Does Timing the Momentum Crowd Pay Off? An Analysis of Hedge Fund Performance*

Ralitsa Petkova†

September, 2019

Abstract

This paper shows that some hedge funds have the ability to adjust their exposure to the momentum strategy by gauging the size of the momentum crowd. Funds that scale back their exposure to momentum at times when short sellers are actively crowded into the strategy earn 8% more annually than funds that increase their exposure at such times. Funds that increase their exposure to momentum at times when positive-feedback trading by mutual funds is expected to intensify earn 5% more annually than funds that decrease their exposure at such times. Funds that scale back their exposure to momentum at times when noise traders are expected to be active earn 7% more annually than funds that increase their loading on momentum at such times. The evidence suggests that some hedge funds are skilled at timing the actions of other active investors and they adjust the direction of their exposure to momentum accordingly. This skill is rewarded by economically significant average returns.

*I thank seminar participants at Case Western Reserve University for helpful comments and suggestions. All remaining errors are mine.

†Weatherhead School of Management, Case Western Reserve University, Cleveland, OH 44106. E-mail: rgp9@case.edu. Tel: 216-368-4778.

"...the key to investment success is not just predicting future fundamentals, but also predicting the movement of other active investors."

Shleifer and Summers (1990)

1 Introduction

Crowding has become a major concern for sophisticated traders, particularly in quantitative strategies. Crowding is the tendency of investors to implement similar strategies and trade in the same direction at the same time.[1] Since sophisticated market participants have access to the same datasets and statistical tools, they are likely to end up with overlapping positions. Such crowded positions can lead to the alpha decay of a strategy as more capital is deployed to exploit it. Furthermore, crowding can increase the tail risk of a strategy. If there are many market participants crowded into a trade, a shock to the system may force everyone to rush to the exits at the same time, exacerbating the risk of falling prices, margin calls, and vanishing liquidity (Brunnermeier and Pedersen (2009)).[2]

There is evidence that practitioners are aware of the importance of how many traders are simultaneously entering the same strategy space and to what extent their signals are correlated. Some firms provide tools to institutional investors that identify the crowdedness of various trading strategies. For example, MSCI offers a "crowding scorecard" which is a standardized measure of the trading activity in a given strategy. Novus publishes a Crowding Index based on the percent of shares owned by hedge funds for each stock. Regulators have also been concerned about crowded trades posing a threat to financial institutions in cases when many market participants exit similar positions at the same time.[3]

[1] There is no formal definition of crowding in the finance literature. The concept of crowding broadly refers to the tendency of a large group of investors to trade in a similar way in response to the same signals, leading to overlapping portfolio positions.

[2] During the Quant Meltdown of August 2007, the simultaneous liquidation of very similar, highly-levered positions by quant equity hedge funds caused massive losses as funds were forced up against their margin limits (Khandani and Lo (2010), Pedersen (2009)).

[3] "*While there may well be more diversity in the types of strategies hedge funds follow, there is also considerable clustering, which raises the prospect of larger moves in some markets if conditions lead to a general withdrawal from these "crowded" trades.*" Timothy Geithner (2004).

While academic research shows that crowding has a negative effect on the profitability of certain equity strategies[4], little is known about the ability of sophisticated investors to respond to crowding by strategically adjusting their investment positions. Do sophisticated investors scale back their exposure to a strategy at times when the strategy is expected to underperform as a result of being overcrowded? Do some investors increase their positions in a trade before it becomes crowded, profiting from the price pressure induced by the herding behaviour of others who join the trade in later periods? In this paper I try to fill the gap by studying whether sophisticated investors strategically respond to crowding by other investors and whether this skill is rewarded by higher future performance. I investigate this issue in the context of the hedge fund industry. Hedge funds are the quintessential arbitrageurs in financial markets and timing the actions of other investors is crucial since arbitrage is risky and expensive.[5] The trades of other active investors in the market can affect the persistence of mispricing, the optimal time of arbitraging it away, and the amount of capital deployed in trading on it. In addition, while the number of hedge funds has exceeded 8,000, the number of publicly listed companies in the US has declined to about 4,000 in 2017. Therefore, a larger set of sophisticated investors is facing a shrinking universe of stocks and alpha is being squeezed from all directions. In such a market, the ability to time the crowd of other active investors in the same strategy space becomes even more valuable.

To investigate hedge funds' ability to time the investor crowd, I use the momentum strategy as a case study. Momentum is a natural candidate for this analysis since it is one of the most popular strategies in the asset-pricing literature and it is considered to be the premier return anomaly.[6] Furthermore, there is evidence that a large percentage of hedge funds implement momentum strategies.[7] In addition, theoretical and empirical work explicitly links momentum to the issue of

[4]See, for example, Lou and Polk (2014).

[5]See, for example, Shleifer and Vishny (1997).

[6]See Fama and French (2008). Jegadeesh and Titman (1993) show that, based on performance over the last three months to one year, past losers continue to be losers and past winners continue to be winners over the next three to twelve months. Several behavioral models explain the existence of return momentum with investor underreaction or overreaction. The debate on the driving forces behind momentum is still ongoing.

[7]According to PREQIN's 2018 Global Hedge Fund Report, 31% of hedge funds offer momentum strategies to their investors.

crowding. While some studies argue that crowding by momentum investors is negatively correlated with momentum profitability (e.g., Stein (2009), Hanson and Sunderam (2013), Lou and Polk (2014)), other studies suggest that crowding may temporarily drive prices in favour of momentum traders (e.g., De Long, Shleifer, Summers, and Waldmann (1990b)).

In this study I ask the following questions: Can hedge fund managers strategically adjust their momentum exposure based on information about the size of the momentum crowd? If so, how much economic value does this timing skill bring to hedge fund investors? These issues are essential to understanding the role of strategy crowding in professional asset management and the ability of hedge funds to continue to deliver alpha. To answer these questions, I proceed in three steps. First, I use empirical measures of the crowd of investors who follow momentum signals. Second, I use a timing model for hedge fund returns that examines whether fund managers adjust their exposure to the momentum strategy conditional on the size of the momentum crowd. Finally, I explore the economic significance of momentum-crowd-timing by examining the out-of-sample performance of funds with different levels of momentum-crowd-timing skill. Furthermore, I perform a number of tests that show that momentum-crowd-timing behavior is related to manager skill rather than random chance.

The theoretical literature on momentum suggests that the crowd of active investors who follow momentum signals is not homogeneous, but consists of various groups of traders with different beliefs and motivations. In line with that literature, I focus on three groups of momentum traders: (i) the crowd of arbitrageurs who trade against mispricing, (ii) the crowd of positive-feedback traders who behave like return chasers, and (iii) the crowd of noise traders who drive security overpricing as a result of behavioral biases. I use empirical measures for the different segments of the momentum crowd. Specifically, as in Hanson and Sunderam (2013), I use the quantity of short-side capital devoted to momentum as a proxy for the crowd of arbitrageurs trading on momentum. It is measured by the strength of the cross-sectional relationship between short interest and the momentum signal at each point in time. The advantage of using this measure is that short interest is mostly associated with the trades of sophisticated investors such as hedge funds.[8] Following Lakonishok, Shleifer, and Vishny (1992), I use the strength of the cross-sectional relationship between excess demand by mutual funds and past return performance to measure the crowd of positive-feedback traders. While there may be other investors who engage in positive-feedback trading, I focus on mutual funds since they represent a crowd of investors that is different from hedge funds and is large enough to have a material market impact when they act simultaneously. Finally, I use the investor sentiment index of Baker and Wurgler (2006) as a proxy for the size of the crowd of noise traders that are active in the market. Previous studies have shown that high levels of the sentiment index are associated with greater participation of sentiment-driven traders in the market.[9] Even though these measures capture different aspects of the crowd of investors who follow momentum signals, they share a common notion that a crowding measure should represent the trading of many investors acting in the same way, at the same time, following the same signals.

I first document that the three momentum crowd measures have predictive ability for future momentum returns. More specifically, the future returns of the momentum strategy are negatively correlated with the quantity of short arbitrage capital devoted to momentum and positively correlated with the intensity of positive-feedback trading in the strategy. The results also reveal that higher levels of investor sentiment tend to predict negative momentum skewness over the following year. Since the three crowding measures contain significant information about future momentum strategy performance, they represent valuable signals to be used in a momentum-crowd-timing model.

I use a timing model for hedge fund returns, based on the classic framework of Treynor and Mazuy (1966), that examines whether fund managers adjust their exposure to the momentum strategy depending on the size of the momentum crowd. The three separate measures that distinguish the crowd of arbitrageurs, positive-feedback traders, and noise traders are the timing

[8] Boehmer, Jones, and Zhang (2012) and Ben-David, Franzoni, and Moussawi (2012) argue that hedge funds account for most short interest in the United States.

[9] See, for example, Stambaugh, Yu, and Yuan (2011).

variables in the model.[10] The timing model controls for hedge funds' exposures to other relevant factors that have been shown to drive hedge fund returns.[11]

Using a large sample of equity market neutral and long-short equity funds from the Lipper TASS database over the period from 1994 to 2017, I show that for a large percentage of funds, exposure to momentum is significantly related to the extent of arbitrage activity in momentum as measured by short selling (16% of all funds), the intensity of positive-feedback trading by mutual funds (21% of all funds), and to the presence of sentiment traders as measured by the investor sentiment index (23% of all funds).

The advantage of using hedge fund returns for the analysis is that I can directly test how considerations associated with crowded trades may propagate to portfolio returns. In addition, an important feature of hedge funds is the speed with which they can alter their investments in response to changing conditions. Hedge fund returns, which are available at higher frequencies (i.e., monthly) than hedge fund holdings are, therefore, especially suitable to study the crowd-timing abilities of hedge funds.[12]

Next, I test whether differences in momentum-crowd-timing behavior across funds can predict hedge fund performance. Specifically, in each month I sort funds into quintiles based on their momentum-crowd-timing coefficients estimated from the previous 36 months. Then, I measure the out-of-sample performance of the portfolios over the next month. I find that hedge funds that decrease their exposure to momentum at times when short sellers have crowded into the strategy earn 8% more annually than funds that increase their exposure at such times. Funds that increase their exposure to momentum at times when positive-feedback trading by mutual funds is expected to intensify earn 5% more annually than funds that decrease their exposure at such times. Funds

[10] The three groups of momentum traders are examined separately since their actions have different implications about the timing behavior of hedge funds. More details on this are provided in Section 2.

[11] The factors are based on the model of Fung and Hsieh (2004).

[12] Grinblatt, Jostova, Petrasek, and Philipov (2016) use institutional 13F holdings to study whether hedge funds trade on momentum. They find that the majority of hedge fund managers are contrarian, although their tendency to sell recent winners is less pronounced. My study differs from Grinblatt et al. (2016) since I use hedge fund returns and assume that hedge funds' exposure to momentum varies over time as the crowd of momentum traders varies as well.

that scale back their exposure to momentum at times when noise traders are expected to be active in the market, and investor sentiment is high, earn 7% more annually than funds that increase their loading on momentum at such times. In all cases, the spread in out-of-sample returns between the top and bottom crowd-timing funds remains significant for about 6 months after portfolio formation.

Furthermore, I investigate whether hedge funds' ability to time the momentum crowd is due to random chance using a bootstrap analysis. The findings strongly suggest that the momentum-crowd-timing abilities of hedge funds cannot be attributed purely to luck. I also find evidence of persistence in momentum-crowd-timing skill.[13] Taken together, the results suggest that momentum-crowd-timing represents managerial skill adding value to hedge fund investors.

When drawing inference about the crowd-timing skills of hedge funds, I try to minimize the impact of biases in the data documented by Fung and Hsieh (1997, 2000), including survivorship bias and backfill bias. I use both live and defunct funds to mitigate the effect of survivorship bias. To alleviate the impact of backfill bias, I discard the return observations before the funds are added to the database. The results suggest that the inference about momentum-crowd-timing is robust to various hedge fund data biases.

Across funds, there is a wide difference in timing ability. Therefore, I investigate whether crowd-timing is related to certain fund attributes. The results indicate that the top hedge fund timers in terms of future performance tend to be smaller funds, with a higher tendency to use leverage, and longer payout periods. These findings seem reasonable as we would expect smaller funds to be more flexible in adjusting their positions. In addition, funds with more managerial discretion (i.e., longer payout period) are more likely to have the flexibility of implementing a timing strategy. Finally, funds that use more leverage are likely to have a higher sensitivity to the risk of margin calls in a crowded market.

Finally, I show that hedge funds identified for their ability to strategically time the momentum crowd possess other skills that are different from momentum investing. Namely, I show that hedge funds that have strategic crowd-timing abilities are more likely to pursue unique investment

[13] Jagannathan, Malakhov, and Novikov (2010) document performance persistence among superior hedge funds.

strategies relative to their style peers, as measured by their strategy distinctiveness index. In addition, I show that the performance gap between top and bottom quintile funds in each momentum-crowd-timing category is larger in times of greater investment opportunities in the hedge fund industry, as measured by the cross-sectional dispersion of stock returns.

The results in this paper contribute to our understanding of the behavior of sophisticated investors and their ability to exploit changes in crowding conditions. The evidence in this study is consistent with hedge funds being aware of crowding issues and making an effort to take them into account. When the momentum strategy is crowded by other arbitrageurs, hedge funds that avoid the crowd perform better. In contrast, when the momentum strategy is about to be exploited by positive-feedback traders, hedge funds that front-run the crowd have better performance. Finally, hedge funds that scale back on momentum exposure ahead of periods of high sentiment have higher future performance.

These empirical results are broadly consistent with several theoretical papers that examine the optimal behavior of arbitrageurs in the presence of other active investors. Stein (2009) points out that in the case of momentum, arbitrageurs' demand for an asset is an increasing function of the recent asset return rather than fundamental value. Therefore, a high past return could mean that the firm received good news and, therefore, is a good candidate for a long position. On the other hand, a high past return could also mean that many other arbitrageurs have already exploited this opportunity to the extent that the asset is now overvalued and, therefore, should be shorted. If arbitrageurs are forced to withdraw capital from the strategy when it is crowded, their collective unwinding of levered positions can lead to abrupt negative returns. Therefore, avoiding momentum at times when it is crowded by other arbitrageurs may be optimal.

Stein's (2009) argument focuses on the notion that crowding by arbitrageurs may dislocate prices and eventually lead to return reversal. However, an argument could be made that some level of crowding is welcomed by arbitrageurs who trade on momentum signals in order to monetize their trades. An arbitrageur might attempt to buy into a momentum position before it becomes too popular, stimulating the interest of positive-feedback traders who buy when prices have risen in the past and sell when prices have fallen. The herding behaviour of positive-feedback traders to the strategy may temporarily drive prices in favor of momentum traders. This is in line with De Long, Shleifer, Summers, and Waldman (1990b), who argue that when some noise traders follow positive-feedback strategies like momentum, it may be optimal for arbitrageurs to "jump on the bandwagon" themselves and feed the bubble in the short run.[14]

Finally, several studies suggest that limits to arbitrage would prevent hedge funds from fully exploiting the momentum anomaly in certain states of the world. Specifically, during high sentiment periods, the views of not-fully-rational noise traders tend to drive security overpricing.[15] To the extent that momentum losers are overpriced during periods of high investor sentiment when noise traders are active in the market, rational arbitrageurs cannot completely eliminate this overpricing due to impediments to short selling.[16] One such impediment comes from the unpredictability of the future resale price (De Long, Shleifer, Summers, and Waldmann (1990a)). As long as the arbitrageur has to liquidate their position in the short-term, they must bear the risk that future overpricing might be more severe or not eliminated at all. Therefore, arbitrageurs' aggressiveness in trading on momentum signals is likely to be limited in the presence of a crowd of noise traders in the market who drive asset mispricing.[17]

[14]Brunnermeir and Nagel (2004) show that hedge funds did not trade against the NASDAQ technology bubble during the 1998-2000 period. Instead, they were heavily invested in technology stocks and they were actually riding the bubble. Griffin, Harris, Shu, and Topaloglu (2011) find that institutional investors dominate the active buys of high tech stocks before the technology bubble burst and the active sell-offs during the technology bubble burst period.

[15]One such group of irrational traders are investors who exhibit the disposition effect, i.e., the tendency to hold on to losing stocks. This type of behavior has been explained by the prospect theory of Kahneman and Tversky (1979), and the disposition effect has been documented among retail investors (e.g., Shefrin and Statman (1985), Odean (1998), Grinblatt and Keloharju (2001)) and money managers (e.g., Wermers (2003), Frazzini (2006), O'Connell and Teo (2009), Jin and Scherbina (2011)). Another group of irrational traders are investors who exhibit cognitive dissonance (Festinger (1957)). This group of noise traders tend to ignore news about stocks that contradict their own sentiment. This phenomenon tends to cause overpricing of losers in high sentiment periods (Antoniou, Doukas, Subrahmanyam (2013)).

[16]Miller (1977) suggests that impediments to short selling play a significant role in limiting the ability of rational traders to exploit overpricing. A growing body of literature contends that investor sentiment could drive asset mispricing (Baker and Wurgler (2006) and Shleifer and Summers (1990)), and that sentiment-induced mispricing may be asymmetrical between high- and low-sentiment environments due to short-sale constraints (Stambaugh, Yu and Yuan (2012)).

[17]Note that the arguments in this paragraph implicitly differentiate between noise traders, such as positive-feedback investors, who trade in the direction of momentum signals, and other noise traders who, as a result of behavioral biases, do not trade in the direction of momentum signals or do so very slowly.

Overall, the theoretical literature implies that the strategic response of arbitrageurs to the crowd of momentum investors is more nuanced than simply going against the crowd. The empirical results in the paper show that understanding these nuances is a valuable skill that some hedge funds posses. Identifying that skill could be beneficial for hedge fund investors since strategic crowd-timing behavior is associated with better performance in the future.

This paper contributes to the literature on hedge fund performance. Previous studies present strong evidence that top hedge funds deliver alpha.[18] In trying to understand the sources of such superior performance, several papers have shown that hedge funds are skilled at predicting stock fundamentals, timing market liquidity and volatility, and hedging macroeconomic risk, among others.[19] In this paper, I examine a new aspect of the investment skill of hedge funds, namely, their ability to tactically adjust strategy exposure by timing the crowd of active investors who are present in the same strategy space.

On the other hand, this paper contributes to the empirical literature on crowding. The notion that momentum crowding affects the future profitability of the momentum strategy has been studied before in the literature (e.g., Hanson and Sunderam (2013) and Lou and Polk (2014)). However, the effect of momentum crowding on the investment decisions of hedge funds implementing momentum strategies has not received much attention. The main innovation in this paper is to provide empirical evidence that hedge funds pursuing momentum respond significantly to the crowd of other investors trading on momentum. These results contribute to our understanding of the timing skills of the most sophisticated investors when faced with crowding in the quantitative-equity-strategy space.

The rest of the paper is organized as follows. Section 2 develops the main hypotheses tested in the paper, motivated by theoretical arguments on strategy crowding. Section 3 describes the

[18] For example, Kosowski, Naik, and Teo (2007) show that the performance of the top hedge funds ranked by the t-statistic of alpha cannot be attributed to random chance.

[19] Chen (2007) and Chen and Liang (2007) find evidence of market-timing and volatility-timing in hedge funds. Cao, Chen, Liang, and Lo (2013) show that hedge funds' market exposure varies with aggregate liquidity conditions. Bali, Brown, and Caglayan (2014) find that hedge funds can time economic uncertainty. Superior performance by hedge funds could also come from exposures to risks commanding positive premiums. For example, Sadka (2010) and Teo (2011) link liquidity risk exposure to the cross section of hedge fund returns. Bali, Brown, and Caglayan (2011, 2012, 2014) show that exposures to fundamental risks such as macroeconomic factors help explain hedge fund returns.

data on hedge fund returns and factors. Section 4 presents the main momentum-crowd timing model, and the construction of the momentum crowd variables. The main results in the paper regarding the economic value of timing the momentum crowd are in Section 5. Section 6 examines the characteristics of top hedge funds in terms of crowd-timing skills. Section 7 identifies other skills that momentum-crowd-timing funds posses, different from their ability to time the crowd. Some robustness tests are presented in Section 8 and Section 9 concludes.

2 Hypothesis Development

In this section, I revisit existing theoretical arguments about the behavior of arbitrageurs in the presence of other active investors, to develop testable hypothesis about the momentum-crowd-timing skills of equity hedge funds. I am particularly interested in hedge funds' ability to time the crowd of other arbitrageurs, the crowd of positive-feedback traders, and the crowd of noise traders who base their decisions on sentiment rather than fundamentals. Although I may not capture all traders whose activities involve momentum strategies, examining these three groups of investors provides valuable insights into the actions of hedge funds under strategy crowding. In addition, the choice of these particular groups of momentum traders is guided by existing theoretical arguments that sophisticated arbitrageurs' trading decisions are affected by other arbitrageurs, positive-feedback traders, and noise traders.

One theoretical foundation behind the importance of timing the crowd of other arbitragers comes from the work of Hong and Stein (1999) and Stein (2009). In Hong and Stein (1999), newswatchers underreact to private signals about stock fundamentals due to slow information diffusion. When only newswatchers are present in the market, prices adjust slowly to new information leading to underreaction and, therefore, momentum. The continuation in returns creates an arbitrage opportunity for momentum traders who base their trades only on price changes over a recent interval. Momentum traders' attempts to profit from the newswatchers' underreaction leads to accelerated price movements in the direction of fundamentals. As more momentum traders join, prices eventually overshoot their long-run equilibrium values, which leads to overreaction and price correction. Hong and Stein (1999) predict that momentum investing will lose money late in the cycle when prices have already overshot long-run equilibrium values as a result of many momentum traders having crowded into the strategy.

Stein (2009) works within the framework of Hong and Stein (1999) and the boundedly-rational arbitrageurs in his model are simple momentum traders. Stein (2009) points out that if arbitrageurs only condition their trading activity on a stock's past return, such a momentum strategy lacks a natural anchor. Specifically, a high past return could mean that the firm received good news and, given that newswatchers underreact to information, arbitrageurs should bid up the stock price. On the other hand, a high past return could mean that other arbitrageurs have already bid the stock price up to the extent that it now reflects fundamental value. Since these two possible scenarios are hard to distinguish by observing past stock returns alone, sometimes there is too much activity in momentum and the initial mispricing is overcorrected. From an individual arbitrageur's perspective, implementing a momentum strategy carries the risk of getting into the strategy when it is already crowded: if arbitrageurs are forced to withdraw capital from the momentum strategy, their collective unwinding of positions can lead to abrupt momentum crashes.

The key feature of Stein's (2009) model is that the stock price (i.e., momentum signal) is not necessarily a summary statistic for fundamental news since it is also influenced by the level of arbitrage activity in the stock. Stein (2009) concludes that if a given price realization reflects a high level of arbitrage activity, then each individual arbitrageur would be better off taking the opposite position in the stock.[20]

At least two empirical studies have examined whether the crowding problem discussed in Stein (2009) affects the time-series variation of momentum profitability. For example, Lou and Polk (2014) derive a measure of arbitrage activity in momentum based on abnormal return correlations among typical momentum strategy stocks. They show that when their measure indicates that

[20] In Stein (2009), each arbitrageur cannot be certain about the number of other arbitrageurs who are active in the same strategy. In Abreu and Brunnermeier (2002), on the other hand, each arbitrageur cannot be certain about the timing of other arbitrageurs, i.e., when they will act on the same strategy.

arbitrageurs are crowded into the strategy, momentum tends to crash and revert, reflecting overreaction as a result of prior rounds of trading pushing prices away from fundamentals. Hanson and Sunderam (2013) develop a different method of inferring the amount of arbitrage capital deployed on momentum, which relies on time variation in the cross section of short interest. They show that higher levels of short arbitrage capital in momentum are negatively related to the profitability of the momentum strategy.

Motivated by the theoretical arguments on the crowding of arbitrageurs, and the empirical success of recent studies in measuring arbitrage capital, I formulate the first hypothesis:

Hypothesis 1: To the extent that hedge funds can infer the size of the crowd of other arbitrageurs that are trading on momentum, decreasing their exposure to momentum when it is crowded should be associated with better performance. Scaling down momentum loadings at such times would help funds mitigate the effect of decreasing momentum returns as a result of crowding. Therefore, a savvy manager who can correctly infer that other arbitrageurs have crowded into momentum would naturally wish to reduce their fund's momentum exposure in anticipation of return reversal.

Another theoretical underpinning of the importance of timing the momentum crowd comes from models of positive-feedback trading. In these models, positive-feedback traders exert price pressure as they buy past winners and sell past losers, thereby generating initial momentum and subsequent reversal in the cross section of stocks. For example, De Long et al. (1990b) show that rational speculators have the incentive to front run positive-feedback traders in order to stimulate a price increase (decline) and take advantage of their subsequent trading.[21] If there is good news today, rational traders buy and push the price higher because feedback traders are willing to take up the position at a higher price in the next period. Therefore, the incentive to ride the bubble stems from predictable feedback trader demand.

[21] Several models that generate cross-sectional momentum feature agents that behave as positive-feedback traders, but the underlying mechanisms differ from each other. In Daniel, Hirshleifer, and Subrahmanyam (1998) the underlying mechanism is based on self-attribution bias, in Barberis et al. (1998) it is based on representativeness, while in Hong and Stein (1999) it is bounded-rationality. In this paper, I take positive-feedback trading as given and study how its intensity affects the momentum exposure of hedge funds.

In Barberis and Shleifer (2003), investors with extrapolative beliefs behave like positive-feedback traders, buying styles that have performed well in the past and selling styles that have performed poorly. Extrapolators, seeing the past performance of an asset, become more optimistic (pessimistic) about the future prospects of the asset and push the price higher (lower) over subsequent trading periods. Eventually, the price of the asset moves significantly above (below) its fundamental value. This is a sign that extrapolators have been buying (selling) it aggressively, causing it to become overpriced (underpriced). The overvaluation (undervaluation) is then followed by low (high) returns. One natural prediction of this model is that the effect on prices is stronger when there are more extrapolators in the economy. Barberis and Shleifer (2003) state that if an arbitrageur is clever enough to anticipate the behavior of extrapolators, the optimal strategy would be a momentum-like strategy that rides with the crowd of extrapolators as long as their investment flows are heavily dependant on styles' past performance.

In the models of De Long et al. (1990b) and Barberis and Shleifer (2003) there is a key difference between the time when arbitrageurs trade on momentum signals and the time when positive-feedback traders do. If there is a positive (negative) price change between time $t-1$ and time t, an arbitrageur buys (shorts) immediately at time t, while a positive-feedback trader buys (sells) at time $t+1$.[22] In this type of framework, momentum trading by arbitrageurs is profitable because it front runs the positive-feedback traders or return extrapolators.

Barberis and Shleifer (2003) suggest that the extrapolators in their model could be institutional investors who chase the best-performing styles. Several empirical studies document that institutions engage in positive-feedback trading. For example, Lakonishok, Shleifer, and Vishny (1992) show evidence from pension managers, while Grinblatt, Titman, and Wermers (1995) and Carhart (1997) show evidence for mutual funds. More recently, Frijns, Gilbert, and Zwinkels (2015) find strong evidence of positive-feedback trading among US mutual funds in a way that is consistent with the assumptions of Barberis and Shleifer (2003). Grinblatt, Jostova, Petrasek, and Philipov (2016) find

[22] Some possible explanations for why positive-feedback traders behave this way include frictions or constraints that delay portfolio adjustment, or the possibility that some investors observe past price changes with a delay (Barberis (2018))

that about 2/3 of mutual fund managers follow momentum strategies, both for purchases and sales. Finally, Peng and Wang (2019) show that mutual funds contribute to cross-sectional momentum through positive-feedback trading.

Motivated by the theoretical arguments on the interaction between rational arbitrageurs and positive-feedback traders, and the empirical evidence that institutions, such as mutual funds, that have a profound effect on stock prices engage in positive-feedback trading, I formulate the second hypothesis:

Hypothesis 2: To the extent that hedge funds can anticipate the size of the crowd of positive-feedback traders, increasing their exposure to momentum at times when positive-feedback traders are expected to actively trade on momentum signals should be associated with better performance. Increasing momentum loadings at such times would help funds benefit from the temporary price pressure from positive-feedback trading. Therefore, a smart manager who can correctly forecast positive-feedback trader demand would naturally wish to increase their fund's momentum exposure in advance of the demand.

Positive-feedback traders are not the only type of noise traders who are prone to irrational behavior. Their behavior is of special interest to arbitrageurs since their trading decisions go in the direction of momentum signals. Previous studies have described the behavior of other noise traders who do not necessarily trade in the direction of momentum signals. For example, some investors' behavior is consistent with the disposition effect explained by the prospect theory of Kahneman and Tversky (1979). That is, they have the tendency to hold onto losing stocks and sell winning stocks. This effect has been documented not only among retail investors (e.g., Shefrin and Statman (1985), Odean (1998), Grinblatt and Keloharju (2001)) but also among more sophisticated investors such as money managers (e.g., Wermers (2003), Frazzini (2006), O'Connell and Teo (2009), Jin and Scherbina (2011)).[23] Another type of investors, prone to cognitive dissonance (Festinger (1957)), tend to ignore information about stocks that contradicts their overall sentiment. For example,

[23] Grinblatt and Han (2005) and Frazzini (2006) note that the disposition effect should lead to underreaction to news for which they find supporting empirical evidence.

Antoniou, Doukas, and Subrahmanyam (2013) argue that investors who are affected by cognitive dissonance do not react promptly to negative information about losers in periods when they are overly optimistic. Both types of investors described above tend not to sell losing stocks. As long as these investors are present in the market in large numbers, negative information about losers will diffuse extremely slowly.

The examples of investors prone to the disposition effect and to cognitive dissonance are just two examples of noise traders who make investment decisions based on sentiment rather than fundamental information.[24] It is reasonable to expect that arbitrageurs' intensity in trading on momentum signals may be affected by the presence of noise traders in the market. This is the case since, as previous studies have argued, when many sentiment traders are active in the market, their collective sentiment can exert an effect on stock prices (e.g., Shleifer and Summers (1990), Baker and Wurgler (2006)). To the extent that momentum losers become severely overpriced during periods of high investor sentiment when the crowd of noise traders active in the market is predicted to be large, rational arbitrageurs cannot completely eliminate this overpricing due to impediments to short selling.[25]

One impediment to short selling comes from the unpredictability of the future resale price (De Long, Shleifer, Summers, and Waldmann (1990a)). As long as the arbitrageur has to liquidate their position in the short-term, they must bear the risk that future overpricing might be more severe or not eliminated at all. Traders who short a security in the belief that its price is too high can be correct, in that the price will eventually fall, but they face the risk that the price will go up before it goes down. Such a price move, requiring additional capital, can force the traders to liquidate at a loss. Shleifer and Vishny (1997) argue that such arbitrage risk looms particularly large for institutional managers, whose career paths depend heavily on recent performance. Fear of this risk limits the size of the arbitrageur's initial position and prevents him from driving the price all the

[24] Sentiment is typically defined as the difference between the beliefs of sentiment-driven traders and correct objective beliefs conditional on available information (e.g., Delong, Shleifer, Summers, and Waldman (1990a)).

[25] Miller (1977) suggests that impediments to short selling play a significant role in limiting the ability of rational traders to exploit overpricing. Even investors who do not face institutional constraints or high shorting costs can nevertheless be deterred by the risks in arbitrage, as discussed by Shleifer and Vishny (1997).

way down to fundamentals. This leads to the third hypothesis:

Hypothesis 3: To the extent that a large crowd of noise traders drives the overpricing of losers in times of high investor sentiment, arbitrageurs' exposure to momentum, controlling for the separate effects described in hypotheses 1 and 2, is predicted to be smaller in periods of high investor sentiment. A smaller loading on momentum at times of high sentiment would reflect arbitrageurs' unwillingness to short losers as aggressively as they would otherwise since they have short horizons and the overpricing may take a long time to correct.

As the arguments above suggest, the crowd of investors who trade on momentum signals is diverse and subject to different beliefs, motivations, and timing of trades. Figure 1 ties it all together by presenting a time line for momentum signals, momentum strategy returns, and the trading demand of various investors. If, for example, the price of an asset increased over the period from $t-1$ to t, Figure 1 shows that arbitrageurs will buy the asset at time t.[26] If at time t smart hedge fund managers can infer that the price move from $t-1$ to t reflects the presence of a large crowd of arbitrageurs who have already established momentum positions, then they would be better off shorting the asset at time t. Alternatively, if at time t smart hedge fund managers can anticipate that positive-feedback traders will buy the asset at time $t+1$ as a result of extrapolating past price moves into the future, then they would be better off to increase their position in the asset at time t and profit from future extrapolative demand. Similar arguments hold in the case in which the asset's price decreases over the period from $t-1$ to t. Finally, arbitrageurs who can forecast that the t to $t+1$ period will be characterised by the presence of many sentiment-driven traders, may lower their short exposure at t to assets that have gone down over $t-1$ to t as a result of impediments to short selling.

To test the three hypotheses developed earlier, I use empirical proxies for the crowd of arbitrageurs, positive-feedback traders, and noise traders who trade on momentum signals. The next section contains details about the constructions of these proxies and their use in the

[26] In the empirical implementation of the momentum strategy, the duration of the period from $t-1$ to t is 12 months, where the last month before t is skipped to avoid the short-run reversal effect documented by Jegadeesh (1990).

momentum-crowd-timing model for hedge funds.

3 Data

In this section, I describe the hedge fund data sample and the factors used in explaining hedge fund returns.

3.1 Hedge Fund Data

The hedge fund data is from the Lipper TASS database, which is one of the most widely-used databases in the hedge fund literature. The data includes monthly fund returns and various fund characteristics. To minimize the impact of sample bias summarized in Fung and Hsieh (2000), I follow the steps suggested by Cao, Chen, Liang and Lo (2013). Specifically, I analyze funds with assets under management (AUM) of $5 million or more and funds that report net returns on a monthly basis. To address survivorship bias, I follow the hedge fund literature and include both live and defunct funds over the period from January 1994 through December 2017. To control for the impact of backfilling bias, the first 12 observations from the time series of each fund are deleted.[27] In addition, I only use funds that report their returns in USD.

Because my focus is on the timing ability of hedge fund managers who implement momentum, the main analysis in the paper includes equity market neutral and long-short equity funds. In subsequent tests, I include other equity investment categories of funds: multi-strategy, convertible arbitrage, event-driven, global macro, and funds of funds. These tests produces similar results to the ones obtained with the main analysis.

The final sample of equity market neutral and long-short equity funds contains 2,116 funds over the sample period of 1994-2017, of which 1,807 are long-short equity and 309 are equity market neutral. Among the sample funds, 305 are alive as of the end of the sample period and 1,811 became defunct during the period. To obtain meaningful results, I require each fund to have at least 36 monthly returns.

[27] The empirical results in the paper are robust to excluding the first 18 or 24 observations from the time series of each hedge fund.

Panel A of Table 1 summarizes the monthly returns for the sample of equity market neutral and long-short equity funds. Over the sample period, the average monthly return for equity market neutral funds is 0.47% (5.64% per year) with a standard deviation of 2.35%. The average monthly return for long-short equity funds is higher at 0.78% (9.36% per year) with a higher standard deviation of 6.11%. The long-short equity category of funds exhibits a wider range of realized monthly returns with the 25th percentile of the distribution at -124% and the 75th percentile at 275%. For both types of funds, the average monthly return appears to be close to the median monthly return.

3.2 Hedge Fund Factors

It is well known that hedge funds often follow dynamic trading strategies (e.g., Fung and Hsieh (1997)) and use derivatives (e.g., Chen (2011)). To control for these previously established patterns, the empirical analysis in the paper controls for the seven-factor model proposed by Fung and Hsieh (2004). The seven factors include both linear and option-like factors that have been shown to explain the variations in hedge fund returns. Panel B of Table 1 presents summary statistics for the Fung-Hsieh seven factors. These factors are the market excess return ($MKTRF$), a size factor (SMB), the monthly change in the 10-year Treasury constant maturity yield (Δ $TERM$), the monthly change in Moody's Baa yield minus the 10-year Treasury constant maturity yield (Δ $CREDIT$), and three trend-following factors: $PFTSBD$ (bond), $PFTSFX$ (currency), and $PFTSCOM$ (commodity). The monthly data for the hedge fund risk factors comes from David A. Hsieh's data library.[28] The data for the other factors comes from CRSP and the Federal Reserve database.

4 Timing Model

The timing model that I use is designed to measure the ability of hedge funds to adjust their exposure to momentum depending on the size of the momentum crowd. Therefore, I first show the

[28] https://faculty.fuqua.duke.edu/~dah7/HFRFData.htm

general framework of the timing model and then I describe the measures that capture the size of the momentum crowd.

4.1 Timing Model Framework

The momentum timing model builds on the work of Treynor and Mazuy (1966). For simplicity, assume that a hedge fund manager generates portfolio returns according to the following process:

$$R_{i,t} = \alpha_i + \beta_{i,t-1}^{mom} R_t^{mom} + u_{i,t}, \tag{1}$$

where $R_{i,t}$ is the return of the fund in excess of the risk-free rate and R_t^{mom} is the return of the momentum strategy. In Equation (1), the fund's exposure to momentum, $\beta_{i,t-1}^{mom}$, varies over time. The fund's momentum exposure is set by the manager at time $t-1$ based on a forecast about momentum conditions at time t.

Timing models that have been used in the hedge fund literature (e.g., Cao, Chen, Liang, and Lo (2013)) represent a time-varying factor exposure as a linear functions of a forecast about factor realizations. The linear functional form can be justified from a Taylor expansion by ignoring higher-order terms (e.g., Shanken (1990)). Following previous models, the specification used here is:

$$\beta_{i,t-1}^{mom} = \gamma_{0i} + \gamma_i E(mom_t^{cond} | I_{t-1}), \tag{2}$$

where I_{t-1} is the information set available to the fund manager at time $t-1$. The coefficient γ_i is the key to measuring momentum-timing skill, i.e., how momentum exposure varies with forecasts about future momentum conditions (mom^{cond}).

Various momentum-timing models could be constructed depending on the momentum conditions they concentrate on. Motivated by the literature on strategy crowding and momentum, I focus on variables that proxy for the size of the crowd of active investors who are likely to trade on momentum. Namely, I assume that fund mangers follow three timing signals to adjust their exposure to momentum. The signals include the crowd of arbitrageurs active in momentum, the crowd of mutual funds implementing positive-feedback trading, and the crowd of noise traders who are affected by sentiment. More specifically, Equation (2) is further represented as:

$$\beta_{i,t-1}^{mom} = \gamma_{0i} + \gamma_{1i}(SH_t - \overline{SH} + v_{1,t}) + \gamma_{2i}(PF_t - \overline{PF} + v_{2,t}) + \gamma_{3i}(BW_t - \overline{BW} + v_{3,t}), \quad (3)$$

where the expressions in parentheses represent the fund's timing signals associated with the momentum crowd. The three timing signals are SH, which stands for the crowd of arbitrageurs, represented by short sellers, PF, which measures the crowd of positive-feedback traders, and BW, which represents the aggregate measure of investor sentiment used in Baker and Wurgler (2006). The terms v_1, v_2, and v_3 are independent and zero mean variables that denote forecast noise realized at time t. They capture the notion that funds have imperfect timing signals. Following the timing literature (e.g., Ferson and Schadt (1996) and Busse (1999)), the three timing signals are demeaned. Accordingly, γ_{0i} captures the fund's average exposure to momentum.

By substituting Equation (3) in Equation (1) and incorporating the forecast noise within the error term, I obtain the following momentum-crowd-timing model:

$$R_{i,t} = \alpha_i + \gamma_{0i}R_t^{mom} + \gamma_{1i}R_t^{mom}(SH_t - \overline{SH}) + \gamma_{2i}R_t^{mom}(PF_t - \overline{PF}) + \gamma_{3i}R_t^{mom}(BW_t - \overline{BW}) + e_{i,t}. \quad (4)$$

A positive (negative) timing coefficient γ_{1i} indicates that the fund has a higher (lower) momentum exposure during times when the strategy is crowded by arbitrageurs as measured by extensive shorting. A positive (negative) timing coefficient γ_{2i} indicates that the fund has a higher (lower) momentum exposure during times when positive-feedback traders' demand is large. A positive (negative) timing coefficient γ_{3i} indicates that the fund has a higher (lower) momentum exposure during times when aggregate investor sentiment is high and more sentiment traders are present in the market.

Motivated by the literature on the interaction between arbitrageurs and other momentum traders, I examine the magnitude and significance of the coefficients γ_{1i}, γ_{2i}, and γ_{3i} for a large cross section of equity market neutral and long-short equity hedge funds. These are the funds that are most likely to implement momentum strategies. The hypotheses developed in Section 2 suggest that funds with lower γ_{1i} coefficients and higher γ_{2i} coefficients will be the funds that are aware of the positions of other active investors in momentum and what these positions imply about the future performance of momentum. Such funds are able to time their exposure to momentum in such a way as to capitalize on the actions of other active investors, i.e., avoid periods of overcrowded short positions and ride the wave of positive-feedback trading.

Funds that view the presence of sentiment traders as an impediment to arbitrage are expected to have lower γ_{3i} coefficients. The literature does not have a specific prediction about the relation between γ_{3i} and future performance. Funds who increase their exposure to momentum when investor sentiment is high could benefit from riding an optimistic wave when it comes to winners. On the other hand, funds who decrease their exposure to momentum when investor sentiment is high could be concerned about the overpricing of losers as a result of overly optimistic sentiment traders. It is an empirical question as to which timing ability is rewarded by higher future performance.

Taking into account the Fung and Hsieh (2004) hedge fund factors, the baseline momentum-crowd-timing model has the following specification:

$$R_{i,t} = \alpha_i + \gamma_{0i} R_t^{mom} + \gamma_{1i} R_t^{mom}(SH_t - \overline{SH}) + \gamma_{2i} R_t^{mom}(PF_t - \overline{PF}) + \gamma_{3i} R_t^{mom}(BW_t - \overline{BW}) + \sum_{j=1}^{J} \beta_j f_{j,t} + e_{i,t}, \tag{5}$$

where f stands for the Fung-Hsieh factors that are different from the momentum factor.

A key observation from the timing model in (5) is that the timing signals are measured over the same period as the realization of the momentum strategy return rather than over previous periods. This feature of the model is important since it allows us to measure hedge funds' ability to anticipate the size of the momentum crowd. This is different than measuring their ability to react to the size of the momentum crowd which would be measured by including lagged values of the timing signals in the model.[29]

[29]A robustness check shows that the ability of hedge funds to forecast the size of the momentum crowd remains significant after controlling for lagged values of the timing signals in equation (5).

4.2 Variable Construction

The model presented above contains factors and timing signals. Below I describe in detail the construction of the momentum factor and the signals that proxy for the size of the momentum crowd. In calculating the variables that proxy for the size of the momentum crowd, I follow the intuition that the measures of crowding should be computed at the stock-level and reflect the demand of various investors associated with the past performance of the stock.

The momentum factor is initially constructed as the return of past winners minus the return of past losers. The winner portfolio consists of stocks in the top decile according to the distribution of cumulative returns from month $t-12$ to $t-2$. The losers portfolio includes the stocks in the bottom decile of the same distribution.[30] The momentum factor is constructed at the beginning of month t and it is rebalanced every month. Month $t-1$ is skipped to avoid short-run reversal (Jegadeesh (1990)). The factor invests \$1 in the winner portfolio and shorts \$1 in the loser portfolio. This implies that there is a constant amount in the long and short leg, so that the volatility of the portfolio is time-varying and the portfolio runs constant leverage. However, it is common practice in the industry to target ex-ante portfolio volatility and let leverage vary over time. To account for this feature of momentum portfolio construction, I follow Barroso and Santa Clara (2015) and scale the long-short momentum portfolio by its realized volatility in the previous six months, targeting a momentum strategy with constant annual volatility of 12%.[31] The scaled version of the momentum strategy is used as the main momentum factor in model (5).[32] Using the scaled version of momentum could be viewed as a way of adjusting momentum exposure conditional on the realized variance of the momentum strategy. If regular momentum has been very volatile in the recent past, scaled momentum decreases momentum exposure, and vice versa.

To estimate the size of arbitrage capital allocated to momentum, I follow the method of Hanson and Sunderam (2013). Their insight is that the cross section of short interest reveals how intensely

[30] These returns are downloaded from Kenneth French's website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[31] For further details about the scaled version of the momentum strategy refer to Barroso and Santa Clara (2015).

[32] All results are robust to using the regular unscaled version of the momentum strategy.

arbitrageurs are using the momentum strategy at a given point in time. For example, when short interest is especially weighted toward past losers, more arbitrage capital is being devoted to momentum strategies. They formalize this idea in a regression setting in which the cross section of short interest every month is explained by the cross section of relative momentum rankings and other control variables,

$$SR_{i,t} = a_t + k_t^{mom} 1_{i,t}^{mom} + \delta_t' x_{i,t} + \epsilon_{i,t}. \tag{6}$$

The regression includes a set of momentum decile dummies, $1_{i,t}^{mom}$, in which the omitted dummy variable is for decile 5. In addition, the vector $x_{i,t}$ contains control variables that are correlated with short interest: a set of size and book-to-market decile dummies (excluding decile 5), institutional ownership, three-month turnover, trailing twelve-month return volatility, dummies for the exchange on which a stock trades, and a dummy that indicates whether a firm has convertible securities outstanding. The regression in (6) is estimated every month using the cross section of stock-level short interest ratios, SR, defined as the number of shares shorted divided by number of shares outstanding.

I use monthly data on short interest for NYSE and AMEX stocks obtained from Compustat. For NASDAQ stocks, short interest data is only available from Compustat beginning in July 2003, so I obtain data directly from the exchange prior to this date. Short interest for each stock in month t is the total number of uncovered shares sold short for transactions settling on or before the 15th of the month. Momentum deciles in month t are based on cumulative returns from months $t-12$ to $t-2$ and they are refreshed each month. For example, short interest observations for July are associated with momentum sorts performed over July (of the previous year) through May. The short interest data is supplemented with data on stock characteristics from CRSP and Compustat. All continuous variables are winsorized in each cross section at the 0.5 and 99.5 percentiles.

The coefficient on the dummy for the lowest momentum decile, $k_t^{mom(1)}$, reflects the increase in short interest at time t associated with being an extreme loser relative to the omitted decile 5 category. Thus, the time series of $k_t^{mom(1)}$ coefficients is a proxy for the quantity of short-side capital devoted to momentum at time t.[33] Since shorting capital is more likely to be associated with hedge fund trades, the time series of $k_t^{mom(1)}$ coefficients will capture the extent of hedge fund crowding in momentum. The first timing signal used in (5) is, therefore, set equal to the time series of arbitrage capital devoted to momentum

$$SH_t = \widehat{k_t^{mom(1)}}. \tag{7}$$

The second timing signal I use is derived from mutual fund holdings. I follow Lakonishok, Shleifer, and Vishny (1992) who argue that the extent of positive-feedback trading in a given security could be measured by the excess demand for this security associated with its past performance. Lakonishok, Shleifer, and Vishny (1992) calculate excess demand for a stock by the current quarter's net buying, aggregated across all money managers who hold that stock. More specifically, the excess demand for a given stock-quarter i is:

$$Dratio(i) = [\$buys(i) - \$sells(i)]/[\$buys(i) + \$sells(i)], \tag{8}$$

where $\$buys(i)$ is the total dollar increases by all money managers in the given stock-quarter (evaluated at the average price during the quarter) and $\$sells(i)$ is the total dollar decreases in holdings. Therefore, the excess demand for a given stock-quarter captures the difference between dollar buys and dollar sells scaled by total activity. Positive-feedback trading strategies are characterized by an excess of purchases over sales for past winners and an excess of sales over purchases for past losers.

I estimate $Dratio(i)$ using quarterly mutual fund holdings data from Thompson-Reuters. Most mutual funds in the database report their holdings on a quarterly basis, even though for a large part of the sample period they are only required to report semi-annually. I impose several filters on the sample of mutual funds, using data from the CRSP Survivor-Bias-Free Mutual Fund Database. Following previous literature, I focus on domestic equity mutual funds and exclude sector funds,

[33] Hanson and Sunderam (2013) point out that their proxy for the arbitrage capital devoted to momentum is the strength of the cross-sectional relationship between short interest and the momentum signal, not simply the total quantity of short interest in stocks in the lowest momentum decile.

index funds, and funds that, on average, have less than 80% of their holdings in stocks. I exclude fund observations before a fund passes the \$5 million threshold for assets under management (AUM). I require the date on which holdings information is recorded and the date on which a holdings report is filed to be no more than six months apart. This ensures that holdings data are relatively recent.

Since the main analysis in the paper uses monthly hedge fund returns to infer the dynamic timing abilities of hedge fund managers, I convert the $Dratio(i)$ into a monthly-frequency variable using the following method. I assume that any dollar increase or decrease in holdings for a given stock occurs uniformly throughout each quarter so that my estimate of the dollar increase or decrease every month within a quarter is based on linear interpolation between two adjacent quarters.[34] Therefore, the numerator in Equation (8) is calculated on a monthly basis and every month within the quarter it is scaled by total activity that quarter.

Following the approach of Lakonishok, Shleifer, and Vishny (1992), I relate the cross section of excess demand at time t to the cross section of relative momentum rankings. Instead of a simple sorting I use a multiple regression framework in order to control for other variables:

$$Dratio_{i,t} = c_t + \lambda_t^{mom} mom_{i,t} + \theta_t' x_{i,t} + \eta_{i,t}, \tag{9}$$

where $mom_{i,t}$ stands for the relative momentum rank of stock i in month t, based on cumulative performance over months $t-12$ to $t-2$,[35] and the vector $x_{i,t}$ contains several control variables. The control variables include firm-level characteristics that might influence money managers' trading decisions: size, book-to-market, turnover, idiosyncratic volatility, share issuance, Amihud illiquidity, industry membership, and the change in holdings induced by fund flows in the current quarter (e.g., Lou (2012)). The key to model (9) is that $Dratio_{i,t}$ and $mom_{i,t}$ are non-overlapping, which means that the coefficient λ reflects how mutual fund managers react to stocks' past one-year returns in

[34] Harris, Hartzmark, and Solomon (2015) use a similar linear interpolation to derive mutual fund holdings within a quarter.

[35] There are several positive feedback trading strategies, depending on the look-back horizon for calculating returns. I focus on past one-year returns, which are typically used in a momentum strategy.

making their portfolio decisions.

The coefficient in front of past performance, λ_t^{mom}, measures the extent to which positive-feedback trading is present in the data for mutual fund holdings. Therefore, the time series of λ_t^{mom} coefficients is a proxy for the intensity of positive-feedback trading related to past performance.[36] The second timing signal used in (5) is, therefore, set equal to the time series of positive-feedback trading intensity

$$PF_t = \widehat{\lambda_t^{mom}}. \tag{10}$$

Intuitively, PF measures the sensitivity of trading decisions to past stock returns. A larger (smaller) PF corresponds to a greater (smaller) tendency of positive-feedback trading.

Finally, the third timing signal used in (5) is simply the value of the aggregate sentiment index of Baker and Wurgler (2006), BW_t. Baker and Wurgler (2006) construct investor sentiment using six variables: closed-end fund premium, NYSE share turnover, the number and average first-day returns of IPOs, the equity share in new issues, and the dividend premium. The sentiment index is the first principal component of these variables that are first orthogonalized against several macroeconomic variables to reduce the potential impact of the business cycle.[37] The monthly data for the investor sentiment index comes from Jeffrey Wurgler's website.[38]

Previous studies have examined alternative measures for the crowd of investors who follow momentum. For example, Lou and Polk (2014) develop a measure which is based on the strength of the correlations among the stocks in the winner or loser portfolios. Pojarliev and Levich (2011) define crowdedness as the percentage of funds with significant positive exposure to a given style less the percentage of funds with significant negative exposure to the same style. Since these measure

[36] Lakonishok, Shleifer, and Vishny (1992) use an alternative measure of positive-feedback trading defined as $Nratio(i) = \#buys(i)/\#active(i)$, where $\#buys(i)$ is the number of managers increasing their holding of the stock in quarter i and $\#active(i)$ is the number of managers changing their holdings. I also use $Nratio(i)$ in place of $Dratio(i)$ and find similar results.

[37] Chen, Han, and Pan (2016) use the sentiment index of Baker and Wurgler (2006) to examine whether exposure to sentiment risk helps explain the cross-sectional variation in hedge fund returns. This paper is different since I use the sentiment index as a timing variable for the momentum strategy rather than a stand-alone factor for hedge fund returns. The results in the paper are robust to controlling for the sentiment risk factor in the main model for hedge fund returns.

[38] http://people.stern.nyu.edu/jwurgler/

are based on returns and not holdings, they do not necessarily distinguish among the various types of traders who follow momentum signals. For the analysis in this paper it is important to distinguish between the different types of momentum traders since their crowding into the strategy has different implications for momentum profitability.

For the rest of the paper, I refer to the three variables SH, PF, and BW as crowd-timing signals. The crowd refers specifically to investors who trade on momentum since the signals are interacted with the returns of the momentum strategy. The SH signal is designed to measure the hedge fund crowd of investors who are actively trading on momentum, PF proxies for the crowd of positive-feedback traders among active mutual funds, and BW proxies for the crowd of noise traders who are likely to be more active during times of high sentiment. All three variables capture the idea that the momentum strategy becomes crowded due to active involvement by various groups of investors and this crowding has implications for future strategy performance.

4.3 Predictive Ability of the Timing Signals

Figure 2 presents plots of all three timing signals for the period from January 1997 to December 2017. The plot of arbitrage capital in momentum in Panel A of Figure 2 (SH) displays an upward trend over the sample period.[39] Therefore, in all subsequent calculations, this variable has been detrended. The plot of the intensity of positive-feedback trading in Panel B of Figure 2 (PF) does not reveal any significant trends over time. Finally, Panel C of Figure 2 shows that investor sentiment (BW) was at an all-time high in February of 2001 which is during the Technology bubble period.

Panel A of Table 2 reports summary statistics for the momentum factor and its scaled counterpart, as well as the three timing signals. The results show that the scaled version of momentum has a lower average monthly return, but its monthly volatility is also much lower than regular momentum. The range of values between the 25th and 75th percentile of the distribution of scaled momentum is tighter than that for regular momentum.

[39] Similar results are reported in Hanson and Sunderam (2013).

Panel B of Table 2 presents the pairwise correlations among the three momentum-crowd-timing signals. The correlation between SH and PF is positive at 5.8%, while the correlation between SH and BW is negative at -28.48%. The correlation between PF and BW is positive at 6.3%. The correlations are relatively small which indicates that the three measures represent different components of the crowd of investors who follow momentum signals.

Before analyzing hedge funds' ability to time the momentum crowd, I examine the predictive power of the level of crowding in momentum for future momentum returns. Based on the notion that capital is slow moving[40] and, therefore, crowding should have a longer lasting predictive power for strategy returns, I examine one-year ahead momentum returns. Panel C of Table 2 presents a forecasting time-series regression of the average return of the scaled momentum strategy over the period from t to $t+12$ on the three timing signals measured at time t. Due to overlapping returns, the t-statistics are adjusted using Newey-West (1987) standard errors. The results show that the return to the momentum strategy from t to $t+12$ is negatively correlated with the size of short arbitrage capital at time t and positively correlated with the intensity of positive feedback trading at time t. The level of investor sentiment does not have a significant effect on future momentum returns in the presence of the other two signals.[41]

Panel C of Table 2 also reports a time-series forecasting regression of the Sharpe Ratio of the scaled momentum strategy from month t to month $t+12$ on the three timing signals measured at time t. The Sharpe Ratio is calculated as the average return from month t to month $t+12$ divided by the standard deviation of return from month t to month $t+12$. The results show that the Sharpe Ratio of momentum over the period from t to $t+12$ is negatively correlated with the size of short arbitrage capital at time t and positively correlated with the intensity of positive-feedback trading at time t. Finally, Panel C of Table 2 reports a time-series forecasting regression of the skewness of the scaled momentum strategy from month t to month $t+12$ on the three timing signals measured

[40] See, for example, Mitchell, Pedersen, and Pulvino (2007) and Duffie (2010).

[41] Stambaugh, Yu, and Yuan (2011) show that momentum returns are higher following high levels of investor sentiment. The results in this paper differ since I also control for the size of arbitrage capital deployed in momentum and the intensity of positive-feedback trading by mutual funds.

at time t. The results from the regression reveal that higher levels of investor sentiment tend to predict negative momentum skewness over the following year. High intensity of positive-feedback trading tends to predict positive momentum skewness in the future. The amount of arbitrage capital in momentum is not significantly related to the future skewness of the strategy. Overall, Table 2 shows that the three timing signals that are motivated from the theoretical literature on arbitrage predict future momentum profits, Sharpe Ratio, or skewness. These results suggest that it is worthwhile for hedge fund managers to monitor these signals since they contain information about future strategy performance.

5 Empirical Results on Momentum Crowd Timing Ability

In this section, I first present the cross-sectional distribution of t-statistics for the crowd-timing coefficients across individual funds. Then, I use a bootstrap analysis to examine the statistical significance of crowd-timing ability. Next, I show that the direction of crowd-timing skill is associated with economically significant returns out-of-sample and it is persistent over time.

5.1 Estimation of Momentum Crowd Timing Coefficients and Distribution of t-statistics

I begin by estimating model (5) for each fund in the sample, using all available time series data for the fund. Only funds with at least 36 monthly observations are included in the analysis. From now on, the scaled momentum strategy return is used in all empirical tests. The regression parameter γ_{0i} measures funds' average exposure to momentum, while the parameters γ_{1i}, γ_{2i}, and γ_{3i} measure the crowd-timing ability of the funds. It is important to first gauge the significance of the γ-coefficients in order to examine whether realized hedge fund returns are consistent with the conjecture that funds pay attention to the actions of active investors who trade on momentum.

Table 3 summarizes the cross-sectional distribution of t-statistics for the γ-coefficients across hedge funds. The first two rows of the table provide a general idea about the sign of the γ-coefficients. The majority of hedge funds (65.38%) have a positive exposure to momentum on average. Approximately half of the funds (50.65%) increase their exposure to momentum when shorting activity in momentum is higher than average, while the other half (49.35%) decrease it. When positive-feedback trading in momentum is higher than average, slightly more than half of the funds (53.46%) decrease their exposure to momentum. When investor sentiment is higher than average, more than half of the funds (58.35%) tend to increase their exposure to momentum.

The other rows of Table 3 specify the percentage of t-statistics that exceed a specified critical value under the assumption of normality. For example, 15.77% of the funds have γ_{0i} t-statistics greater than 1.96 (2.5% significance level in the right tail), whereas only 4.79% have γ_{0i} t-statistics smaller than 1.96 (2.5% significance level in the left tail). Therefore, the right tail of the γ_{0i} distribution appears thicker than the left tail, which suggests that a larger percentage of funds have positive exposure to momentum on average.

In addition, 5.58% of the funds have γ_{1i} t-statistics greater than 1.96, while 5.07% have γ_{1i} t-statistics smaller than 1.96. Therefore, a large proportion of funds show significant timing relative to the crowd of other arbitrageurs in momentum, with the sign of the timing coefficient being equally balanced between positive and negative. The t-statistics associated with the γ_{2i} coefficients show that 5.49% of funds have t-statistics greater than 1.96, while 8.53% have t-statistics smaller than 1.96. This implies that the left tail of the γ_{2i} distribution is thicker and more funds tend to decrease their exposure to momentum when positive-feedback trading is intensified. Finally, 10.24% of the funds have γ_{3i} t-statistics greater than 1.96, while 5.44% have γ_{3i} t-statistics smaller than 1.96. Therefore, a larger percentage of funds tend to increase their exposure to momentum when investor sentiment is high.

Overall, the cross-sectional distribution of t-statistics in Table 3 shows that a substantial fraction of long-short equity and market neutral hedge funds are exposed to momentum and that exposure varies significantly with the size of the momentum crowd. The pattern of the t-statistics suggests that the number of funds that follow the crowd of arbitrageurs in momentum and the number that go against it are very close to each other. Furthermore, more funds tend to bet against positive-feedback traders than jump on the bandwagon with them. Finally, a bigger fraction of funds tends to have a positive exposure to momentum when the crowd of sentiment traders is large.

The inference above is drawn based on a normality assumption. To allow for the possibility that hedge fund returns are not normal and the random chance that some funds will appear to have significant t-statistics even if their true timing coefficients are zero, I use a bootstrap analysis to examine momentum-crowd-timing ability. The bootstrap analysis does not rely on a normality assumption and helps determine whether the significant crowd timing documented in Table 3 occurs by chance.

5.2 Bootstrap Analysis of Momentum Crowd Timing Ability

In this section I describe the bootstrap procedure used to examine the statistical significance of momentum-crowd-timing coefficients for individual hedge funds. The bootstrap procedure used here is similar to that of Kosowski, Timmermann, White, and Wermers (2006), Kosowski, Naik, and Teo (2007), Jiang, Yao, and Yu (2007), and Fama and French (2010), all building on Efron (1979). The basic idea of the bootstrap analysis is to simulate hypothetical fund returns that are constructed in a way to have the same factor loadings as the actual fund returns but no momentum-crowd-timing ability. If model (5) is estimated over hypothetical fund returns, the comparison between the bootstrapped distribution of t-statistics associated with momentum-crowd-timing signals and the actual distribution of t-statistics from Table 3 would reveal whether the momentum-crowd-timing ability documented previously is purely due to statistical chance. More specifically, the bootstrap procedure consists of the following steps:

1. Estimate the momentum-crowd-timing model in (5) for fund i and store the estimated coefficients as well as the time series of residuals.

2. Resample the residuals with replacement and obtain a randomly resampled residual time series. Then generate monthly excess returns for a pseudo fund that has no momentum-crowd-timing skill since the coefficients on the momentum-crowd-timing terms have been set to zero.

3. Estimate the momentum-crowd-timing model in (5) using the pseudo-fund returns from Step 2,

and store the estimates of the timing coefficients and their t-statistic. Since the pseudo fund has true timing coefficients of zero by construction, any non-zero timing coefficients (and t-statistic) come from sampling variation.

4. Complete steps 1-3 across all funds in the sample, so that the whole cross section of crowd-timing coefficients and their t-statistics can be derived.

5. Repeat steps 1-4 for 10,000 iterations to generate the empirical distributions of various cross-sectional statistics for the pseudo fund returns, e.g., t-statistics.

For a given cross-sectional statistic, I calculate its empirical p-value as the frequency that the values of the bootstrapped cross-sectional statistic for the pseudo funds from 10,000 simulations exceed the actual value of the cross-sectional statistic. I conduct this analysis mainly for the t-statistics of the crowd-timing coefficients, because t-statistics are pivotal statistics and, therefore, they have favorable sampling properties in a bootstrap analysis (Horowitz (2001)).

Table 4 reports the t-statistics for both bottom and top crowd-timing coefficient percentiles (1%, 2.5%, 5%, and 10%), as well as empirical p-values from the bootstrap analysis. Panel A corresponds to t-statistics associated with timing the crowd of arbitrageurs, Panel B reports t-statistics for timing the crowd of positive-feedback traders, and Panel C corresponds to t-statistics related to timing sentiment traders. The results reveal that for all extreme percentiles reported in the table, the significance of the crowd-timing coefficients documented in Table 3 is unlikely to be driven by random chance. Overall, the evidence from the bootstrap analysis in Table 4 suggests that realized hedge fund returns are consistent with managers timing the momentum crowd. Even in cases in which the timing goes in the opposite direction of what is predicted to be optimal, the significance of the crowd-timing coefficients cannot be attributed purely to random chance. To further explore whether crowd timing truly reflects skill, I examine the economic significance of momentum-crowd timing.

5.3 Economic Significance of Momentum Crowd Timing: Portfolio Sorts

The previous results show that a large proportion of hedge funds engage in timing the momentum crowd. In this section, I analyze the economic significance of this timing. In particular, I examine whether crowd-timing ability is associated with superior future performance. More importantly, the sign of the timing coefficients in model (5) is key in identifying funds that get the timing of the momentum crowd right. For example, the theoretical literature on crowding predicts that it may be optimal for arbitrageurs to scale away from momentum at times when the strategy is crowded by other arbitrageurs rather than follow the crowd.

Each month starting from January 1997, I estimate the crowd-timing coefficients for each fund using model (5) and the past 36-month estimation period. Then I form five quintile portfolios based on each crowd-timing coefficient. The portfolios are equally-weighted and held for a month after portfolio formation. This process is repeated until the end of the sample in December 2017. Based on the resulting time series of portfolio returns, I compute the average return of each portfolio and its alpha from the 7-factor model of Fung and Hsieh (1997).

Table 5 presents results about the economic value of crowd-timing ability over the period from January 1997 to December 2017.[42] Specifically, the bottom quintile of funds with respect to the γ_1 coefficient has an average monthly return of 1.15%, while the top quintile has an average monthly return of 0.49%. The average γ_1 coefficient if quintile 1 is negative, while that of quintile 5 is positive. Therefore, funds that decrease their exposure to momentum the most at times when other arbitrageurs are crowding into momentum, earn 0.66% more per month (7.92% per year) than funds that increase their exposure to momentum. The difference in average returns is economically and statistically significant.

In the case of timing the positive-feedback-trading crowd, the bottom quintile of funds with respect to γ_2 has an average monthly return of 0.59%, while the top quintile has an average monthly return of 1.04%. The difference is 0.45% per month (5.40% per year) and statistically significant.

[42] The sample period is from January 1997 to December 2017 since the observations of the first three years from January 1994 to December 1996 of each fund are used for the estimation of the timing coefficients in January 1997.

The average γ_2 coefficient if quintile 1 is negative, while that of quintile 5 is positive. Therefore, funds that increase their exposure to momentum the most at times when positive-feedback traders are active outperform funds that decrease their exposure in such times.

In the case of sentiment timing, the bottom quintile of funds with respect to γ_3 has an average monthly return of 1.13%, while the top quintile has an average monthly return of 0.54%. The difference between the two, 0.59% per month (7.08% per year), is economically and statistically significant. Therefore, funds that decrease their exposure to momentum the most at times when investor sentiment is high outperform funds that increase their exposure in such times.

The pattern in the Fung-Hsieh alphas for all portfolios sorted by crowd-timing coefficients is very similar to the pattern in raw average returns. Overall, the results in Table 5 are in line with arguments from various theoretical papers on arbitrage. Monitoring the size of the crowd of investors who are involved in momentum pays off in terms of higher future performance. The results do not necessarily imply that the better future performance is entirely due to the momentum exposure of hedge funds. It could still be the case that funds who are capable of timing the momentum crowd possess other skills that ultimately drive their superior performance. It is reasonable to assume that funds that are good at timing the momentum crowd might be good at timing other market conditions as well.

If the performance gap between quintiles 1 and 5 in each momentum-crowd-timing coefficient group is related to managerial skill, we would expect a certain degree of persistence after portfolio formation. To test this conjecture, I examine the persistence of the crowd-timing ability of hedge funds by tracking quintile portfolio returns for several months after portfolio formation (without rebalancing). More specifically, after forming five quintile portfolios based on crowd-timing coefficients, I estimate model (5) for post-formation returns and record fund managers' crowd-timing coefficients after portfolio formation. Untabulated results show that there is persistence in hedge funds' crowd-timing coefficients: for each γ coefficient, funds that were grouped into quintile 1 (quintile 5) based on ranking-period crowd-timing ability continued to be classified into quintile

1 (quintile 5) in terms of crowd-timing skill after portfolio formation. This effect persists for 12 months after portfolio formation.

The portfolio statistics reported in Table 5 correspond to a one month holding period after portfolio formation. Next, I extend the holding period to 3, 6, 9, or 12 months after portfolio formation. This examines the robustness of the previous results with respect to the rebalancing frequency of the portfolios, and also reveals whether the value of timing the momentum crowd extends to more distant periods after portfolio formation. Table 6 reports the average monthly returns for 3-, 6-, 9-, and 12-month holding periods for portfolios sorted by crowd-timing coefficients. In the case of portfolios sorted by γ_1, the results show that the quintile of portfolios with the most negative γ_1 coefficients continues to have higher average performance than the quintile of portfolios with the most positive γ_1 coefficients. The difference in performance is statistically significant up to 6 months after portfolio formation. The economic significance of the difference in performance continues to be large, ranging from 0.46% per month (5.04% per year) for a 3-month holding period to 0.29% per month (2.28% per year) for a 12-month holding period.

For portfolios sorted by γ_2, the quintile of portfolios with the most negative γ_2 coefficients continues to have lower average performance than the quintile of portfolios with the most positive γ_2 coefficients. The difference in performance is statistically significant up to 6 months after portfolio formation and economically significant up to a 12-month holding period (ranging from 0.55% per month (6.60% per year) to 0.47% per month (5.64% per year)). Finally, for portfolios sorted by γ_3, the quintile of portfolios with the most negative γ_3 coefficients continues to have higher average performance than the quintile of portfolios with the most positive γ_3 coefficients. The difference in performance is not statistically significant, but its economic significance continues to be large (ranging from 0.49% per month (5.88% per year) to 0.45% per month (5.40% per year)).

In summary, Table 6 shows that the results in Table 5 are robust to increasing the length of the holding period for portfolios sorted by timing ability. Furthermore, the value of crowd-timing continues to be relevant for several periods after portfolio formation. This high degree of persistence lends support to the hypothesis that the relation between strategic momentum-crowd-timing and future performance is related to managerial skill.

The pattern of the returns in Table 6 also reveals that timing the momentum crowd tends to pay off in statistically significant terms at relatively shorter frequencies (3 to 6 months). By the time 12 months have passed since portfolio formation, the statistical significance of superior timing ability tends to diminish. This is consistent with momentum being a medium-term strategy that requires frequent evaluation of the stocks that are classified as winners and losers.

In Figure 3, I report additional information about the economic significance of crowd-timing ability. The figure tracks the cumulative performance of quintiles 1 and 5 for each crowd-timing coefficient. The results in Panel A show that \$1 invested in funds that decrease (increase) their momentum exposure when the crowd of arbitrageurs is large grows to \$10.58 (\$3.28) at the end of the sample period. Further, \$1 invested in funds that increase (decrease) their momentum exposure at times of intensive positive-feedback trading grows to \$6.81 (\$3.78) at the end of the sample period. Finally, \$1 invested in funds that decrease (increase) their momentum exposure when investor sentiment is high grows to \$9.16 (\$3.54) at the end of the sample period. The difference in cumulative returns between the top and bottom group of funds in each case is very substantial. All funds experience their largest drawdown during the recent Great Recession of 2008-2009. Overall, Figure 3 shows that higher cumulative returns from investing in hedge funds are associated with decreasing exposure to momentum when other arbitrageurs are crowded into the strategy, increasing exposure to momentum when positive-feedback traders are active, and decreasing momentum exposure when the crowd of sentiment traders is large.

The results reported in Tables 5 and 6 examine one aspect of crowd-timing ability at a time. However, it is possible that hedge funds follow multiple signals about investors who are likely to be trading on momentum. Therefore, in Table 7, I report the average returns of hedge fund portfolios that have been double-sorted on crowd-timing coefficients. The main goal of the table is to examine the economic value of following multiple timing signals. In Panel A of Table 7, returns are sorted into three groups based on γ_1 and three groups based on γ_2. This results in 9 portfolios. The panel shows that funds that decrease their exposure to momentum when other arbitrageurs are crowding into the strategy and increase their exposure to momentum when positive-feedback traders are active generate the highest average return and alpha among the 9 portfolios. In Panel B, returns are sorted into three groups based on γ_1 and three groups based on γ_3. Funds that decrease their exposure to momentum when other arbitrageurs are crowding into the strategy and decrease their exposure to momentum when investor sentiment is high generate the highest average return and alpha among these 9 portfolios. In Panel C, returns are sorted into three groups based on γ_2 and three groups based on γ_3. Funds that increase their exposure to momentum when positive-feedback traders are active and decrease their exposure to momentum when investor sentiment is high generate the highest average return and alpha among these 9 portfolios. In summary, the results in Table 7 show that a combination of momentum-crowd-timing skills results in superior future performance relative to following just one timing signal. Therefore, the information contained in each timing signal is different and contributes separately to future performance.

5.4 Economic Significance of Momentum Crowd Timing: Fama-MacBeth Regressions

The results so far indicate that the crowd-timing ability of hedge funds is associated with higher future returns. In addition, the direction of the timing coefficients is also important since it indicates whether hedge funds trade with or against the momentum crowd. In general, it pays off to trade against the hedge fund crowd, with the positive-feedback crowd, and against the sentiment traders crowd.

Next, I control for fund characteristics that are correlated with returns, in order to analyze whether crowd-timing ability provides independent information for future performance. In particular, I use Fama-MacBeth (1973) regressions of fund excess returns on crowd-timing coefficients together with fund characteristics. To avoid look-ahead bias, excess returns at time t are matched with timing coefficients estimated from time $t-36$ to $t-1$ and characteristics measured at time $t-1$. The fund characteristics include lagged fund return, management fee, incentive fee, redemption notice period, lockup period, a dummy for whether the fund is leveraged, minimum investment, high water mark, fund age, and fund size as measured by AUM. When used in the Fama-MacBeth regressions, the size of each fund has been orthogonalized to its age to control for the high correlation between the two variables.

Table 8 reports results from the Fama-MacBeth regressions. Model 1 includes the timing coefficients as the only independent variables in the regressions. The results show that a lower γ_1 coefficient is associated with significantly higher average returns in the future. Additionally, a higher γ_2 coefficient has a significantly positive effect on future returns. The γ_3 coefficient does not have a significant impact on returns in the presence of the other two. Model 2 includes various fund characteristics relative to Model 1. The results show that the explanatory power of γ_1 and γ_2 remains significant in the presence of the fund characteristics. Other variables that stand out as being significant determinants of average fund returns are redemption notice period, high water mark, and fund age. Overall, momentum-crowd-timing ability contains independent information about future fund performance in the sample.

5.5 Correlation between Momentum Component of Hedge Fund Returns and Momentum Strategy

The previous results show that funds that strategically adjust their exposure to momentum in anticipation of the momentum crowd earn higher future returns. In this section I compute the correlation between the momentum component of hedge fund returns and future returns of the momentum strategy. If funds time their exposure to momentum in ways that try to anticipate future momentum performance, then the momentum-related part of their portfolios should reflect that. The analysis is performed for quintiles 1 and 5 within each crowd-timing coefficient group. In particular, for the quintile of funds with the lowest γ_1 coefficient, I run the following regression:

$$Q1_t - RF_t = \alpha + \gamma_0 R_t^{mom} + \gamma_1 R_t^{mom}(SH_t - \overline{SH}) + \gamma_2 R_t^{mom}(PF_t - \overline{PF}) + \gamma_3 R_t^{mom}(BW_t - \overline{BW}) + \sum_{j=1}^{J} \beta_j f_{j,t} + e1_t. \quad (11)$$

The momentum-related component of the return of quintile 1 is defined as:

$$\widehat{q1}_t = \widehat{\gamma}_0 R_t^{mom} + \widehat{\gamma}_1 R_t^{mom}(SH_t - \overline{SH}) + \widehat{\gamma}_2 R_t^{mom}(PF_t - \overline{PF}) + \widehat{\gamma}_3 R_t^{mom}(BW_t - \overline{BW}). \tag{12}$$

Next, I compute correlations of the form:

$$Corr(R_t^{mom}, \widehat{q1}_{t-k}), for\ t = 0, ..., 12. \tag{13}$$

I repeat the analysis for the quintile of funds with the highest γ_1 coefficient. The same calculations are performed for the extreme quintiles of funds sorted by γ_2 and γ_3.

Results are reported in Table 9. In the case of timing the crowd of arbitrageurs (γ_1), the momentum part of the return of quintile 1 is positively correlated with the momentum strategy, up to 7 months in the future. After that, the correlation becomes negative. The momentum component of the return of quintile 5 is negatively correlated with the momentum strategy, up to 12 months in the future. Thus, funds that are identified by their ability to decrease their momentum exposure when the hedge fund crowd is predicted to be large set up their portfolios in a way that covaries positively with the momentum strategy, i.e., the momentum component of their portfolio is profitable on average. The opposite conclusion holds for funds that increase their exposure to momentum when hedge funds are crowding into the strategy.

In the case of positive-feedback crowd-timing ability (γ_2), the momentum part of the return of quintile 1 is negatively correlated with the momentum strategy, up to 12 months in the future. The momentum component of the return of quintile 5 is positively correlated with the momentum strategy, up to 7 months in the future. After that, the correlation becomes negative. Therefore, funds that increase their exposure to momentum when positive-feedback trading is intensified have more profitable momentum portfolios than funds that decrease their exposure to momentum at such times. Finally, in the case of timing sentiment traders (γ_3), the momentum part of the return of quintile 1 is positively correlated with the momentum strategy, up to 10 months in the future. After that, the correlation becomes negative. The momentum component of the return of quintile 5 is negatively correlated with the momentum strategy, up to 12 months in the future.

6 Which Hedge Funds Anticipate the Momentum Crowd?

In this section, I use fund characteristics to examine which hedge funds are more likely to time momentum in anticipation of the size of the momentum crowd.

Hedge funds use leverage and, therefore, they are exposed to the risk of sudden margin calls that can lead to forced liquidations (e.g., Lo (2008) and Ang, Gorovyy, and van Inwegen (2011)). Hedge funds will be especially vulnerable to this risk during periods of crowding when liquidations can occur at the same time across multiple funds. Therefore, it is reasonable to assume that more highly levered hedge funds will be more prone to scale down their exposure to momentum when other arbitrageurs are crowded into the strategy. In addition, to the extent that timing the momentum crowd involves more effort and frequent adjustments to the strategy of the fund, we would expect that the ability of hedge funds to time the momentum crowd would be decreasing in the size of the funds' assets under management. Finally, it is reasonable to assume that funds with a higher degree of managerial discretion, approximated by a longer payout period, will be more actively involved in timing the momentum crowd.

I collect data on three fund characteristics from the TASS Lipper database: assets under management (AUM), dummy variable for the usage of leverage, and payout period. For each crowd-timing coefficient from model (5), I calculate the median fund characteristic within each portfolio quintile. The results are reported in Table 10. When funds are sorted by their ability to time the crowd of arbitrageurs, the results show that funds in the extreme quintiles (1 and 5) tend to have smaller AUM. The difference in size relative to quintiles 2, 3, and 4 is significant for both quintiles 1 and 5. Therefore, smaller hedge funds tend to have both more positive and more negative exposure to momentum when the strategy is crowded by arbitrageurs. Funds that both decrease and increase their exposure to momentum the most when it is crowded by hedge funds tend to use leverage more than the other funds, although the difference relative to all other groups is not statistically significant. Funds with the lowest γ_1 coefficients also have the longest payout period, which is statistically different than the payout period of funds with the highest γ_1

coefficients.

When funds are sorted by their ability to time the crowd of positive-feedback traders, the results show that hedge funds that front run mutual fund demand for momentum stocks tend to be smaller, use leverage more, and have longer payout periods. Their characteristics are statistically different from those of the other funds sorted by the γ_2 timing coefficient. Hedge funds that trade against the crowd of positive-feedback traders also tend to be smaller and have relatively longer payout periods, but they tend to use leverage less.

Finally, funds that decrease their exposure to momentum the most at times when sentiment traders are present in the market are smaller than other funds and this effect is statistically significant. However, they not differ from other funds sorted by the γ_3 timing coefficient in terms of the usage of leverage or the length of their payout period.

7 Does Momentum-Crowd-Timing Behavior Reveal Other Hedge Fund Skills?

The predictive power of momentum-crowd-timing for hedge fund performance suggests that this timing behavior could reveal other skills that are not directly observable. In this section I present further analysis of the relation between momentum-crowd-timing and managerial skill. The analysis contains two steps.

First, I test whether hedge funds that have strategic crowd-timing abilities are more likely to pursue unique investment strategies, thus revealing that they are generally more skilled. Second, I test whether the performance gap between top and bottom quintile funds in each crowd-timing category widens in times of greater investment opportunities in the hedge fund industry, when managerial ability is more valuable.

7.1 Strategy Distinctiveness

If managers are skillful, they are likely to engage in innovative and unique trading strategies, thereby delivering performance that co-moves less with the overall performance of the hedge fund sector, or with the performance of the specific style to which their fund belongs. Previous studies (e.g., Sun, Wang, and Zheng (2012)) define the distinctiveness of a hedge fund strategy as 1 minus the sample correlation of a fund's return with the return of the style the fund belongs to:

$$SDI_i = 1 - Corr(R_{i,t}, R_{style,t}), \tag{14}$$

where SDI stands for strategy distinctiveness index, $R_{i,t}$ is the return of the hedge fund, and $R_{style,t}$ is the return of a style index to which it belongs. The SDI of a fund can be viewed as a measure of how far the fund is from its style cluster. The higher SDI, the more distinctive the fund's strategy is. I further define upside SDI as 1 minus the correlation of a fund's return with the return of the style the fund belongs to conditional on the style return being positive. The idea behind this measure is to capture the tendency of funds to follow unique strategies in states when the overall style is performing well. Alternatively, I define downside SDI as 1 minus the correlation of a fund's return with the return of the style index conditional on the style return being negative. This captures the fund's tendency to pursue distinct strategies at times when the overall style is underperforming.

I compute the out-of-sample SDI measures of hedge funds sorted by momentum-crowd-timing ability. The SDI measures are calculated relative to the returns of three Credit Suisse hedge fund indices: the Long/Short Equity Index, the Equity Market Neutral Index, and the Overall Equity Hedge Fund Index. Monthly return data for the indices comes from the TASS Lipper Hedge Fund Index Database. Table 11 reports the results. Panel A reports the strategy distinctiveness index of hedge funds with respect to the Credit Suisse Long/Short Equity Hedge Fund Index. Funds that decrease their exposure to momentum when short sellers are crowded in the strategy (Q1 of category γ_1) have higher SDI, upside SDI, and downside SDI than funds that increase their exposure to momentum when short sellers are crowded in the strategy (Q5 of category γ_1). This suggests that hedge funds identified for their ability to strategically time the crowd of short sellers in momentum also tend to reveal skill in terms of pursuing distinct strategies from their style cohort. Panel A also shows that funds identified for their ability to front run positive-feedback traders in momentum (Q5 of category γ_2) implement strategies that are more distinct than their style cohort relative to funds that trade against positive-feedback traders (Q1 of category γ_2). In addition, funds that scale back their exposure to momentum when sentiment traders are active (Q1 of category γ_3) have higher SDI measures than funds that increase their momentum exposure in the presence of many sentiment traders (Q5 of category γ_3).

Panel B of Table 11 reports the strategy distinctiveness index of hedge funds with respect to the Credit Suisse Equity Market Neutral Hedge Fund Index. The SDI measures of all funds in Panel B tend to be larger than the ones reported in Panel A. This indicates that the funds in quintiles 1 and 5 of momentum-crowd-timing ability tend to be less correlated with the equity market neutral style. The results still show that the funds with more distinctive overall investment strategies tend to be funds that scale back momentum exposure in the presence of a large crowd of short sellers, increase momentum exposure to front-run the crowd of positive-feedback traders, and decrease momentum exposure when sentiment traders are active.

Finally, Panel C of Table 11 examines the distinctiveness of hedge fund strategies relative to all hedge funds that belong to the equity category. The results in Panel C are similar to the ones reported above. Overall, the SDI analysis reveals that funds that strategically adjust their exposure to momentum in anticipation of the momentum crowd pursue more distinct strategies in their overall investment portfolio.

7.2 Taking Advantage of Investment Opportunities

To the extent that the return difference between quintiles 1 and 5 within each momentum-crowd-timing coefficient group is driven by differences in skill, it should increase in times of greater investment opportunities in the hedge fund industry. This is the case since managerial skill is expected to be more valuable when more investment opportunities are available to the hedge fund manager. The presence of cross-sectional return dispersion is a well-known indicator of greater investment opportunities in the hedge fund industry. Active hedge fund trades are likely to generate returns that differ from the market return when individual stock returns are more dispersed.[43] As dispersion increases, so does the potential for a skilled manager to outperform, due to the payoff from identifying future winners/losers.

There are other arguments that suggest that managers with superior ability could have an advantage in states of higher dispersion. Several studies find a positive relation between return dispersion and future volatility at both the market (e.g., Bekaert and Harvey (1997), Stivers (2003)) and firm (e.g., Connolly and Stivers (2006)) levels. Bessembinder, Chan, and Seguin (1996) find a positive relation between dispersion and trading volume for individual stocks. Loungani, Rush, and Tave (1990) and Brainard and Cutler (1993) show that high dispersion is a significant predictor of unemployment caused by sectoral shifts when many company revaluations occur. Therefore, high dispersion periods represent a natural setting to examine managerial skill.

Table 12 reports the out-of-sample performance gap between quintiles 5 and 1 within each momentum-crowd-timing coefficient category, conditional on high and low cross-sectional dispersion states. A given month is classified as a high cross-sectional dispersion month if current return dispersion is above its median up to that month. Similarly, a month is classified as a low cross-sectional dispersion month if current return dispersion is below its median up to that month. The results show that in high cross-sectional dispersion states, funds that decrease their exposure to momentum the most at times when other arbitrageurs are crowding into momentum earn 0.80% more per month than funds that increase their exposure to momentum. In low cross-sectional dispersion states the average return difference is 0.58% more per month. The difference between the two states is statistically significant.

In addition, in high cross-sectional dispersion states, funds that increase their exposure to momentum the most at times when positive-feedback traders are expected to be active earn 0.54% more per month than funds that decrease their exposure to momentum. In low cross-sectional dispersion states the average return difference is 0.30% more per month. The difference between

[43] Gorman, Sapra, and Weigand (2010) argue that cross-sectional return dispersion is highly relevant in active portfolio management. Return dispersion can be used as an effective alpha dispersion signal for investors who pursue long-short strategies.

the two states is statistically significant.

Finally, in high cross-sectional dispersion states, funds that decrease their exposure to momentum the most when sentiment traders are expected to be active earn 0.68% more per month than funds that decrease their exposure to momentum. In low cross-sectional dispersion states the average return difference is 0.54% more per month. The difference between the two states is statistically significant.

In summary, managers who strategically time the momentum crowd tend to outperform their style peers, especially in states of high cross-sectional dispersion in individual stock returns. This evidence indicates that they are not only skilled at timing the momentum crowd, but also at taking advantage of investment opportunities in the market.

Overall, the specific momentum timing model in the main section of the paper identifies hedge funds that have the ability to correctly anticipate the actions of the momentum crowd. This section complements these results by showing that momentum-crowd-timing skill is related to other valuable managerial skills such as pursuing distinctive strategies and recognizing investment opportunities.

8 Robustness Tests

The goal of this section is to examine whether the economic significance of crowd-timing ability is robust to several new tests. The new tests involve excluding from the sample prominent periods such as the Technology Bubble and the Financial Crisis, incorporating all equity investment categories of funds, and including control variables that have been shown to be related to hedge fund returns.

8.1 Excluding 1999-2001 and 2008-2009

To ensure that the previous results are not driven by the inclusion of bubble or crisis periods in the sample, I exclude data from 1999 to 2001 (i.e., the Technology Bubble) and from 2008 to 2009 (i.e., Financial Crisis). I examine whether the results are robust to the exclusion of these periods.

Table 13 presents the economic value of crowd-timing ability, excluding the 1999-2000 and 2008-

2009 periods. The results show that the premium for decreasing exposure to momentum when the hedge fund crowd is large is 0.60% per month (0.66% for full sample). The risk-adjusted premium is 0.54% per month (0.68% for full sample). Both are statistically significant at the 5% level. The premium for increasing exposure to momentum when the positive-feedback trading crowd is large is 0.47% per month (0.45% for full sample), and its alpha is 0.49% per month (0.54% for full sample). Both are statistically significant at the 5% level. Finally, the premium for decreasing exposure to momentum when investor sentiment is high is 0.55% per month (0.59% for full sample). The risk-adjusted premium is 0.65% per month (0.71% for full sample). Both are statistically significant at the 5% level. The difference in spreads between the top and bottom quintiles in each γ category for the sample when the crisis periods are excluded and for the full sample is not very substantial. Therefore, the results are robust to excluding the Technology Bubble period and the Financial Crisis period from the sample.

8.2 Including All Equity-Oriented Hedge Funds

In this section I include all other equity investment categories of funds reported in the TASS Lipper database and repeat the analysis in Section 5. The categories include multi-strategy (408 funds), convertible arbitrage (187 funds), event-driven (561 funds), global macro (301 funds), and funds of funds (1791 funds). To save space, I report the main results based on sorting in Table 14. The results show that the premium for decreasing exposure to momentum when the hedge fund crowd is large is 0.36% per month (4.32% per year). The risk-adjusted premium is 0.35% per month (4.20% per year). Both are statistically significant at the 5% level. The premium for increasing exposure to momentum when the positive-feedback trading crowd is large is 0.26% per month (3.12% per year), and its alpha is 0.29% per month (3.48% per year). Both are statistically significant at the 5% level. Finally, the premium for decreasing exposure to momentum when investor sentiment is high is 0.25% per month (3.00% per year). The risk-adjusted premium is 0.32% per month (3.84% per year). Both are statistically significant at the 5% level.

The premia associated with momentum crowd-timing ability reported in Table 14 are smaller than the premia reported in the main analysis. This is to be expected since equity market neutral and long-short equity hedge funds are more likely to implement quantitative momentum strategies. Nevertheless, the premia are economically significant and they are in line with the idea that there is value in being able to time the size of the momentum crowd.

8.3 Including Other Control Variables

In this section I include several additional variables in the analysis in order to test the robustness of the crowd-timing skills of hedge funds discussed earlier. First, to the extent that signals such as the intensity of short interest, positive-feedback trading, and investor sentiment are serially correlated, their values in month t may contain information from previous months. Thus, a fund manager may adjust momentum exposure based on lagged values of momentum crowd signals. However, according to by Ferson and Schadt (1996), lagged signals are public information and adjusting fund momentum loadings based on public information does not reflect true timing skill.

Second, previous studies have shown that hedge funds can time market liquidity and volatility. Therefore, I test the crowd-timing ability of hedge funds controlling for market liquidity and volatility timing. Finally, to allow for the possibility that funds simply chase previous returns, I control for the past performance of the momentum strategy.

The updated timing model for hedge fund returns becomes:

$$\begin{aligned} R_{i,t} = \alpha_i &+ \gamma_{0i} R_t^{mom} + \gamma_{1i} R_t^{mom} (\widetilde{SH}_t) + \gamma_{2i} R_t^{mom} (\widetilde{PF}_t) + \gamma_{3i} R_t^{mom} (\widetilde{BW}_t) \\ &+ \delta_{1i} R_t^{mom} (SH_{t-1} - \overline{SH}) + \delta_{2i} R_t^{mom} (PF_{t-1} - \overline{PF}) + \delta_{3i} R_t^{mom} (BW_{t-1} - \overline{BW}) \\ &+ \beta_{0i} R_t^{mktrf} + \beta_{1i} R_t^{mktrf} (Vol_t - \overline{Vol}) + \beta_{2i} R_t^{mktrf} (Liq_t - \overline{Liq}) + \beta_{3i} R_{t-1}^{mom} + e_{i,t}, \end{aligned} \tag{15}$$

In this model, $\widetilde{SH}_t$, $\widetilde{PF}_t$, and $\widetilde{BW}_t$ are the innovations in the momentum timing signals computed from an AR(2) process. The terms $SH_{t-1} - \overline{SH}$, $PF_{t-1} - \overline{PF}$, and $BW_{t-1} - \overline{BW}$ are the demeaned lagged values of the momentum timing signals. In addition, Vol_t is realized market volatility in month t, $\overline{Vol}$ is the volatility mean, Liq_t is aggregate market liquidity based on Amihud, and $\overline{Liq}$ is the mean of liquidity. Finally, R_{t-1}^{mom} refers to the lagged return of the momentum strategy.

Using the updated model in (14), I repeat the analysis in Sections 5.3 and 5.4. I find that the crowd-timing ability of hedge funds continues to have economic and statistical significance after controlling for all the variables mentioned above. These results are not tabulated to save space, but they are available upon request.

9 Conclusion

This paper investigates the link between hedge funds' ability to time the momentum crowd and their future performance. I use three empirical measures for the crowd of investors who trade on momentum signals. The measures separate the crowd of arbitrageurs, the crowd of mutual funds engaged in positive-feedback trading, and the crowd of sentiment traders who tend to be more active when investor sentiment is high. It is important to have separate measure of these various groups of traders since their actions have different implications about the performance of the momentum strategy. The three measures of crowding in momentum predict the performance of the momentum strategy in the future in terms of returns, Sharpe ratio, and skewness.

I find that a large proportion of hedge funds have time-varying momentum exposures that are significantly related to the size of the momentum crowd. A bootstrap analysis reveals that the significance of the time-variation in momentum loadings is not accidental but rather reflects managers' decisions to take the momentum crowd into account. The momentum-crowd-timing ability of hedge funds significantly predicts the cross section of hedge fund returns. In particular, hedge funds that decrease their exposure to momentum at times when other arbitrageurs have crowded into the strategy earn 8% more annually than funds that increase their exposure at such times. Funds that increase their exposure to momentum when positive-feedback traders are active earn 5% more annually than funds that decrease their exposure at such times. Funds that scale back their exposure to momentum at times when many sentiment traders are present in the market earn 7% more annually than funds that increase their loading on momentum at such times. Additional regression results show that the predictive ability of crowd-timing skill is different from the effect of other fund characteristics and other measures of timing skill.

The analysis in this paper is inspired by the theoretical literature on strategy crowding, especially in the context of momentum, and anecdotal evidence from the industry about investors' concerns related to crowded trades. Consistent with the implications of the theoretical studies, my findings document a link between cross-sectional differences in crowd-timing skill and fund performance. The evidence presented in this paper suggests that a set of smart arbitrageurs are aware of the implications of crowding for strategy performance and they implement dynamic adjustments to their risk exposures. Furthermore, the results are consistent with the notion that arbitrageurs sometimes choose to trade with the crowd to take advantage of temporary price pressure.

References

Abreu, Dilip, and Markus K. Brunnermeier, 2002, Synchronization risk and delayed arbitrage, *Journal of Financial Economics* 66, 341-360.

Andreassen, Paul and Stephen Kraus, 1988, Judgmental prediction by extrapolation, Unpublished paper, Department of Psychology, Harvard University.

Antoniou, Constantinos, Doukas, John A., and Avanidhar Subrahmanyam, 2013, Cognitive dissonance, sentiment, and momentum, *Journal of Financial and Quantitative Analysis* 48, 245-275.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645-1680.

Bali, Turan, Stephen Brown, and Mustafa Caglayan, 2011, Do hedge funds' exposures to risk factors predict their future returns?, *Journal of Financial Economics* 101, 36-68.

Bali, Turan, Stephen Brown, and Mustafa Caglayan, 2012, Systematic risk and the cross-section of hedge fund returns, *Journal of Financial Economics* 106, 114-131.

Bali, Turan, Stephen Brown, and Mustafa Caglayan, 2014, Macroeconomic risk and hedge fund returns, *Journal of Financial Economics* 114, 1-19.

Barberis, Nicholas and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 161-199.

Barberis, Nicholas, Shleifer, Andrei, and R. W. Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307-343.

Barroso, Pedro and Pedro Santa-Clara, 2015, Momentum has its moments, *Journal of Financial Economics* 116, 111-120.

Bekaert, G., and C Harvey, 1997, Emerging equity market volatility, *Journal of Financial Economics* 43, 29-77.

Ben-David, I., F. Franzoni, and R. Moussawi, 2012, Hedge fund stock trading in the financial crisis of 2007-2008, *Review of Financial Studies* 25, 1-54.

Bessembinder, H., Chan, K., and P. Seguin, 1996, An empirical examination of information, differences of opinion, and trading activity, *Journal of Financial Economics* 40, 105-134.

Boehmer, E., C. Jones, and X. Zhang, 2013, Shackling short sellers: The 2008 shorting ban, *Review of Financial Studies* 26, 1363-1400.

Brainard, S., and D. Cutler, 1993, Sectorial shifts and cyclical unemployment reconsidered, *The Quarterly Journal of Economics* 108, 219-243.

Brunnermeier, Markus, and Stefan Nagel, 2004, Hedge funds and the technology bubble, *Journal of Finance* 59, 2013-2040.

Brunnermeier, Markus, and Lasse H. Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201-2238.

Busse, J., 1999, Volatility timing in mutual funds: evidence from daily returns, *Review of Financial Studies* 12, 1009-1041.

Cao, Charles, Yong Chen, Bing Liang, and Andrew Lo, 2013, Can hedge funds time market liquidity?, *Journal of Financial Economics* 109, 493-516.

Carhart, Mark, 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.

Chen, Yong, 2007, Timing ability in the focus market of hedge funds, *Journal of Investment Management* 5, 66-98.

Chen, Yong, 2011, Derivatives use and risk taking: evidence from the hedge fund industry, *Journal of Financial and Quantitative Analysis* 46, 1073-1106.

Chen, Yong, Han, Bing, and Jing Pan, 2016, Sentiment risk, sentiment timing, and hedge fund returns, Working Paper, Mays Business School.

Chen, Yong, and Bing Liang, 2007, Do market timing hedge funds time the market?, *Journal of Financial and Quantitative Analysis* 42, 827-856.

Connolly, R., and C. Stivers, 2006, Information content and other characteristics of the daily cross-sectional dispersion in stock returns, *Journal of Empirical Finance* 13, 79-112.

Cutler, David, Poterba, James, and Lawrence Summers, 1990, Speculative dynamics and the role of feedback traders, *American Economic Review* 80, 63-68.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam, 1998, Investor psychology and security market under-and overreaction, *Journal of Finance* 53, 1839-1885.

DAvolio, Gene, 2002, The market for borrowing stock, *Journal of Financial Economics* 66, 271-306.

De Long, Bradford, Shleifer, Andrei, Summers, Lawrence, and Robert Waldmann, 1990a, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703-738.

De Long, Bradford, Shleifer, Andrei, Summers, Lawrence, and Robert Waldmann, 1990b, Positive feedback investment strategies and destabilizing rational speculation, *Journal of Finance* 45, 379-395.

Duffie, Darrell, 2010, Asset price dynamics with slow-moving capital, *Journal of Finance* 65, 1237-1267.

Engelberg, Joseph E., Reed, Adam, and Mathhew Ringgenberg, 2018, ShortSelling Risk, *Journal of Finance* 73, 755-786.

Fama, Eugene and Kenneth French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55-84.

Fama, Eugene and Kenneth French, 2008, Dissecting anomalies, *Journal of Finance* 63, 1653-1678.

Ferson, W. and R. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425-460.

Festinger, L., 1957, A theory of cognitive dissonance, Stanford University Press, Stanford, CA.

Frazzini, Andrea, 2006, The disposition effect and underreaction to news, *Journal of Finance* 61, 2017-2046.

Frijns, Bart, Gilbert, Aaron, and Remco C. J. Zwinkels, 2016, On the style-based feedback trading of mutual fund managers, *Journal of Financial and Quantitative Analysis* 51, 771-800.

Fung, William, and David Hsieh, 1997, Empirical characteristics of dynamic trading strategies: the case of hedge funds, *Review of Financial Studies* 10, 275-302.

Fung, William, and David Hsieh,2000, Performance characteristics of hedge funds and CTA funds: natural versus spurious biases, *Journal of Financial and Quantitative Analysis* 35, 291-307.

Fung, William, and David Hsieh, 2004, Hedge fund benchmarks: A risk-based approach, *Financial Analysts Journal* 60, 65-80.

Gorman, Larry R., Sapra, Steven G., and Robert A. Weigand, 2010, The cross-sectional dispersion of stock returns, alpha, and the information ratio, *Journal of Investing* 19, 113-127.

Griffin, John, Harris, Jeffrey, Shu, Tao, and Selim Topaloglu, 2011, Who drove and burst the tech bubble?, *Journal of Finance* 66, 1250-1290.

Grinblatt, Mark and Bing Han, 2005, Prospect theory,mental accounting, and momentum, *Journal of Financial Economics* 78, 311-339.

Grinblatt, Mark and Matti Keloharju, 2001, What makes investors trade?" , *Journal of Finance* 56, 589-616.

Grinblatt, Mark, Jostova, Gergana, Petrasek, Lubomir, and Alexander Philipov, 2016, Style and skill: Hedge funds, mutual funds, and momentum, Working Paper, University of California, Los Angeles (UCLA).

Grinblatt, Mark, Titman, Sheridan, and Russ Wermers, 1995, Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior, *American Economic Review* 85, 1088-1105.

Hanson, Samuel and Adi Sunderam, 2013, The growth and limits of arbitrage: Evidence from short interest, *Review of Financial Studies* 27, 1238-1286.

Harris, Lawrence, Hartzmark, Samuel M., and David H. Solomon, 2015, Juicing the dividend yield: Mutual funds and the demand for dividends, *Journal of Financial Economics* 116, 433-451.

Hong, Harrison and Jeremy Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance* 54, 2143-2184.

Huang, Simon, 2019, The momentum gap and return predictability, Working Paper, Isenberg School of Management, University of Massachusetts Amherst.

Jagannathan, R., Malakhov, A., and D. Novikov, 2010, Do hot hands exist among hedge fund managers? An empirical evaluation, *Journal of Finance* 65, 217-255.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881-898.

Jegadeesh, Narasimhan and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jegadeesh, Narasimhan and Sheridan Titman, 2001, Profitability of momentum strategies: An evaluation of alternative explanations, *Journal of Finance* 56, 699-720.

Jin, L., and A. Scherbina, 2011, Inheriting losers, *Review of Financial Studies* 24, 786-820.

Kahneman, Daniel and Amos Tversky, 1979, Prospect theory: An analysis of decision under risk, *Econometrica* 47, 263-291.

Khandani, Amir E. and Andrew W. Lo, 2007, What happened to the Quants in August 2007?, *Journal of Investment Management* 5, 29-78.

Kosowski, Robert, Narayan Naik, and Melvyn Teo, 2007, Do hedge funds deliver alpha? A bayesian and bootstrap analysis, *Journal of Financial Economics* 84, 229-264.

Lakonishok, Josef, Shleifer, Andrei, and Robert W. Vishny, 1992, The impact of institutional trading on stock prices, *Journal of Financial Economics* 32, 22-43.

Lou, Dong, and Christopher Polk, 2014, Comomentum: Inferring arbitrage activity from return correlations, Working Paper, London School of Economics.

Loungani, P., Rush, M., and W. Tave, 1990, Stock market dispersion and unemployment, *Journal of Monetary Economics* 25, 367-388.

Miller, E. M., 1977, Risk, uncertainty and divergence of opinion, *Journal of Finance* 32, 1151-1168.

Mitchell, Mark, Pedersen, Lasse Heje, and Todd Pulvino, 2007, Slow moving capital, *American Economic Review* 97, 215-220.

O'Connell, P. G. J., and M. Teo, 2009, Institutional investors, past performance, and dynamic loss aversion, *Journal of Financial and Quantitative Analysis* 44, 155-188.

Odean, Terrance, 1998, Are investors reluctant to realize their losses?, *Journal of Finance* 53, 1775-1798.

Pedersen, Lasse H., 2009, When everyone runs for the exit, *The International Journal of Central Banking* 5, 177-199.

Peng, Cameron and Chen Wang, 2019, Positive feedback trading and stock prices: Evidence from mutual funds, Working Paper, London School of Economics.

Pojarliev, Momtchil and Richard M. Levich, 2011, Detecting crowded trades in currency funds, *Financial Analysts Journal* 67, 26-39.

Sadka, Ronnie, 2010, Liquidity risk and the cross-section of hedge fund returns, *Journal of Financial Economics* 98, 54-71.

Shanken, Jay, 1990, Intertemporal asset pricing: An empirical investigation, *Journal of Econometrics* 45, 99-120.

Shefrin, Hersh and Meir Statman, 1985, The disposition to sell winners too early and ride losers too long: Theory and evidence, *Journal of Finance* 40, 777-790.

Shleifer, Andrei and Laurence H. Summers, 1990, The noise trader approach to finance, *Journal of Economic Perspectives* 4, 19-33.

Shleifer, Andrei and Robert W. Vishny, 1997, The limits of arbitrage, *The Journal of Finance* 52, 35-55.

Stambaugh, Robert F., Yu, Jianfeng, and Yu Yuan, 2011, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104, 288-302.

Stein, Jeremy, 2009, Sophisticated investors and market Efficiency, *Journal of Finance* 64, 1517-1548.

Stivers, C., 2003, Firm-level return dispersion and the future volatility of aggregate stock market returns, *Journal of Financial Markets* 6, 389-411.

Sun, Zheng, Wang, Ashley, and Lu Zheng, 2012, The road less traveled: Strategy distinctiveness and hedge fund performance, *Review of Financial Studies* 25, 96-143.

Teo, Melvyn, 2011, The liquidity risk of liquid hedge funds, *Journal of Financial Economics* 100, 24-44.

Treynor,J. and K. Mazuy, 1966, Can mutual funds outguess the market?, *Harvard Business Review* 44, 131-136.

Wermers, Russ, 2003, Is money really 'smart'? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence, Working Paper, University of Maryland.

Table 1. **Summary Statistics of Hedge Fund Returns and Hedge Fund Factors**

Panel A reports summary statistics of hedge fund returns obtained from Lipper TASS. The data covers both active and defunct hedge funds. For each fund, the first 12 months of return are excluded to address backfilling bias. The sample includes equity market neutral and long-short equity funds that report net-of-fee returns in USD on a monthly basis and have assets under management of at least \$5 million. The descriptive statistics include number of funds in the sample, mean return, median return, standard deviation of return, and the 25th and 75th percentiles of the return distribution. All returns are in % per month. Panel B reports summary statistics for the Fung-Hsieh seven factors obtained from David A. Hsieh's data library. These factors are the market excess return ($MKTRF$), a size factor (SMB), the monthly change in the 10-year treasury constant maturity yield ($\Delta\ TERM$), the monthly change in Moody's Baa yield minus the 10-year treasury constant maturity yield ($\Delta\ CREDIT$), and three trend-following factors: $PFTSBD$ (bond), $PFTSFX$ (currency), and $PFTSCOM$ (commodity). The sample period in both panels is from January 1997 to December 2017.

	N	Mean	Median	StDev	25%	75%
Panel A: Summary of average fund returns						
Equity Market Neutral	309	0.47	0.46	2.35	-0.48	1.45
Long-Short Equity	1807	0.78	0.71	6.11	-1.24	2.75
Panel B: Summary of average factor returns						
$MKTRF$		0.40	1.18	4.93	-2.57	3.88
SMB		0.28	0.06	3.88	-2.07	2.59
$\Delta\ TERM$		-0.02	-0.04	0.23	-0.16	0.12
$\Delta\ CREDIT$		0.01	-0.01	0.22	-0.08	0.07
$PTFSBD$		-1.74	-3.70	14.73	-12.72	3.62
$PTFSFX$		1.09	-2.45	18.60	-11.74	9.46
$PTFSCOM$		0.20	-2.72	13.70	-8.49	6.76

Table 2. **Summary Statistics of Momentum Crowd Signals and Momentum Strategy Returns**

Panel A reports summary statistics for the momentum factor (MOM), the scaled momentum factor of Barroso and Santa Clara (MOM^*), and three momentum timing signals. The timing signals are the intensity of short selling in momentum (SH), the extent of positive-feedback trading in momentum (PF), and the Investor Sentiment Index of Baker and Wurgler (2006) (BW). The descriptive statistics include mean, median, standard deviation, and the 25th and 75th percentiles of the sample distribution. The sample period in Panel A is from January 1997 to December 2017. Panel B shows time-series forecasting regressions of the form:

$$R^{mom^*}_{t,t+12} = b_0 + b_1(SH_t - \overline{SH}) + b_2(PF_t - \overline{PF}) + b_3(BW_t - \overline{BW}) + u_{t,t+12}, \tag{16}$$

and

$$Sharpe\,Ratio^{mom^*}_{t,t+12} = b_0 + b_1(SH_t - \overline{SH}) + b_2(PF_t - \overline{PF}) + b_3(BW_t - \overline{BW}) + v_{t,t+12}, \tag{17}$$

and

$$Skewness^{mom^*}_{t,t+12} = b_0 + b_1(SH_t - \overline{SH}) + b_2(PF_t - \overline{PF}) + b_3(BW_t - \overline{BW}) + v_{t,t+12}, \tag{18}$$

where $R^{mom^*}_{t,t+12}$ is the average monthly return of the scaled momentum strategy from month t to month $t+12$, $Sharpe\,Ratio^{mom^*}_{t,t+12}$ is the Sharpe Ratio of the scaled momentum strategy calculated as average return from month t to month $t+12$ divided by standard deviation of return from month t to month $t+12$, and $Skewness^{mom^*}_{t,t+12}$ is the skewness of the scaled momentum strategy calculated from month t to month $t+12$. The three timing signals are the size of the arbitrage capital in momentum based on the intensity of short interest, SH, the size of the positive-feedback crowd, PF, and the Investment Sentiment Index of Baker and Wurgler (2006). All signals are demeaned. T-statistics are shown in parentheses and are computed using Newey-West (1987) standard errors allowing for serial correlation of up to twelve monthly lags. The sample period is from June 1986 to December 2017.

Panel A: Descriptive Statistics					
	Mean	Median	SD	25%	75%
MOM	1.50	2.71	12.26	-2.62	7.38
MOM^*	1.04	1.43	4.53	-1.36	3.81
SH	0.0064	0.0050	0.0076	0.0008	0.0125
PF	0.2939	0.3074	0.1526	0.1859	0.3994
BW	0.2713	0.2000	0.7645	-0.2700	0.5800

Panel B: Correlation Matrix of Timing Signals			
	SH	PF	BW
SH	1		
PF	0.0580	1	
BW	-0.2848	0.0630	1

Panel C: Predictive Regressions					
	b_0	b_1	b_2	b_3	R^2
$R^{mom^*}_{t,t+12}$	0.16	-2.4	0.29	0.04	0.09
	(15.14)	(-2.17)	(4.40)	(1.13)	
$Sharpe\,Ratio^{mom^*}_{t,t+12}$	0.28	-1.93	0.51	0.07	0.08
	(15.83)	(-1.99)	(4.44)	(1.42)	
$Skewness^{mom^*}_{t,t+12}$	-0.36	-0.24	1.02	-0.21	0.08
	(-9.62)	(-0.17)	(4.18)	(-3.36)	

Table 3. **Cross-sectional Distribution of t-statistics for Momentum Crowd Timing Coefficients across Hedge Funds**

This table summarizes the distribution of t-statistics for the momentum-crowd-timing coefficients of hedge funds. For each fund with at least 36 monthly return observations, the coefficients come from the following model estimated over the whole sample period:

$$R_{i,t} = \alpha_i + \gamma_{0i} R_t^{mom^*} + \gamma_{1i} R_t^{mom^*} (SH_t - \overline{SH}) + \gamma_{2i} R_t^{mom^*} (PF_t - \overline{PF}) + \gamma_{3i} R_t^{mom^*} (BW_t - \overline{BW}) + \sum_{j=1}^{J} \beta_j f_{j,t} + e_{i,t}, \tag{19}$$

where $R_{i,t}$ is the excess return of the fund, $R_t^{mom^*}$ is the return of the scaled momentum strategy, $f_{j,t}$ stands for the Fung-Hsieh hedge fund factors, SH is a measure of the crowd of arbitrageurs trading on momentum, PF measures the crowd of positive-feedback traders, and BW is the investor sentiment index of Baker and Wurgler (2006). The t-statistics are computed using Newey-West (1987) standard errors allowing for serial correlation of up to twelve monthly lags. The numbers in the table report the percentage of funds with t-statistics of the crowd-timing coefficient exceeding the indicated values. The sample period is from January 1997 to December 2017.

	Percentage of Funds			
t-stat	γ_0	γ_1	γ_2	γ_3
t<0	34.62%	50.65%	53.46%	41.65%
t>0	65.38%	49.35%	46.54%	58.35%
t≤-2.326	3.09%	2.81%	5.62%	3.00%
t≤-1.96	4.79%	5.07%	8.53%	5.44%
t≤-1.645	6.73%	7.38%	11.94%	7.84%
t≤-1.282	10.65%	12.96%	18.40%	11.62%
t≥1.282	27.52%	12.86%	14.38%	21.76%
t≥1.645	20.01%	8.53%	9.08%	15.17%
t≥1.96	15.77%	5.58%	5.49%	10.24%
t≥2.326	10.97%	3.69%	3.37%	6.59%

Table 4. **Bootstrap Analysis of Momentum Crowd Timing Ability**

This table presents the results of the bootstrap analysis for the momentum-crowd-timing coefficients of hedge funds. For each fund with at least 36 monthly return observations, the coefficients come from the following model:

$$R_{i,t} = \alpha_i + \gamma_{0i} R_t^{mom^*} + \gamma_{1i} R_t^{mom^*}(SH_t - \overline{SH}) + \gamma_{2i} R_t^{mom^*}(PF_t - \overline{PF}) + \gamma_{3i} R_t^{mom^*}(BW_t - \overline{BW}) + \sum_{j=1}^{J} \beta_j f_{j,t} + e_{i,t} \quad , \tag{20}$$

where $R_{i,t}$ is the excess return of the fund, $R_t^{mom^*}$ is the return of the scaled momentum strategy, $f_{j,t}$ stands for the Fung-Hsieh hedge fund factors, SH is a measure of the crowd of arbitrageurs trading on momentum, PF measures the crowd of positive-feedback traders, and BW is the investor sentiment index of Baker and Wurgler (2006). The t-statistics are computed using Newey-West (1987) standard errors allowing for serial correlation of up to twelve monthly lags. The γ coefficients measure momentum-crowd-timing ability. In the table, the first row of each panel reports the sorted t-statistics of momentum-crowd timing coefficients across individual hedge funds, and the second row is the empirical p-values from bootstrap simulations. The number of resampling iterations is 10,000. The details of the bootstrap analysis are in Section 5.2.

	Panel A							
	Bottom t-statistics for γ_1				Top t-statistics for γ_1			
	1%	2.5%	5%	10%	10%	5.0%	2.5%	1%
t-stat	-2.99	-2.43	-1.96	-1.44	1.49	2.04	2.62	3.09
p-value	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00
	Panel B							
	Bottom t-statistics for γ_2				Top t-statistics for γ_2			
	1%	2.5%	5%	10%	10%	5.0%	2.5%	1%
t-stat	-4.02	-2.97	-2.44	-1.81	1.57	2.02	2.54	3.13
p-value	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00
	Panel C							
	Bottom t-statistics for γ_3				Top t-statistics for γ_3			
	1%	2.5%	5%	10%	10%	5.0%	2.5%	1%
t-stat	-3.09	-2.46	-2.04	-1.42	1.98	2.57	3.13	3.76
p-value	0.00	0.00	0.00	0.00	0.04	0.02	0.00	0.00

Table 5. **Hedge Fund Returns Sorted by Momentum Crowd Timing Ability: Short Holding Period**

This table presents the out-of-sample average returns and Fung-Hsieh alphas for portfolios of hedge funds sorted by momentum-crowd-timing coefficients. Each month, and for each timing coefficient, I form 5 quintile portfolios based on hedge funds' momentum-crowd-timing coefficients estimated from the past 36 months using this model:

$$R_{i,t} = \alpha_i + \gamma_{0i} R_t^{mom^*} + \gamma_{1i} R_t^{mom^*}(SH_t - \overline{SH}) + \gamma_{2i} R_t^{mom^*}(PF_t - \overline{PF}) + \gamma_{3i} R_t^{mom^*}(BW_t - \overline{BW}) + \sum_{j=1}^{J} \beta_j f_{j,t} + e_{i,t} \quad , \tag{21}$$

where $R_{i,t}$ is the excess return of the fund, $R_t^{mom^*}$ is the return of the scaled momentum strategy, $f_{j,t}$ stands for the Fung-Hsieh hedge fund factors, SH is a measure of the crowd of arbitrageurs trading on momentum, PF measures the crowd of positive-feedback traders, and BW is the investor sentiment index of Baker and Wurgler (2006). The portfolios are held for 1 month and then they are rebalanced. The t-statistics are reported below the returns and they are calculated based on Newey-West heteroskedasticity and autocorrelation-consistent standard errors. All returns are in % per month. The last two rows of the table present the return difference, and its significance, between the top and bottom quintiles in each γ coefficient category. The sample period is from January 1997 to December 2017.

Quintiles	γ_1		γ_2		γ_3	
	Ave Ret	Alpha	Ave Ret	Alpha	Ave Ret	Alpha
Q1	1.15	0.92	0.59	0.29	1.13	0.94
	(3.66)	(2.87)	(2.75)	(1.74)	(3.51)	(2.99)
Q2	0.67	0.41	0.58	0.34	0.69	0.44
	(4.90)	(3.46)	(3.24)	(2.92)	(4.90)	(3.70)
Q3	0.63	0.37	0.60	0.36	0.59	0.32
	(3.25)	(2.93)	(4.43)	(3.17)	(3.25)	(2.93)
Q4	0.58	0.33	0.71	0.47	0.55	0.32
	(3.45)	(2.44)	(4.68)	(3.45)	(3.57)	(2.35)
Q5	0.49	0.24	1.04	0.83	0.54	0.23
	(2.34)	(1.50)	(3.35)	(2.63)	(2.55)	(1.26)
Q5-Q1	-0.66	-0.68	0.45	0.54	-0.59	-0.71
	(-2.29)	(-2.33)	(2.63)	(2.91)	(-2.11)	(-2.48)

Table 6. **Hedge Fund Returns Sorted by Momentum Crowd Timing Ability: Long Holding Period**

This table presents the out-of-sample average returns for portfolios of hedge funds sorted by momentum-crowd-timing coefficients. Each month, and for each timing coefficient, I form 5 quintile portfolios based on hedge funds' momentum-crowd-timing coefficients estimated from the past 36 months using this model:

$$R_{i,t} = \alpha_i + \gamma_{0i} R_t^{mom^*} + \gamma_{1i} R_t^{mom^*}(SH_t - \overline{SH}) + \gamma_{2i} R_t^{mom^*}(PF_t - \overline{PF}) + \gamma_{3i} R_t^{mom^*}(BW_t - \overline{BW}) + \textstyle\sum_{j=1}^{J} \beta_j f_{j,t} + e_{i,t}, \quad (22)$$

where $R_{i,t}$ is the excess return of the fund, $R_t^{mom^*}$ is the return of the scaled momentum strategy, $f_{j,t}$ stands for the Fung-Hsieh hedge fund factors, SH is a measure of the crowd of arbitrageurs trading on momentum, PF measures the crowd of positive-feedback traders, and BW is the investor sentiment index of Baker and Wurgler (2006). The portfolios are held for K=3,6,9,or 12 months and then they are rebalanced. The t-statistics are reported below the returns and they are calculated based on Newey-West heteroskedasticity and autocorrelation-consistent standard errors. All returns are in % per month. The last two rows of the table present the return difference, and its significance, between the top and bottom quintiles in each γ coefficient category. The sample period is from January 1997 to December 2017.

Quintiles	γ_1				γ_2				γ_3			
	K=3	6	9	12	K=3	6	9	12	K=3	6	9	12
Q1	1.02	1.01	0.98	0.95	0.57	0.59	0.59	0.60	1.10	1.09	1.07	1.06
	(4.26)	(4.31)	(4.21)	(4.07)	(2.75)	(2.92)	(2.95)	(3.04)	(3.31)	(3.21)	(3.09)	(2.99)
Q2	0.84	0.90	0.86	0.84	0.61	0.63	0.61	0.59	0.71	0.69	0.64	0.63
	(3.84)	(2.56)	(2.78)	(2.90)	(1.88)	(2.55)	(2.09)	(1.85)	(4.84)	(3.02)	(3.07)	(2.60)
Q3	0.57	0.58	0.83	0.74	0.63	0.64	0.63	0.65	0.52	0.65	0.67	0.63
	(2.90)	(3.37)	(2.96)	(2.97)	(4.61)	(3.40)	(3.53)	(3.56)	(2.83)	(3.87)	(3.30)	(3.48)
Q4	0.53	0.59	0.90	0.68	0.72	0.71	0.69	0.67	0.63	0.62	0.61	0.61
	(2.73)	(2.87)	(2.87)	(2.61)	(4.50)	(4.26)	(4.09)	(3.92)	(2.96)	(3.44)	(3.48)	(3.48)
Q5	0.56	0.60	0.65	0.66	1.12	1.09	1.07	1.07	0.61	0.61	0.60	0.61
	(2.68)	(2.88)	(2.77)	(2.80)	(3.42)	(3.31)	(3.19)	(3.10)	(2.91)	(3.01)	(3.00)	(3.06)
Q5-Q1	-0.46	-0.41	-0.33	-0.29	0.55	0.51	0.48	0.47	-0.49	-0.48	-0.47	-0.45
	(-2.19)	(-1.96)	(-1.50)	(-1.26)	(2.23)	(1.91)	(1.75)	(1.64)	(-1.69)	(-1.59)	(-1.55)	(-1.46)

Table 7. **Hedge Fund Returns Sorted by Momentum Crowd Timing Ability: Double Sorts**

This table presents the out-of-sample average returns and Fung-Hsieh alphas for portfolios of hedge funds sorted by momentum-crowd-timing coefficients. Each month, and for each timing coefficient, I form 3 tercile portfolios based on hedge funds' momentum-crowd-timing coefficients estimated from the past 36 months using this model:

$$R_{i,t} = \alpha_i + \gamma_{0i} R_t^{mom^*} + \gamma_{1i} R_t^{mom^*}(SH_t - \overline{SH}) + \gamma_{2i} R_t^{mom^*}(PF_t - \overline{PF}) + \gamma_{3i} R_t^{mom^*}(BW_t - \overline{BW}) + \textstyle\sum_{j=1}^{J} \beta_j f_{j,t} + e_{i,t} \quad (23)$$

where $R_{i,t}$ is the excess return of the fund, $R_t^{mom^*}$ is the return of the scaled momentum strategy, $f_{j,t}$ stands for the Fung-Hsieh hedge fund factors, SH is a measure of the crowd of arbitrageurs trading on momentum, PF measures the crowd of positive-feedback traders, and BW is the investor sentiment index of Baker and Wurgler (2006). In Panel A, returns are sorted independently into three groups based on γ_{1i} and three groups based on γ_{2i}. In Panel B, returns are sorted independently into three groups based on γ_{1i} and three groups based on γ_{3i}. In Panel C, returns are sorted independently into three groups based on γ_{2i} and three groups based on γ_{3i}. The portfolios are held for 1 month and then they are rebalanced. The t-statistics are reported below the returns and they are calculated based on Newey-West heteroskedasticity and autocorrelation-consistent standard errors. All returns are in % per month. The sample period is from January 1997 to December 2017.

Panel A: Sorting on SH and PF Timing						
	Ave Ret			Alpha		
	Low γ_2	Mid γ_2	High γ_2	Low γ_2	Mid γ_2	High γ_2
Low γ_1	0.86	0.70	1.27	0.55	0.46	1.04
	(4.50)	(4.88)	(2.83)	(3.45)	(3.49)	(2.23)
Mid γ_1	0.42	0.59	0.76	0.33	0.38	0.52
	(2.15)	(5.11)	(4.58)	(1.88)	(3.76)	(3.39)
High γ_1	0.38	0.58	0.65	0.08	0.34	0.45
	(1.67)	(3.19)	(3.33)	(0.64)	(2.37)	(2.67)

Panel B: Sorting on SH and BW Timing						
	Ave Ret			Alpha		
	Low γ_3	Mid γ_3	High γ_3	Low γ_3	Mid γ_3	High γ_3
Low γ_1	1.10	0.66	0.84	0.83	0.39	0.62
	(2.82)	(4.73)	(4.10)	(2.07)	(3.02)	(3.16)
Mid γ_1	0.79	0.30	0.56	0.50	0.24	0.32
	(4.11)	(2.25)	(3.44)	(3.22)	(1.93)	(2.11)
High γ_1	0.85	0.62	0.52	0.65	0.35	0.26
	(2.77)	(3.26)	(2.34)	(2.73)	(2.38)	(1.36)

Panel C: Sorting on PF and BW Timing						
	Ave Ret			Alpha		
	Low γ_3	Mid γ_3	High γ_3	Low γ_3	Mid γ_3	High γ_3
Low γ_2	0.81	0.40	0.44	0.56	0.32	0.13
	(3.77)	(2.17)	(2.12)	(3.20)	(1.88)	(0.76)
Mid γ_2	0.71	0.61	0.57	0.50	0.38	0.32
	(4.77)	(5.38)	(3.27)	(3.93)	(3.90)	(2.04)
High γ_2	1.17	0.74	0.72	0.99	0.49	0.46
	(3.52)	(4.45)	(3.14)	(2.91)	(3.30)	(2.18)

Table 8. **Fama-MacBeth Regressions of Hedge Fund Returns on Momentum Crowd Timing Coefficients and Controls**

This table reports results from Fama-MacBeth regressions of hedge fund excess returns on momentum-crowd-timing coefficients with controls for various fund characteristics. In each month t for each hedge fund, crowd-timing coefficients are estimated using model (5) over the period $t-36$ to $t-1$. The coefficients are used as independent variables in cross-sectional regressions that also control for fund characteristics measured at time $t-1$: lagged return, management fee, incentive fee, redemption notice period, lockup period, leveraged dummy, minimum investment, high water mark, age, and size (AUM). The t-statistics are based on Newey-West standard errors. The sample period is from January 1997 to December 2017.

Variable	Model 1	Model 2
γ_1	-0.0060	-0.0010
	(-2.04)	(-2.62)
γ_2	0.2420	0.0967
	(2.87)	(3.19)
γ_3	0.4444	0.1399
	(1.03)	(1.09)
Lagged Return		2.5755
		(1.29)
Management Fee		0.0544
		(0.87)
Incentive Fee		0.0043
		(0.94)
Redemption Notice Period		0.0027
		(3.25)
Lockup Period		0.0057
		(1.58)
Leveraged		0.0280
		(0.63)
Min Investment		0.0000
		(-0.10)
High Water Mark		0.2069
		(4.13)
Age		0.0014
		(4.00)
AUM		0.0000
		(-0.92)

Table 9. **Correlation between Momentum Component of Hedge Fund Returns and Momentum Strategy**

This table presents time-series correlations between the scaled momentum strategy and lags of the momentum components of hedge fund returns. The analysis is performed for quintiles 1 and 5 within each crowd-timing coefficient group from model (5). For each quintile, its momentum-related component is derived by first running model (5) and then extracting the fitted value of the regression corresponding to all terms that involve the momentum strategy return. The sample period is from January 1997 to December 2017.

	γ_1		γ_2		γ_3	
k	$Corr(R_t^{mom}, \widehat{q1}_{t-k})$	$Corr(R_t^{mom}, \widehat{q5}_{t-k})$	$Corr(R_t^{mom}, \widehat{q1}_{t-k})$	$Corr(R_t^{mom}, \widehat{q5}_{t-k})$	$Corr(R_t^{mom}, \widehat{q1}_{t-k})$	$Corr(R_t^{mom}, \widehat{q5}_{t-k})$
0	18.31%	-38.21%	-23.82%	14.39%	55.47%	-65.08%
1	17.25%	-34.77%	-21.18%	13.53%	53.21%	-60.47%
2	15.75%	-33.06%	-19.94%	12.02%	49.22%	-56.51%
3	13.69%	-30.99%	-19.16%	10.44%	43.52%	-51.66%
4	11.75%	-28.47%	-17.68%	8.76%	38.13%	-46.49%
5	8.46%	-27.48%	-17.57%	5.80%	30.49%	-41.96%
6	6.09%	-25.90%	-16.80%	3.62%	24.32%	-37.42%
7	2.36%	-24.75%	-16.40%	0.32%	17.01%	-33.28%
8	-1.71%	-23.92%	-16.09%	-3.26%	10.47%	-30.18%
9	-7.17%	-24.60%	-17.36%	-7.93%	1.84%	-27.88%
10	-11.24%	-23.24%	-16.60%	-11.44%	-4.61%	-24.38%
11	-15.50%	-21.70%	-15.79%	-15.02%	-11.14%	-20.79%
12	-19.13%	-19.67%	-14.23%	-18.19%	-16.22%	-17.33%

Table 10. **Fund Characteristics**

This table presents characteristics for portfolios of hedge funds sorted by momentum-crowd-timing coefficients. Each month, and for each timing coefficient, I form 5 quintile portfolios based on hedge funds' momentum-crowd-timing coefficients estimated from the past 36 months using this model:

$$R_{i,t} = \alpha_i + \gamma_{0i} R_t^{mom^*} + \gamma_{1i} R_t^{mom^*}(SH_t - \overline{SH}) + \gamma_{2i} R_t^{mom^*}(PF_t - \overline{PF}) + \gamma_{3i} R_t^{mom^*}(BW_t - \overline{BW}) + \sum_{j=1}^{J} \beta_j f_{j,t} + e_{i,t}, \quad (24)$$

where $R_{i,t}$ is the excess return of the fund, $R_t^{mom^*}$ is the return of the scaled momentum strategy, $f_{j,t}$ stands for the Fung-Hsieh hedge fund factors, SH is a measure of the crowd of arbitrageurs trading on momentum, PF measures the crowd of positive-feedback traders, and BW is the investor sentiment index of Baker and Wurgler (2006). The portfolios are held for 1 month and then they are rebalanced. The characteristics are assets under management (AUM), a dummy variables that equals 1 if the fund uses leverage (Leveraged), and the payout period. The sample period is from January 1997 to December 2017.

Quintiles	AUM	Leveraged	Payout Period
		γ_1	
Q1	38,880,529	0.6256	11.8704
Q2	51,427,664	0.5777	10.5202
Q3	55,069,694	0.5710	10.0769
Q4	58,588,457	0.5996	9.3340
Q5	36,034,931	0.6458	8.8522
		γ_2	
Q1	37,586,592	0.6154	10.6599
Q2	54,070,141	0.5879	9.7915
Q3	56,233,019	0.5746	9.7389
Q4	52,934,049	0.6020	9.3947
Q5	38,096,877	0.6395	11.4858
		γ_3	
Q1	34,335,059	0.6307	10.5243
Q2	51,220,218	0.5729	9.1538
Q3	57,999,871	0.5655	9.6275
Q4	54,242,351	0.6082	11.3846
Q5	42,087,973	0.6425	10.8644

Table 11. **Strategy Distinctiveness of Hedge Fund Returns Sorted by Momentum Crowd Timing Ability**

This table reports the out-of-sample strategy distinctiveness index (SDI) for portfolios of hedge funds sorted by momentum-crowd-timing coefficients. Each month, and for each timing coefficient, 5 quintile portfolios are formed based on hedge funds' momentum-crowd-timing coefficients estimated from the past 36 months using this model:

$$R_{i,t} = \alpha_i + \gamma_{0i} R_t^{mom^*} + \gamma_{1i} R_t^{mom^*} (SH_t - \overline{SH}) + \gamma_{2i} R_t^{mom^*} (PF_t - \overline{PF}) + \gamma_{3i} R_t^{mom^*} (BW_t - \overline{BW}) + \sum_{j=1}^{J} \beta_j f_{j,t} + e_{i,t} \quad (25)$$

where $R_{i,t}$ is the excess return of the fund, $R_t^{mom^*}$ is the return of the scaled momentum strategy, $f_{j,t}$ stands for the Fung-Hsieh hedge fund factors, SH is a measure of the crowd of arbitrageurs trading on momentum, PF measures the crowd of positive-feedback traders, and BW is the investor sentiment index of Baker and Wurgler (2006). The portfolios are held for 1 month and then they are rebalanced. The strategy distinctiveness index for each portfolio is computed as 1 minus the sample correlation of a portfolio's return with the return of a certain style. Three styles are examined using the corresponding Credit Suisse Index style returns: long/short equity (Panel A), equity market neutral (Panel B), and all equity hedge funds combined (Panel C). Upside SDI measures strategy distinctiveness at times when the returns of the corresponding style are positive, while downside SDI measures strategy distinctiveness at times when the returns of the corresponding style are negative. The p-values are reported below the SDI numbers. Only results for quintiles 1 and 5 within each momentum-crowd-timing coefficient are reported. The sample period is from January 1997 to December 2017.

Panel A: Credit Suisse Long/Short Equity Hedge Fund Index									
Quintiles		γ_1			γ_2			γ_3	
	SDI	upside SDI	downside SDI	SDI	upside SDI	downside SDI	SDI	upside SDI	downside SDI
Q1	45.38%	65.35%	48.78%	15.05%	30.15%	25.65%	38.16%	54.32%	33.99%
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q5	16.60%	29.14%	29.66%	42.69%	64.54%	32.88%	22.61%	47.91%	28.36%
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Panel B: Credit Suisse Equity Market Neutral Hedge Fund Index									
Quintiles		γ_1			γ_2			γ_3	
	SDI	upside SDI	downside SDI	SDI	upside SDI	downside SDI	SDI	upside SDI	downside SDI
Q1	80.71%	79.14%	85.37%	78.30%	76.30%	81.18%	82.83%	81.45%	90.51%
	0.00	0.00	0.26	0.00	0.01	0.34	0.01	0.01	0.46
Q5	77.34%	73.93%	85.14%	79.39%	80.39%	87.72%	72.56%	67.27%	78.82%
	0.00	0.00	0.25	0.00	0.00	0.14	0.00	0.00	0.10
Panel C: Credit Suisse Hedge Fund Index									
Quintiles		γ_1			γ_2			γ_3	
	SDI	upside SDI	downside SDI	SDI	upside SDI	downside SDI	SDI	upside SDI	downside SDI
Q1	50.71%	67.23%	59.95%	28.99%	43.20%	41.13%	45.60%	60.80%	52.41%
	0.00	0.00	0.0003	0.00	0.00	0.00	0.00	0.00	0.00
Q5	26.49%	41.60%	49.23%	46.82%	65.86%	56.82%	29.44%	50.09%	45.94%
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 12. **Hedge Fund Returns Sorted by Momentum Crowd Timing Ability: Different States of Cross-Sectional Return Dispersion**

This table presents the out-of-sample performance gap between quintiles 5 and 1 within each momentum-crowd-timing coefficient category, conditional on high and low cross-sectional dispersion states. A given month is classified as a high cross-sectional dispersion month if current return dispersion is above its median up to that month. Similarly, a month is classified as a low cross-sectional dispersion month if current return dispersion is below its median up to that month. Each month, and for each timing coefficient, I form 5 quintile portfolios based on hedge funds' momentum-crowd-timing coefficients estimated from the past 36 months using this model:

$$R_{i,t} = \alpha_i + \gamma_{0i} R_t^{mom^*} + \gamma_{1i} R_t^{mom^*}(SH_t - \overline{SH}) + \gamma_{2i} R_t^{mom^*}(PF_t - \overline{PF}) + \gamma_{3i} R_t^{mom^*}(BW_t - \overline{BW}) + \sum_{j=1}^{J} \beta_j f_{j,t} + e_{i,t} \quad (26)$$

where $R_{i,t}$ is the excess return of the fund, $R_t^{mom^*}$ is the return of the scaled momentum strategy, $f_{j,t}$ stands for the Fung-Hsieh hedge fund factors, SH is a measure of the crowd of arbitrageurs trading on momentum, PF measures the crowd of positive-feedback traders, and BW is the investor sentiment index of Baker and Wurgler (2006). The portfolios are held for 1 month and then they are rebalanced. The t-statistics are reported below the returns and they are calculated based on Newey-West heteroskedasticity and autocorrelation-consistent standard errors. All returns are in % per month. The sample period is from January 1997 to December 2017.

Quintiles	γ_1	γ_2	γ_3	γ_1	γ_2	γ_3
	High CS Dispersion			Low CS Dispersion		
Q5-Q1	-0.80%	0.54%	-0.68%	-0.58%	0.30%	-0.54%
	(-2.33)	(2.76)	(-2.85)	(-2.43)	(2.14)	(-2.40)

Table 13. **Hedge Fund Returns Sorted by Momentum Crowd Timing Ability: Excluding 1999-2000 and 2008-2009**

This table presents the out-of-sample average returns and Fung-Hsieh alphas for portfolios of hedge funds sorted by momentum-crowd-timing coefficients. Each month, and for each timing coefficient, I form 5 quintile portfolios based on hedge funds' momentum-crowd-timing coefficients estimated from the past 36 months using this model:

$$R_{i,t} = \alpha_i + \gamma_{0i} R_t^{mom^*} + \gamma_{1i} R_t^{mom^*}(SH_t - \overline{SH}) + \gamma_{2i} R_t^{mom^*}(PF_t - \overline{PF}) + \gamma_{3i} R_t^{mom^*}(BW_t - \overline{BW}) + \sum_{j=1}^{J} \beta_j f_{j,t} + e_{i,t} \quad (27)$$

where $R_{i,t}$ is the excess return of the fund, $R_t^{mom^*}$ is the return of the scaled momentum strategy, $f_{j,t}$ stands for the Fung-Hsieh hedge fund factors, SH is a measure of the crowd of arbitrageurs trading on momentum, PF measures the crowd of positive-feedback traders, and BW is the investor sentiment index of Baker and Wurgler (2006). The portfolios are held for 1 month and then they are rebalanced. The t-statistics are reported below the returns and they are calculated based on Newey-West heteroskedasticity and autocorrelation-consistent standard errors. All returns are in % per month. The last two rows of the table present the return difference, and its significance, between the top and bottom quintiles in each γ coefficient category. The sample period is from January 1997 to December 2017, but it excludes the periods January 1999-December 2000 and January 2008-December 2009.

Quintiles	γ_1		γ_2		γ_3	
	Ave Ret	Alpha	Ave Ret	Alpha	Ave Ret	Alpha
Q1	1.14	0.94	0.59	0.40	1.09	0.97
	(3.16)	(2.55)	(3.02)	(2.36)	(3.12)	(2.71)
Q2	0.66	0.44	0.61	0.43	0.70	0.49
	(4.76)	(3.61)	(4.67)	(3.88)	(5.25)	(4.22)
Q3	0.61	0.41	0.58	0.40	0.62	0.43
	(4.87)	(3.78)	(4.47)	(3.56)	(5.17)	(4.19)
Q4	0.57	0.40	0.68	0.49	0.58	0.40
	(3.80)	(3.15)	(4.62)	(3.75)	(3.73)	(2.99)
Q5	0.54	0.40	1.07	0.89	0.54	0.32
	(2.92)	(2.62)	(3.06)	(2.53)	(2.66)	(1.86)
Q5-Q1	-0.60	-0.54	0.47	0.49	-0.55	-0.65
	(-2.86)	(-2.59)	(2.55)	(2.54)	(-2.76)	(-2.01)

Table 14. **Hedge Fund Returns Sorted by Momentum Crowd Timing Ability: All Equity Category Funds**

This table presents the out-of-sample average returns and Fung-Hsieh alphas for portfolios of hedge funds sorted by momentum-crowd-timing coefficients. Each month, and for each timing coefficient, I form 5 quintile portfolios based on hedge funds' momentum-crowd-timing coefficients estimated from the past 36 months using this model:

$$R_{i,t} = \alpha_i + \gamma_{0i} R_t^{mom^*} + \gamma_{1i} R_t^{mom^*}(SH_t - \overline{SH}) + \gamma_{2i} R_t^{mom^*}(PF_t - \overline{PF}) + \gamma_{3i} R_t^{mom^*}(BW_t - \overline{BW}) + \sum_{j=1}^{J} \beta_j f_{j,t} + e_{i,t} \quad , \tag{28}$$

where $R_{i,t}$ is the excess return of the fund, $R_t^{mom^*}$ is the return of the scaled momentum strategy, $f_{j,t}$ stands for the Fung-Hsieh hedge fund factors, SH is a measure of the crowd of arbitrageurs trading on momentum, PF measures the crowd of positive-feedback traders, and BW is the investor sentiment index of Baker and Wurgler (2006). The portfolios are held for 1 month and then they are rebalanced. The t-statistics are reported below the returns and they are calculated based on Newey-West heteroskedasticity and autocorrelation-consistent standard errors. All returns are in % per month. The last two rows of the table present the return difference, and its significance, between the top and bottom quintiles in each γ coefficient category. The categories of hedge funds that are included in the analysis are long-short equity, equity market neutral, multi-strategy, convertible arbitrage, event-driven, global macro, and funds of funds. The sample period is from January 1997 to December 2017.

Quintiles	γ_1		γ_2		γ_3	
	Ave Ret	Alpha	Ave Ret	Alpha	Ave Ret	Alpha
Q1	0.80	0.56	0.52	0.26	0.78	0.56
	(4.75)	(3.39)	(3.39)	(2.05)	(4.50)	(3.50)
Q2	0.55	0.30	0.44	0.27	0.54	0.31
	(5.82)	(3.85)	(2.51)	(2.02)	(5.79)	(3.88)
Q3	0.54	0.29	0.48	0.24	0.47	0.25
	(2.57)	(2.06)	(5.24)	(3.22)	(5.43)	(3.44)
Q4	0.49	0.26	0.55	0.33	0.48	0.24
	(4.45)	(2.90)	(5.18)	(3.64)	(2.51)	(2.02)
Q5	0.45	0.21	0.78	0.55	0.52	0.25
	(2.69)	(1.59)	(4.41)	(3.24)	(3.13)	(1.73)
Q5-Q1	-0.36	-0.35	0.26	0.29	-0.25	-0.32
	(-2.59	(-2.49)	(2.96)	(2.12)	(-1.96)	(-2.41)

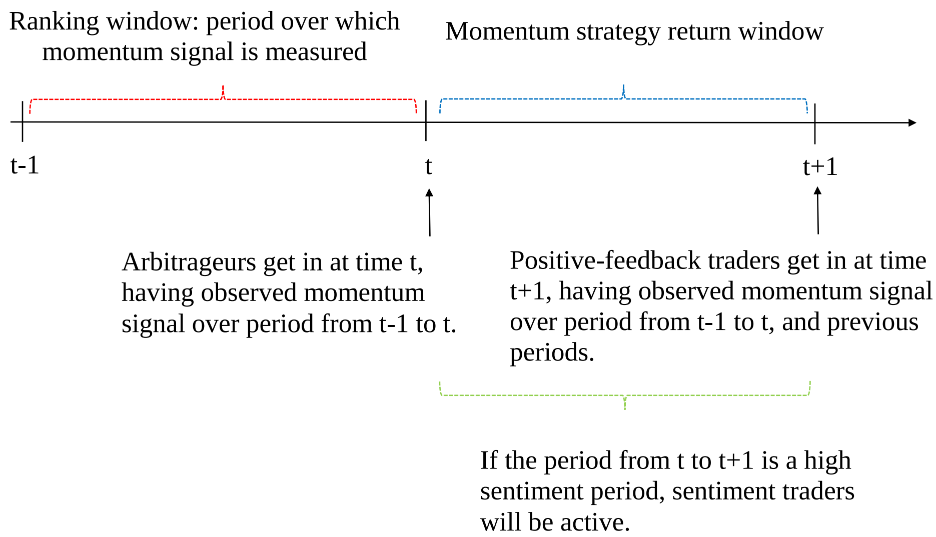

Figure 1. **Time Line of Momentum.**
This figure presents a time line for momentum signals, momentum strategy returns, and the trading demand of various investors.

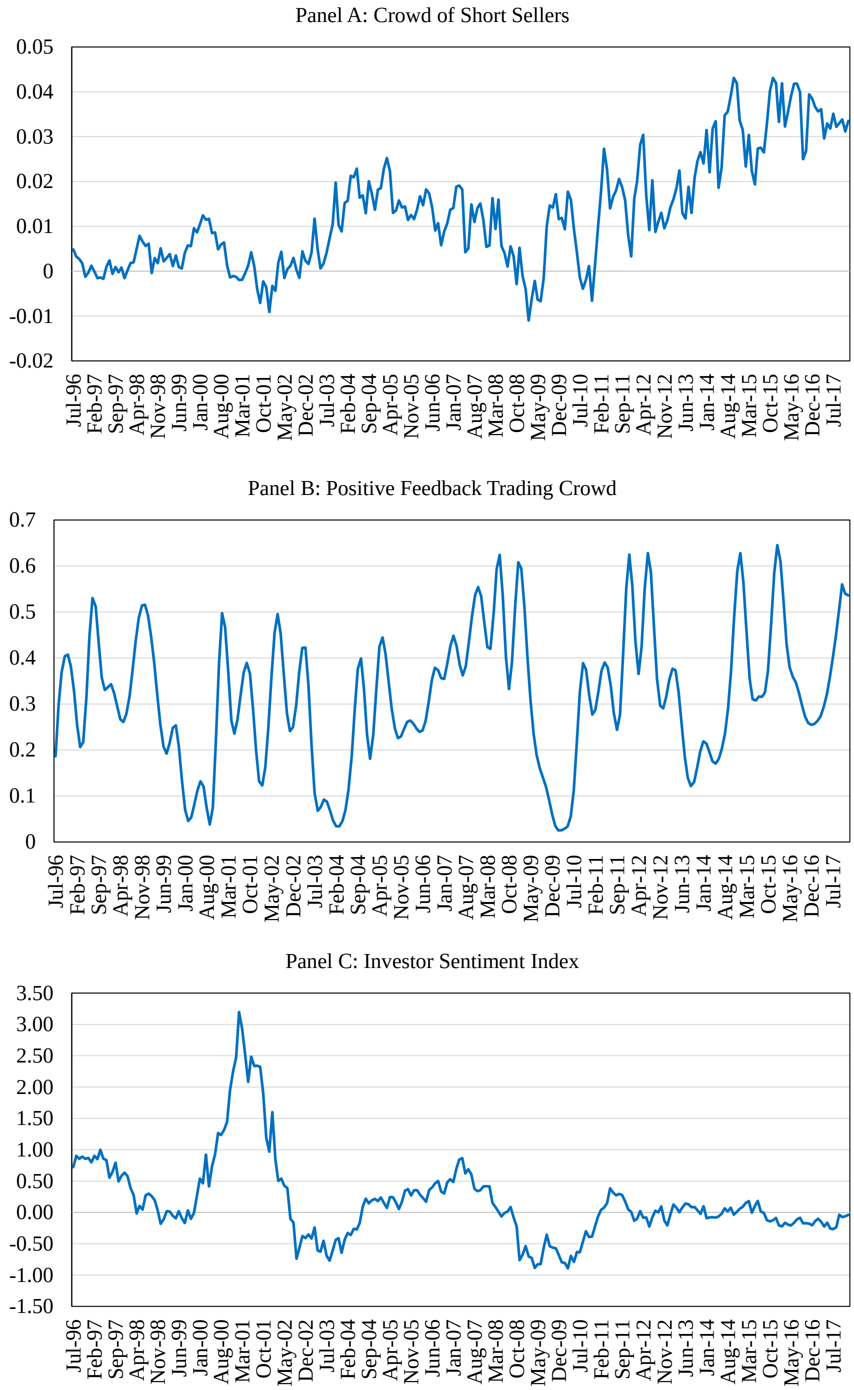

Figure 2. **Momentum Crowd Timing Signals.**

This figure plots the time series of three measures that proxy for the size of the momentum crowd. Panel A plots the time series of the crowd of arbitrageurs who are trading on momentum. It is measured by the strength of the cross-sectional relationship between short interest and the momentum signal at each point in time (model (6) in the paper). Panel B plots the time series of the crowd of positive-feedback traders. It is measured by the strength of the cross-sectional relationship between excess demand by mutual funds and past return performance (model (9) in the paper). Panel C plots the time series of the investor sentiment index of Baker and Wurgler (2006), which is used as a proxy for the size of the crowd of sentiment traders. The sample period is from January 1997 to December 2017.

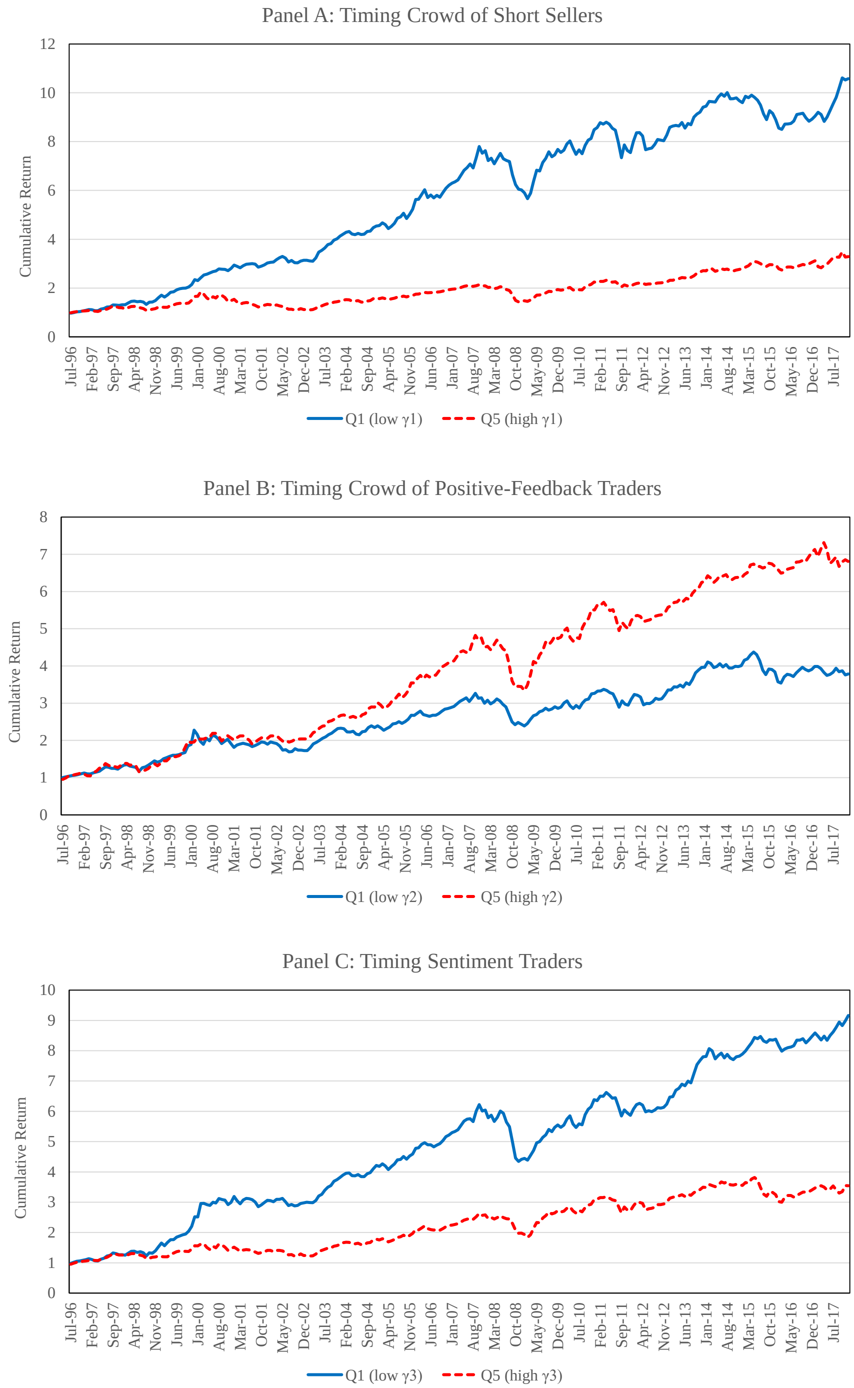

Figure 3. **Cumulative Returns of Investing in Hedge Funds with Different Timing Ability.**
This figure plots cumulative returns of hedge fund portfolios sorted by momentum-crowd-timing ability. Each month starting from January 1997, and for each timing coefficient, 5 quintile portfolios are formed based on hedge funds' momentum-crowd-timing coefficients estimated from the past 36 months using the following model:

$$R_{i,t} = \alpha_i + \gamma_{0i} R_t^{mom^*} + \gamma_{1i} R_t^{mom^*}(SH_t - \overline{SH}) + \gamma_{2i} R_t^{mom^*}(PF_t - \overline{PF}) + \gamma_{3i} R_t^{mom^*}(BW_t - \overline{BW}) + \sum_{j=1}^{J} \beta_j f_{j,t} + e_{i,t}, \quad (29)$$

where $R_{i,t}$ is the excess return of the fund, $R_t^{mom^*}$ is the return of the scaled momentum strategy, $f_{j,t}$ stands for the Fung-Hsieh hedge fund factors, SH is a measure of the crowd of arbitrageurs trading on momentum, PF measures the crowd of positive-feedback traders, and BW is the investor sentiment index of Baker and Wurgler (2006). The portfolios are held for 1 month and then they are rebalanced. Each panel presents the cumulative returns of quintiles 1 and 5 within each γ coefficient category. The sample period is from January 1997 to December 2017.