Journal of Financial Economics 122 (2016) 544–567

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Short selling meets hedge fund 13F: An anatomy of informed demand[☆]

Yawen Jiao[a], Massimo Massa[b], Hong Zhang[c,*]

[a] School of Business Administration, University of California, Riverside, Anderson Hall 0129, 900 University Avenue, Riverside, CA 92521, USA
[b] INSEAD, Boulevard de Constance, 77305 Fontainebleau Cedex, France
[c] PBC School of Finance, Tsinghua University, 43 Chengfu Road, Haidian District, Beijing 100083, PR China

ARTICLE INFO

Article history:
Received 16 August 2015
Revised 30 November 2015
Accepted 29 January 2016
Available online 16 September 2016

JEL codes:
G20
G14

Keywords:
Short selling
Hedge funds
13F
Informed demand
Hedging

ABSTRACT

The existing literature treats the short side (i.e., short selling) and the long side of hedge fund trading (i.e., fund holdings) independently. The two sides, however, complement each other: opposite changes in the two are likely to be driven by information, whereas simultaneous increases (decreases) of the two may be motivated by hedging (unwinding) considerations. We use this intuition to identify informed demand and document that it exhibits highly significant predictive power over returns (approximately 10% per year). We also find that informed demand forecasts future firm fundamentals, suggesting that hedge funds play an important role in information discovery.

1. Introduction

Hedge funds are known to both buy and short sell stocks on a large scale. Changes in hedge fund holdings and fluctuations in short interest are therefore largely two facets of the same phenomenon: hedge fund trading.[1] The literature, however, treats these two variables separately, as if they were independent phenomena. For instance, the short selling literature investigates whether short sellers are informed and whether they can help improve price efficiency (e.g., Senchack and Starks, 1993; Asquith and Meulbroek, 1995; Aitken, Frino, McCorry, and Swan, 1998;

☆ We would like to thank Douglas Breeden, Charles Cao, Hui Chen, Bernard Dumas, Philip Dybvig, Paul Gao, Zhiguo He, John Griffin, Mariassunta Giannetti, Zhiguo He, Pierre Hillion, Soren Hivkiar, Jennifer Huang, Marcin Kacperczyk, Bing Liang, Dong Lou, Stavros Panageas, Lubos Pastor, Neil Pearson, Lasse Pedersen, Philip Strahan, Nagpurnanand Prabhala, David Reeb, Matthew Spiegel, Hao Zhou, Lu Zheng and seminar participants at CKGSB, INSEAD, National University of Singapore, and PBC School of Finance.

* Corresponding author. Fax: +86-10 6279 9557.
E-mail addresses: yawenj@ucr.edu (Y. Jiao), massimo.massa@insead.edu (M. Massa), zhangh@pbcsf.tsinghua.edu.cn (H. Zhang).

[1] The recent attack of foreign-listed Chinese stocks, which leads the Bloomberg China Reverse Merger Index tracking 82 NYSE-listed Chinese companies to have "tumbled 68% from its peak at the start of 2010" (*Financial Times*, April 10, 2012), illustrates the important role that hedge funds play in the short selling market (see Ljungqvist and Qian, 2014, for details). Indeed, the notion that hedge funds take both long and short positions at a large scale is widely supported by practitioners. BarclayHedge, for instance, claims that the long–short strategy of "taking long positions in stocks that are expected to increase in value and short positions in stocks that are expected to decrease in value" is "used primarily by hedge funds" and that "hedge funds … simply do this on a grander scale" (www.barclayhedge.com). Goldman Sachs has estimated in its report "Hedge fund trend monitor" that hedge funds accounted for 85% of total short interest positions, or $361 billion, as of December 31, 2009.

http://dx.doi.org/10.1016/j.jfineco.2016.09.001

Cohen, Diether, and Malloy, 2007; Boehmer, Jones, and Zhang, 2008; Boehmer, Huszar, and Jordan, 2010; Saffi and Sigurdsson, 2011; Hirshleifer, Teoh, and Yu, 2011; Akbas, Boehmer, Erturk, Sorescu, 2013; Boehmer and Wu, 2013). The question of whether managers are informed and whether they can deliver superior performance is also at the core of the analysis of the hedge fund industry (e.g., Fung and Hsieh, 1997; Ackermann, McEnally, and Ravenscraft, 1999; Agarwal and Naik, 2004; Getmansky, Lo, and Makarov, 2004; Kosowski, Naik, and Teo, 2007; Agarwal, Daniel, and Naik, 2009, 2011; Aragon and Nanda, 2012; Sun, Wang, and Zheng, 2012; and Cao, Chen, Liang, and Lo, 2013, just to name a few).

However, a *joint* analysis of hedge funds and short selling—for instance, regarding changes in both hedge fund holdings and short interest—is lacking in the literature. This inattention is surprising because joint information is needed in many situations to understand motivations for hedge fund trading. Consider, for instance, the case in which aggregate hedge fund ownership of a specific stock increases. While such a “net buy” may be driven by private information that predicts positive changes in stock prices, it may also arise because of hedging—e.g., hedge fund managers use the long position to hedge the systematic risk of their arbitrage strategy. It is not surprising, therefore, that changes in hedge fund ownership have not been found to be informative ex ante (e.g., Griffin and Xu, 2009), which may simply reflect the prevalence of the second (hedging) effect. In the presence of both hedging and information-driven trading motivations, therefore, assessments of the informational content of hedge fund trading can hardly be complete if we focus only on one class of trades.

In this paper, we bridge this gap by proposing a novel approach that *jointly* considers short selling and hedge fund holdings to differentiate between various trading motivations. Returning to the previous example, if short interest decreases over the *same* period in which aggregate hedge fund ownership increases, hedge funds as a whole are likely to trade on a positive signal, which we refer to as *informed long demand*. When the opposite trading pattern occurs, i.e., when short interest increases over the *same* period in which hedge fund ownership decreases, the trading reflects *informed short demand*. By contrast, a simultaneous increase (decrease) in both short interest and hedge fund ownership may occur when hedge funds use both the long and the short sides to form arbitrage portfolios (or to unwind existing arbitrage positions), which we can loosely refer to as *hedging* (*unwinding*) demand.[2] Given that the direction of the signals for *hedging/unwinding* demand cannot be easily identified ex ante, it is critical to focus on *informed long/short demand* to properly assess the informativeness of hedge fund trading.

This novel identification strategy allows us to shed new light on the informational content of hedge fund trading using information from both hedge fund 13F filings and short selling information for the complete list of U.S. stocks for the period from 2000 to 2012. Because we observe only aggregate information regarding short selling activities for each stock (rather than how each hedge fund conducts short selling), we aggregate hedge fund ownership at the stock level, so that the two sides of information can be used jointly to infer informed demand at the stock level. We proceed in three steps.

In the first step, we examine the predictive power of informed demand for out-of-sample abnormal returns. We find strong evidence that informed long (short) demand is associated with positive (negative) out-of-sample abnormal stock returns, suggesting that such demand is indeed informative. The economic magnitude is sizable. For instance, if we define informed long (short) demand as a dummy variable that takes a value of one when changes in short interest and hedge fund holdings belong to the most positive (negative) quintiles of stocks in the same period, we find that this proxy is related to a 6.6% (−3.2%) annualized abnormal return in the next quarter under the traditional Fama-MacBeth specifications. In other words, stocks characterized by informed long demand outperform stocks characterized by informed short demand by as much as 9.8% per year. If we directly construct portfolios, rebalanced at quarterly frequency, that buy/sell stocks with the top 20% informed long/short demand, the abnormal return over the entire sample period is approximately 10.5% per year. This magnitude is on par with that obtained in the regression analysis.

We further confirm that return predictability identified in this way is not affected by the value premium, the size premium, or momentum. Neither is it spuriously generated by various more recently documented anomalies associated with the ratio of gross profit to assets (Novy-Marx, 2013), operating profit (Fama and French, 2015), asset growth (Cooper, Gulen, and Schill, 2008), investment growth (Hou, Xue, and Zhang, 2015), net stock issuance (Xing, 2008), accruals (Fama and French, 2008), and the logarithm of net operating assets (Hirshleifer, Hou, Teoh, and Zhang, 2004). Rather, the return predictability of informed demand appears to arise from very different economic considerations than known asset pricing anomalies.[3]

In the second step, therefore, we investigate potential economic channels through which informed demand achieves its predictive power. For this purpose, we first split the sample into two subgroups based on a list of firm characteristics, including market capitalization, turnover ratio, analyst coverage, and dispersion of analyst forecasts. We then perform the return predictability test for

[2] Alternatively, one can also view the long side and short side of trading as coming from two different groups of traders and interpret *hedging* demand as a situation in which the two groups have different opinions regarding expected stock returns. The interpretation of our main results, however, remains the same.

[3] Our results are also robust to the use of different cutoff points for the definition of positive (negative) short interest and hedge fund holding changes, the use of different out-of-sample windows, and the inclusion of controls for hedging (unwinding) demand. In addition, placebo tests show that hedging and unwinding demands, unlike informed demand, do not exhibit consistent predictive power for returns, especially over a one-year horizon—and asset pricing anomalies absorb the remaining significance. Finally, we observe that, consistent with Griffin and Xu (2009), hedge fund holding information alone has insignificant return predictive power, which further confirms the empirical importance of combining hedge fund information with short selling information.

each subsample of stocks. We find that return predictability is more significant for stocks with high market capitalization, a high turnover ratio, high analyst coverage, and high analyst dispersion. The association with the first three characteristics suggests that our findings are unlikely to be driven by (small) size-related firm characteristics, (low) liquidity-related market conditions, or (low) analyst coverage-related public information, whereas the association with the last characteristic suggests that improved information processing (e.g., Kim and Verrecchia, 1994; Engelberg, Reed, and Ringgenberg, 2012) could play a role in the predictive power of informed demand.

Motivated by this finding, we further explore the informational content of informed demand by examining the extent to which it can predict firm fundamentals, especially those unexpected by the market. Following Akbas, Boehmer, Erturk, Sorescu (2013), we consider several types of proxies for firm fundamentals. The first is a proxy for the future (real) performance of firms, which is proxied by future returns on assets (ROA) or future changes in ROA, where ROA can be either adjusted or unadjusted by industry peers. The second relates to the unexpected component of earnings, measured by standardized unexpected earnings (SUE). In addition, we investigate whether informed demand can predict the future behavior of other market participants, including revisions of recommendations by analysts (Analyst revision) and responses of the public to unexpected news about firm-level fundamentals, where the latter is proxied by cumulative abnormal returns around earnings announcements (CARs).

We find that informed demand has significant forecasting power for all of the above measures, suggesting that the savvy traders behind such demand are not only well informed about firm-level financial information (ROA, SUE) but also sufficiently sophisticated to predict analyst revisions and market reactions to firm-level information. Jointly, these results imply that the predictive power of informed demand may come from the discovery of information about firm fundamentals above and beyond what the market or even analysts know. Hence, return predictability documented in previous tests could be directly related to hedge fund managers' superior ability to process firm-level information. Hedging demand and unwinding demand, by contrast, do not exhibit similar forecasting power.

However, if return predictability arises from managerial skill, skillful managers should be able to deliver persistent performance at the fund level. Indeed, persistent performance is the key way to validate managerial skill. Our next task, therefore, is to examine whether performance associated with informed demand is persistent at the fund level. To this end, we quantify the performance of informed trading of a particular fund as the abnormal return that can be generated by stocks characterized by informed long demand, as implied by fund-specific holdings. We find that hedge funds that have higher in-sample performance rankings (measured, e.g., over the past 12-month period) also deliver better performance or have higher performance rankings out of sample (e.g., over the next quarter). Moreover, the top 10% (20%) of funds significantly outperform the bottom 10% (20%) of funds in the future. In other words, the performance of informed demand is highly persistent for the subset of the top-performing funds, suggesting that the return predictability of informed demand may indeed be associated with managerial skill. By contrast, hedging demand is incapable of generating out-of-sample performance. This insignificance implies that persistent hedge fund skills could be more associated with superior firm-specific information than other sources of information.

In the final step of our analysis, we implement several additional tests to further enrich our economic intuition and to rule out several alternative explanations for this return predictability. We first note that short interest changes are measured at the quarterly frequency to match that of hedge fund holding changes. Although this design suffices to demonstrate the value of a *joint* analysis of long and short positions, in practice information from the long side may have a longer duration than that of the short side.[4] Especially, quarterly short interest changes may not fully capture the informativeness of the shorts. Consistent with this notion, we document that monthly information can be used to further enhance the predicting power of the short side.

We then validate the importance of the hedge fund industry in processing information by conducting a placebo test in which we replace hedge fund holdings with mutual fund holdings. We find little evidence of return predictability in this case, confirming that it is reasonable to combine short selling with hedge fund holdings rather than holdings of other institutional investors, such as mutual funds, in analyzing informed demand.

Next, we consider whether return predictability can be related to hedge fund strategies that target mutual fund flows rather than firm-level information. Indeed, hedge funds may trade to take advantage of the price impact of mutual fund inflows/outflows on stocks (e.g., Shive and Yun, 2013; Arif, Ben-Rephael, and Lee, 2014), which may lead to return predictability. We find that informed long (short) demand appears to be negatively (positively) correlated with mutual fund flows, which is the opposite of what a strategy of riding the price impact of mutual fund flows would suggest. On the other hand, we do find evidence that hedge funds *unwind* their positions—likely due to risk management motives—before the occurrence of mutual fund fire sales, suggesting that hedge funds may use different strategies to benefit from mutual fund flows and private information.

Another potential alternative explanation is that return predictability may arise due to the compensation associated with liquidity provision. However, we find, for instance, that informed short demand is related neither to turnover nor to the Amihud illiquidity measure. This finding, together with the previous finding that return predictability is more prominent for larger and more liquid stocks, suggests that the return predictability of hedge fund trading is unlikely to be related to liquidity provision.

Overall, these findings support the intuition that the joint analysis of short selling and hedge fund holding changes is crucial in revealing the trading motivations of

[4] We thank the anonymous referee for this insightful comment.

perhaps the most sophisticated/informed investors in the market. Building on this intuition, our tests further document that a key component of their informativeness may arise from information discovery regarding firm fundamentals, which subsequently affects dissemination of information in the financial markets.

To the best of our knowledge, we are the first to propose such a joint analysis of short selling and hedge fund holdings and to link it to fundamental stock analysis. Our findings shed new light on the informational content of both short sellers (e.g., Senchack and Starks, 1993; Asquith and Meulbroek, 1995; Aitken, Frino, McCorry, and Swan, 1998; Cohen, Diether, and Malloy, 2007; Boehmer, Jones, and Zhang, 2008; Saffi and Sigurdsson, 2011; Akbas, Boehmer, Erturk, and Sorescu, 2013; Boehmer and Wu, 2013) and hedge fund managers (e.g., Fung and Hsieh, 1997; Ackermann, McEnally, and Ravenscraft, 1999; Agarwal and Naik, 2004; Getmansky, Lo, and Makarov, 2004; Kosowski, Naik, and Teo, 2007; Agarwal, Daniel, and Naik, 2009, 2011; Aragon and Nanda, 2012; Sun, Wang, and Zheng, 2012; Cao, Cheng, Liang, and Lo, 2013).

Our paper is closely related to Griffin and Xu (2009), which we extend by proposing that the use of information from short selling is necessary to complement holdings-based information in order to identify informed demand shocks that are otherwise hidden among various trading motivations. Chen, Da, and Huang (2015) link the difference between abnormal hedge fund holdings and abnormal short interest to the profitability of anomalies, finding that the former reduces mispricing. We differ in proposing a more flexible empirical framework to understand various trading motivations and in documenting that the return predictability of informed demand may arise from its predictive power vis-à-vis firm fundamentals. Such return predictability can be interpreted as an explicit type of managerial skill in the hedge fund industry.

The remainder of the paper is organized as follows. Section 2 presents the data that we employ and the main variables constructed for the analysis. Section 3 describes the main empirical findings. Section 4 relates informed demand to firm fundamentals and discusses the implications of the findings. Section 5 presents additional tests and robustness tests, and a brief conclusion follows.

2. Data and construction of the variables

The data that we use are compiled from various databases. We first retrieve hedge fund holding information from 13F filings from the Securities and Exchange Commission (SEC). Since 1978, institutional investors with at least one hundred million U.S. dollars under management have been required to file 13F forms with the SEC each quarter for U.S. equity holdings of more than two hundred thousand dollars or more than ten thousand shares. This regulation allows us to construct holding or ownership data for each stock based on aggregations of various types of institutional investors.

The identities of the hedge funds, which are collected from the Thomson Reuters Institutional Holdings (13F) database, are cross-referenced with 13F filings from the FactSet LionShares database. As noted by Ben-David, Franzoni, Landier, and Moussawi (2013), the hedge fund list identified in the Thomson Reuters 13F database is consistent with the FactSet LionShares identifications of hedge fund companies. We identify hedge funds in the Thomson Reuters 13F database as follows. Institutional investors are divided into five types in this database: 1) bank trust departments, 2) insurance companies, 3) investment companies and their managers, 4) independent investment advisers, and 5) others. We exclude institutions classified as type 1 or type 2.[5] For each remaining institution, we manually check its SEC ADV forms. Following Brunnermeier and Nagel (2004) and Griffin and Xu (2009), we require an institution to have more than 50% of its investments listed as “other pooled investment vehicles,” including private investment companies, private equity, and hedge funds, or more than 50% of its clients listed as “high net worth individuals” for inclusion in our hedge fund sample. We also require that institutions charge performance-based fees to be included in the hedge fund sample. Finally, we manually check the website of each institution satisfying the above requirements to confirm that its primary business is hedge fund-related activity.[6]

Although our sample can be extended to earlier periods, we focus on the post-2000 period because the number of hedge funds in 13F filings became reasonably large only toward the end of the 1990s. Furthermore, the destabilizing effects of hedge funds on stock prices during the tech-bubble period of the late 1990s are well documented by Brunnermeier and Nagel (2004) and Griffin, Harris, Shu, and Topaloglu (2011). We must therefore avoid the confounding effects associated with the tech-bubble period.

With regard to stocks, we start with all the publicly listed companies for which we have accounting and stock market information from Center for Research in Security Prices (CRSP)/Compustat. We then exclude American depositary receipts (ADRs) and stocks with incomplete information to construct control variables (as detailed below).[7] Finally, we match the remaining stocks with the hedge fund holdings and short interest data. Our final sample includes 5,357 stocks for the period from 2000 to 2012, invested in by 1,397 hedge fund holding companies that report quarterly equity holdings in 13F filings.

Our main variables are constructed as follows. First, to construct our main dependent variable for the return predictability tests, we obtain the quarterly return, $r_{i,t}$, for stock i in a given quarter t as the compound monthly returns reported by CRSP. Following Daniel, Grinblatt, Titman, and Wermers (DGTW, 1997), we compute the

[5] It is well-known that the type classification in the 13F database is inaccurate after 1998. However, the classification errors are almost entirely driven by misclassifying type 3 or 4 institutions as type 5 institutions (Lewellen, 2011); therefore, they do not affect our sample.

[6] Some of these institutions do not have websites. However, for most of them, we were able to determine whether they are hedge funds through a news search. The remaining institutions are included in the hedge fund sample because discussions with hedge fund managers indicate that some hedge funds are reluctant to maintain websites. Excluding these funds does not lead to qualitative changes in our results.

[7] Excluding penny stocks (stocks priced at less than $1/share) does not change our results. See Ince and Porter (2006) for a more detailed discussion of these screening criteria.

abnormal performance of a stock, which we refer to as the DGTW-adjusted return, as the return of the stock net of the return of its style benchmark based on its size, book-to-market, and prior-period return characteristics.[8] We then compute the quarterly DGTW-adjusted return for each stock, denoted as $DGTW3_{i,\,t}$, as the compound monthly DGTW-adjusted return of the stock in the quarter. We also compute the abnormal return over a 1-year horizon, denoted $DGTW12_{i,\,t}$, in a similar way.

Next, to construct our main independent variables, we compute short interest (*SI*) as the average monthly dollar value of short interest scaled by the total dollar value of all outstanding shares of the stock in the month, both of which are obtained from Compustat. Because our hedge fund holding data are available at quarter-end, we use $SI_{i,\,t}$ at the end of the quarter to extract quarterly changes in short interest. The use of the average short interest within a quarter leads to very similar results. Moreover, because our ultimate goal is to retrieve informed trading from both the long and short sides of trading at the stock level, we also aggregate hedge fund holdings to compute hedge fund ownership for each stock, which we label $HFOwn_{i,\,t}$ for stock *i* in a given quarter *t*.

We define *informed long demand* as a dummy variable, $DLong_{i,\,t}$, that takes a value of one when hedge fund ownership increases from quarter $t-1$ to quarter t and short interest decreases over the same period and zero otherwise. That is,

$$DLong_{i,t} = I\{\Delta HFOwn_{i,t} > 0\} \times I\{\Delta SI_{i,t} < 0\},$$

where $I\{.\}$ is an indicator function, and $\Delta HFOwn_{i,\,t} = HFOwn_{i,\,t} - HFOwn_{i,\,t-1}$ and $\Delta SI_{i,\,t} = SI_{i,\,t} - SI_{i,\,t-1}$ denote changes in hedge fund holdings and short interest, respectively.

Similarly, *informed short demand* is defined as a dummy variable, $DShort_{i,\,t}$, that takes a value of one when hedge fund ownership decreases from quarter $t-1$ to quarter t and short interest increases over the same period and zero otherwise, i.e., $DShort_{i,\,t} = I\{\Delta HFOwn_{i,\,t} < 0\} \times I\{\Delta SI_{i,\,t} > 0\}$.

In addition to informed demand, we also define *hedging* (*unwinding*) *demand* as a simultaneous increase (decrease) in both hedge fund ownership and short interest, denoted $DHedge_{i,\,t} = I\{\Delta HFOwn_{i,\,t} > 0\} \times I\{\Delta SI_{i,\,t} > 0\}$ ($DUnwind_{i,\,t} = I\{\Delta HFOwn_{i,\,t} < 0\} \times I\{\Delta SI_{i,\,t} < 0\}$). Unwinding demand can be triggered by the need to liquidate existing trading positions to lock in profits or by fire sales. These two variables can not only provide placebo tests to validate the informational content of *informed demand*, but also enrich our understanding regarding various strategies adopted by the hedge fund industry, as later sections will show.

A second, alternative, way to define informed demand is to sort stocks into terciles according to $\Delta HFOwn_{i,\,t}$ or $\Delta SI_{i,\,t}$ and then to define informed long (short) demand as a dummy variable that takes a value of one if the stock's $\Delta HFOwn_{i,\,t}$ belongs to the top (bottom) tercile and its $\Delta SI_{i,\,t}$ belongs to the bottom (top) tercile and zero otherwise. In other words, *informed long demand* can be defined as the simultaneous occurrence of both the "highest" increase in hedge fund holdings and the "highest" decrease in short interest, where the "highest" increase or decrease is defined on the basis of tercile values of $\Delta HFOwn_{i,\,t}$ and $\Delta SI_{i,\,t}$ in a given period. To avoid confusion, we refer to tercile-based *informed demand* variables as $DLong^{Ter}_{i,t}$ and $DShort^{Ter}_{i,t}$. Similarly, we also define *informed long (short) demand* on the basis of quintiles of $\Delta HFOwn_{i,\,t}$ and $\Delta SI_{i,\,t}$ values and denote it as $DLong^{Quin}_{i,t}$ ($DShort^{Quin}_{i,t}$). Unreported tests using quartile-based variables yield very similar results.

Tercile- or quintile-based proxies enable sharper identification based on more profitable information in the case of informed demand and stronger hedging motivations in the case of hedging demand. However, the previous proxies based on positive or negative changes in short interest and holdings (e.g., $DLong_{i,\,t}$ and $DShort_{i,\,t}$) are likely to be more representative—as more stocks are involved—yet less informative. We will therefore mainly rely on $DLong_{i,\,t}$ and $DShort_{i,\,t}$ to establish our main results. We will then verify that these results are robust to alternative definitions of informed demand and use quintile-based partitions to illustrate the economic magnitude of return predictability.

We also construct a set of control variables following Gompers and Metrick (2001). *DIV* is the dividend yield calculated as dividends divided by market capitalization; *Age* is the number of months since the stock first appeared in CRSP; and *Price* refers to the stock price per share. *Turnover* is the stock turnover rate (volume divided by shares outstanding) in the last month prior to the beginning of the quarter. *Vol* is the standard deviation of returns over the past 24 months. Finally, *SP500* is a dummy equal to one for stocks in the Standard & Poor's (S&P) 500 index and zero otherwise. We use natural log transformations of several of these variables (i.e., *LgAge, LgPrc, LgTurn, LgVol*) in the regressions to reduce the impact of outliers.

Table 1 provides summary statistics. First of all, to further demonstrate the notion that the long side (13F holdings) and the short side (short interest) could be largely the two facets of the same hedge fund industry, in Panel A, we report the year-by-year average positions and changes in hedge fund ownership and short interest (both are scaled by the total number of shares outstanding), and compare these summary statistics with those of mutual fund ownership. The comparison is based on the sample of stocks that have nonzero hedge fund ownership, short interest, mutual fund ownership, and non-missing price information. Models (1)–(3) tabulate the average positions of these variables; Models (4)–(6) report their changes, which essentially describe the capital flows in and out of their positions. Fig. 1 visualizes the time series of the average positions in Panel A and those for their changes in Panel B. There, levels and changes of hedge fund ownership and short selling are plotted in solid lines (left scale), while mutual fund ownership/ownership changes are plotted in dashed lines (right scale).

From both Panel A of Table 1 and Panel A of Fig. 1, we can see that both hedge fund ownership and short interest display a very similar pattern. Both of them had grown dramatically before the 2008–2009 global financial crisis.

[8] A detailed description and data are available at http://www.rhsmith.umd.edu/faculty/rwermers/ftpsite/DGTW/coverpage.htm.

Table 1
Summary statistics.

This table provides summary statistics for the main variables. Panel A tabulates the year-by-year information of hedge fund ownership and short interest. More specifically, for stocks that have nonzero hedge fund (HF) ownership, short interest, and mutual fund (MF) ownership and non-missing price information, the first three columns report the average ownership of hedge funds (in % with respect to the total number of shares outstanding), average short interest (in % with respect to the total number of shares outstanding), as well as the average ownership of mutual funds (in % with respect to the total number of shares outstanding) of stocks. The next three columns tabulate the year-by-year changes in these variables. Panel B reports the mean, median, standard deviation, and 10% and 90% quantile values for main variables. Panel C reports the correlation matrix for these variables. A detailed definition of these variables is provided in Appendix A.

Panel A: Hedge fund ownership and short-selling activities by years

Year	Hedge fund ownership	Short interest	Mutual fund ownership	HF ownership changes	Short interest changes	MF ownership changes
2000	2.38%	1.78%	16.80%	0.40%	0.07%	0.53%
2001	2.78%	2.31%	18.63%	0.40%	0.54%	1.83%
2002	3.19%	2.83%	21.24%	0.41%	0.51%	2.61%
2003	3.72%	3.25%	21.85%	0.53%	0.42%	0.61%
2004	5.15%	3.60%	21.50%	1.43%	0.35%	−0.35%
2005	6.45%	3.99%	22.00%	1.30%	0.39%	0.50%
2006	7.65%	4.78%	22.90%	1.20%	0.79%	0.89%
2007	9.04%	6.00%	23.45%	1.39%	1.22%	0.55%
2008	8.34%	6.61%	24.50%	−0.70%	0.61%	1.06%
2009	6.73%	4.60%	25.19%	−1.61%	−2.00%	0.69%
2010	6.79%	4.73%	24.87%	0.06%	0.12%	−0.32%
2011	7.10%	4.82%	25.65%	0.31%	0.09%	0.78%
2012	7.45%	4.56%	25.50%	0.34%	−0.26%	−0.16%

Panel B: Summary statistics of major variables

	Mean	Std Dev	10%	Median	90%
DLong	0.2465	0.4310	0	0	1
DShort	0.2405	0.4274	0	0	1
DHedging	0.2640	0.4408	0	0	1
DUnwinding	0.2418	0.4282	0	0	1
DGTW 3m	0.0044	0.2342	−0.2216	−0.0072	0.2254
DGTW 12m	0.0146	0.5298	−0.4404	−0.0291	0.4591
Div	0.0161	0.0398	0	0	0.0427
LgAge	234.71	198.89	53.00	171.00	480.00
LgPrc	25.93	40.64	3.58	19.20	51.99
LgTurn	0.1637	0.1536	0.0258	0.1186	0.3559
LgVol	0.1277	0.0806	0.0553	0.1105	0.2168
SP500	0.1586	0.3653	0	0	1

Panel C: Correlation matrix

	DLong	*DShort*	*DHedging*	*DUnwinding*	DGTW 3m	DGTW 12m	*Div*	*LgAge*	*LgPrc*	*LgTurn*	*LgVol*	*SP500*
DLong	1											
DShort	−0.3219	1										
	(0.0000)											
DHedging	−0.3426	−0.3371	1									
	(0.0000)	(0.0000)										
DUnwinding	−0.323	−0.3178	−0.3382	1								
	(0.0000)	(0.0000)	(0.0000)									
DGTW 3m	0.0171	−0.015	0.0088	−0.0109	1							
	(0.0000)	(0.0000)	(0.0023)	(0.0002)								
DGTW 12m	0.0179	−0.0154	0.0027	−0.0056	0.4649	1						
	(0.0000)	(0.0000)	(0.3547)	(0.0572)	(0.0000)							
Div	−0.0012	−0.0035	−0.003	0.0033	0.0053	0.0079	1					
	(0.6685)	(0.2216)	(0.2902)	(0.2467)	(0.0665)	(0.0075)						
LgAge	0.0162	0.0033	−0.0041	−0.0113	0.0038	0.0068	0.0846	1				
	(0.0000)	(0.2500)	(0.1572)	(0.0001)	(0.1834)	(0.0212)	(0.0000)					
LgPrc	0.0026	0.0041	0.0066	−0.0072	−0.0028	−0.0086	−0.0357	0.1599	1			
	(0.3585)	(0.1493)	(0.0225)	(0.0126)	(0.3337)	(0.0035)	(0.0000)	(0.0000)				
LgTurn	−0.0238	0.0107	−0.0173	0.0458	−0.015	−0.0104	−0.0276	−0.0219	0.0344	1		
	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0000)	(0.0004)	(0.0000)	(0.0000)	(0.0000)			
LgVol	−0.0071	−0.0076	−0.0233	0.0371	0.0093	0.0193	−0.0618	−0.1987	−0.2103	0.2294	1	
	(0.0134)	(0.0080)	(0.0000)	(0.0000)	(0.0012)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)		
SP500	0.0249	0.0101	−0.014	−0.0134	0.0063	0.0091	0.0298	0.4418	0.1882	0.1038	−0.1851	1
	(0.0000)	(0.0005)	(0.0000)	(0.0000)	(0.0281)	(0.0020)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	

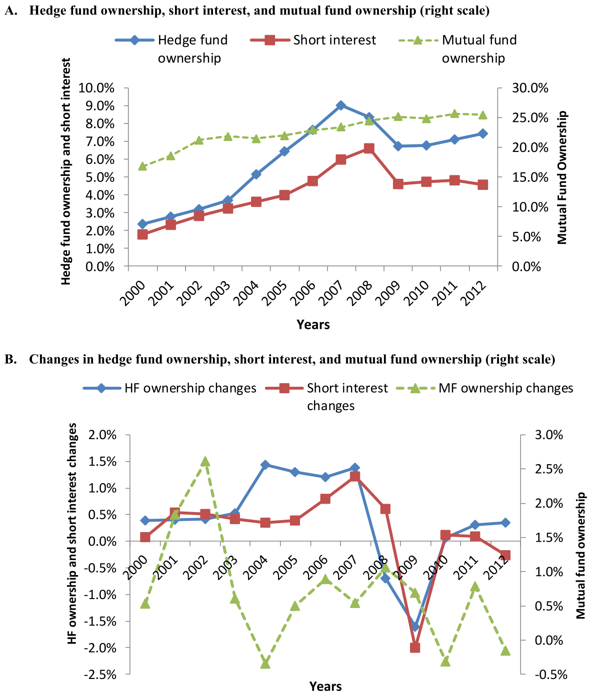

Fig. 1. Hedge fund ownership and short interest over time.
This figure demonstrates the year-by-year information of hedge fund ownership, short interest, and mutual fund ownership. More specifically, for stocks that have nonzero hedge fund ownership, short interest, and mutual fund ownership and non-missing price information, Panel A plots the average ownership of hedge funds (in % with respect to the total number of shares outstanding; left scale), average short interest (in % with respect to the total number of shares outstanding; left scale), as well as the average ownership of mutual funds (in % with respect to the total number of shares outstanding; right scale) of stocks. Panel B plots the year-by-year changes in these variables. The two plots are based on our final sample of 5,357 stocks for the period from 2000 to 2012, invested in by 1,397 hedge fund holding companies.

The occurrence of the crisis, however, almost completely disrupted the upward trend of both of them with both curves declining sharply during the year 2009. After 2009, both curves exhibited much smaller fluctuations. In contrast, the growth path of mutual fund ownership has been much smoother and very different from that of hedge fund ownership and short interest.

Panel B of Fig. 1 further illustrates that changes in hedge fund ownership and short interest closely trace each other, while changes in mutual fund ownership behave in a very different way. Mathematically, the correlation between changes in hedge fund ownership and short interest is as high as 0.75, while the correlation between changes in mutual fund ownership and short interest is only 0.18. These patterns and numbers suggest that hedge fund and short selling activities are highly interconnected to each other, whereas activities of other institutional investors, such as mutual funds, are not as closely related.

In addition to time series information, we also examine the extent to which the pool of stocks that short sellers trade overlaps with either hedge fund holdings or mutual fund holdings in the cross section. In any given year, short selling activities typically occur in 91% of the stocks that have nonzero hedge fund ownership. By contrast, short selling happens in only 58% of the stocks that have nonzero mutual fund ownership. These numbers are highly stable over the last decade, suggesting that hedge funds and short sellers pay more attention to a similar list of stocks than mutual funds and short sellers do.

The resemblance between hedge funds and short selling both in the time series and in the cross section suggests that hedge funds are among the most important players in the short selling market. This notion is indeed widely supported by practitioners. Goldman Sachs, for instance, has estimated that the short selling market as of December 31, 2009 amounts to $425 billion (bn) in the U.S., in which hedge funds account for 85%, or $361 bn. Among the $361 bn short positions, $310 bn come from stock-level short selling activities (the rest are about short positions in Exchange-traded funds (ETFs) and market indices). These numbers suggest a dominant role that the hedge fund industry plays in the short selling market. In a similar manner, BarclayHedge claims that the long-short strategy of "taking long positions in stocks that are expected to increase in value and short positions in stocks that are expected to decrease in value" is "used primarily by hedge funds."

Panel B of Table 1 reports the summary statistics for our main variables in our final sample of 5,357 stocks for the period from 2000 to 2012. Our main proxies for informed demand, *DLong* and *DShort*, have an average value of approximately 25%, suggesting that informed long and short demand occur approximately 25% of the time. Next, the average abnormal quarterly return is approximately 0.44%, with a standard deviation of 23%. This distribution is similar to what is reported in the literature.

Panel C of Table 1 presents the correlation coefficients of these variables. The most interesting pattern is that abnormal returns are positively associated with informed long demand and negatively associated with informed short demand. This observation provides some initial evidence that informed demand could be related to stock returns. Of course, this pattern provides only preliminary and in-sample evidence of such a relation. In the next section, we examine the out-of-sample return predictability of informed demand.

3. Return predictability

In this section, we examine whether informed long demand or informed short demand can predict out-of-sample abnormal returns. We mainly rely on a multivariate specification and provide a portfolio-based analysis to verify the robustness of our findings.

3.1. Baseline specifications

We estimate the following baseline Fama-MacBeth regression at a quarterly frequency to detect the return predictability of informed demand:[9]

[9] In our main tests, all *t*-statistics are Newey and West adjusted with lagged quarterly or yearly information when quarterly and 12-month performance are used as the dependent variable, respectively. We have also examined various ways of computing the optimal number of lags—for

$$DGTW_{i,t+1} = \alpha_i + \beta_i \times Informed\ Demand_{i,t} + C \times M_{i,t} + \epsilon_{i,t+1}, \quad (1)$$

where $DGTW_{i,\ t+1}$ refers to the out-of-sample DGTW-adjusted abnormal return of stock *i* accumulated over quarter $t+1$; $Informed\ Demand_{i,\ t}$ is a vector of informed demand variables, including $DLong_{i,\ t}$ and $DShort_{i,\ t}$ in the lagged quarter; and $M_{i,\ t}$ stacks a list of control variables, including *DIV, LgAge, LgPrc, LgTurn, LgVol*, and *SP500*.[10]

The results are reported in Table 2. In Panel A, Models (1)–(3) provide the results of the baseline regression on the quarterly return predictability of $DLong_{i,\ t}$ and $DShort_{i,\ t}$. We find that, independently or jointly, $DLong_{i,\ t}$ forecasts positive abnormal returns and $DShort_{i,\ t}$ forecasts negative abnormal returns in the next quarter. The predictive power is highly statistically significant, which is consistent with the idea that these two variables capture the informed demand of hedge fund managers/short sellers.

Models (4)–(6) provide a preliminary analysis of the impact of hedging demand and unwinding demand ($DHedging_{i,\ t}$ and $DUnwind_{i,\ t}$). We find that hedging demand is typically associated with positive abnormal returns and that unwinding demand is typically associated with negative abnormal returns. However, the return predictability of hedging/unwinding demand is not as robust as that of informed demand. To illustrate this point, we report the return predictability of tercile-based (in Model 7) and quintile-based (Model 8) informed and hedging/unwinding demand. As mentioned above, demand variables that are defined in this way proxy for the more profitable information or more extreme hedging motivations of the hedge funds. In these two models, although (more profitable) informed demand still forecasts abnormal returns at an enlarged magnitude, the predictive power of (more extreme) hedging and unwinding demand becomes marginal, if not insignificant. This result confirms the minimal informational content of hedging demand.

Note that, while the informational content of hedging and unwinding demand is marginal, the directions of predicted future return appear more consistent with the long-side information than the short side. For instance, in Model (7), $DHedging_{i,\ t}$ and $DUnwind_{i,\ t}$ predict positive and negative abnormal return, respectively. Since the two variables are defined as simultaneous increases and decreases in both long and short positions, changes in long positions seem to dominate the short side in predicting future return. For instance, increases in long and short positions of $DHedging_{i,\ t}$ should predict positive and negative return, respectively, on their own. The empirical observation, however, is that $DHedging_{i,\ t}$ predicts positive return, which can happen only when the return predictive power of the long side dominates that of the short side. Does this observation imply that short selling is in general less informative than changes in hedge fund holdings? The answer is no and we will come back to this issue in later sections.

We also note that the economic magnitude of the return predictability of $DLong_{i,\ t}$ and $DShort_{i,\ t}$ is sizable. To provide an example, consider Model (8) of Panel A, in which the regression parameters for $DLong^{Quin}_{i,t}$ and $DShort^{Quin}_{i,t}$ are 0.016 and −0.008, respectively. Hence, informed long demand and short demand are generally associated with an annualized abnormal return of 6.6% (annualized as $(1+0.016)^4-1$) and −3.2% (annualized as $(1-0.008)^4-1$) in the following quarter, respectively. If we add the two parameters together, we find that stocks characterized by informed long demand outperform stocks characterized by informed short demand by as much as 9.8%. The corresponding return difference implied by Model (3) is 4.4%. Although smaller in magnitude, the economic impact is no less impressive, given that the stocks covered by such long and short demand are approximately 25% of the stocks in a given period.

Panel B provides further insights by relating informed, hedging, and unwinding demand to cumulative abnormal returns over the next 12-month period. We observe that informed long (short) demand still predicts positive (negative) abnormal returns over this longer forecasting period. The return predictability of hedging and unwinding demand, however, disappears completely, suggesting that the 3-month return predictability of hedging and unwinding demand may reflect only a short-term price impact that is diluted over the longer horizon of 12 months. The next section further demonstrates that the short-term predictability of hedging and unwinding demand could be associated with trading strategies related to asset pricing anomalies. By contrast, the return predictability of $DLong_{i,\ t}$ and $DShort_{i,\ t}$ remains significant over the 12-month horizon, suggesting that it is unlikely to be driven by short-term price impacts. Thus, hedging demand and unwinding demand provide a sort of placebo test with respect to the informativeness of $DLong_{i,\ t}$ and $DShort_{i,\ t}$.

3.2. Informed demand vs. anomalies

It is important to ask whether informed demand could be spuriously related to asset pricing anomalies, as the latter involve only public information. To conduct this test, we control for ten asset pricing anomalies—i.e., firm characteristics that could be associated with future returns above and beyond what traditional asset pricing models would predict. They are: book-to-market ratio (for value premium), the logarithm of firm size (for size premium), lagged return in the previous 12 months (for momentum—using 6-month lagged return does not change our results), gross profit to assets ratio, operating profit, asset growth, investment growth, net stock issuance, accruals, and the logarithm of net operating assets. Recent literature identifies these factors as the most important anomalies that could predict stock returns (see, among others, Jegadeesh and Titman, 1993; Hirshleifer, Hou, Teoh, and Zhang, 2004; Cooper, Gulen, and Schill, 2008; Fama and French, 2008; Xing, 2008; Novy-Marx, 2013; Fama and French, 2015; Hou, Xue, and Zhang, 2015).

instance, based on the third root of the number of quarters or overlapping periods in independent variables—and then Newey and West adjusted the *t*-statistics accordingly. These adjustments do not change our results.

[10] Because the dependent variable already nets out the size, book-to-market, and momentum characteristics of similar stocks, we do not explicitly include these three characteristics in the current regression. Adding these additional control variables does not affect our results.

Table 2
Results of the baseline regression.

Panel A reports the results of the following baseline Fama-MacBeth regression at a quarterly frequency:

$$DGTW_{i,t+1} = \alpha_i + \beta_i \times Informed\ Demand_{i,t} + C \times M_{i,t} + \epsilon_{i,t+1},$$

where $DGTW_{i,\,t+1}$ refers to the out-of-sample DGTW-adjusted abnormal return of stock *i* accumulated over quarter $t+1$; $Informed\ Demand_{i,\,t}$ refers to a vector of informed demand variables, including $DLong_{i,\,t}$ and $DShort_{i,\,t}$ in the lagged quarter; and $M_{i,\,t}$ stacks a list of control variables, including *DIV*, the dividend yield calculated as dividends divided by market capitalization, *LgAge*, the logarithm of number of months since the stock first appeared in CRSP, *LgPrc*, the logarithm of the stock price per share, *LgTurn*, the logarithm of stock turnover rate prior to the beginning of the quarter, *LgVol*, the logarithm of the standard deviation of returns over the past 24 months, and *SP500*, a dummy equal to one for stocks in the S&P 500 index and zero otherwise. Panel B replaces the dependent variable with the out-of-sample DGTW-adjusted abnormal return of stock *i* accumulated over 1 year starting from quarter *t*. A detailed definition of these variables is provided in Appendix A. The superscripts ***, **, and * refer to the 1%, 5%, and 10% levels of statistical significance, respectively. The sample period is from 2000 to 2012.

Panel A: Out-of-sample quarterly abnormal return (DGTW-adjusted) regressed on informed demand variables

	DLong by positive/negative changes in long/short positions						DLong by terciles	DLong by quintiles
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DLong	0.008***		0.006***				0.012***	0.016***
	(4.06)		(3.49)				(3.91)	(2.97)
DShort		−0.007***	−0.005***				−0.007**	−0.008**
		(−4.11)	(−3.47)				(−2.37)	(−2.31)
DHedging				0.005**		0.003*	0.005*	0.006
				(2.51)		(1.98)	(1.80)	(1.46)
DUnwinding					−0.005**	−0.004**	−0.005*	−0.007*
					(−2.56)	(−2.09)	(−1.98)	(−1.81)
Div	−0.017	−0.017	−0.017	−0.016	−0.017	−0.017	−0.016	−0.017
	(−0.40)	(−0.39)	(−0.40)	(−0.37)	(−0.41)	(−0.39)	(−0.37)	(−0.39)
LgAge	0.002*	0.002*	0.002*	0.002*	0.002*	0.002*	0.002*	0.002*
	(1.88)	(1.94)	(1.91)	(1.93)	(1.93)	(1.94)	(1.88)	(1.79)
LgPrc	−0.005*	−0.005*	−0.005*	−0.006*	−0.006*	−0.006*	−0.006*	−0.006*
	(−1.76)	(−1.78)	(−1.76)	(−1.83)	(−1.84)	(−1.86)	(−1.81)	(−1.82)
LgTurn	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(0.54)	(0.55)	(0.57)	(0.48)	(0.52)	(0.50)	(0.54)	(0.50)
LgVol	−0.006	−0.006	−0.006	−0.006	−0.006	−0.006	−0.006	−0.006
	(−1.02)	(−1.03)	(−1.03)	(−1.01)	(−1.01)	(−1.01)	(−1.03)	(−1.06)
SP500	0.004	0.005	0.004	0.005	0.004	0.005	0.004	0.005
	(1.30)	(1.33)	(1.29)	(1.40)	(1.32)	(1.37)	(1.20)	(1.29)
Constant	−0.003	0.000	−0.002	−0.003	0.000	−0.001	−0.002	−0.002
	(−0.18)	(0.01)	(−0.09)	(−0.15)	(0.01)	(−0.06)	(−0.11)	(−0.10)
Observations	121,216	121,216	121,216	121,216	121,216	121,216	121,216	121,216
R-square	0.022	0.022	0.023	0.022	0.022	0.023	0.024	0.024

Panel B: Out-of-sample annual abnormal return (DGTW-adjusted) regressed on informed demand variables

	DLong by positive/negative changes in long/short positions						DLong by terciles	DLong by quintiles
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DLong	0.017***		0.013**				0.022***	0.029***
	(3.57)		(2.46)				(3.07)	(3.98)
DShort		−0.016***	−0.012***				−0.015**	−0.021**
		(−4.16)	(−2.73)				(−2.39)	(−2.09)
DHedging				0.009**		0.006*	0.010	0.011
				(2.36)		(1.93)	(1.19)	(1.03)
DUnwinding					−0.008	−0.005	−0.003	0.007
					(−1.47)	(−1.00)	(−0.51)	(0.85)
Div	−0.115	−0.115	−0.115	−0.116	−0.114	−0.115	−0.116	−0.115
	(−1.26)	(−1.27)	(−1.27)	(−1.29)	(−1.28)	(−1.29)	(−1.29)	(−1.28)
LgAge	0.011**	0.011**	0.011**	0.011**	0.011**	0.011**	0.011**	0.011**
	(2.59)	(2.65)	(2.63)	(2.65)	(2.61)	(2.64)	(2.59)	(2.62)
LgPrc	−0.034**	−0.034**	−0.033**	−0.034**	−0.034**	−0.034**	−0.034**	−0.034**
	(−2.46)	(−2.46)	(−2.45)	(−2.47)	(−2.46)	(−2.47)	(−2.44)	(−2.46)
LgTurn	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.003
	(0.51)	(0.52)	(0.53)	(0.48)	(0.53)	(0.51)	(0.49)	(0.42)
LgVol	−0.016	−0.016	−0.016	−0.016	−0.016	−0.016	−0.017	−0.017
	(−0.89)	(−0.90)	(−0.89)	(−0.89)	(−0.88)	(−0.88)	(−0.91)	(−0.91)
SP500	0.024**	0.024**	0.024**	0.025**	0.024**	0.024**	0.024**	0.025**
	(2.24)	(2.26)	(2.23)	(2.29)	(2.27)	(2.28)	(2.16)	(2.23)
Constant	0.025	0.032	0.029	0.026	0.033	0.030	0.027	0.024
	(0.54)	(0.69)	(0.62)	(0.57)	(0.68)	(0.63)	(0.59)	(0.52)
Observations	114,713	114,713	114,713	114,713	114,713	114,713	114,713	114,713
R-square	0.021	0.021	0.021	0.020	0.020	0.021	0.023	0.022

Table 3
Informed demand vs. asset pricing anomalies.
This table extends the baseline quarterly Fama-MacBeth regression of Table 2 as follows:

$$DGTW_{i,t+1} = \alpha_i + \beta_i \times Informed\ Demand_{i,t} + C \times M_{i,t} + D \times Anomaly_{i,t} + \epsilon_{i,t+1},$$

where $DGTW_{i,\,t+1}$ refers to the out-of-sample DGTW-adjusted abnormal return of stock i accumulated over quarter $t+1$; $Informed\ Demand_{i,\,t}$ refers to a vector of informed demand variables, including $DLong_{i,\,t}$ and $DShort_{i,\,t}$ in the lagged quarter; $M_{i,\,t}$ stacks a list of control variables, including *DIV, LgAge, LgPrc, LgTurn, LgVol*, and *SP500*, and $Anomaly_{i,\,t}$ stacks a list of firm characteristics that could be associated with asset return, including book-to-market ratio (for value premium), the logarithm of firm size (for size premium), lagged return in the previous 12 months (for momentum), gross profit to assets ratio, operating profit, asset growth, investment growth, net stock issuance, accruals, and the logarithm of net operating assets. We focus on tercile-based informed demand variables, and tabulate the regression results here. The corresponding baseline regression without anomalies is reported in Model (7) in Panel A of Table 2. Using quintile-based informed demand variables leads to very similar results. A detailed definition of these variables is provided in Appendix A. The superscripts ***, **, and * refer to the 1%, 5%, and 10% levels of statistical significance, respectively.

Out-of-sample abnormal return (quarterly) regressed on informed demand (by terciles)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
DLong	0.011***	0.012***	0.011***	0.011***	0.012***	0.012***	0.012***	0.011***	0.011***	0.012***	0.009***
	(3.74)	(3.77)	(3.67)	(3.75)	(3.60)	(3.80)	(3.58)	(3.73)	(3.92)	(3.84)	(3.41)
DShort	−0.007***	−0.007***	−0.008***	−0.008***	−0.007**	−0.008***	−0.007**	−0.008***	−0.008***	−0.008***	−0.007**
	(−2.83)	(−2.82)	(−2.89)	(−2.92)	(−2.30)	(−3.05)	(−2.66)	(−3.08)	(−2.83)	(−2.92)	(−2.17)
DHedging	0.003	0.003	0.003	0.003	0.000	0.002	0.003	0.003	0.001	0.003	−0.002
	(0.98)	(1.03)	(0.99)	(0.90)	(0.05)	(0.57)	(1.01)	(0.86)	(0.37)	(0.98)	(−0.38)
DUnwinding	−0.006**	−0.006**	−0.006**	−0.006**	−0.003	−0.006**	−0.006**	−0.006**	−0.006**	−0.006**	−0.003
	(−2.22)	(−2.19)	(−2.17)	(−2.25)	(−1.00)	(−2.17)	(−2.42)	(−2.22)	(−2.03)	(−2.26)	(−0.93)
Value (B/M)	−0.000										0.005
	(−0.12)										(1.18)
Size (Log_size)		0.003									0.001
		(1.64)									(0.50)
Momentum (Lag Ret)			0.006*								−0.008***
			(1.77)								(−3.86)
Gross profit to assets				0.028***							0.015
				(3.89)							(1.61)
Operating profit					0.032***						0.033***
					(7.44)						(6.94)
Asset growth						0.040***					0.078***
						(4.46)					(4.68)
Investment growth							0.000				−0.003
							(0.13)				(−1.32)
Net stock issuance								0.024**			0.005
								(2.23)			(0.28)
Accruals									0.001		−0.005
									(0.18)		(−0.79)
Net operating assets										0.007	−0.042***
										(1.45)	(−3.84)
Controls and constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	118,418	118,418	118,418	118,397	76,881	118,413	108,724	118,358	93,047	118,401	73,542
R-square	0.027	0.028	0.027	0.030	0.034	0.033	0.027	0.028	0.027	0.028	0.067

Table 3 tabulates the results. To better understand the potential influence of asset pricing anomalies, we start with Model (7) in Panel A of Table 2 as the baseline model (which uses tercile-based informed demand variables to predict next-quarter abnormal returns). The above anomalies are included one by one in Models (1)–(10) and then combined in Model (11). We first observe that *value premium, size premium, investment growth*, and *accruals* insignificantly affect performance. Because these anomalies do not directly affect returns, they also have little impact on the return predictability of the informed demand variables.

Next, Model (3) suggests that, consistent with the existing literature, *momentum*, when included on its own, positively predicts future abnormal returns. However, the sign of the impact flips when all other anomalies are included, as in Model (11). At the same time, *gross profit to assets* and *net stock issuance* each have significant effects when they are included on their own—but the significance of these variables dissipates in the joint model. The reverse occurs for *net operating assets*: the return impact is significant in the joint model but not when included alone. The inconsistencies across the stand-alone and joint models suggest that the effects of these variables on returns are not robust to alternative specifications—as a result, they also have little impact on the return predictability of the informed demand variables.

Finally, the remaining two types of anomalies, *operating profit* and *asset growth*, exert consistent return effects across the stand-alone and joint models and may affect the power of the informed demand variables. In particular, *operating profit* in Model (5) absorbs the return predictability of unwinding demand, suggesting that the latter variable may be associated with public information related to the *operating profit* of a firm rather than private information processed by skilled hedge fund managers.

The impact of *operating profit* on the return predictability of informed long/short demand, however, is minor. Compared with Model (7) in Panel A of Table 2, the return predictability of $DLong_{i,\,t}$ and $DShort_{i,\,t}$ in Model (5) maintains a similar economic magnitude and level of statistical significance. Indeed, the economic magnitude and statistical significance of the two variables drop very little even when all anomalies are included, as in Model (11), suggesting that the return predictability of informed demand arises for very different economic reasons than the asset pricing anomalies discussed above.[11]

Before proceeding further, it is also worth noting that Table 3 highlights the aforementioned intuition that it is crucial to use short selling information to distinguish information-motivated trading from general changes in hedge fund holdings. Let us take positive changes in hedge fund holdings (i.e., holding-implied "net buy") as an example. Mathematically, positive changes in holdings are equivalent to the summation of $DLong_{i,\,t}$ and $DHedging_{i,\,t}$.[12] Hence, the return predictability of the former can be inferred from our regressions as the summation of the regression coefficients, $DLong_{i,\,t}$ and $DHedging_{i,\,t}$. From Model (11), while the return predictability of $DLong_{i,\,t}$ is significantly positive (with a regression coefficient of 0.009 and a t-statistic of 3.41), that of $DHedging_{i,\,t}$ is negative (with a regression coefficient of −0.002 and a t-statistic of −0.38), leaving the summation of the two coefficients—or the return predictability of a holding-implied net buy—not only smaller in magnitude than $DLong_{i,\,t}$ but also statistically insignificant (with an F-statistic of 2.63 and a p-value of 11.2%). This insignificant result confirms the importance of our proposal to use short selling information to further extend the analysis of Griffin and Xu (2009).[13]

3.3. Portfolio analysis

A better way to illustrate the economic magnitude of our results is to construct portfolios based on our informed demand variables and then compute long-term performance based on these portfolios. To implement such a strategy, we go long (short) in stocks characterized by substantial informed long (short) demand.

Model (1) in Table 4 displays the results of this empirical strategy. At the beginning of each quarter, we focus on stocks characterized by the highest informed long (short) demand constructed on the basis of either terciles of short interest and hedge fund ownership changes (Panel A) or quintiles of short interest and hedge fund ownership changes (Panel B). We report the long-run abnormal returns that can be generated by this strategy. The results show that stocks characterized by informed long (short) demand generate significant positive (negative) abnormal returns in the long run and that the difference between these two groups of stocks is highly significant, both statistically and economically. For example, stocks characterized by quintile-based informed long demand can generate a quarterly DGTW-adjusted return of 2.12% in an equal-weighted portfolio, which translates into an annual return of 8.76%. These stocks outperform those characterized by quintile-based informed short demand by 2.53% per quarter or 10.5% per year. This magnitude is close to the magnitude obtained in the multivariate analysis. Value-weighted portfolios yield consistent results.

In Models (2)–(4) of Table 4, we further explore how the return predictability of informed demand decays over time. To do so, we report the average out-of-sample abnormal returns generated over the two, three, and four quarters following the quarter in which we construct informed demand. Models (2)–(4) report the cumulative DGTW-adjusted returns over these longer holding periods. The most important finding is that performance does not revert in the future: the cumulative return differences between equal-weighted long and short portfolios in Panel B, for instance, amount to 4.54%, 5.25%, and 5.8% over the next two, three, and four quarters, respectively. The first-quarter abnormal return is 2.53%, which decays to 2%, 0.7%, and 0.5% in the second, third, and fourth quarters, respectively. The decay in abnormal returns is not accompanied by a price reversal.

This evidence is therefore generally consistent with a slow diffusion of information in the market. Of course, the type of information underlying this return predictability is unclear. To address this issue, in the next section, we relate informed demand to various types of firm fundamentals.

4. Informed demand and firm fundamentals

Thus far, our results demonstrate that informed demand retrieved from the joint analysis of short selling and hedge fund holding information has significant forecasting power for out-of-sample stock returns. The next step is to investigate the channels through which informed demand achieves such predictive power. In this section, therefore, we explore the hypothesis that the source of the return

[11] Another way to demonstrate the difference between informed and hedging/unwinding demand here is to examine whether high *operating profit* can directly lead hedge funds to increase hedging demand. When we regress hedge fund demand variables on lagged *operating profit* and other lagged control variables, we find that *operating profit* significantly increases (decreases) hedge fund hedging (unwinding) demand, and that such predicting power concentrates in periods with higher economic policy uncertainty (Baker, Bloom, and Davis, 2015). By contrast, *operating profit* does not predict informed demand. To save space, the results are tabulated Table IN1 of the Internet Appendix.

[12] As is easily seen, $DLong_{i,\,t}+DHedge_{i,\,t}=I\{\Delta HFOwn_{i,\,t}>0\}\times I\{\Delta SI_{i,\,t}<0\}+I\{\Delta HFOwn_{i,\,t}>0\}\times I\{\Delta SI_{i,\,t}>0\}=I\{\Delta HFOwn_{i,\,t}>0\}\times(I\{\Delta SI_{i,\,t}<0\}+I\{\Delta SI_{i,\,t}>0\})=I\{\Delta HFOwn_{i,\,t}>0\}$.

[13] We obtain very similar results when we use quartile- or quintile-based informed demand variables to conduct the same test: among the ten anomalies, *value premium, size premium, investment growth*, and *accruals* have insignificant effects on returns; *momentum, gross profit to assets, net stock issuance*, and *net operating assets* have inconsistent effects across the stand-alone and joint models. All these variables, when used alone, do not affect the return predictability of informed demand. Finally, *operating profit* and *asset growth* continue to exhibit consistent return effects across the stand-alone and joint models. In particular, *operating profit* also absorbs the return predictability of unwinding demand but not that of informed demand. Finally, the predictive power of $DLong_{i,\,t}$ and $DHedging_{i,\,t}$ are also significantly positive and insignificantly negative, leaving the summation of the coefficients smaller than that of $DLong_{i,\,t}$ and marginally insignificant. Because all the patterns are very similar to tercile-based informed demand, in the interest of brevity, we do not tabulate them here.

Table 4
Portfolio-based analyses.
In Panel A, we first independently double sort stocks into terciles based on hedge fund 13F holding changes and short interest changes. We then focus on two portfolios of stocks that have experienced the largest net-long and net-short demand shocks. We then report the DGTW-adjusted return that can be generated by the two portfolios over the entire sample period (2000–2012). A detailed definition of these variables is provided in Appendix A.

Panel A: Cumulative DGTW return of tercile information-based informed demand

	DGTW return of equal-weighted portfolio				DGTW return of value-weighted portfolio			
	$t+1$	$t+1$ to $t+2$	$t+1$ to $t+3$	$t+1$ to $t+4$	$t+1$	$t+1$ to $t+2$	$t+1$ to $t+3$	$t+1$ to $t+4$
Dlong= 1 in t	1.750%	3.260%	3.776%	4.041%	1.319%	2.398%	3.100%	3.413%
Dshort= 1 in t	−0.186%	−0.328%	−0.123%	−0.185%	−0.732%	−0.439%	0.320%	0.115%
Dlong-minus-Dshort	1.936%	3.588%	3.899%	4.226%	2.050%	2.836%	2.780%	3.298%
t-stat	(5.41)	(6.48)	(5.92)	(5.80)	(4.79)	(5.00)	(4.64)	(3.89)

Panel B: Cumulative DGTW return of quintile information-based informed demand

	DGTW return of equal-weighted portfolio				DGTW return of value-weighted portfolio			
	$t+1$	$t+1$ to $t+2$	$t+1$ to $t+3$	$t+1$ to $t+4$	$t+1$	$t+1$ to $t+2$	$t+1$ to $t+3$	$t+1$ to $t+4$
Dlong= 1 in t	2.122%	3.532%	4.249%	4.741%	1.729%	2.758%	3.100%	3.696%
Dshort= 1 in t	−0.409%	−1.008%	−1.003%	−1.053%	−0.330%	−0.055%	0.414%	0.595%
Dlong-minus-Dshort	2.531%	4.540%	5.252%	5.795%	2.059%	2.813%	2.686%	3.101%
t-stat	(4.75)	(5.76)	(5.50)	(4.98)	(3.58)	(3.25)	(2.59)	(2.55)

predictability of informed demand lies in the ability of the hedge funds to forecast firm fundamentals.

4.1. Subsample analysis

We start by splitting our sample into two subgroups based on a list of firm characteristics such as market capitalization, turnover ratio, analyst coverage, and dispersion of analyst forecasts. Splitting the sample into these subgroups allows us to better understand the effects of size and liquidity on return predictability. We therefore conduct return predictability tests, as in Model (3) of Table 2, for each subsample of stocks.

Table 5 reports the corresponding regression results, Panels A and B for quarterly and yearly return predictive power for informed demand, and Panels C and D for hedging/unwinding demand. To save space, in Panels B–D, we only tabulate the coefficients for the main variables. The full specifications of the regression parameters can be found in the Internet Appendix. Specifically, Models (1) and (2) of the first two panels apply the baseline specification to the subsamples of firms characterized by small and large market capitalization, employing a 50–50 split. We find that the return predictability of both informed long demand and informed short demand remains significant in the subsample of large firms but that informed short demand loses its predictive power in cases of small firms. This pattern suggests that return predictability is unlikely to be driven by firm characteristics or stock returns related to (small) size, as otherwise, we would find stronger return predictability for smaller stocks. Our results suggest, on the contrary, that the stronger short sale constraint faced by small stocks may distort information discovery and dissemination.

Consistent with this intuition, Models (3)–(8), which tabulate the regression results for stocks with different turnover ratios, analyst coverage, and analyst dispersion, illustrate that the return predictability of both informed long demand and informed short demand remains significant mostly for stocks with high liquidity (i.e., high turnover ratios) and high analyst coverage/dispersion. The results for turnover suggest that the abnormal returns predicted by informed demand are not merely a compensation for illiquidity-related market conditions. The findings regarding analyst coverage/dispersion suggest that $DLong_{i,\,t}$ and $DShort_{i,\,t}$ are informed even in the presence of analysts, suggesting that hedge fund managers process information more effectively than analysts (e.g., Kim and Verrecchia, 1994; Engelberg, Reed, and Ringgenberg, 2012).

When, in Panels C and D, we extend our subsample analysis to hedging and unwinding demand based on Model (6) of Table 2, we find that the predicting power of these two variables does not concentrate in the same subsamples of stocks. Indeed, in many cases, the marginal significance of the two variables as observed in Table 2 disappears in the subsamples, confirming that the return predictability of informed demand arises for very different economic reasons than hedging and unwinding demands.

4.2. Forecasting firm fundamentals

Our previous tests illustrate that the predictive power of informed demand does not stem from firm characteristics or market conditions related to (small) size, (low) liquidity, and (low) analyst attention. Hence, our next task is to explore the extent to which informed traders can predict firm fundamentals. In the spirit of Akbas, Boehmer, Erturk, Sorescu (2013), we explore several proxies for unexpected shocks to firm fundamentals. The first is related to the real performance of firms, proxied by return on assets (ROA) or changes in ROA, adjusted or unadjusted by industry peers. We calculate *ROA* by dividing the firm's income before extraordinary items by total assets. We define Δ*ROA* as the difference between *ROA* in the current quarter and *ROA* four quarters ago (i.e., the same quarter in the previous year to account for seasonality in operating performance), and we define *Ind-adj ROA* (*Ind-adj* Δ*ROA*) as *ROA* minus the industry mean during the quarter, where

Table 5
Subsample analyses.

This table applies the baseline regression from Table 2 to subsamples of stocks constructed based on different stock characteristics, including market capitalization, turnover ratio, analyst coverage, and dispersion of analyst forecasts. For each of the characteristics, we split the sample in any given quarter into two subsamples based on the median value. We then apply the baseline regression to each subsample of stocks and tabulate the regression results. Panels A and B apply the subsample analysis to the baseline regressions involving informed demand (Model 3 in Panel A and Panel B of Table 2, respectively). Panels C and D apply the subsample analysis to the baseline regressions regarding hedging/unwinding demand (Model 6 in Panel A and Panel B of Table 2, respectively). To save space, for Panels B–D, we only tabulate the coefficients for the main variables. A detailed definition of these variables is provided in Appendix A. The full specifications of the regression parameters can be found in the Internet Appendix. The superscripts ***, **, and * refer to the 1%, 5%, and 10% levels of statistical significance, respectively.

Panel A: Subsample analyses for out-of-sample quarterly abnormal return (DGTW-adjusted)

	Firm size		Turnover		Analyst coverage		Dispersion of analysts	
	(1) Small	(2) Large	(3) Low	(4) High	(5) Low	(6) High	(7) Low	(8) High
DLong	0.008***	0.005***	0.003	0.009***	0.006*	0.007***	0.011***	0.007***
	(2.71)	(2.69)	(1.25)	(3.88)	(1.98)	(3.36)	(3.47)	(2.76)
DShort	−0.004	−0.007***	−0.006**	−0.005***	−0.007**	−0.004**	−0.006	−0.005***
	(−1.02)	(−4.39)	(−2.13)	(−2.73)	(−2.35)	(−2.54)	(−1.58)	(−2.76)
Div	−0.016	0.020	0.036	−0.053	0.011	−0.079	−0.062	−0.025
	(−0.29)	(0.45)	(0.64)	(−0.97)	(0.20)	(−1.60)	(−0.70)	(−0.44)
LgAge	0.007***	−0.003	0.003**	0.000	0.006***	−0.001	0.001	−0.000
	(4.16)	(−1.50)	(2.10)	(0.23)	(4.38)	(−0.59)	(0.27)	(−0.14)
LgPrc	−0.008*	−0.004	−0.005*	−0.005	−0.005	−0.010**	−0.010***	−0.005
	(−1.80)	(−1.42)	(−1.77)	(−1.38)	(−1.54)	(−2.50)	(−2.70)	(−1.05)
LgTurn	−0.000	0.000	0.004*	−0.009*	−0.001	−0.000	0.001	0.000
	(−0.01)	(0.10)	(1.76)	(−1.95)	(−0.45)	(−0.09)	(0.21)	(0.00)
LgVol	−0.004	−0.008	0.000	−0.010	−0.005	−0.007	−0.005	−0.004
	(−0.73)	(−0.92)	(0.09)	(−1.44)	(−1.00)	(−0.88)	(−0.72)	(−0.51)
SP500	0.029**	0.004	0.003	0.003	0.009	0.005*	0.003	0.006
	(2.01)	(1.46)	(0.76)	(0.92)	(1.25)	(1.79)	(0.78)	(1.51)
Constant	−0.022	0.018	0.016	−0.022	−0.031	0.027	0.024	0.007
	(−1.06)	(0.88)	(0.84)	(−1.13)	(−1.54)	(1.43)	(1.15)	(0.34)
Observations	60,596	60,620	60,596	60,620	56,834	64,382	31,713	57,436
R-square	0.027	0.037	0.027	0.030	0.026	0.036	0.043	0.034

Panel B: Subsample analyses for out-of-sample annual abnormal return (DGTW-adjusted)

	Firm size		Turnover		Analyst coverage		Dispersion of analysts	
	(1) Small	(2) Large	(3) Low	(4) High	(5) Low	(6) High	(7) Low	(8) High
DLong	0.016	0.009**	0.010	0.017***	0.013	0.014***	0.011*	0.016***
	(1.45)	(2.52)	(0.87)	(3.46)	(1.13)	(3.80)	(1.88)	(3.58)
DShort	−0.012	−0.011***	−0.014*	−0.009*	−0.011	−0.012***	−0.008	−0.014**
	(−1.30)	(−3.72)	(−1.82)	(−1.92)	(−1.08)	(−2.93)	(−1.10)	(−2.39)

Panel C: Subsample analyses for out-of-sample quarterly abnormal return (DGTW-adjusted)

	Firm size		Turnover		Analyst coverage		Dispersion of analysts	
	(1) Small	(2) Large	(3) Low	(4) High	(5) Low	(6) High	(7) Low	(8) High
DHedging	0.002	0.004*	0.005*	0.001	0.004	0.002	0.001	0.003
	(0.80)	(1.99)	(1.94)	(0.48)	(1.66)	(0.82)	(0.37)	(1.19)
DUnwinding	−0.006**	−0.002	−0.001	−0.006***	−0.001	−0.005**	−0.004	−0.006**
	(−2.09)	(−0.91)	(−0.42)	(−2.70)	(−0.40)	(−2.16)	(−0.86)	(−2.26)

Panel D: Subsample analyses for out-of-sample annual abnormal return (DGTW-adjusted)

	Firm size		Turnover		Analyst coverage		Dispersion of analysts	
	(1) Small	(2) Large	(3) Low	(4) High	(5) Low	(6) High	(7) Low	(8) High
DHedging	0.011*	0.004	0.013**	−0.000	0.010	0.003	0.000	0.006
	(1.79)	(1.00)	(2.53)	(−0.07)	(1.63)	(0.65)	(0.08)	(1.18)
DUnwinding	−0.011	−0.001	−0.006	−0.005	−0.008	−0.003	0.002	−0.009
	(−1.19)	(−0.26)	(−0.63)	(−0.93)	(−0.85)	(−0.83)	(0.27)	(−1.62)

industries are defined by two-digit Standard Industrial Classification (SIC) codes. We repeat the return predictability regression as specified in Eq. (1) but replace the dependent variable, out-of-sample abnormal returns, with out-of-sample average ROA in the 12-month (four-quarter) period following the construction of informed demand.

The results are reported in Models (1)–(4) of Table 6. Across these specifications, we find that $DLong_{i,\,t}$ and $DShort_{i,\,t}$ forecast positive and negative ROA of firms, respectively. The predictive power is again highly significant, which is consistent with the notion that informed demand reflects capable traders' abilities to forecast firm fundamentals.

Although ROA reflects the long-term profitability of firms, the financial market typically pays special attention to short-term cash flows, such as earnings. Thus, another way to achieve return predictability is to process earnings-related information more effectively than the market. Hence, our second proxy for (unexpected changes in) firm fundamentals is related to the portion of earnings that is unpredicted by the market, namely, standardized unexpected earnings (*SUE*). If informed demand predictability is truly driven by information, we expect informed demand to forecast *SUE*. Following Hirshleifer, Myers, Myers, and Teoh (2008), we compute *SUE* as the seasonal difference in split-adjusted earnings per share scaled by the split-adjusted end-of-quarter price (i.e., the price at the end of the quarter prior to the earnings announcement).

We also supplement *SUE* with another important variable that may help us understand the informativeness of capable traders, namely, analyst revisions. *Analyst revision* is the change in the consensus analyst earnings estimate, computed as the difference in mean estimates from the previous month divided by the stock price at the end of the previous month. If informed demand forecasts not only *SUE* but also analyst revisions, then hedge fund managers behind the demand have the ability to process earnings-related information, and their informational advantage would exceed that of analysts. Models (5) and (6)

Table 6

Forecasting firm fundamentals.

This table explores the predictability of net demands on out-of-sample firm fundamentals. In Models (1) and (2) of Panel A, we regress firm ROA or changes in ROA in the following year on informed long or short demand. Models (3) and (4) further adjust ROA or changes in ROA by the industry average. Models (5) and (6) tabulate the results for similar predictive regressions when the dependent variables are next-period SUE and analyst revisions. Model (7) reports the predictability of informed demand for next-period CARs. Control variables include *BM*, the book-to-market ratio, *DIV*, the dividend yield calculated as dividends divided by market capitalization, *LgAge*, the logarithm of number of months since the stock first appeared in CRSP, *LgPrc*, the logarithm of the stock price per share, *LgSize*, the logarithm of market capitalization, *LgTurn*, the logarithm of stock turnover rate, *LgVol*, the logarithm of the standard deviation of returns over the past 24 months, *Ret3*, stock return in the last quarter, *Ret9*, stock return of the three quarters prior to the last quarter, and *SP500*, a dummy equal to one for stocks in the S&P 500 index and zero otherwise. Panel B applies the same tests to hedging and unwinding demand. A detailed definition of these variables is provided in Appendix A. The superscripts ***, **, and * refer to the 1%, 5%, and 10% levels of statistical significance, respectively.

Panel A. Predictability of informed demands

	(1) Out-of-sample ROA or changes in ROA	(2)	(3)	(4)	(5)	(6) SUE or analyst revision	(7) Mkt response
Dependent variable =	ROA	ΔROA	Ind-adj ROA	Ind-adj ΔROA	SUE	Analyst revision	CAR
DLong	0.001***	0.001**	0.001**	0.001**	0.001***	0.010***	0.001***
	(3.18)	(2.40)	(2.44)	(2.44)	(2.76)	(2.94)	(3.09)
DShort	−0.001**	−0.001***	−0.001*	−0.001**	−0.001**	−0.006*	−0.001**
	(−2.29)	(−2.72)	(−1.84)	(−2.41)	(−2.62)	(−1.95)	(−2.31)
BM	0.002***	0.002***	0.003***	0.001***	0.004***	−0.041**	0.001*
	(3.05)	(3.49)	(4.73)	(3.25)	(3.86)	(−2.50)	(1.74)
Div	0.036***	0.002	0.028***	0.002	0.017	−0.105	−0.019***
	(5.19)	(0.49)	(3.82)	(0.48)	(0.75)	(−1.57)	(−3.08)
LgAge	0.003***	0.001***	0.003***	0.001***	0.001***	0.002	0.000*
	(11.21)	(3.27)	(10.15)	(3.26)	(3.37)	(0.56)	(1.96)
LgPrc	0.016***	−0.001**	0.014***	−0.001**	−0.005***	−0.002	0.001**
	(14.82)	(−2.36)	(15.06)	(−2.14)	(−4.61)	(−0.52)	(2.52)
LgSize	0.000	0.001***	0.002***	0.001***	0.001***	0.005*	−0.000
	(0.48)	(4.86)	(4.62)	(4.74)	(3.93)	(1.74)	(−0.92)
LgTurn	−0.000	−0.001***	−0.001***	−0.000**	−0.000	−0.003	−0.001**
	(−1.03)	(−2.98)	(−3.52)	(−2.28)	(−0.78)	(−0.88)	(−2.43)
LgVol	−0.008***	−0.002**	−0.005***	−0.001**	0.001	−0.017**	0.001*
	(−6.26)	(−2.02)	(−4.05)	(−2.14)	(0.89)	(−2.42)	(1.74)
Ret3	0.011***	0.014***	0.010***	0.012***	0.016***	0.064***	−0.000
	(8.07)	(14.95)	(8.25)	(14.02)	(7.25)	(4.79)	(−0.18)
Ret9	0.007***	0.002**	0.006***	0.001	−0.002	0.018***	−0.000
	(5.96)	(2.28)	(5.56)	(1.67)	(−1.26)	(3.19)	(−0.45)
SP500	−0.004***	−0.001**	−0.006***	−0.001**	0.000	0.002	−0.000
	(−5.47)	(−2.26)	(−8.84)	(−2.20)	(0.35)	(0.19)	(−0.24)
Constant	−0.098***	−0.027***	−0.095***	−0.024***	−0.019*	−0.221***	−0.001
	(−19.40)	(−6.00)	(−20.34)	(−5.96)	(−1.90)	(−8.64)	(−0.43)
Observations	111,513	111,040	111,513	111,040	113,932	95,336	105,846
R-square	0.188	0.032	0.165	0.028	0.040	0.015	0.020

(*continued on next page*)

Table 6
Continued.

Panel B. Predictability of hedging/unwinding demands

Variables	(1) Out-of-sample ROA or changes in ROA	(2)	(3)	(4)	(5)	(6) SUE or analyst revision	(7) Mkt response
	ROA	ΔROA	Ind-adj ROA	Ind-adj ΔROA	SUE	Analyst revision	CAR
DHedging	−0.000	−0.000	−0.000	−0.000	−0.000	−0.002	0.000
	(−0.74)	(−1.28)	(−0.55)	(−1.59)	(−1.00)	(−0.52)	(0.27)
DUnwinding	−0.001**	0.000	−0.001**	0.000	0.000	−0.007	−0.000
	(−2.28)	(0.52)	(−2.12)	(0.35)	(0.65)	(−1.59)	(−1.32)
BM	0.002**	0.002***	0.003***	0.001***	0.004***	−0.041**	0.001
	(2.15)	(3.09)	(3.55)	(2.86)	(3.85)	(−2.51)	(1.57)
Div	0.036***	0.002	0.027**	0.002	0.018	−0.101	−0.019**
	(3.77)	(0.48)	(2.65)	(0.48)	(0.79)	(−1.55)	(−2.49)
LgAge	0.003***	0.001***	0.003***	0.001***	0.001***	0.002	0.000
	(8.43)	(3.27)	(7.79)	(3.04)	(3.38)	(0.55)	(1.52)
LgPrc	0.016***	−0.001**	0.014***	−0.001*	−0.005***	−0.002	0.001*
	(10.04)	(−2.03)	(10.31)	(−1.85)	(−4.56)	(−0.54)	(1.87)
LgSize	0.000	0.001***	0.002***	0.001***	0.001***	0.006*	−0.000
	(0.32)	(4.29)	(3.15)	(4.19)	(3.95)	(1.78)	(−0.72)
LgTurn	−0.000	−0.001***	−0.001**	−0.000**	−0.001	−0.004	−0.001*
	(−0.80)	(−2.73)	(−2.62)	(−2.11)	(−0.85)	(−0.94)	(−1.92)
LgVol	−0.008***	−0.001	−0.004***	−0.001	0.001	−0.016**	0.001
	(−4.57)	(−1.52)	(−3.01)	(−1.63)	(0.93)	(−2.34)	(1.46)
Ret3	0.011***	0.014***	0.010***	0.012***	0.016***	0.062***	−0.000
	(7.75)	(13.96)	(7.84)	(13.14)	(7.26)	(4.69)	(−0.33)
Ret9	0.007***	0.002**	0.006***	0.001	−0.002	0.018***	−0.000
	(4.87)	(2.05)	(4.59)	(1.54)	(−1.22)	(3.21)	(−0.35)
SP500	−0.004***	−0.001*	−0.006***	−0.001*	0.000	0.002	−0.000
	(−3.71)	(−1.76)	(−6.06)	(−1.72)	(0.39)	(0.20)	(−0.13)
Constant	−0.097***	−0.027***	−0.095***	−0.024***	−0.019*	−0.216***	−0.001
	(−14.76)	(−4.77)	(−15.56)	(−4.88)	(−1.87)	(−8.54)	(−0.26)
Observations	111,513	111,040	111,513	111,040	113,932	95,336	105,846
R-squared	0.188	0.032	0.165	0.028	0.040	0.015	0.020

confirm this conjecture: both $DLong_{i,\,t}$ and $DShort_{i,\,t}$ exhibit the proper power in predicting positive and negative *SUE* and analyst revisions, respectively. Hedge fund managers' predictive power regarding analyst revisions is especially noteworthy, as it suggests that the informed demand identified in our paper could be based on information superior to that of professional analysts.

Finally, hedge fund managers may also be sufficiently sophisticated to predict market reactions to firm-level information, which would allow them to benefit from their trading. To explore this channel, we examine whether informed demand forecasts cumulative abnormal returns (CARs) around earnings announcements. The dependent variable in this case is constructed as market-adjusted returns upon earnings announcements over a [−1, 1] window, where the market is defined as the value-weighted portfolio of NYSE/Amex/Nasdaq stocks. We find strong evidence that informed long (short) demand peaks before positive (negative) CARs, confirming that the information-processing ability of hedge funds allows them to benefit from market reactions. The ability of hedge funds to process earnings-related information is consistent with Kim and Verrecchia (1994) in general and Engelberg, Reed, and Ringgenberg (2012) in particular.

Panel B applies the tests of forecasting firm fundamentals to hedging and unwinding demand. The results show that hedging demand does not forecast *ROA* or *SUE*. Nor does it forecast *Analyst Revision* or *CARs*. Hence, this variable is unrelated to the efforts of forecasting firm fundamentals. Next, unwinding demand may appear to forecast *ROA* in the first model. However, when we further adjust for seasonality and use Δ*ROA* as the dependent variable, the forecasting power disappears. A similar pattern can be observed for industry-adjusted *ROA* and Δ*ROA*. The forecasting ability of unwinding demand, in this regard, does not survive the adjustment of seasonality. To the extent that it does not forecast *SUE, Analyst Revision*, or *CARs*, unwinding demand is also unlikely to be associated with the effort of forecasting firm fundamentals.

4.3. Persistence in performance of informed demand

The previous section implies that the return predictability of informed demand, which can directly contribute to fund performance, arises from managerial skill in processing firm-level information. To further validate this interpretation, it is crucial to examine whether the performance delivered by informed demand persists at the fund level. In other words, if managerial skills, rather than good luck or estimation error, is the driving force behind such performance, we expect managers with such skills to persistently deliver this performance at least in the near future.

To examine this question, we quantify the performance of informed demand of a particular fund as the abnormal returns generated by stocks characterized by informed long demand, as implied by fund-specific holdings. In particular, we define *informed long demand for fund f* in any given quarter as $DLong_{f,\,i,\,t} = I\{\Delta HFOwn_{f,\,i,\,t} > 0\} \times I\{\Delta SI_{i,\,t} < 0\}$,

Table 7
Persistence in performance of informed demand (fund-level test).
This table examines the persistence in performance that can be generated by informed demand. To do so, we first define fund-level informed long demand for a particular fund f in any given quarter as $DLong_{f,i,t} = I\{\Delta HFOwn_{f,i,t} > 0\} \times I\{\Delta SI_{i,t} < 0\} = 1$, where $I\{.\}$ is an indicator function, and $\Delta HFOwn_{f,i,t} = HFOwn_{f,i,t} - HFOwn_{f,i,t-1}$ and $\Delta SI_{i,t} = SI_{i,t} - SI_{i,t-1}$ denote the changes in holdings of fund f and short interest, respectively, and quantify the performance for fund f to conduct informed trading as the average DGTW return that can be generated by stocks that have informed long demand as implied by its holdings. In any given quarter, we then sort all funds into ten deciles according to their realized performance in conducting informed trading as exhibited in the 12-month period, and create ten dummy variables to indicate their ranks (Decile 1 to Decile 10 for low to high performance). Models (1) and (2) then regress, in Fama-MacBeth specifications, the out-of-sample quarterly performance or performance rank of informed long demand on the rank dummies of realized performance. We further conduct an F-test on coefficient difference between *Decile 10* and *Decile 1* dummies for each regression model as well as an F-test on the coefficient difference between the summation of *Decile 9* and *Decile 10* and that of *Decile 1* and *Decile 2*. The testing results are reported in the last two lines of the table. Finally, we apply the same analysis to fund-level hedging demand ($DHedging_{f,i,t} = I\{\Delta HFOwn_{f,i,t} > 0\} \times I\{\Delta SI_{i,t} > 0\} = 1$), and tabulate the results in Models (3) and (4). A detailed definition of these variables is provided in Appendix A. The superscripts ***, **, and * refer to the 1%, 5%, and 10% levels of statistical significance, respectively.

Out-of-sample abnormal return or ranks regressed on lagged performance ranks of hedge funds

	A. Persistence test based on informed demand		B. Persistence test based on uninformed demand	
	Out-of-sample return (1)	Out-of-sample ranks (2)	Out-of-sample return (3)	Out-of-sample ranks (4)
Decile 1	0.001	0.029	−0.015	−0.118*
	(0.37)	(0.29)	(−1.04)	(−1.88)
Decile 2	0.001	−0.067	−0.015	−0.059
	(0.29)	(−0.68)	(−0.96)	(−0.63)
Decile 3	0.004*	0.123	−0.013	0.073
	(1.93)	(1.51)	(−0.84)	(0.76)
Decile 4	0.002	−0.005	−0.012	0.040
	(0.74)	(−0.07)	(−0.77)	(0.42)
Decile 5	0.001	0.001	−0.001	−0.070
	(0.25)	(0.01)	(−0.29)	(−0.85)
Decile 6	0.003	0.126	−0.013	−0.020
	(1.13)	(1.41)	(−0.88)	(−0.20)
Decile 7	−0.000	−0.029	−0.018	−0.117
	(−0.12)	(−0.38)	(−1.14)	(−0.97)
Decile 8	0.003	0.096	−0.014	0.007
	(1.40)	(1.17)	(−0.93)	(0.07)
Decile 9	0.008**	0.251***	−0.016	−0.021
	(2.45)	(2.85)	(−1.02)	(−0.23)
Decile 10	0.013***	0.219**	−0.014	0.050
	(3.37)	(2.15)	(−0.87)	(0.50)
Constant	0.003	5.448***	0.019	5.556***
	(1.24)	(90.76)	(1.21)	(79.28)
Observations	23,545	23,545	23,631	23,631
R-square	0.024	0.025	0.024	0.024
F-test on D10 − D1	10.54***	3.99*	0.12	2.47
F-test on (D9 + D10) − (D1 + D2)	14.72***	9.7***	0.01	2.03

where $\Delta HFOwn_{f,i,t} = HFOwn_{f,i,t} - HFOwn_{f,i,t-1}$ denotes changes in holdings of fund f. Compared with our previous definition, the only difference here is that we use fund-level holding changes to construct fund-specific informed long demand. We can then quantify the performance of informed demand for fund f as the average DGTW return that can be generated by stocks characterized by informed long demand, as implied by its holdings.[14]

Next, for each quarter, we sort all the funds into 10 deciles based on the performance of informed demand realized over the 12-month period prior to the quarter, and create ten dummy variables to indicate their in-sample performance rankings (Decile 1 for low and Decile 10 for high performance). We then compute the out-of-sample performance realized in this quarter for each fund—as well as its performance rank—in terms of informed long demand. Finally, we test the relationship between the out-of-sample performance of funds and the ten dummy variables describing their in-sample performance.

Table 7 tabulates the results. More specifically, in Models (1) and (2), we adopt a Fama-MacBeth specification in which we regress the out-of-sample quarterly performance and performance ranks of the funds on their in-sample performance rankings. We observe that hedge funds characterized by top performance rankings in the past (Decile 9 and Decile 10) deliver significantly positive performance as a result of their informed long demand in the next quarter. In contrast, funds characterized by lower in-sample performance rankings typically achieve insignificant future performance. Out-of-sample performance thus increases with in-sample performance rank. Such persistence is consistent with the notion that the return predictability of informed demand reflects managerial skill that enables capable fund managers to consistently generate superior performance in the near future.

[14] Mathematically, we can also construct informed short demand for each fund. The lack of fund-level short selling data, however, makes this variable less attractive, as ideally, we would use fund-level short selling activities to infer fund-level informed short demand in the cross section. Hence, our fund-level tests focus on informed long demand. Likewise, we focus on hedging demand as a placebo test.

To further examine the difference in performance persistency between top and bottom funds, we conduct an *F*-test of the difference between the coefficients for the *Decile 10* and *Decile 1* dummies in Model (1). The results are reported in the line "*F-test on Decile 10-Decile 1*" at the bottom of the table. The *F*-test shows that the difference is statistically significant. In other words, the top 10% of funds significantly outperform the bottom 10% of funds in generating out-of-sample performance reflective of informed long demand. Similarly, we also conduct an *F*-test of the difference between the summation of *Decile 9* and *Decile 10* coefficients and that of *Decile 1* and *Decile 2* coefficients. The result as reported in the last line of the table shows that the top 20% of funds significantly outperform the bottom 20% of funds. Additionally, we get similar *F*-test findings with regard to coefficient differences for the regression reported in Model (2). The results confirm that the top 10% or 20% of funds significantly outrank the bottom 10% or 20% of funds in generating out-of-sample performance. Our results, therefore, imply that a significant fraction of funds (approximately 20%) may indeed be skilled in processing firm-specific information, as suggested in the previous sections.

As a comparison, we also apply the same persistence test to hedging demand which, according to our previous tests, is unrelated to the ability of forecasting firm fundamentals. Models (3) and (4) resemble Models (1) and (2), except that hedge funds are now sorted by the performance that can be generated by their fund-specific hedging demand prior to a quarter. Unlike the case of informed long demand, hedging demand does not generate highly persistent performance in the future. This insignificance, evident from the insignificant *F*-statistics reported in the last two lines of the table, has important implications for the hedge fund industry. It suggests that hedge fund skills and persistent performance are likely to be more associated with superior firm-specific information than other sources of information.

Overall, the results presented in this section support the working hypothesis that informed demand identified in the joint analysis of short selling and hedge fund ownership changes predicts firm fundamentals above and beyond the abilities of the general public to make such predictions. The main reason informed demand predicts returns is that it represents the demand of hedge fund managers skillful in processing firm-level information. Our remaining task is to rule out alternative explanations unrelated to the processing of information. We take up this task in the next section.

5. Extensions and alternative explanations

We now implement several additional tests to further enrich our economic intuition and to rule out several alternative explanations for this return predictability. We first further explore the informativeness of short positions in terms of return predictability, and compare it to that of long positions. Next, we conduct a placebo test in which informed demand is constructed based on holdings of mutual funds. We then explore whether the source of return predictability could be related to the exploitation of mutual fund flows or the provision of liquidity.

5.1. Informativeness of shorts

We recall that in Model (7) of Table 2, the directions of predicted future returns appear more consistent with the long-side information than the short side. The question is whether this observation implies that short selling is in general less informative than changes in hedge fund holdings. This section explores this issue.

Before we start, we first note that short interest changes are measured at the quarterly frequency to match that of hedge fund holding changes. However, information from the long side may have a longer duration than that of the short side (e.g., Diether, Lee, and Werner, 2008; Boehmer and Wu, 2013). Indeed, Cohen, Diether, and Malloy (2007) demonstrate that positive demand shocks in shorts typically predict negative return up to 6 months—but the statistical significance of these predictions lasts only for (the first) 2 months. In this case, although quarterly short interest changes serve well the goal of demonstrating the value of a *joint* analysis of long and short positions, the power could be lower than the case when monthly information is properly used.

To explore the full spectrum of the informativeness of the shorts, we conduct three steps of analysis related to quarterly short interest changes ($\Delta SI_{i,\,t}$). In the first step, we examine separately the return predicting power of quarterly increases and decreases in $SI_{i,\,t}$. That is, we keep the analysis at *quarterly frequency*, but separate the two directions of $\Delta SI_{i,\,t}$. This allows us to examine each demand variable in detail. Accordingly, we define two dummy variables. The first (second) is labeled *ΔSI_Top Tercile* (*ΔSI_Bottom Tercile*) which takes the value of one when the quarterly short interest change of a stock belongs to the top (bottom) tercile among all the stocks during the same period, and zero otherwise. Note that, in our previous tests, the first variable has been used to construct tercile-based informed short and hedging demand ($DShort^{Ter}_{i,t}$ and $DHeding^{Ter}_{i,t}$), while the second variable has been used to construct tercile-based informed long and unwinding demand ($DLong^{Ter}_{i,t}$ and $DUnwinding^{Ter}_{i,t}$).

We then regress the abnormal return on these two dummy variables, and report the results in Models (1) and (4) of Table 8. Interestingly, we find that increases in short interest predict negative return, while decreases in short interest do not exhibit the same degree of predicting power. This asymmetry resembles the difference between the return predictive power of positive and negative demand shocks in the short-selling market as reported in Cohen, Diether, and Malloy (2007), and is consistent with the notion that hedge funds dig out negative information of firms and short sell their stocks. That is, new openings of short positions may indicate the arrival of new (negative) information, which predict future return following standard informed trading models (e.g., Kyle, 1985). The closure of existing short positions, however, may not be as informative. Short sellers may choose to partially close or even fully cover their existing short positions not only when information changes but also when their return

Table 8
The return predictive power of short selling.

This table further explores the return predicting power of short interest. In baseline Model (1), abnormal return is regressed on a dummy variable that takes the value of one when the quarterly short interest change of a stock belongs to the top tercile among all the stocks during the same period and zero otherwise. This dummy variable, labeled "*ΔSI_Top Tercile,*" captures large increases in short interest. In Model (2), we decompose this dummy variable into three components, where each component is represented by a dummy variable. Specifically, conditioning on the occurrence of top-tercile short interest changes (i.e., *ΔSI_Top Tercile= 1*), the first (*DShort, Tercile-based*) and the second (*DHedging, Tercile-based*) dummy variables take the value of one when the contemporaneous change in hedge fund holdings of the same stock belongs to the bottom- and the top-tercile among all the stocks, respectively, and the third dummy variable (*D_Others*) takes the value of one otherwise. In Model (3), we further differentiate two scenarios of quarterly short interest changes, depending on whether short interest changes in each month of the quarter are along the same direction (*Consistent SI changes*) or not (*Inconsistent SI changes*), and apply the two scenarios to further decompose each of the three dummy variables. In baseline Model (4), abnormal return is regressed on a dummy variable that takes the value of one when the quarterly short interest change of a stock belongs to the bottom tercile among all the stocks and zero otherwise. This dummy variable captures large decreases in short interest and is labeled "*ΔSI_Bottom Tercile*." Model (5) further decomposes this dummy variable into three components: i.e., conditioning on the occurrence of bottom-tercile short interest changes (i.e., *ΔSI_Bottom Tercile= 1*), the first (*DLong, Tercile-based*) and the second (*DUnwinding, Tercile-based*) dummy variables take the value of one when the contemporaneous change in hedge fund holdings of the same stock belongs to the top and the bottom tercile among all the stocks, respectively, and the third dummy variable (*D_Others*) takes the value of one otherwise. Model (6) further differentiates, for each of the three dummy variables, two scenarios of quarterly short interest changes depending on whether short interest changes in each month of the quarter take the same direction (*Consistent SI changes*) or not (*Inconsistent SI changes*). A detailed definition of these variables is provided in Appendix A. The superscripts ***, **, and * refer to the 1%, 5%, and 10% levels of statistical significance, respectively.

Out-of-sample quarterly abnormal return (DGTW-adjusted) regressed on short interest and its components

	A. Top tercile short interest (SI) changes				B. Bottom tercile short interest (SI) changes		
	(1)	(2)	(3)		(4)	(5)	(6)
ΔSI_Top Tercile	−0.004*			*ΔSI_Bottom Tercile*	0.001		
	(−1.80)				(0.73)		
DShort, Tercile-based		−0.009***		*DLong, Tercile-based*		0.012***	
		(−3.28)				(4.45)	
DHedging, Tercile-based		0.003		*DUnwinding, Tercile-based*		−0.005**	
		(1.16)				(−2.19)	
D_Others		−0.007**		*D_Others*		−0.003	
		(−2.54)				(−1.25)	
DShort Consistent SI changes			−0.012***	*DLong_Consistent SI changes*		0.010**	
			(−2.76)				(2.29)
DShort _Inconsistent SI changes		−0.008***		*DLong_Inconsistent SI changes*			0.013***
			(−2.89)				(4.33)
DHedge_Consistent SI changes			−0.006	*DUnwinding_Consistent SI changes*		−0.004	
			(−1.27)				(−0.72)
DHedge_Inconsistent SI changes		0.005*		*DUnwinding_Inconsistent SI changes*	−0.006**		
			(1.86)				(−2.24)
D_Others_Consistent SI changes		−0.013***		*D_Others_Consistent SI changes*	−0.008*		
			(−2.82)				(−1.89)
D_Others_Inconsistent SI changes		−0.005*		*D_Others_Inconsistent SI changes*	−0.000		
			(−1.81)				(−0.09)
Div	−0.015	−0.014	−0.010	*Div*	−0.016	−0.016	−0.012
	(−0.35)	(−0.33)	(−0.24)		(−0.38)	(−0.37)	(−0.28)
LgAge	0.002*	0.002*	0.002*	*LgAge*	0.002*	0.002*	0.002*
	(1.92)	(1.95)	(1.80)		(1.94)	(1.93)	(1.70)
LgPrc	−0.005*	−0.005*	−0.006*	*LgPrc*	−0.005*	−0.005*	−0.006*
	(−1.74)	(−1.76)	(−1.90)		(−1.78)	(−1.75)	(−1.90)
LgTurn	0.002	0.002	0.001	*LgTurn*	0.001	0.001	0.001
	(0.76)	(0.73)	(0.67)		(0.50)	(0.44)	(0.36)
LgVol	−0.006	−0.006	−0.006	*LgVol*	−0.006	−0.006	−0.006
	(−0.99)	(−1.05)	(−1.03)		(−1.01)	(−1.01)	(−1.00)
SP500	0.004	0.004	0.004	*SP500*	0.004	0.004	0.005
	(1.09)	(1.14)	(1.21)		(1.29)	(1.28)	(1.42)
Constant	0.001	0.000	0.002	*Constant*	−0.002	−0.003	−0.001
	(0.06)	(0.01)	(0.09)		(−0.12)	(−0.15)	(−0.05)
Observations	121,216	121,216	120,927	*Observations*	121,216	121,216	120,927
R-square	0.022	0.024	0.026	*R-square*	0.023	0.024	0.025

goals are achieved. In the latter case, their private information already gets fully incorporated into stock prices; hence, stock prices more or less follow a random walk until the arrival of new information.[15]

To further understand the informativeness of shorts, our second step of analysis compares the above predicting power of quarterly short interest changes with that

[15] Boehmer, Jones, and Zhang (2008) find that shorting flow in NYSE—the sum of executed shares of short sale orders—is more informative than monthly changes in short interest. Since monthly change in short interest is the sum of shares shorted less all covering transactions plus other short sales (manual NYSE short sales and off-NYSE short sales), the result could suggest that covering of shorts are not as informative as opening.

of contemporaneous hedge fund holding changes. We decompose the top-tercile short interest changes into three cases, depending on whether the concurrent hedge fund holding changes are very positive, very negative, or otherwise. That is, conditioning on $\Delta SI\_Top\ Tercile=1$, we define three dummy variables as follows: the first two dummy variables ($DShort^{Ter}_{i,t}$ and $DHedging^{Ter}_{i,t}$) take the value of one when the contemporaneous changes in hedge fund holdings of the same stock belong to the bottom and top tercile among all the stocks, respectively, and the third dummy variable (D_Others) takes the value of one otherwise. The first two variables are exactly the same as what we have explored in Model (7) of Table 2, while the last variable allows us to examine the informativeness of shorts when hedge funds do not take strong actions in their holdings. Similarly, conditioning on the occurrence of bottom-tercile short interest changes (i.e., $\Delta SI\_Bottom\ Tercile=1$), we also define three dummy variables. The first ($DLong^{Ter}_{i,t}$) and the second ($DUnwinding^{Ter}_{i,t}$) take the value of one when the contemporaneous changes in hedge fund holdings of the same stock belong to the top and bottom tercile among all the stocks, respectively, and the third dummy variable (D_Others) takes the value of one otherwise. This decomposition allows us to understand the return predicting power of short interest changes in detail.

Model (2) reports the return predictability of the three components of positive short interest changes. We can see that, consistent with our previous results, informed short demand significantly predicts negative future return. Its statistical level is higher than that of $\Delta SI\_Top\ Tercile$ in Model (1), which lends support to the notion that the combination of long and short enhances the informativeness of the signal. Also, hedging demand predicts positive, albeit insignificant, return, suggesting that the long side could be more informative than the short side—though the difference is insignificant. Finally, in the remaining case when there are no strong signals from hedge fund holdings, positive changes in short interest (D_Others) predicts negative return. Both the magnitude and statistical power are higher than those of $\Delta SI\_Top\ Tercile$ in Model (1), suggesting that shorts are highly informative in this case.

Model (5) reports the return predictability of positive short interest changes. Informed long significantly predicts positive future return as expected. Unwinding demand predicts negative return, which suggests that the long side is significantly more informative than the short side in this case. Even when there are no strong signals from the long side, negative short interest changes (D_Others) predict insignificant return. Putting together, negative quarterly changes in short interest appear less informative than positive short interest changes or changes in hedge fund holdings.

But do these results, especially those of negative SI changes, imply a general conclusion that short selling is less informative than hedge fund holdings? The answer is no. To see the intuition, we conduct the third step of analysis in which we further incorporate monthly short selling information into the above analysis. More specifically, we use monthly SI information to differentiate the following two different scenarios of quarterly SI changes: in the first scenario, *monthly* SI changes have consistent signs within a same quarter (hereafter *Consistent SI Changes*), while in the second scenario they do not (hereafter *Inconsistent SI Changes*).[16] Since in the first scenario the long-term (quarterly) signals are consistent with the short-term (monthly) signals, we expect short selling thereof to be more informative.

Model (3) reports the results when $DShort^{Ter}_{i,t}$ and $DHedging^{Ter}_{i,t}$ are decomposed into these two scenarios. We can see that the predicting power of informed short demand remains significant in both scenarios. Hedging demand constructed on the basis of *inconsistent SI changes* now predicts significant positive return, suggesting that long-side operation outweighs shorts in terms of return predictive power in this scenario. However, hedging demand constructed on the basis of *consistent SI changes* now predicts negative yet insignificant return, suggesting that the power has now actually shifted to the short side, though the difference between the short and the long side is not statistically significant. Finally, when there are no strong signals from the long side, both *consistent* and *inconsistent* SI changes properly predict negative return, but the statistical power of the former ($t=-2.82$) is much higher than that of the latter ($t=-1.81$). In brief, we can see that *consistent* positive SI changes are more informative than *inconsistent* positive SI changes, and that the return predictive power of *consistent* and positive SI changes is at least at par with that of hedge fund holding changes.

Model (6) applies similar tests to $DLong^{Ter}_{i,t}$ and $DUnwinding^{Ter}_{i,t}$. We find that $DLong^{Ter}_{i,t}$ is informative in both scenarios. Meanwhile, the (negative) return predictability of $DUnwinding^{Ter}_{i,t}$ is statistically significant only in the scenario of *inconsistent SI changes*. For *consistent SI changes*, future return is still negative but no longer significant. Hence, *consistent* SI changes also improve the informativeness of shorts for negative SI changes.

This analysis also allows us to better understand the previous controversial results tabulated in Model (7) of Table 2. In that model, hedge fund holdings appear more informative than short interest changes, which may be at odds with the vast evidence regarding the informativeness of short selling. When we apply the two scenarios of *SI changes* to hedging and unwinding demand, however, we find that hedge fund holdings are more informative only than *inconsistent SI changes. Consistent SI changes*, by contrast, appear to be as informative as the long side, as hedging and unwinding demand lose their predicting power when they are constructed based on *consistent SI changes*. Hence, the result that the informativeness of quarterly SI changes can be further improved based on monthly information helps us to reconcile Table 2 with the short-selling literature. To save space, we tabulate the results in Table IN2 in the Internet Appendix.

Overall, we find that quarterly SI changes may lose some power in predicting stock return. However, our analysis also suggests that this problem does not affect the construction of informed demand (it indeed works against us in finding any significant predictive power of informed

[16] In our sample, 25.6% and 27.5% of positive and negative changes in SI are consistent, respectively.

Table 9
A placebo test using mutual fund holdings.
This table conducts a placebo test for the baseline quarterly Fama-MacBeth regression of Table 2 by replacing hedge fund holdings by mutual fund holdings as follows:

$$DGTW_{i,t+1} = \alpha_i + \beta_i \times Informed\ Demand_{i,t}^{MF} + C \times M_{i,t} + \epsilon_{i,t+1},$$

where $DGTW_{i,\ t+1}$ refers to the out-of-sample DGTW-adjusted abnormal return of stock i accumulated over quarter $t+1$; $Informed\ Demand_{i,t}^{MF}$ refers to a vector of informed demand variables contrasted from mutual fund holdings; and $M_{i,\ t}$ stacks a list of control variables, including *DIV, LgAge, LgPrc, LgTurn, LgVol*, and *SP500*. More specifically, $Informed\ Demand_{i,t}^{MF}$ is constructed in a similar way as before, except that we replace the aggregate hedge fund holdings information by the aggregate mutual fund holdings information. Models (1) and (2) regress quarterly and annual out-of-sample abnormal return on informed demand variables constructed from the aggregate mutual fund holdings, while Models (3) and (4) regress abnormal return on hedging and unwinding demand variables. Models (5)/(6) and Models (7)/(8) report similar regressions for tercile- and quintile-based informed demand variables, respectively. A detailed definition of these variables is provided in Appendix A. The superscripts ***, **, and * refer to the 1%, 5%, and 10% levels of statistical significance, respectively.

	DLong by changes				DLong by terciles		DLong by quintiles	
	Quarterly (1)	Annual (2)	Quarterly (3)	Annual (4)	Quarterly (5)	Annual (6)	Quarterly (7)	Annual (8)
DLong	−0.001	−0.002			0.001	0.004	0.003	0.008
	(−0.66)	(−0.43)			(0.54)	(0.79)	(0.86)	(1.13)
DShort	0.003	0.003			0.003	0.004	0.002	−0.000
	(1.40)	(0.60)			(1.22)	(0.59)	(0.98)	(−0.10)
DHedging			−0.003	−0.003	−0.002	−0.003	−0.005	−0.009
			(−1.60)	(−0.68)	(−0.93)	(−0.42)	(−1.46)	(−1.27)
DUnwinding			0.002	0.008*	0.005*	0.014*	0.010***	0.011
			(1.24)	(1.75)	(2.01)	(1.71)	(2.88)	(1.24)
Div	−0.015	−0.089	−0.014	−0.087	−0.015	−0.087	−0.013	−0.087
	(−0.42)	(−0.96)	(−0.38)	(−0.92)	(−0.41)	(−0.92)	(−0.35)	(−0.92)
LgAge	0.002	0.010***	0.002	0.010***	0.002*	0.010***	0.002*	0.010***
	(1.66)	(3.54)	(1.60)	(3.59)	(1.68)	(3.66)	(1.68)	(3.57)
LgPrc	−0.006**	−0.031**	−0.006**	−0.031**	−0.006**	−0.031**	−0.006**	−0.031**
	(−2.15)	(−2.57)	(−2.12)	(−2.53)	(−2.13)	(−2.57)	(−2.13)	(−2.56)
LgTurn	0.002	0.007	0.002	0.007	0.002	0.007	0.002	0.007
	(0.69)	(0.84)	(0.71)	(0.82)	(0.82)	(0.79)	(0.75)	(0.80)
LgVol	−0.005	−0.013	−0.005	−0.013	−0.005	−0.013	−0.005	−0.013
	(−0.86)	(−0.72)	(−0.85)	(−0.72)	(−0.85)	(−0.73)	(−0.85)	(−0.73)
SP500	0.007*	0.028**	0.006*	0.028**	0.006*	0.028**	0.006*	0.028**
	(1.99)	(2.57)	(1.92)	(2.57)	(1.76)	(2.46)	(1.81)	(2.49)
Constant	0.006	0.038	0.006	0.036	0.005	0.034	0.004	0.036
	(0.31)	(0.88)	(0.32)	(0.83)	(0.26)	(0.81)	(0.23)	(0.84)
Observations	133,417	126,087	133,417	126,087	133,417	126,087	133,417	126,087
R-square	0.021	0.018	0.021	0.018	0.022	0.019	0.022	0.019

demand), and that using monthly information could significantly increase the power of quarterly SI measures.

5.2. A placebo test based on mutual fund holdings

We now validate the importance of the hedge fund industry in processing information by conducting a placebo test in which we replace hedge fund holdings by mutual fund holdings. More specifically, beginning with Models (7) and (8) of Table 2, we replace informed demand variables in these regressions with similar variables constructed using mutual fund holdings. For instance, *informed long demand* is now defined as $DLong_{i,t}^{MF} = I\{\Delta HFOwn_{i,t}^{MF} > 0\} \times I\{\Delta SI_{i,t} < 0\}$, where $\Delta HFOwn_{i,t}^{MF} = HFOwn_{i,t}^{MF} - HFOwn_{i,t-1}^{MF}$ denotes changes in mutual fund holdings rather than hedge fund holdings.

The results are tabulated in Table 9. Models (1) and (2) regress quarterly and annual out-of-sample abnormal returns on informed demand variables. Models (3) and (4) regress abnormal return on hedging and unwinding demand variables. Models (5)/(6) and Models (7)/(8) report similar regressions for tercile- and quintile-based informed demand variables, respectively. We observe that mutual fund holding implied long and short demand variables are not informative. Hence, our previous analyses and conclusions are applicable only to the hedge fund industry. This finding is important in that it validates our motivation to jointly use hedge fund holdings and short selling information rather than combining the latter information with holdings of other institutional investors such as mutual funds.

5.3. Alternative explanations

Because mutual fund holdings are in general less informed than hedge fund holdings, hedge funds may exploit (less-informed) mutual fund trading, especially when such trading is driven by exogenous factors—such as large inflows and outflows, which may generate price pressures. If so, the aforementioned return predictability may be related to information on “dumb money”—i.e., exploitation of mutual fund flows. More specifically, if hedge funds can buy/sell stocks in which there are large mutual fund inflows/outflows, they may profit from the price effects of subsequent mutual fund trading induced by the inflows/outflows (e.g., Shive and Yun, 2013; Arif, Ben-Rephael, and Lee, 2014 provide evidence on daily frequencies). We therefore examine the relationship between

Table 10
Hedge fund demand and large mutual fund flows.

This table explores how informed demand predicts mutual fund flows. In Models (1)–(4) and Models (5)–(8) of Panel A, we regress, in Fama-MacBeth specifications, large mutual fund inflows and outflows on informed demand and hedging/unwinding demand, respectively. In Models (1), (2), (5), and (6), mutual fund flows at the stock level are measured by quarterly aggregate mutual fund holding changes scaled by lagged trading volume. Models (3), (4), (7), and (8) provide an alternative definition of extreme mutual fund flows, in which we compute the "active" part of mutual fund flows. More specifically, active flow is constructed as the difference between actual and expected number of shares held by mutual funds, divided by lagged trading volume. The expected number of shares held by fund f for stock i is computed as the value of the stock held by the fund if it keeps the same portfolio weights as last quarter adjusted for the passive effect of stock price change on portfolio weight change using the method of Kacperczyk, Sialm, and Zheng (2005) divided by stock price; we then sum this measure across all funds. Finally, large inflows and outflows are defined as mutual fund flows within the top and bottom 10% of the distribution in terms of magnitude, respectively. Panel B reports the results of Logit regression. To save space, we only tabulate the coefficients for the main variables. A detailed definition of these variables is provided in Appendix A. The full specifications of the regression parameters can be found in the Internet Appendix. The superscripts ***, **, and * refer to the 1%, 5%, and 10% levels of statistical significance, respectively.

Panel A: Fama-MacBeth regression of large mutual fund flows (top 10%) on hedge fund demand

	Total Inflow/Outflow		Active Inflow/Outflow			Total Inflow/Outflow		Active Inflow/Outflow	
	(1) Inflow	(2) Outflow	(3) Inflow	(4) Outflow		(5) Inflow	(6) Outflow	(7) Inflow	(8) Outflow
DLong	−0.006***	0.014***	−0.007***	0.002	*DHedging*	−0.001	−0.007***	0.002	−0.007***
	(−2.97)	(5.84)	(−3.01)	(0.60)		(−0.72)	(−2.95)	(0.70)	(−2.76)
DShort	0.005**	−0.010***	0.008***	−0.001	*DUnwinding*	0.003	0.003	−0.005**	0.013***
	(2.02)	(−4.86)	(3.17)	(−0.39)		(1.15)	(1.12)	(−2.03)	(4.02)
Controls	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
Observations	119,644	119,644	119,626	119,626		119,644	119,644	119,626	119,626
R-square	0.057	0.025	0.035	0.113		0.057	0.024	0.034	0.114

Panel B: Logit regression of large mutual fund flows (top 10%) on hedge fund demand

	Total Inflow/Outflow		Active Inflow/Outflow			Total Inflow/Outflow		Active Inflow/Outflow	
	(1) Inflow	(2) Outflow	(3) Inflow	(4) Outflow		(5) Inflow	(6) Outflow	(7) Inflow	(8) Outflow
DLong	−0.061**	0.143***	−0.083***	0.065**	*DHedging*	−0.027	−0.063***	0.024	−0.132***
	(−2.34)	(6.08)	(−3.32)	(2.54)		(−1.06)	(−2.58)	(0.99)	(−5.15)
DShort	0.075***	−0.147***	0.098***	−0.018	*DUnwinding*	0.043*	0.052**	−0.030	0.097***
	(2.91)	(−5.78)	(3.95)	(−0.68)		(1.65)	(2.08)	(−1.20)	(3.59)
Controls	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
Observations	119,644	119,644	119,626	119,626		119,644	119,644	119,626	119,626

mutual fund flows and informed demand. To the extent that the price effects are more significant for large flows, especially large outflows (e.g., Coval and Stafford, 2007), we explore whether informed or other demands can forecast large inflows and outflows associated with the mutual fund industry.

In the spirit of Shive and Yun (2013), we use quarterly aggregate mutual fund holding changes (scaled by lagged trading volume) to proxy for flows of capital into and out of stocks. Large stock-level inflows and outflows are subsequently defined as those among the top and bottom 10% in the cross section of mutual fund flows (our results are robust to alternative thresholds, such as 5%). We also provide an alternative definition of extreme mutual fund flows, in which we further compute the "active" part of mutual fund flows inferred from lagged portfolio weights.[17] Then, we regress these measures of large inflows/outflows on lagged hedge fund demand, and report the results in Table 10. We adopt Fama-MacBeth specifications in Panel A and Logit specifications in Panel B. In both panels, Models (1), (2), (5), and (6) and Models (3), (4), (7), and (8) report the results for the main and alternative proxies of flows, respectively.

The results are similar across the two panels. First of all, we find that informed long (short) demand predicts negative (positive) extreme inflows and positive (negative) large outflows. This result is the opposite of what a strategy of riding the price impact of large flows would predict. Hence, if anything, informed demand does not seem to be motivated by exploiting mutual fund flows. Rather, informed demand of hedge funds focuses on firm-specific information which mutual funds are not capable of replicating (e.g., Table 9), and mutual funds in this case may simply supply liquidity for such trades.

More importantly, we find that hedging/unwinding demand does seem to respond to mutual fund flows—i.e., hedge funds trade more on the potential occurrence of large outflows than inflows. If we look at the directions of trading, we see that hedge funds seem to unwind their positions before the occurrence of large outflows (i.e., unwinding demand increases while hedging demand decreases). Hence, hedge funds on average reduce their holdings before mutual fund fire sales, which can help them avoid the associated negative price impact of fire sales.

[17] More specifically, active flow is constructed as the difference between actual and expected number of shares held by mutual funds, divided by lagged trading volume. The expected number of shares held by fund f for stock i is computed as the value of the stock held by the fund if it keeps the same portfolio weights as last quarter [adjusted for the passive effect of stock price change on portfolio weight change using the method of Kacperczyk, Sialm, and Zheng (2005)] divided by stock price; we then sum this measure across all funds.

Table 11
Net demand and liquidity provision.
This table explores the relationship between liquidity and net demand changes. In Panel A, Models (1)–(3) regress the average turnover ratio of the firm in the concurrent period, the next quarter, and the next year with respect to the informed-demand quarter, on informed demand as well as a list of control variables. Models (4)–(6) apply the same analysis to hedging and unwinding demands. Panel B replaces the turnover ratio by the Amihud illiquidity measure of the corresponding period. To save space, we only tabulate the coefficients for the main variables. A detailed definition of these variables is provided in Appendix A. The full specifications of the regression parameters can be found in the Internet Appendix. The superscripts ***, **, and * refer to the 1%, 5%, and 10% levels of statistical significance, respectively.

Panel A: Turnover ratio vs. *hedge fund demands*

	(1) Concurrent	(2) Next quarter	(3) Next year		(4) Concurrent	(5) Next quarter	(6) Next year
DLong	−0.005***	−0.007***	−0.005***	*DHedging*	0.017***	0.017***	0.013***
	(−4.21)	(−5.62)	(−5.05)		(7.14)	(6.65)	(6.89)
DShort	−0.001	−0.000	−0.002	*DUnwinding*	−0.010***	−0.009***	−0.005***
	(−0.31)	(−0.03)	(−1.23)		(−7.25)	(−6.99)	(−4.78)
Controls	Yes	Yes	Yes		Yes	Yes	Yes
Observations	121,220	121,220	115,282		121,220	121,220	115,282
R-square	0.472	0.422	0.462		0.475	0.424	0.464

Panel B: Amihud illiquidity vs. *hedge fund demands*

	(1) Concurrent	(2) Next quarter	(3) Next year		(4) Concurrent	(5) Next quarter	(6) Next year
DLong	0.000	−0.000	−0.000	*DHedging*	0.000	−0.000	0.000
	(0.08)	(−1.19)	(−1.43)		(0.15)	(−0.28)	(0.02)
DShort	−0.000	0.000	−0.000	*DUnwinding*	0.000	−0.000	0.000
	(−1.27)	(0.16)	(−1.42)		(1.14)	(−0.29)	(1.10)
Controls	Yes	Yes	Yes		Yes	Yes	Yes
Observations	120,504	120,457	114,194		120,504	120,457	114,194
R-square	0.078	0.066	0.069		0.079	0.066	0.069

This pattern is consistent with Shive and Yun (2013). Using the 13F data, these authors find that hedge funds profit from trading against mutual fund flows and that hedge funds trade more on expected mutual fund fire sales than on inflows.

The above benefit alone, however, does not differentiate the potential motivations of hedge fund trading. Particularly, hedge funds may reduce their holdings before the occurrence of fire sales either due to risk management incentives (i.e., to reduce the total exposure to a potential risk) or because of profit-chasing reasons (i.e., to maximize the trading profits that can be reaped from mutual fund flows). Given that fire sales constructed at the stock level reflect the power of the entire mutual fund industry—and thus could be treated either as a source of risk or as a source of profit—it is difficult to judge, ex ante, which strategy will be preferred by the hedge fund industry. Short selling information, however, can be used to further differentiate the two. If hedge funds simply try to maximize trading profits, they should also open new short positions to ride on the negative price impact of fire sales. By contrast, risk management would motivate hedge funds to reduce both long and short positions—and thus their total exposure—to the potential source of risk. Since Table 10 demonstrates that hedge funds choose to also unwind their short positions, their trading behavior is more consistent with risk management incentives. Here again a *joint* analysis of hedge fund long and short information may shed new light on the interpretation of known empirical patterns.

Another possible explanation for the return predictability of informed demand is liquidity provision. If $DLong_{i,\,t}$ and $DShort_{i,\,t}$ are related to liquidity supply, i.e., stock purchases (sales) in the presence of selling (buying) pressure, these variables should be associated with a return premium that compensates for liquidity provision. To examine this potential explanation, we regress liquidity measured in different periods—concurrent, next quarter, and next year[18]—on $DLong_{i,\,t}$ and $DShort_{i,\,t}$. We use two different proxies for liquidity, turnover and the Amihud illiquidity measure.[19] We report the results for the turnover ratio in Panel A of Table 11, and the results for the Amihud illiquidity measure in Panel B. In each panel, we tabulate the results side by side for informed demand in Models (1)–(3) and for hedging and unwinding demands in Models (4)–(6).

Models (1)–(3) of Panel A show that informed long demand reduces concurrent and future liquidity, whereas informed short demand is unrelated to liquidity. The reduction in liquidity suggests that informed long demand, if anything, consumes liquidity rather than supplies it to the market. In Panel B, we find that not only does informed short demand remain unrelated to liquidity at any horizon, but informed long demand loses its power as well. Hence, informed demand does not appear to supply liquidity to the market. Furthermore, because the Amihud measure can also be interpreted as a price impact, informed short demand does not even appear to benefit from a price

[18] All periods are considered with respect to the quarter in which informed demand is constructed.

[19] More specifically, liquidity is proxied by turnover, defined as the average stock turnover in the last 3 months or the average Amihud illiquidity measure over the last 12 months, where the Amihud illiquidity measure in month t is the average daily ratio of the absolute stock return to the dollar trading volume over the month.

impact; this conclusion is consistent with our findings in Table 2. Models (4)–(6) in Panel A illustrate that hedging demand and unwinding demand may differ in their relationship with liquidity. However, neither demand is associated with Amihud illiquidity in Panel B, making a clean interpretation difficult to achieve.

Overall, these findings fail to support alternative interpretations of predictability that differ from the discovery of information about firm fundamentals.

6. Conclusion

We investigate the informational content of hedge fund trading through the joint use of information from both the long and short sides. We propose that opposite changes in short interest and hedge fund holdings are likely to be driven by information, whereas simultaneous increases (decreases) in short interest and hedge fund holdings are likely to be motivated by hedging (unwinding) incentives. This intuition allows us to utilize short selling and hedge fund holding information to identify informed long and short demand.

Using this identification strategy, we show that informed demand changes have high predictive power for returns. Furthermore, informed demand predicts out-of-sample firm fundamentals, such as ROA, earnings surprises, analyst revisions, and CARs. By contrast, informed demand does not appear to be driven by mutual fund flows or liquidity provision. These findings suggest that the observed return predictability of informed demand can be explained in terms of the discovery of information about firm fundamentals. This process, in turn, can be interpreted as reflecting a type of managerial skill in the hedge fund industry.

Our results suggest that short selling and hedge fund holdings complement each other in revealing important trading motivations of informed fund managers. More research that integrates short selling and hedge funds could therefore be fruitful in providing insights into information dissemination and asset price formation in the market.

Appendix A. Variable definitions

Informed and hedging (unwinding) demand variables	
DLong	Informed long demand: $DLong_{i,t}=I\{\Delta HFOwn_{i,t}>0\}\times I\{\Delta SI_{i,t}<0\}$, where $I\{.\}$ is an indicator function, and $\Delta HFOwn_{i,t}=HFOwn_{i,t}-HFOwn_{i,t-1}$ and $\Delta SI_{i,t}=SI_{i,t}-SI_{i,t-1}$ denote the changes in hedge fund holdings and short interest, respectively. *DLong* based on tercile or quintile partitions of $\Delta HFOwn$ and ΔSI is used in various tests as indicated in the tables.
DShort	Informed short demand: $DShort_{i,t}=I\{\Delta HFOwn_{i,t}<0\}\times I\{\Delta SI_{i,t}>0\}$. *DShort* based on tercile or quintile partitions of $\Delta HFOwn$ and ΔSI is used in various tests as indicated in the tables.
DHedge	Hedging demand: $DHedge_{i,t}=I\{\Delta HFOwn_{i,t}>0\}\times I\{\Delta SI_{i,t}>0\}$. *DHedge* based on tercile or quintile partitions of $\Delta HFOwn$ and ΔSI is used in various tests as indicated in the tables.
DUnwind	Unwinding demand: $DUnwind_{i,t}=I\{\Delta HFOwn_{i,t}<0\}\times I\{\Delta SI_{i,t}<0\}$. *DUnwind* based on tercile or quintile partitions of $\Delta HFOwn$ and ΔSI is used in various tests as indicated in the tables.
Stock performance and control variables	
$DGTW_i$	Benchmark-adjusted abnormal returns constructed following the method of DGTW (1997). Specifically, $DGTW_i$ is computed as the return of stock *i* net of the return of its style benchmark based on cross-sectional quintile partitions of market capitalization, book-to-market ratio, and prior 12-month returns.
Div	Dividend yield calculated as dividends divided by market capitalization.
Age	Number of months since the stock first appears in CRSP.
Prc	Price per share.
Turn	The average turnover (volume divided by shares outstanding) in the last month prior to the beginning of the quarter.
Vol	The standard deviation of returns over the past 24 months.
SP500	A dummy equal to one for stocks in the S&P 500 index and zero otherwise.
Characteristics related to anomalies	
B/M	Book-to-market ratio, defined as the book value of equity at the fiscal-year-end of the fiscal year ended before the most recent June 30, divided by the market capitalization on December 31 of that fiscal year.
Size	Market capitalization (in $ millions), defined as the product of stock price and the number of shares outstanding.
Lag Ret	Cumulative return from month −11 to month 0.
Gross profit to asset	Gross profit divided by total assets.
Operating profit	Gross profit minus selling, general, and administrative expenses minus interest expense, divided by book value of equity. Stocks with missing or negative book value are dropped.
Asset growth	Total assets divided by total assets of the previous fiscal year and then minus one.
Investment growth	Capital expenditure divided by capital expenditure of the previous fiscal year.
Net stock issuance	The split-adjusted shares outstanding divided by the split-adjusted shares outstanding of the previous fiscal year and then minus one. The split-adjusted shares outstanding are calculated as shares outstanding times the adjustment factor (AJEX).
Accruals	Change in operating working capital per split-adjusted share from last to current fiscal years divided by book value of equity per split-adjusted share. Operating working capital is computed as current assets minus cash and short-term investments minus the difference of current liability and debt in current liabilities.
Net operating assets	Operating assets minus operating liabilities, scaled by total assets at the end of last fiscal year. Operating assets are computed as total assets minus cash and short-term investment. Operating liabilities are computed as total assets minus debt included in current liabilities (filled as zero if missing) minus long-term debt (filled as zero if missing) minus minority interests (filled as zero if missing) minus book value of preferred stocks as described in the definition of book equity (filled as zero if missing), and minus common equity.

References

Ackermann, C., McEnally, R., Ravenscraft, D., 1999. The performance of hedge funds: risk, return, and incentives. Journal of Finance 54, 833–874.

Agarwal, V., Daniel, N.D., Naik, N.Y., 2009. Role of managerial incentives and discretion in hedge fund performance. Journal of Finance 64, 2221–2256.

Agarwal, V., Daniel, N.D., Naik, N.Y., 2011. Do hedge funds manage their reported returns? Review of Financial Studies 24, 3281–3320.

Agarwal, V., Naik, N.Y., 2004. Risks and portfolio decisions involving hedge funds. Review of Financial Studies 17, 63–98.

Aitken, M.J., Frino, A., McCorry, M.S., Swan, P.L., 1998. Short sales are almost instantaneously bad news: evidence from the Australian Stock Exchange. Journal of Finance 53, 2205–2223.

Akbas, F., Boehmer, E., Erturk, B., Sorescu, S., 2013. Short interest, returns, and fundamentals. University of Kansan, Singapore Management University, and Texas A&M University Unpublished working paper.

Aragon, G.O., Nanda, V., 2012. Tournament behavior in hedge funds: high--water marks, fund liquidation, and managerial stake. Review of Financial Studies 25, 937–974.

Arif, S., Ben-Rephael, A., Lee, C., 2014. Do short-sellers profit from mutual funds? Evidence from daily trades. Stanford University and Indiana University Unpublished working paper.

Asquith, P., Meulbroek, L., 1995. An empirical investigation of short interest. Massachusetts Institute of Technology Unpublished working paper.

Baker, S.R., Bloom, N., Davis, S.J., 2015. Measuring economic policy uncertainty. NBER working paper no. 21633.

Ben-David, I., Franzoni, F., Landier, A., Moussawi, R., 2013. Do hedge funds manipulate stock prices? Journal of Finance 68, 2383–2434.

Boehmer, E., Huszar, Z.R., Jordan, B.D., 2010. The good news in short interest. Journal of Financial Economics 96, 80–97.

Boehmer, E., Jones, C.M., Zhang, X., 2008. Which shorts are informed? Journal of Finance 63, 491–527.

Boehmer, E., Wu, J., 2013. Short selling and the price discovery process. Review of Financial Studies 26, 287–322.

Brunnermeier, M., Nagel, S., 2004. Hedge funds and the technology bubble. Journal of Finance 59, 2013–2040.

Cao, C., Chen, Y., Liang, B., Lo, A.W., 2013. Can hedge funds time market liquidity? Journal of Financial Economics 109, 493–516.

Chen, Y., Da, Z., Huang, D., 2015. Arbitrage trading: the long and the short of it. Texas A&M University, University of Notre Dame, and University of North Carolina at Greensboro Unpublished working paper.

Cohen, L., Diether, K.B., Malloy, C.J., 2007. Supply and demand shifts in the shorting market. Journal of Finance 62, 2061–2096.

Cooper, M.J., Gulen, H., Schill, M.J., 2008. Asset growth and the cross-section of stock returns. Journal of Finance 63, 1609–1652.

Coval, J., Stafford, E., 2007. Asset fire sales (and purchases) in equity markets. Journal of Financial Economics 86, 479–512.

Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. Journal of Finance 52, 1035–1058.

Diether, K.B., Lee, K.H., Werner, I.M., 2008. Short-sale strategies and return predictability. Review of Financial Studies 22, 575–607.

Engelberg, J.E., Reed, A.V., Ringgenberg, M.C., 2012. How are shorts informed? Short sellers, news, and information processing. Journal of Financial Economics 105, 260–278.

Fama, E.F., French, K.R., 2008. Dissecting anomalies. Journal of Finance 63, 1653–1678.

Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fung, W., Hsieh, D.A., 1997. Empirical characteristics of dynamic trading strategies: the case of hedge funds. Review of Financial Studies 10, 275–302.

Getmansky, M., Lo, A.W., Makarov, I., 2004. An econometric model of serial correlation and illiquidity in hedge fund returns. Journal of Financial Economics 74, 529–609.

Gompers, P.A., Metrick, A., 2001. Institutional investors and equity prices. Quarterly Journal of Economics 116, 229–259.

Griffin, J.M., Harris, J.H., Shu, T., Topaloglu, S., 2011. Who drove and burst the tech bubble? Journal of Finance 66, 1251–1290.

Griffin, J.M., Xu, J., 2009. How smart are the smart guys? A unique view from hedge fund stock holdings. Review of Financial Studies 22, 2531–2570.

Hirshleifer, D., Hou, K., Teoh, S.H., Zhang, Y., 2004. Do investors overvalue firms with bloated balance sheets? Journal of Accounting and Economics 38, 297–331.

Hirshleifer, D., Teoh, S.H., Yu, J.J., 2011. Short arbitrage, return asymmetry, and the accrual anomaly. Review of Financial Studies 24, 2429–2461.

Hirshleifer, D.A., Myers, J.N., Myers, L.A., Teoh, S.H., 2008. Do individual investors cause post-earnings announcement drift? Direct evidence from personal trades. Accounting Review 83, 1521–1550.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: an investment approach. Review of Financial Studies 28, 650–705.

Ince, O.S., Porter, R.B., 2006. Individual equity return data from Thomson Datastream: handle with care!. Journal of Financial Research 29, 463–479.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48, 65–91.

Kacperczyk, M., Sialm, C., Zheng, L., 2005. On the industry concentration of actively managed equity mutual funds. Journal of Finance 60, 1983–2012.

Kim, O., Verrecchia, R.E., 1994. Market liquidity and volume around earnings announcements. Journal of Accounting and Economics 17, 41–67.

Kosowski, R., Naik, N., Teo, M., 2007. Do hedge funds deliver alpha? A Bayesian and bootstrap analysis. Journal of Financial Economics 84, 229–264.

Kyle, A.S., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.

Lewellen, J., 2011. Institutional investors and the limit of arbitrage. Journal of Financial Economics 102, 62–80.

Ljungqvist, A., Qian, W., 2014. How constraining are limits to arbitrage? Evidence from a recent financial innovation. NBER working paper no. 19834.

Novy-Marx, R., 2013. The other side of value: the gross profitability premium. Journal of Financial Economics 108, 1–28.

Saffi, P.A.C., Sigurdsson, K., 2011. Price efficiency and short selling. Review of Financial Studies 24, 821–852.

Senchack, A.J., Starks, L.T., 1993. Short-sale restrictions and market reaction to short-interest announcements. Journal of Financial and Quantitative Analysis 28, 177–194.

Shive, S., Yun, H., 2013. Are mutual funds sitting ducks? Journal of Financial Economics 107, 220–237.

Sun, Z., Wang, A., Zheng, L., 2012. The road less traveled: strategy distinctiveness and hedge fund performance. Review of Financial Studies 25, 96–143.

Xing, Y., 2008. Interpreting the value effect through the Q-theory: an empirical investigation. Review of Financial Studies 21, 1767–1795.