon_{i,t} \quad (1)$$

where R_i is the return on stock i in excess of the risk-free rate (one-month Treasury-bill rate); R_M is the return of the market (the CRSP value-weighted all stock index) in excess of the risk-free rate; SMB is the return on a portfolio of small stocks minus the return on a portfolio of large stocks; HML is the return on a portfolio of high book-to-market stocks minus the return on a portfolio of low book-to-market stocks; UMD is a momentum factor (the difference between the value-weighted return of a portfolio of stocks with high returns over the months t-12 to t-2 and the value-weighted portfolio of stocks with low returns over the same period); and ε_i is the regression residual. I estimate the above regression at the end of June each year, using the previous twelve months of daily stock returns data. To mitigate the problem of thin trading, I require that each firm has at least 180 trading days in which both the daily return and trading volume are nonzero as reported by CRSP. The standard deviation of the daily returns residuals is multiplied by the square root of the number of trading days used in the yearly regression to obtain an annualized measure of idiosyncratic volatility.

Idiosyncratic skewness

Following Harvey and Siddique [2000] and Kumar [2009], I measure idiosyncratic skewness as the third moment of the residuals obtained by regressing daily stock returns on a two factor model:

$$R_{i,t} = \alpha_i + \beta_{1i} R_{M,t} + \beta_{2i} R_{Mt}^2 + \varepsilon_{i,t} \quad (2)$$

As with idiosyncratic volatility, this regression is estimated at the end of June each year using the previous 12 months of daily returns data. The idiosyncratic skewness of stock i is defined as the skewness of the residual ε_i.

Portfolio Construction

I form value-weighted portfolios comprising lottery and non-lottery stocks at the end of June each year from June 1987

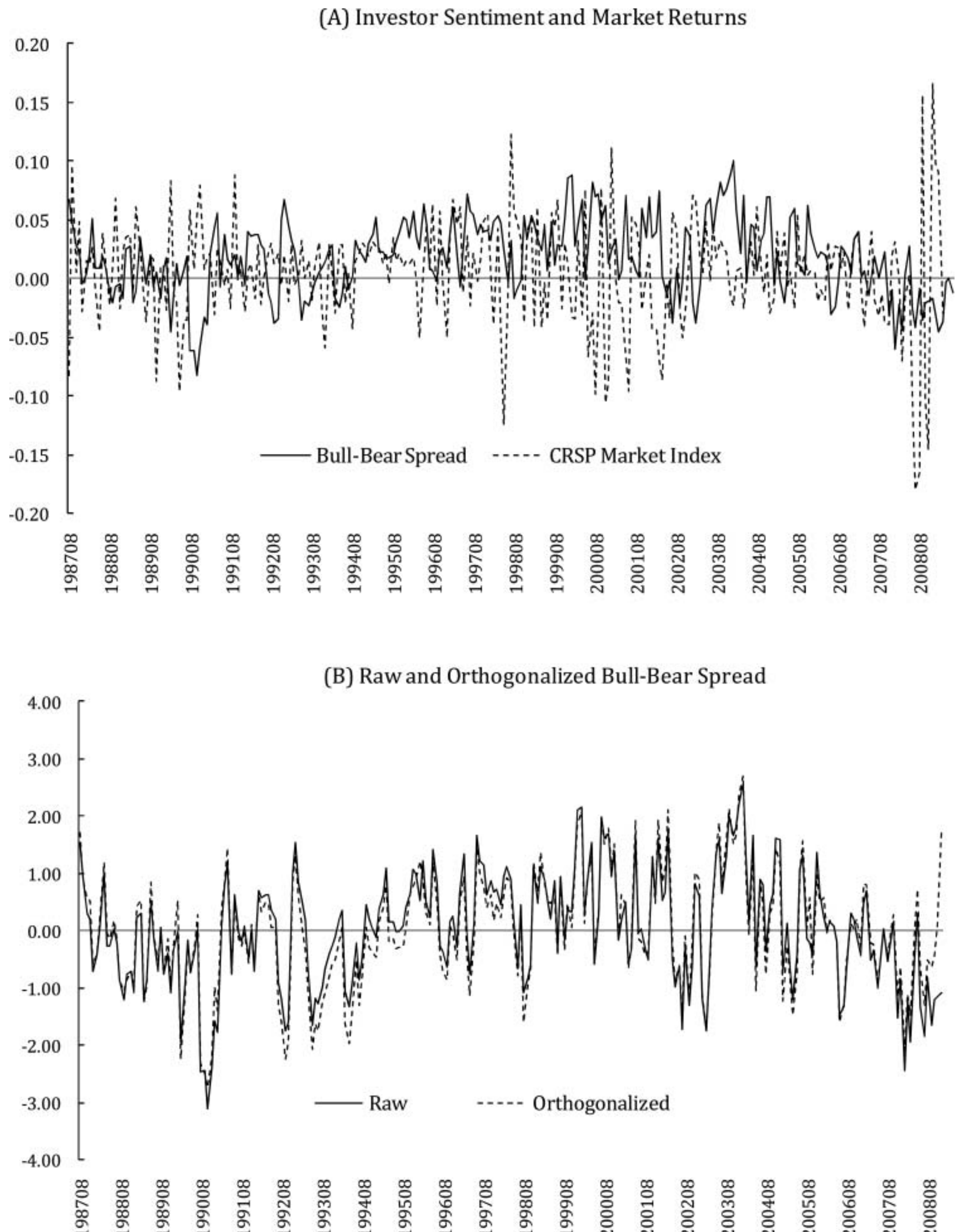

FIGURE 1 Investor Sentiment.

through June 2009 by sorting stocks based on their stock price, idiosyncratic volatility, and idiosyncratic skewness. Stocks in the highest k^{th} percentile of idiosyncratic volatility, the highest k^{th} percentile of idiosyncratic skewness, and the lowest k^{th} percentile of price are assumed to be stocks that investors are likely to identify as lottery stocks. Similarly, I define nonlottery stocks to be those that are in the lowest k^{th} percentile of idiosyncratic volatility, the lowest k^{th} percentile of idiosyncratic skewness, and the highest k^{th} percentile of price. Other stocks that are not placed in these two categories are grouped under "others." I choose $k = 33$ to have a considerable number of lottery-type stocks for each tertile but obtain similar results with $k = 50$.

PRELIMINARY ANALYSIS

Time Series of the Bull-Bear Spread

Figure 1A shows the time series of the bull-bear spread along with the market index (the CRSP value-weighted all stock index). The bull-bear spread in the figure is calculated by averaging the weekly bull-bear spreads for each month.

Sentiment is positively correlated with market returns but is smoother.[3] The correlation between the bull-bear spread and low price stocks (those in the lowest tertile based on market capitalization at June 30 each year) is 0.28 compared with 0.17 with the market index. Reflecting the high level of noise

TABLE 1
Characteristics of Lottery and Nonlottery Stocks

Stock Characteristic	No. of stocks	*Ivol*	*ISkew*	Price	Firm size (in million $)	Book-to-market ratio	Institutional ownership
			Stock Characteristics (1987–2009)				
Lottery							
Mean	491	103.98	1.722	$3.21	65.05	0.804	11.11%
Median	459	105.70	1.674	$3.16	44.93	0.738	8.12%
Non-lottery							
Mean	385	24.40	−0.382	$39.09	7514.86	0.612	47.25%
Median	389	23.62	−0.304	$38.53	6661.16	0.594	42.13%
Others							
Mean	3,242	50.39	0.408	$18.51	1453.83	0.707	35.07%
Median	3,205	51.33	0.411	$17.64	987.43	0.689	29.78%

This table presents descriptive statistics of lottery stocks, nonlottery type stocks, and other stocks that do not belong to either of these categories. The data are obtained from the CRSP database and include common stocks that are traded on NYSE, AMEX, and NASDAQ (share codes 10 and 11). Stocks are sorted at June 30 each year from June 1987 through June 2009 based on stock price, idiosyncratic skewness (*Iskew*) and idiosyncratic volatility (*Ivol*). Idiosyncratic volatility is defined as the standard deviation of the residuals from the Fama-French three-factor model plus a momentum factor where daily returns over the previous twelve months are used to estimate the model. *Ivol* is reported as an annualized percentage. Idiosyncratic skewness is defined as the scaled third moment of residuals from a model that contains the market excess return and the square of the market excess return as factors. The market portfolio is the CRSP value-weighted (all stock) index. Lottery stocks are defined as stocks in the lowest 33rd price percentile, highest 33rd idiosyncratic volatility percentile and highest 33rd idiosyncratic skewness percentile. Nonlottery stocks are those in the highest 33rd price percentile, lowest 33rd idiosyncratic volatility percentile, and lowest 33rd idiosyncratic skewness percentile. Firm size is market capitalization as of June 30 each year. Book-to-market is the ratio of book equity value from the fiscal year ending in calendar year t-1 divided by the market equity value at June 30 of year t. Firms with negative book-to-market ratios are excluded from the sample. Institutional ownership is the percentage of common stocks owned by institutions that submit 13f filings and are obtained from Thomson Reuters. The third column reports the time series mean and median number of stocks in each group. The other columns report the time series average of the yearly cross-sectional mean and median value of each stock characteristic.

in monthly returns, correlations between the bull-bear spread and returns based on annual data (averages of monthly sentiment and returns) are significantly higher: 0.37 for the market index and 0.48 for the low-price portfolio. The higher correlation between sentiment and low price firms is consistent with smaller firms being the preferred habitat of individual investors as documented by Kumar [2009] and others.

How do we know that the bull-bear spread measure investor sentiment and not economic fundamentals? While we can never be sure, we can remove from the bull-bear spread, observable macroeconomic variables that are highly correlated with business cycle variations. The residuals from this regression should be a cleaner proxy for individual investor sentiment. I choose five economic factors for this regression based on the follow the widely cited approach of Chen, Roll, and Ross [1986]. The factors are: (a) changes in industrial production, (b) unexpected inflation, (c) changes in expected inflation, (d) term structure, and (e) bond default risk premium. Details of the construction of these variables are described in Chen et al. Figure 1(B) shows that the raw and orthogonalized bull-bear spread are similar, suggesting that the bull-bear spread is largely independent of business cycle fluctuations.

Table 1 presents descriptive statistics of lottery and nonlottery stocks in the sample. The numbers in the table are time series mean and median of yearly cross-sectional averages of each stock characteristic. By construction, lottery stocks exhibit significantly lower price, higher idiosyncratic volatility, and higher idiosyncratic skewness than nonlottery stocks. Consistent with the findings of Kumar [2009], lottery stocks are smaller and have higher book-to-market ratios than nonlottery stocks. Furthermore, they have much lower levels of institutional ownership (11%) than nonlottery stocks (47.3%), indicating that many institutions are averse to owning these stocks.

Table 2 summarizes the unconditional and conditional performance of lottery and nonlottery stock portfolios. Panel A reports average returns and Fama-French four-factor (FF4) alphas calculated using data for the whole sample period. All return numbers in the table are annualized. Lottery stocks underperform nonlottery stocks in terms of both raw returns and alphas. L-N denotes the portfolio that is long the lottery stock portfolio and short the nonlottery stock portfolio. This portfolio earned an average return of −6.9%, which is significant at 5%. Risk-adjustment does not make much difference to this result. The FF4 alpha is −7.5% (t-statistic: 2.34), which is quite close to the annualized alpha of 7.1% reported by Kumar [2009], albeit for a different sample period (1980–2005). The CAPM alpha is −7.2% (t-statistic: −2.08). Both alphas are economically large.

Panel B reports average returns and alphas, conditional on sentiment. The results are obtained using a simple, model-free approach whereby I sort portfolio returns into three sentiment states of low, medium, and high using the bull-bear spread. The sorts are based on daily returns, which are later used in the stochastic dominance tests. Since AAII data are only available weekly, I set daily returns to the weekly AAII sentiment data as follows: if the AAII survey results are reported on day t for week s, that week's bull-bear spread is matched to daily returns from days t-1 to t-5. This is to

TABLE 2
Portfolio Return Statistics

	Panel A. Unconditional Returns		
Portfolio:	*Lottery (L)*	*Non-lottery (NL)*	*L-NL*
Mean return	0.038	0.103	−0.069
T-statistic	(0.57)	(2.00)**	(−2.08)**
FF4 alpha	−0.030	0.046	−0.075
T-statistic	(−0.75)	(1.82)*	(−2.24)**
CAPM alpha	−0.017	0.054	−0.072
T-statistic	(−0.32)	(1.53)	(−2.08)**
	Panel B. Returns in Sentiment States		
Sentiment State	*Lottery (L)*	*Non-lottery (NL)*	*L-NL*
Low Sentiment	Nobs: 854		
Mean return	−1.055	−0.717	−0.338
T-statistic	(−5.23)***	(−4.18)***	(−3.38)***
FF4 alpha	−0.389	−0.125	−0.257
T-statistic	(−3.73)***	(−1.76)*	(−2.66)***
CAPM alpha	−0.305	−0.267	−0.305
T-statistic	(−3.19)***	(−2.92)***	(−3.19)***
Medium Sentiment	Nobs: 3870		
Mean return	0.098	0.169	−0.071
T-statistic	(1.56)	(3.12)***	(−1.96)**
FF4 alpha	−0.017	0.054	−0.072
T-statistic	(−0.42)	(2.13)**	(−1.99)**
CAPM alpha	0.000	0.076	−0.077
T-statistic	(−0.007)	(2.09)**	(−2.12)**
High sentiment	Nobs: 812		
Mean return	0.873	0.649	0.224
T-statistic	(5.99)***	(5.42)***	(2.69)***
FF4 alpha	0.199	0.057	0.139
T-statistic	(1.86)*	(0.74)	(1.51)
CAPM alpha	0.550	0.374	0.172
T-statistic	(3.95)***	(3.65)***	(1.79)*

This table reports summary statistics of returns of three value-weighted portfolios. *L* is a portfolio consisting of lottery-type stocks, *N* is a portfolio consisting of nonlottery type stocks and *L-N* is a portfolio that longs *L* and shorts *N*. Portfolios are formed at June 30 each year from June 1987 through June 2009. Details of the portfolio construction are described in Table 1. Panel A reports unconditional average of daily returns and Fama-French four-factor (FF4) alphas for each portfolio. The factors include three factors related to the market (*MKT*), firm size (*SMB*), and book-to-market ratio (*HML*) used by Fama and French [1992, 1993] and a momentum factor (*UMD*). Daily returns and alphas are reported annualized. Panel B reports average returns and FF4 alphas within three sentiment states based on the bull-bear spread from the American Association of Individual Investors (AAII). The results are obtained by matching daily returns to weekly values of the bull-bear spread as described in the text. The low (high) sentiment state comprises days in which the bull-bear spread is at least one standard deviation below (above) its sample mean. The rest of the observations are classified as the medium sentiment state. T-statistics for alphas are calculated using Newey-West standard errors. Statistics marked bold are significant at 10% or lower.

account for the fact that the AAII survey results for each week are based on all responses received over the previous five days. After constructing this sample, I sort the returns into the three sentiment states, where low (high) sentiment states are days in which the bull-bear spread is at least one standard deviation below (above) the sample mean. The rest of the observations are placed in the medium sentiment state. The number of days in the low, medium, and high sentiment states is 854, 3,878, and 812, respectively. The corresponding mean bull-bear spreads are −0.21, 0.09, and 0.38. Thus, individual investors were optimistic on average.

As before, all average returns and alphas are shown in Panel B are annualized. We see that average returns increase monotonically with sentiment. For most of the sample period, lottery stocks underperform nonlottery stocks. The underperformance is 33.8% in the low sentiment state is 7.1% in the medium sentiment states. As these two states account for 85% of the sample, lottery stocks are clearly a "bad deal" for risk-averse investors, a result I confirm later using stochastic dominance tests. However, risk-seeking investors may find lottery stocks attractive because in the high sentiment state, lottery stocks compensate for their generally poor performance with higher average returns than nonlottery stocks. The mean return spread in the high sentiment state is a significantly positive 22.4%, and the CAPM alpha gives a hint of positive abnormal returns (though the FF4 alpha is not statistically significant even at 10%). More importantly perhaps is that lottery stocks have much better odds of delivering huge returns compared with nonlottery stocks in the high sentiment state. More concretely, if we define a huge return as one that is at least one standard deviation above the overall sample mean for lottery stocks, our data show that the odds of getting such a return in the high sentiment state is 48% higher for lottery stocks than nonlottery stocks; that is, the odds ratio is 1.48. In contrast, the odds ratio in the low and medium sentiment state is 1.05 and 1.37, respectively. Defining a huge return based on nonlottery stocks does not change the ranking very much (the respective odds ratios are 1.47, 0.98, and 1.27). These comparisons suggest that risk seekers may find lottery stocks to be an attractive bet in high sentiment periods.

INVESTOR SENTIMENT AND LOTTERY STOCK PREFERENCES

Overview of Stochastic Dominance Theory

My main conjectures are that (a) lottery stock investors are risk seekers and (b) demand for lottery stocks by risk seekers is positively related to investor sentiment. I now test these conjectures formally. Since risk preferences cannot be directly observed, I infer them indirectly from market data based on the returns of lottery and nonlottery stocks. I employ a stochastic dominance (SD) approach to conduct the statistical inferences.

SD theory provides an expected utility framework for evaluating choices under risk.[4] The advantage of using SD to study risk preferences is that SD criteria exist for a wide range of risk preferences. In particular, the theory provides a one-to-one mapping of an investor's choice between two risky assets, which is determined by comparing their return distributions, and his risk preference as summarized by the general shape of his utility function. The range of utility functions that SD theory encompasses is also wide (see Starmer [2000]).

Stochastic Dominance Criteria

The most common SD rules are first, second, and third order stochastic dominance. First order stochastic dominance (FSD) is the most stringent since it only assumes that investors display nonsatiation. To illustrate the concept, suppose an investor who exhibits nonsatiation is choosing between two risky assets, X and Y, whose distributions may be nonnormal. If the cumulative distribution function of X is everywhere below that of Y, the investor will always choose X since it has a higher probability of exceeding any return than Y. In the language of SD, X stochastically dominates Y at first order.

In this article, I focus on second order stochastic dominance (SSD) and third order stochastic dominance (TSD). These rules apply to choices made by nonsatiated and risk-averse individuals. If asset X is preferred to asset Y for all nonsatiated and risk-averse investors, then X must second order stochastically dominate Y. The converse is also true: if X second order stochastically dominates Y, then any investor with concave utility function will prefer X to Y. TSD also assumes that risk-averse individuals prefer return distributions that are more positively skewed. This makes TSD an appealing rule for analyzing lottery-type stocks. If asset X and asset Y have the same mean but X has a more positively skewed return distribution, a TSD-type investor will prefer X to Y.

Since lottery stock investors may be risk seeking, a generalization of SD theory to risk seekers is useful. In the mean-variance framework, a risk seeker might prefer the riskier of two assets if that asset's expected returns is "not too low." This intuition can be extended to the general framework of stochastic dominance. In particular, Meyer [1977, Theorem 4) generalizes the SSD relation to individuals with unspecified convex utility functions. Going forward, I shall refer to this concept as risk seeking stochastic dominance (RSSD).

Stochastic Dominance Test

Empirically, propositions such as "X stochastically dominate Y at *j*th order" are testable using market data. Specifically, we can use SD tests to compare the returns of a pair of assets to infer investors' risk preferences for one asset over another. For example, if tests show that lottery-type stocks are third order stochastically dominated by nonlottery type stocks, then risk aversion and skewness preference cannot explain investors' preference for lottery-type stocks. However, we might be able to rationalize this preference if we assume that investors are locally risk seeking using RSSD tests.

I use a test for stochastic dominance developed by Linton et al. [2005] to examine risk preferences. An attractive feature of the Linton et al. test is that it allows general forms of mutual dependence between prospects and non i.i.d. data, making it well suited for financial data. The test strategy and test statistics are described below.

Let F and G denote the cumulative distribution functions for two risky assets. The following inequality defines first, second, and third order stochastic dominance of F over G:

$$I_j(z; F) \leq I_j(z; G) j = 1, 2, \text{ and } 3 \tag{3}$$

where z is the joint support of F and G, j is the order of stochastic dominance and

$$I_j(z; F) = F(z) \tag{4}$$

$$I_2(z; F) = \int_{-\infty}^{z} F(t)dt = \int_{-\infty}^{z} I_1(t; F)dt \tag{5}$$

$$I_3(z; F) = \int_{-\infty}^{z} \int_{-\infty}^{t} F(s)dsdt = \int_{-\infty}^{z} I_2(t; F)dt \tag{6}$$

Define $I_j(z; G)$ analogously for distribution G.

For an investor with utility function satisfying $U' \geq 0$, F first order stochastically dominates G if and only if Equation (1) for $j = 1$ holds. This is intuitive since $F(z)$ lies everywhere to the right of$G(z)$. As mentioned, FSD is a very stringent decision rule.

Turning to SSD, for an investor with utility function satisfying $U' \geq 0$ and $U'' \leq 0$, F second order stochastically dominate G if and only if Equation (1) for $j = 2$ holds. SSD implies that the area under$F(z)$ is everywhere smaller than the area under $G(z)$.

TSD assumes risk aversion and a preference for positive skewness. Intuitively, such an investor will accept some losses in exchange for a (small) chance of large gains. More formally, for any investor with utility functions satisfying $U' \geq 0$,$U'' \leq 0$, and$U''' \geq 0$, F third order stochastically dominate G if and only if Equation (1) with $j = 3$ holds.

Meyer [1977] generalized the SSD criterion to risk seekers. F is preferred to G for all investors with utility functions satisfying $U' \geq 0$ and $U'' \geq 0$ if and only if

$$\int_{z}^{\infty} (F(t) - G(t))dt \leq 0 \text{ for all } z \tag{7}$$

Thus, risk-seeking stochastic dominance is also an SSD rule since a second degree restriction is imposed on the utility function.

The null and alternative hypotheses for $j = 1$, 2, and 3 may be stated as follows:

$$H_0 : I_j(z; F) \leq I_j(z; G) \tag{8}$$

for all z, with strict inequality for at least one value of z, and

$$H_1 : I_j(z; F) > I_j(z; G) \text{ for some } z, \tag{9}$$

Here, the null hypothesis is that F (weakly) stochastically dominates G at the jth order. The alternative hypothesis is that stochastic dominance fails at some point. The Linton et al. test uses the following test statistic to evaluate stochastic dominance:

$$\overline{SD}_j^{F,G} = \sup_z \sqrt{N}(I_j(z; \hat{F}) - I_j(z; \hat{G})) \tag{10}$$

where N is the sample size. Reversing the order of comparison gives a similar test statistic for the null hypothesis that G (weakly) stochastically dominates F at the jth order.

$$\overline{SD}_j^{G,F} = \sup_z \sqrt{N}(I_j(z;\hat{G}) - I_j(z;\hat{F})) \tag{11}$$

I use Equations (10) and (11) for $j = 2$, and 3 to test the direction of stochastic dominance for risk-averse investors. The corresponding test statistics for risk-seekers under RSSD are:

$$\overline{SD}_{2^*}^{F,G} = \sup_z \sqrt{N}[(I_2(\infty;\hat{F}) - I_2(z;\hat{F})) - (I_2(\infty;\hat{G}) - I_2(z;\hat{G}))] \tag{12}$$

$$\overline{SD}_{2^*}^{G,F} = \sup_z \sqrt{N}[(I_2(\infty;\hat{G}) - I_2(z;\hat{G})) - (I_2(\infty;\hat{F}) - I_2(z;\hat{F}))] \tag{13}$$

I infer that F stochastically dominates G for risk averters if $\overline{SD}_j^{F,G}$ is "small" and $\overline{SD}_j^{G,F}$ is "large." I perform a similar test for risk seekers using Equations (12) and (13). For both risk averters and risk seekers, I perform the tests separately for low and high sentiment states as defined previously. If demand for lottery stocks is sentiment driven, lottery stocks should dominate nonlottery stocks for risk seekers when sentiment is high. This preference should weaken or reverse when sentiment is low.

Statistical inference for the test is complicated by the fact that the distributions of the test statistics are nonstandard. Linton et al. [2005] propose a subsampling bootstrap approach that can be used to compute empirical p-values. They show via Monte Carlo simulations that the bootstrap approach leads to good test performance in terms of size and power given large sample sizes ($N > 1{,}000$). I use this bootstrap approach to perform statistical inference. Details of the bootstrap method are given in the appendix.

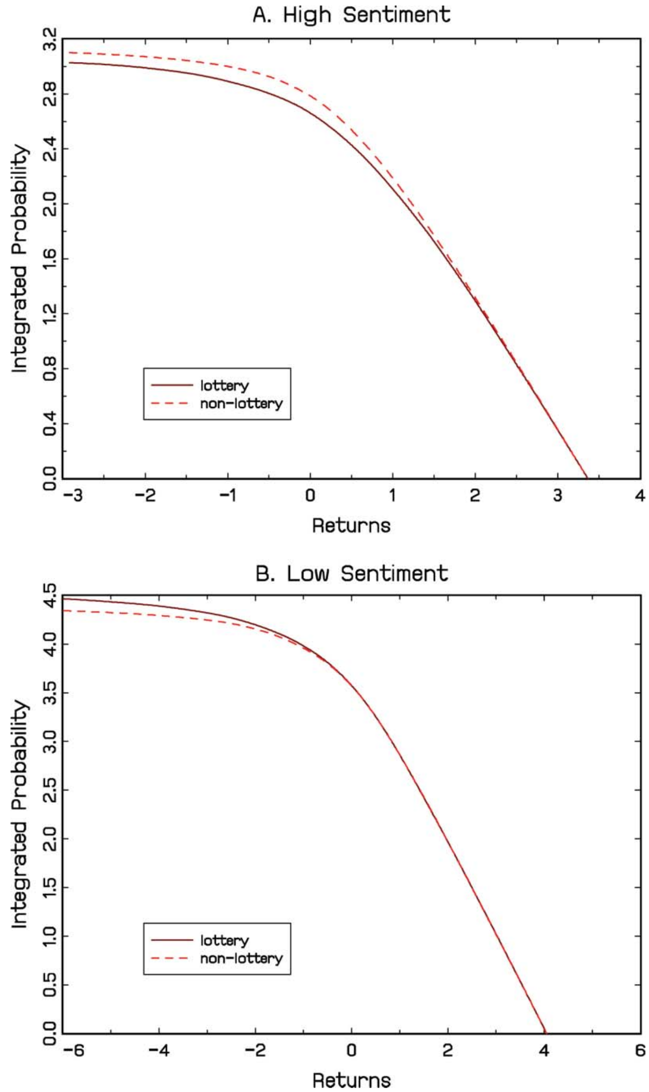

FIGURE 2 Integrated Cumulative Distribution Functions and Investor Sentiment: Risk Seekers. (color figure available online).

RESULTS OF STOCHASTIC DOMINANCE ANALYSIS

Before discussing the results of the SD test, it is instructive to examine the integrated cdfs for lottery and nonlottery stocks. These are shown in Figure 2 risk seekers and Figure 3 for risk averters. For each type of investor, I plot the integrated cdfs separately for high and low sentiment states.

Figure 2(A) shows that when sentiment is high, the distribution function for lottery stocks is everywhere below that of nonlottery stocks. This indicates that risk seekers prefer lottery to nonlottery stocks in high sentiment periods. Figure 2(B) shows that this preference is reversed during low-sentiment periods, which is consistent with reduced risk tolerance. Figure 3 shows that in contrast to risk seekers, risk averters always prefer nonlottery stocks.

Table 3 presents the results of the Linton et al. test. Results for risk averters are in Panel A while results for risk seekers are in Panel B. The numbers in the tables are bootstrap p-values for two null hypotheses. The first null hypothesis is that lottery stocks stochastically dominate nonlottery stocks. This is denoted by $L \succeq NL$ in Table 3. To complete the test, I test a second null hypothesis, which is that nonlottery stocks stochastically dominate lottery stocks. This is denoted by $N \succeq LL$.

Table 3 shows three results. First, risk averters generally avoid lottery stocks. This can be seen from the p-values which are mostly small for the null hypothesis that $L \succeq NL$ and large for the reverse null hypothesis. The exception is for the high sentiment state where p-values for both null hypotheses are small, indicating that investors are indifferent between lottery and nonlottery stocks. Second, in high sentiment periods, risk seekers clearly prefer lottery stocks to nonlottery stocks. This

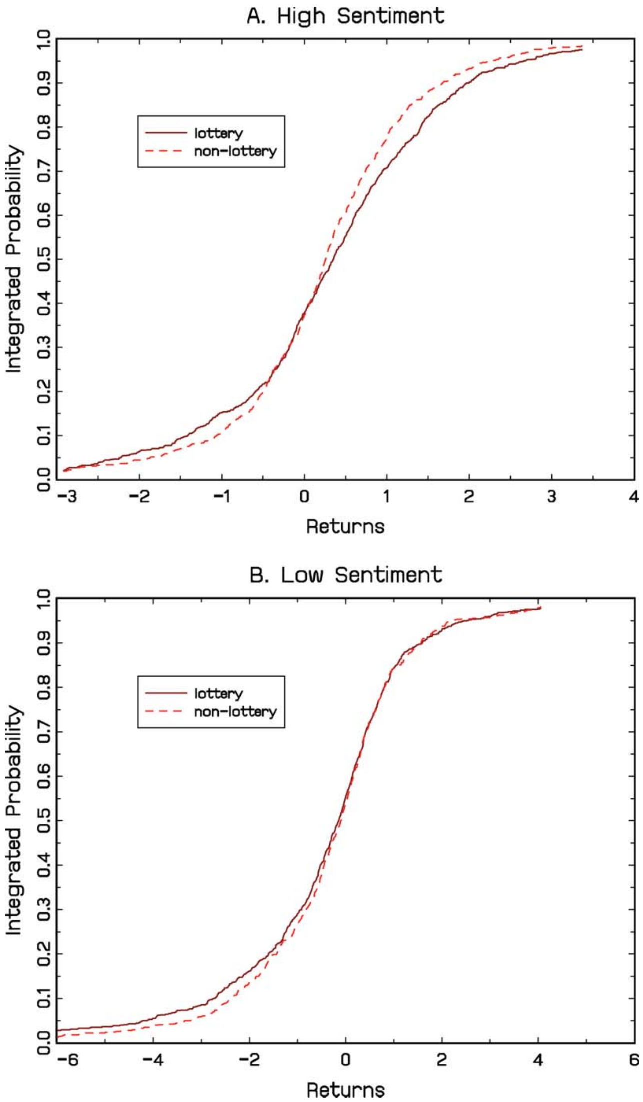

FIGURE 3 Integrated Cumulative Distribution Functions and Investor Sentiment: Risk averters. (color figure available online).

TABLE 3
Results of Stochastic Dominance Test

Risk Profile	Sentiment State	*SD* Test *p*-values		Preference
		$L \succ NL$	$NL \succ L$	
Panel A. *Risk-averters*	High	0.0130	0.0100	Neither
	Low	0.0000	0.9960	Non-lottery
Risk-averters with skewness preference	High	0.0820	0.3630	Non-lottery
	Low	0.0000	0.9960	Non-lottery
Panel B. *Risk Seekers*	High	0.9840	0.0000	Lottery
	Low	0.0000	0.6740	Non-lottery

This table reports results of stochastic dominance tests using daily returns of value-weighted portfolios of lottery stocks (L) and nonlottery stocks (N). Portfolios are formed at June 30 each year from June 1987 through June 2009 as described in Table 1. Returns are matched to sentiment as follows: each week's bull-bear spread from the AAII is matched to the prior five day's returns to account for the one-week delay between survey response and reporting. High (low) sentiment periods are those in which the bull-bear spread is at least one standard deviation above (below) the sample mean bull-bear spread. The stochastic dominance test is that of Linton, Maasoumi, and Whang [2005]. Two null hypotheses are tested to assess the direction of stock preference for risk-averters (Panel A) and risk-seekers (Panel B). The first null hypothesis ($L \succeq N$) is that lottery stocks stochastically dominate non-lottery stocks. The second null hypothesis is the reverse of the first null ($N \succeq L$). The numbers reported in the table are *p*-values computed using bootstrap simulations with 10,000 replications (see appendix).

can be seen from the *p*-values in Panel B, which are large for the null hypothesis that $L \succeq NL$ and small for the reverse null. Third, this preference for lottery stocks is reversed in low-sentiment periods. Together, the results suggest that high sentiment reinforces risk-seeking behavior and demand for lottery stocks while low sentiment dampens investors' risk appetite for volatile stocks. Overall, the results of the SD tests support the conjecture that the clientele for lottery stocks are risk seekers and that their demand for such stocks are highly sensitive to shifts in investor sentiment.

That lottery stocks attract risk seekers is consistent with the predictions of the realization utility model. As Barberis and Xiong [2010] demonstrated, stocks with low average returns do not deter risk-seeking investor as long as they are sufficiently volatile to provide occasional bursts of extreme positive utility when investors realize large gains. On a more general note, risk seeking is not necessarily irrational when it is viewed in the proper context. Friedman and Savage [1948] noted long ago that people buy both insurance and lotteries. Canner, Mankiw, and Weil [1997] observe that financial advisers typically advice their clients to layer their portfolios with cash at the bottom of the asset pyramid and riskier assets in the upper layers. This layering of assets is also central to behavioral portfolio theory (BPT) developed by Shefrin and Statman [2000] under which investors are rationally risk seeking because they only risk what they afford to lose.

Our results also indicate risk seeking combined with positive investor sentiment explains why lottery stocks can occasionally produce stellar returns. At the same time, low sentiment can negate the demand for lottery stocks even among risk seekers. Since investor sentiment is positively correlated with overall market returns, the effects of sentiment on risk preferences is consistent with a wealth effect in which investors' risk aversion varies inversely with market state, as suggested by Campbell and Cochrane [1999] and Barberis, Huang, and Santos [2001]. In unreported results, we also find that neither lagged market returns nor lagged return spread predicts the current return spread. Thus, the sentiment effect that we document in this article is not due to momentum trading but rather is more consistent with the view that individual investors tend to trade on noise as if it were information, as in the classic description by Black [1986].

Finally, our study shows that simply focusing on unconditional returns mask interesting return patterns that are sentiment related. The lottery puzzle is a puzzle if we assume that all investors are risk averse and ignore the role of investor sentiment. It may not be that puzzling from the viewpoint of risk seekers who believe they can profit from sentiment. In particular, while gambling-motivated investors are aware that while lottery stocks generally underperform, they also know that they are compensated for this underperformance during high-sentiment periods, when individuals as a group are net buyers of lottery stocks. In short, to better understand the lottery stock phenomenon, we need account for both unusual risk preferences and the propensity for individual investors to trade on sentiment.

CONCLUSION

Lottery stocks provide a laboratory-type setting to study the role of risk preferences and investor sentiment. I conjecture that lottery stocks attract mainly risk seekers. Moreover, given that individual investors are the main clientele of lottery stocks, I posit that demand for lottery stocks are likely to be affected by investor sentiment. I test these conjectures using a stochastic dominance approach. The main findings of this study are easy to summarize. Consistent with the predictions of the realization model proposed by Barberis and Xiong [2010], I find that risk averters generally avoid lottery stocks. Risk seekers are strongly attracted to lottery stocks when investor sentiment is positive, but this preference is reversed when sentiment wanes. Sentiment has a significant positive impact on the return spread between lottery and nonlottery stocks. Although the unconditional average return spread is negative, lottery stocks outperform nonlottery stocks on average during high-sentiment periods. Moreover, in high-sentiment periods, lottery stocks provide much better odds of gaining huge rewards compared with nonlottery stocks. The sentiment effect combined with risk-seeking predisposition may explain why many individual investors find lottery stocks attractive.

NOTES

1. Results of each week's survey are reported on the AAII website and Barron's. Finance commentators also discuss the survey findings regularly on Bloomberg and Reuters.
2. Another well-known sentiment survey is the weekly Advisors Sentiment (AS) survey conducted by Investor Intelligence (II) which polls the number of market newsletters that are bullish, neutral and bearish (no specified time horizon for all cases). Since newsletter writers are market professionals, the AS survey can be considered as a proxy for institutional investor sentiment. The survey data are provided on a subscriber basis.
3. The first order autocorrelation for the monthly bull-bear spread is 0.55.
4. Hadar and Rusell [1969], Hanoch and Levy [1969], Rothschild and Stiglitz [1970, 1971], and Whitmore [1970] lay the utility foundations of stochastic dominance theory. Starmer [2000] explains that stochastic dominance criteria are also meaningful for a wide class of nonexpected utility models such as the Generalized Expected Utility model of Machina [1982]. Levy and Wiener [1998] and Alibes and Heukamp [2006] generalize the second order stochastic dominance condition to cumulative prospect theory (Tversky and Kahnemann [1992]).

REFERENCES

Alibes, M. B. and F. H. Heukamp. "Stochastic Dominance and Cumulative Prospect Theory." *Management Science*, 52, (2006), pp. 1409–1423.

Baker, M. and J. Stein. "Market Liquidity as a Sentiment Indicator." *Journal of Financial Markets*, 7, (2004), pp. 271–300.

Baker, M. and J. Wurgler. "Investor Sentiment and the Cross-section of Stock Returns." *Journal of Finance*, 61, (2006), pp. 1645–1680.

Bange, M. M. "Do the Portfolios of Small Investors Reflect Positive Feedback Trading?" *Journal of Financial and Quantitative Analysis*, 35, (2000), pp. 239–255.

Barber, B., T. Odean, and N. Zhu. "Do Retail Trades Move Markets?" *Review of Financial Studies*, 22, (2009), pp. 151–186.

Barber, B., T. Odean, and N. Zhu. "Systematic Noise." *Journal of Financial Markets*, forthcoming.

Barberis, N., M. Huang, and T. Santos. "Prospect Theory and Asset Prices." *Quarterly Journal of Economics*, 116, (2001), pp. 1–53.

Barberis, N. and W. Xiong. "Realization Utility." Working Paper, Yale University and Princeton University, 2010.

Black, F. "Noise." *Journal of Finance*, 41, (1986), pp. 529–543.

Border, K. "Revealed Preference, Stochastic Dominance and the Expected Utility Hypothesis." *Journal of Economic Theory*, 56, (1992), pp. 20–42.

Brown, G. W. "Volatility, Sentiment and Noise Traders." *Financial Analysts Journal*, 55, (1999), pp. 82–90.

Campbell, J. Y. and J. H. Cochrane. "By Force of Habit: A Consumption-based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy*, 107, (1999), pp. 205–251.

Canner, N., N. G. Mankiw, and D. Weil. "An Asset Allocation Puzzle." *American Economic Review*, 87, (1997), pp. 181–191.

Chen, N. F., R. Roll, and S. A. Ross. "Economic Forces and the Stock Market." *Journal of Business*, 59, (1986), pp. 383–403.

D'Avolio, G. "The Market for Borrowing Stock." *Journal of Financial Economics*, 66, (2002), pp. 271–306.

De Long, B., A. Shleifer, L. Summers, and R. Waldmann. "Positive Feedback Investment Strategies and Destabilizing Rational Speculation." *Journal of Finance*, 45, (1990), pp. 379–395.

Fisher, K. L. and M. Statman. "Investor Sentiment and Stock Returns." *Financial Analysts Journal*, 56, (2000), pp. 16–22.

Fama, E. F. and K. R. French. "The Cross-section of Expected Stock Returns." *Journal of Finance*, 47, (1992), pp. 427–465.

Fama, E. F. and K. R. French. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, 33, (1993), pp. 3–56.

Frazanni, A. and O. A. Lamont. "Dumb Money: Mutual Fund Flows and the Cross-Section of Stock Returns." *Journal of Financial Economics*, 88, (2008), pp. 299–322.

Friedman, M. and L. J. Savage. "The Utility Analysis of Choices Involving Risk." *Journal of Political Economy*, 56, (1948), pp. 279–304.
Hanoch, G. and H. Levy. "The Efficiency Analysis of Choices Involving Risk." *Review of Economic Studies*, 36, (1969), pp. 335–346.
Hadar, J. and W. R. Russell. "Rules for Ordering Uncertain Prospects." *American Economic Review*, 59, (1969), pp. 25–34.
Harvey, C. R. and A. Siddique. "Conditional Skewness in Asset Pricing Tests." *Journal of Finance*, 55, (2000), pp. 1263–1295.
Kumar, A. "Who Gambles in the Stock Market?" *Journal of Finance*, 64, (2009), pp. 1889–1933.
Lee, C., A. Shleifer, and R. H. Thaler. "Investor Sentiment and the Closed-end Fund Puzzle." *Journal of Finance*, 46, (1991), pp. 75–109.
Levy, H. and Z. Wiener. "Stochastic Dominance and Prospect Dominance with Subjective Weighting Functions." *Journal of Risk and Uncertainty*, 16, (1998), pp. 147–163.
Lim, G., E. Maasoumi, and V. Martin. "A Re-examination of the Equity Premium Puzzle: A Robust Nonparametric Approach." *North American Journal of Economics and Finance*, 17, (2006), pp. 173–189.
Linton, O., E. Maasoumi, and Y. J. Whang. "Consistent Testing For Stochastic Dominance Under General Sampling Schemes." *Review of Economic Studies*, 72, (2005), pp. 735–765.
Machina, M. "Expected Utility Theory Without the Independence Axiom." *Econometrica*, 50, (1982), pp. 277–323.
Meyer, J. "Second Degree Stochastic Dominance With Respect to a Function." *International Economic Review*, 18, (1977), pp. 477–487.
Rothschild, M. and J. Stiglitz. "Increasing Risk I: A Definition." *Journal of Economic Theory*, 2, (1970), pp. 225–243.
Rothschild, M. and J. Stiglitz. "Increasing Risk II: Its Economic Consequences." *Journal of Economic Theory*, 3, (1971), pp. 66–84.
Shefrin, H. M. and M. Statman. "Behavioural Portfolio Theory." *Journal of Financial and Quantitative Analysis*, 35, (2000), pp. 127–151.
Shleifer, A. and L. H. Summers. "The Noise Trader Approach to Finance." *Journal of Economic Perspective*, 4, (1990), pp. 19–33.
Starmer, C. "Developments in Non-Expected Utility Theory: The Hunt for A Descriptive Theory of Choice Under Risk." *Journal of Economic Literature*, 38, (2000), pp. 332–382.
Tversky, A. and D. Kahneman. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty*, 5, (1992), pp. 297–323.
Whitmore, G. A. "Third Degree Stochastic Dominance." *American Economic Review*, 60, (1970), pp. 457–459.

Appendix: Details of Bootstrap Simulations

Following Linton, Maasoumi, and Whang [2005], I use a paired moving block bootstrap technique to approximate the unknown sampling distribution of the stochastic dominance test statistics. Let L denote the length of blocks used in the procedure satisfying $L \propto T^{\gamma}$ for some $0 \le \gamma \le 1$. Therefore, I obtain a total of $T - L + 1$ overlapping blocks of paired observations from the original sample. For each bootstrap, B blocks are drawn randomly with replacement and stacked in the order sampled to produce a bootstrap sample of T observations, that is, $BL = N$. I set the number of bootstrap repetitions at $M = 10{,}000$. Following Lim et al. [2006], I set $L = \alpha\sqrt{N}$ where $\alpha = 4$. The results are qualitatively similar using $\alpha = 2$ or 6.

Define the recentered bootstrap statistics by:

$$I^{c}_{j,i}(z;\hat{F}) = I_{j,i}(z;\hat{F}) - I_{j,OB}\,(z:\hat{F})\ \text{for}\ i = 1, 2, ..., M$$

where $j = 1, 2$ and 3 for first, second and third order stochastic dominance, and

$$I_{j,OB}(z:\hat{F}) = \frac{1}{N(j-1)!} \times \sum_{i=1}^{N} w(k, L, N) 1(X_i \le z)(z - X_i)^{j-1}$$

$$w(i, L, N) = \left\{ \begin{array}{c} i/L \\ 1 \\ (N-i+1)/L \end{array} \right\} \begin{array}{l} \text{if } i \in [1, L-1] \\ \text{if } i \in [L, N-L+1] \\ \text{if } i \in [N-L+2, N] \end{array}$$

Define $I^{c}_{j,i}(z;\hat{G})$ and $I^{c}_{j,OB}(r;\hat{G})$ analogously. The recentered bootstrap test statistics are then given by

$$SD^{F,G}_{j,i} = \sqrt{N} \sup_{z} \left(I^{c}_{j,i}(z;\hat{F}) - I^{c}_{j,i}(z;\hat{G})\right)\ \text{for}\ i = 1, 2, ..., M,$$

$$SD^{G,F}_{j,i} = \sqrt{N} \sup_{z} \left(I^{c}_{j,i}(z;\hat{G}) - I^{c}_{j,i}(z;\hat{F})\right)\ \text{for}\ i = 1, 2, ..., M$$

where $SD^{F,G}_{j,i}$ is the test statistic for the null hypothesis that F stochastically dominates G at the jth order and $SD^{F,G}_{j,i}$ is the test statistic for the reverse null. Test statistics for the risk-seeking case can be defined analogously using Equations (12) and (13). The distributions of these subsample test statistics are used to approximate the distributions of the test statistics given in the text. By randomly sampling paired blocks of observations instead of individual observations, this technique allows for serial correlations and general dependence.