The Determinants of Increased Short-Selling Activity

Hung Wan Kot
Department of Finance and Decision Sciences
Hong Kong Baptist University
Kowloon Tong, Hong Kong
Phone: (852) 3411 7558
Fax: (852) 3411 5585
E-mail: hwkot@hkbu.edu.hk

This draft: March 2014

I would like to thank Kalok Chan, Sudeep Ghosh, Zsuzsa Huszar, Joseph Lee (SFC Hong Kong), Lewis Tam and the participants in the finance seminars at Curtin University, Fudan University, Hong Kong Polytechnic University and the AsianFA/TFA conference in Taipei for their helpful comments and suggestions. Thanks also to Harry K.M. Leung for his capable research assistance. Any errors are my own.

The Determinants of Increased Short-Selling Activity

Abstract

Using hedge fund assets and institutional ownership (IO) as proxies for short-selling demand and supply, this paper investigates the determinants of increased short-selling activity on the NYSE/Amex and Nasdaq from 1988 to 2011. Hedge fund assets, IO and liquidity are significantly related to short-selling activity, and the relationships are stronger in the 2001-2011 period than in the 1988-2000 period. The ownership of short-term hold institutions has a stronger relationship with short-selling activity than ownership of long-term hold institutions. Further, both the expected and unexpected components of short selling have stronger relationships with short-term IO than with long-term IO.

Key words: short selling, hedge funds, institutional ownership, liquidity, demand and supply

JEL code: G01, G11, G14

1. Introduction

Short selling has increased significantly over the last two decades. For example, short interest grew by 15% on the NYSE/Amex and 16% on the Nasdaq annually from 1988 to 2011. The short interest ratio (SIR), calculated as the short interest scaled by the number of outstanding shares, has also increased sharply. The median SIR on the NYSE was 0.84% in January 1988 and 4.61% in December 2011. Diether, Lee and Werner (2009a) report that short sales accounted for 24% of NYSE trading volume and 31% on the Nasdaq in 2005. Stock prices are more accurate when short sellers are more active and market quality decreases when short selling is banned (Boehmer and Wu, 2013; Boehmer, Jones and Zhang, 2013). Evidence shows that short selling has played an increasingly important role in equity trading.

Studies have focused on whether short sellers are skilled investors and can produce abnormal returns, and on determining the information advantage of short sellers. Another stream of recent short-selling studies has focused on the 2008 short-sale ban and intraday short selling of regulation SHO pilot program stocks. However, the significantly increased short-selling activity itself has received limited attention in academic research. No paper has investigated why short-selling activity has increased so much over the past two decades.[1] The increase could be due to the rapid growth of the hedge fund industry, the increased availability of shares to lend in the market or other reasons.

In this paper, I investigate the increased short-selling activity from both the demand and supply sides. The hedge fund industry has expanded rapidly since the 2000s. Goldman Sachs (2009) assumes that hedge funds account for 85% of all of the short interest in the market. Using the short portfolio value reported by Goldman Sachs (2009) and matched with the short interest

[1] Hanson and Sunderam (2013) mention that short-selling activity has increased significantly over the past two decades. However, the focus of their paper is on the growth of arbitrage capital and the implications for strategy returns.

reported by the exchanges, I find that hedge funds accounted for around 80% of all of the short interest from the second quarter of 2007 to the second quarter of 2009, varying from 85% at the end of the third quarter of 2008 to 74% in the fourth quarter of 2008. I hypothesize that the increased short-selling activity had a stronger positive relationship with the shorting demand from the hedge funds in the later period.

Institutional ownership (IO) has also increased significantly, making more stocks available to borrow and short. The proportion of listed shares owned by institutions was 19% in the first quarter of 1988 and 45% in the last quarter of 2011. I hypothesize that there is a stronger positive relationship between IO and short-selling activity in the later period. Following Yan and Zhang (2009), I further classify the institutions into long- and short-term hold institutions. Long-term hold institutions such as index funds and pension funds intend to lend shares in the market over the long term, and provide a stable source for shares to be borrowed. Short-term hold institutions such as active traded mutual funds also have the incentive to lend shares to increase fund performance; however, their supply to the lending market is less stable. An increase or decrease in the shares supplied by short-term hold institutions shifts the supply curve of the shares available to short. Therefore, a change in supply from short-term hold institutions relates more to the changes in short-selling demand. I hypothesize that short-selling activity is more strongly related to the ownership of short-term hold institutions than that of long-term institutions.

I further investigate the relationship between liquidity and short selling, as the overall trading volume has also increased in the past two decades. I specifically investigate the bid-ask spread (SPREAD) and Amihud's (2002) illiquidity measures (ILLIQ). These two variables measure different aspects of liquidity. SPREAD captures the cost of immediacy, and ILLIQ captures how much a given trading volume moves the price (Corwin and Schultz, 2012). The minimum tick size decreased from one eighth to one sixteenth and from one sixteenth to one penny in 1997 and 2001, respectively (Corwin and Schultz, 2012). If liquidity is an important determinant of the increased short-selling activity, I hypothesize that the relationships between SPREAD and SIR and between ILLIQ and SIR are stronger in the later period.

Using short-interest data taken from the NYSE/Amex and Nasdaq, the total assets of the hedge funds and the IO data from 1988 to 2011, I find strong evidence in support of these hypotheses. Hedge fund assets, IO and liquidity are each significantly related to short-selling activity, and the relationships are stronger in 2001-2011 than in 1988-2000. Short-term IO has a stronger relationship with short-selling activity than long-term IO. Both the expected and unexpected components of short selling have stronger relationships with short-term IO than with long-term IO. The results are robust after controlling for the firm characteristics, other short-selling potentials and market-wide information.

This paper contributes to the literature in the following way. First, to the best of my knowledge, it is the first paper to investigate why short selling increased so much from the demand and supply sides. I provide direct evidence of a strong relationship between hedge fund markets and short-selling activity. Further, I find new evidence that short-term IO provides the main supply of shares rather than long-term IO. These results help us better understand the recent short-selling increase.

Asquith, Pathak and Ritter (2005) use short interest as a proxy for demand and IO as a proxy for supply to investigate future abnormal returns. Cohen, Diether and Malloy (2007) use stock loan fees and quantities to isolate shifts in the shorting demand and supply and their relationships with future returns. This paper extends the sample period to 2011, covering the 2007-2009 financial crisis and post-crisis period, and complements the short-selling literature from a large panel with 24 years of data to understand how demand and supply are related to increased short selling.

The remainder of this paper is organized as follows. In Section 2, I provide a short literature review and develop hypotheses. In Section 3, I describe the data source and construction of the variables. In Section 4, I test the three hypotheses. In Section 5, I test the joint hypotheses with control variables in a multivariate regression framework. In Section 6, I further investigate the expected and unexpected components of short interest. I conclude the paper in Section 7.

2. Recent literature and hypotheses

2.1 Short selling

Short-selling studies have usually focused on the relationship between the level of short interest and future returns and perhaps expectedly find that it is negative (Seneca, 1976; Senchack and Starks, 1993; Desai et al., 2002).[2] Some studies have investigated the role of short-sale constraints and future returns and found that short-sale constrained stocks tend to have lower future returns (Chen et al., 2002; Arnold et. al., 2005; Asquith et. al., 2005). Other studies have attempted to determine whether short sellers short stocks based on fundamental financial characteristics or firm-specific information (Dechow et. al., 2001; Christophe et al., 2004). Although most of the early studies are based on monthly short-interest data, recent studies have relied more heavily on daily short-selling volume. Edwards uses daily volumes to investigate the shorting of IPOs and finds that short selling does not explain IPO underpricing (Edwards et al. 2010). Henry and Koski (2010) investigate the short selling of SEOs and find no evidence of informed short selling, but rather patterns consistent with manipulative trading. No weekend effect in short selling has been found for NYSE or Nasdaq stocks (Blau et al., 2009; Christophe et al., 2009).

Using Regulation SHO data, Diether et al. (2009a) demonstrate that short sellers both

[2] For the institutional aspects of short selling, refer to studies by D’Avolio (2002) and Geczy et al. (2002).

increase trading following positive returns and correctly predict abnormal negative returns, indicating that short sellers trade on short-term overreactions to stock prices. In a further study using the same data, Diether et al. (2009b) find that short-selling activity increases for both NYSE- and Nasdaq-listed pilot stocks and that daily stock returns and volatility are unaffected.

Cohen et al. (2007) use proprietary data related to stock loan fees and quantities to examine the link between shorting and stock prices. They find that shorting demand is an important predictor of future stock returns. The prediction is stronger in environments with less public information flow, suggesting that shorting is an important mechanism for signaling private information. Boehmer et al. (2008) analyze proprietary NYSE order data from 2000 to 2004 and find that short sellers are well informed, especially institutional and non-program short sellers. In a further study, Boehmer et al. (2012) find that heavier shorting occurs the week before negative earnings surprises, analyst downgrades and downward revisions in analyst earnings forecasts, and suggest that short sellers know more than analysts about firm fundamentals. Using the Regulation SHO short-selling data and news events from the Dow Jones archive, Engelberg et al. (2012) show that the trading advantage of short sellers arrives largely from their ability to analyze publicly available information.

2.2 Hypotheses

I develop three hypotheses to understand why short selling has increased in the last two decades.

The hedge fund market has expanded rapidly in the last few decades. It is well known that hedge funds rely heavily on long-short strategies and high-frequency trading. Hedge funds also trade on private information (Massoud et al., 2011). While it is widely believed that hedge funds contribute the most to short selling, the evidence is limited. Goldman Sachs (2009) assumes that hedge funds account for 85% of all of the short interest in the market. Using the short portfolio value reported by Goldman Sachs (2009) and matched with the short-interest value reported by the exchanges, I also find that hedge funds count for around 80% of all of the short interest from the second quarter of 2007 to the second quarter of 2009, varying from 85% in the third quarter of 2008 to 74% in the fourth quarter of 2008. Therefore, I make the following hypothesis.

H1: Short-selling activity has a positive relationship with the hedge funds market, and the relationship is stronger in the later period.

Short sellers usually need to borrow shares of interest, and they usually borrow from an institutional investor through a broker. IO has increased significantly over the past two decades in the U.S. market, making it easier to borrow shares for short sales. A higher IO implies a lower level of short-sale constraints and induces more short-selling activity (Nagel, 2005). In addition, long-term hold institutions such as index funds and pension funds intend to lend the shares in the market over the long term, and provide the stable supply for shares to be borrowed. Meanwhile, while short-term hold institutions such as active traded mutual funds also have the incentive to lend shares to increase fund performance, their supply to the lending market is less stable. The increase or decrease of the shares supplied from the short-term hold institutions shift the supply curve of the shares available to short. Therefore, changes in the supply from the short-term hold institutions relate more to changes in the short-selling demand. Therefore, I make the following hypotheses.

H2a: Short-selling activity is positively related to institutional ownership, and the relationship should be stronger in the later period.

H2b: Short-selling activity has a stronger relationship with the ownership of short-term hold institutions than that of long-term hold institutions.

Liquidity must be considered in any type of securities trading, especially in short-term trading. The overall trading volume has also increased in the past two decades. Short sellers presumably take liquidity risk into account, and greater liquidity in a stock should promote short selling. The minimum tick size decreased from one eighth to one sixteenth and from one sixteenth to one penny in 1997 and 2001, respectively (Corwin and Schultz, 2012). If liquidity is an important determinant of increased short-selling activity, I predict that the relationship between the bid-ask spread and short selling is stronger in the later period. Therefore, I make the following hypothesis.

H3: Short-selling activity is positively related to liquidity, and the relationship is stronger in the later period.

3. Methods

3.1 Data

The monthly short-interest data were obtained directly from the exchanges for 1988-2008 and from Compustat for 2009-2011.[3] The daily and monthly stock return data, daily closing bid-ask spreads and CRSP value-weighted index were obtained from the CRSP database. The accounting data were obtained from the annual Compustat time series. The quarterly IO data were obtained from Thomson Financial's institutional database. The monthly hedge fund data were obtained directly from Hedge Fund Research Inc. The information on options availability was obtained from the Berkeley Options Database for 1988-1995 and from the Option Metrics Database for 1996-2011. The S&P 500 Index constituents were obtained from Compustat's monthly updates,

[3] Short interest has been published in the middle and at the end of each month since 2007. It had previously been published only in the middle of the month. For the sake of consistency, the short interest recorded in the middle of each month was used for the entire period. The short interest for the Nasdaq began in June 1998.

which are available at WRDS. The CBOE VIX index was obtained directly from the CBOE website.[4]

Figure 1 shows the time-series trend for the aggregate short interest and the SIR on the NYSE/Amex and Nasdaq from 1988 to 2011. The figure shows a similar pattern for both the aggregate short interest and SIR, i.e., both have increased significantly over the past two decades. Both the aggregate short interest and SIR on the NYSE/Amex are consistently higher than those on the Nasdaq from 1988 to 2011, indicating more short-selling activity on the NYSE/Amex stocks. The sample period covers two major events that affected the U.S. markets, including the Internet bubble burst around 2000 and the financial crisis from 2007 to 2009. Figure 1 shows that the short-selling activity does not decrease much after the Internet bubble burst around 2000. However, short-selling activity decreases significantly during the financial crisis. For example, on the NYSE/Amex, the SIR increases from 1.63% in January 2000 to a peak of 9.7% in August 2008, and then decreases sharply to 4.7% in December 2009. Despite the large variation in short-selling activity around the financial crisis period, short-selling activity increases significantly from the late 1980s to 2011.

[Figure 1 here]

Because the short positions of hedge funds are not required for disclosure, it is impossible to investigate the relationship between short selling and the short sale of hedge funds at the stock level. Further, not all hedge funds report their long and short positions. It is also impossible to use the short sale value of hedge funds to estimate the total value of short sales at the fund level. The best option is to use the total value of the hedge funds asset as a proxy for the short-selling demand.

[4] http://www.cboe.com/micro/VIX/historical.aspx.

I use Hedge Fund Research Inc.'s hedge fund database, which is one of the main hedge fund databases, in combination with active and dead funds. The database covers funds dating back to 1980, and includes more than 19,600 funds in total. Up to April 2013, the database includes more than 7,200 active funds and 12,300 dead funds (excluding funds of funds (4,878) and index funds (163)). The main strategies of a hedge fund include equity hedge, event-driven, macro and relative value. I further restrict the fund investment region to the North America and the fund asset currency to U.S. dollars.

Figure 2 reports the number of hedge funds and the total assets from 1988 to 2011. The graph shows a rapid growth in the hedge funds market from 1988 to 2006, whereupon the market shrinks significantly along with the subprime mortgage crisis and then the financial crisis. At the beginning of 1988, there are 25 hedge funds with $310 million in assets. At the end of 2007, there are 2252 hedge funds with $355919 million in assets. The asset growth rate is 45% per year. However, at the end of 2011, there are 1655 hedge funds with $256813 million in assets. The asset values show a "V" pattern, and the number of hedge funds shows a downward sloping line from 2008 to 2011.

[Figure 2 here]

3.2 Main variables

An SIR is constructed to track the short-selling activity on a monthly basis. It is defined as the short interest scaled by the number of outstanding shares.

I use LNHFA to proxy for the short-selling demand, which is defined as the natural logarithm of the total hedge fund assets on a monthly basis, excluding funds of funds and index funds.

I use IO to proxy for the short-selling supply, which is defined as the fraction of a firm's outstanding shares owned by institutions on a quarterly basis.

Following Yan and Zhang (2009), I further classify IO into long-term hold institutions (IO_LONG) and short-term hold institutions (IO_SHORT) based on their portfolio turnovers over the previous four quarters. In each quarter, each institution's aggregate purchases and sales are computed as

$$CR\_buy_{k,t} = \sum_{\substack{i=1 \\ S_{k,i,t} > S_{k,i,t-1}}}^{N_k} \left| S_{k,i,t} P_{i,t} - S_{k,i,t-1} P_{i,t-1} - S_{k,i,t-1} \Delta P_{i,t} \right| \tag{1}$$

$$CR\_sell_{k,t} = \sum_{\substack{i=1 \\ S_{k,i,t} \leq S_{k,i,t-1}}}^{N_k} \left| S_{k,i,t} P_{i,t} - S_{k,i,t-1} P_{i,t-1} - S_{k,i,t-1} \Delta P_{i,t} \right| \tag{2}$$

Here, $P_{i,t-1}$ and $P_{i,t}$ are the share prices of stock *i* at the end of quarters *t*-1 and *t*, respectively, and $S_{k,i,t-1}$ and $S_{k,i,t}$ are the numbers of shares of stock *i* held by investor *k* at the end of quarters *t*-1 and *t*, respectively. $CR\_buy_{k,t}$ and $CR\_sell_{k,t}$ are institution *k*'s aggregate purchases and sales for quarter *t*, respectively. Institution k's churn rate for quarter *t* is defined as

$$CR_{k,t} \equiv \frac{min\left(CR_{buy_{k,t}}, CR_{sell_{k,t}}\right)}{\sum_{i=1}^{N_k} \frac{S_{k,i,t} P_{i,t} + S_{k,i,t-1} P_{i,t-1}}{2}} \tag{3}$$

Each institution's average churn rate over the previous four quarters is then quantified as

$$AVG\_CR_{k,t} = \frac{1}{4} \sum_{j=0}^{3} CR_{k,t-j} \tag{4}$$

At the end of each year, the institutional investors are then sorted into three groups based on $AVG\_CR_{k,t}$. Those with the highest $AVG\_CR_{k,t}$ are classified as short-term institutional investors, and those ranked in the bottom third are classified as long term. These classifications are then used in the next year to compute the proportion of short-term and long-term IO. IO_LONG (IO_SHORT) is defined as the number of shares held by long-term (short-term) institutional investors scaled by the total number of shares in the outstanding firms.

Figure 3 shows the IO, IO_LONG and IO_SHORT from 1988 to 2011. There is a significant increase in IO. The IO values are 0.19 and 0.45 in the first quarter of 1988 and final quarter of

2011, respectively. The ownership of short-term hold institutions also increases from 0.1 in the first quarter of 1988 to 0.195 in the final quarter of 2011. The ownership of long-term hold institutions does not increase before 2000, and then increases significantly afterwards. The corresponding IO_LONGs are 0.04, 0.05 and 0.14 in the first quarter of 1988, first quarter of 2000 and final quarter of 2011, respectively. Figure 3 shows that the level of IO_SHORT is consistently higher than that of IO_LONG, indicating that most of the shares are owned by short-term-trading financial institutions.

[Figure 3 here]

I use two variables to proxy for liquidity. SPREAD is a monthly average of the daily closing bid-ask spread, which is calculated as the bid-ask difference scaled by the average of the bid and ask stock prices. ILLIQ is the measure of firm illiquidity used by Amihud (2002) and is calculated as follows:

$$ILLIQ_{it} = \frac{1}{D_{im}} \sum_{t=1}^{D_{im}} \frac{|R_{imd}|}{VOLD_{imd}} \tag{5}$$

where D_{im} is the number of days for which data are available for stock *i* in month *m*, R_{imd} is the return on stock *i* on day *d* of month *t* and $VOLD_{imd}$ is the respective daily volume in dollars.

3.3 Control variables

Based on the literature, I construct three sets of control variables. I begin by including two variables related to firm characteristics. The first variable is LNCAP, which is the natural logarithm of a firm's market capitalization calculated monthly. Market capitalization could be a proxy for borrowing cost, which should have a positive effect; it could also be a proxy for information transparency, and thus be expected to have a negative effect on short-selling activity. The relationship between LNCAP and SIR could be positive or negative. The second variable is BM, which is the book-to-market ratio of the firm's assets calculated monthly. BM is adjusted by the industry average using two-digit SIC codes and serves as a proxy for overvaluation. This should predict a negative relationship between BM and SIR.

The control variables in the second set are related to other potential short-selling reasons. I include five variables here. The first variable is PR1Y, which is the cumulative return on a firm's stock in the previous year from month *t*-12 to *t*-1. Jegadeesh and Titman (1993) document a price continuation effect over a 3- to 12-month period and show that a momentum strategy of buying past winners and selling past losers can earn a return of 1% per month. Some short sellers are trend traders who close their position if the stock price has increased recently and open a new short position if it has decreased. This faith in momentum should manifest in a negative relationship between PR1Y and SIR.

I construct two variables related to arbitrage and hedging. D_DEBT is a dummy variable representing the existence of convertible debt (1 for yes, 0 for no) evaluated using the Compustat annual data. D_OPT is a dummy variable representing the existence of options on a firm's stock (1 for yes, 0 for no) evaluated monthly. An investor may short a stock to take advantage of a price differential between a convertible security and the underlying stock, or short an acquirer's stock based on an announcement of a merger or an acquisition (Arnold et al., 2005). Market makers in options must also short stocks to hedge their positions (Grundy et al., 2012). Both D_DEBT and D_OPT should be positively related to SIR.

The fourth variable is VOLAT, which is the standard deviation of the daily stock returns over the past 3 months and a proxy for the difference in investor opinions.[5] Given the same level of short-sale constraints, more investors will short a stock if there is a large difference of opinion. Boehme et al. (2006) show that a dispersion in investor opinion in the presence of short-sale

[5] Another popular proxy is the dispersion in analysts' earnings forecast (DISP), which requires that the stock be covered by at least three analysts. Many small stocks would be excluded if I were to use DISP as a proxy.

constraints leads to stock price overvaluation. I expect a positive relationship between VOLAT and SIR.

The fifth variable is D_SP500, which is a dummy variable indicating whether the stock is included in the S&P 500 Index in a given month (1 for yes, 0 for no). Most S&P 500 Index constituent stocks have large capitalization, are very liquid and have a high proportion of IO. When a stock is added to the S&P 500 Index, its liquidity improves and its opportunities for arbitrage between the index futures markets and the underlying stock increase. All of these factors should promote short selling. However, index constituent stocks are covered by more financial analysts, have been listed longer and have a lower level of information asymmetry, and are therefore less likely to be overpriced. Investors have less of an incentive to short such highly transparent stocks despite the weaker short-sale constraints. The information advantage effect may or may not dominate the short-sale constraint effect, suggesting a negative or positive relationship between S&P 500 Index membership and short-selling activity.

The control variables in the third set are related to market return and volatility. First, MKT represents the monthly returns of the CRSP value-weighted index. Second, LNVIX is the natural logarithm of VIX, which is the 15th close CBOE VIX index.[6] I use these two variables to control any potential short selling related to the market condition.

3.4 Preliminary statistics

Before I reveal the regression results, I report the summary statistics and correlation matrix among the variables.

Table 1 reports the summary statistics. Consistent with Figure 1, the average SIR on the NYSE/Amex is 3%, which is higher than the 2% SIR average on the Nasdaq. The IO of the

[6] I use VXO for the pre-1990 volatility indexes.

NYSE/Amex stocks (45%) is higher than that of the Nasdaq stocks (29%). The ownership of long-term hold institutions (IO_LONG) is lower than that of short-term hold institutions (IO_SHORT) on both the NYSE/Amex and Nasdaq. For example, on the NYSE/Amex, the IO_LONG and IO_SHORT are 10% and 22%, respectively. The NYSE/Amex stocks also have higher liquidity as measured by the bid-ask spread and Amihud illiquidity measure.

[Table 1 here]

Consistent with the notion that the NYSE/Amex stocks are relatively large in terms of market capitalization, the stocks have more underlying options listed on exchanges and are included more often in the S&P 500 Index.

Table 2 reports the correlation coefficients. The correlative signs are consistent with the predictions. For example, SIR is positively correlated with LNHFA and IO_SHORT, and negatively correlated with SPREAD and ILLIQ. Among the control variables, some have high correlations as expected. For example, in the NYSE/Amex markets, market capitalization is highly correlated with IO (0.5), highly negatively correlated with the bid-ask spread (-0.47), highly negatively correlated with ILLIQ (-0.54), highly correlated with the options availability dummy (0.59) and highly correlated with the S&P 500 Index constituent stock dummy (0.58).

[Table 2 here]

4. Hypotheses tests

4.1 Hedge funds and short selling

To test H1 on the relationship between short-selling activity and the hedge fund market, I perform the following univariate regressions:

$$LNSI_t = \alpha + \beta LNHFA_t + \varepsilon \quad (6)$$

$$SIR_t = \alpha + \beta LNHFA_t + \varepsilon \quad (7)$$

LNSI is the natural logarithm of the short-interest dollar value of the common stocks, SIR is the average SIR value of the common stocks and LNHFA is the natural logarithm of the total hedge fund assets.

Table 3 reports the results. Consistent with H1, I find strong evidence of a relationship between the hedge funds market and short-selling activity. Equation (6) tests the relationship between hedge fund assets and short-interest value. On the NYSE/Amex, the coefficients are 0.1076 (with a t-value of 3.66) for 1988-2000 and 1.0475 (with a t-value of 7.38) for 2001-2011. On the Nasdaq, the coefficients are 0.1467 (with a t-value of 3.86) for 1988-2000 and 1.0714 (with a t-value of 7.85) for 2001-2011. The relationship between the hedge funds market and short-selling activity is significantly positive. More importantly, the difference in coefficients as shown in Panel C is also statistically significant. This means that the relationship between the hedge funds market and short-selling activity is stronger for the 2001-2011 period.

[Table 3 here]

Equation (7) tests the relationship between hedge fund assets and the general short-interest level as measured by the average SIR. The results are also presented in Table 3, and are similar to the results for Equation (6). For example, on the NYSE/Amex, the coefficients are 0.0008 (with a t-value of 2.14) for 1988-2000 and 0.0419 (with a t-value of 4.17) for 2001-2011. The difference in the coefficients for the two sub-periods is 0.0411 (with a t-value of 4.09).

4.2 Institutional ownership and short selling

To test H2 on the relationship between IO and short-selling activity, I perform the following two regressions:

$$SIR_{it} = \alpha + \beta IO_{it}(or\ IO\_LONG, IO\_SHORT) + \varepsilon \quad (8)$$

$$SIR_{it} = \alpha + \beta_1 IO\_LONG_{it} + \beta_2 IO\_SHORT_{it} + \varepsilon \quad (9)$$

Equation (8) is used to test H2a and Equation (9) is used to test H2b. The results are reported in Table 4. The t-values are clustered by the firm-plus-year dummy (Peterson, 2009).

The results in Panel A of Table 4 support H2a. For example, on the NYSE/Amex, the coefficients are 0.0124 (with a t-value of 9.91) for 1988-2000 and 0.0403 (with a t-value of 24.22) for 2001-2011. On the Nasdaq, the coefficients are 0.0292 (with a t-value of 23.87) for 1988-2000 and 0.0659 (with a t-value of 43.31) for 2001-2011. The difference in coefficients between the two sub-periods is 0.0279 (with a t-value of 14.90) on the NYSE/Amex and 0.0367 (with a t-value of 19.83) on the Nasdaq, respectively. I also report the relationship between IO_LONG and IO_SHORT using SIR in the sub-periods and find results similar to those of IO. The results clearly show that IO and SIR are positively related, and that the relationship is stronger in the 2001-2011 period.

[Table 4 here]

Panel B of Table 4 reports the results for Equation (9). The results show that the coefficient of IO_LONG is 0.0202 (with a t-value of 4.04) and that of IO_SHORT is 0.0615 (with a t-value of 18.04) on the NYSE/Amex. The difference in coefficients is 0.0412 (with a t-value of 6.40). On the Nasdaq, the coefficient of IO_LONG is 0.0526 (with a t-value of 10.59) and that of IO_SHORT is 0.0886 (with a t-value of 36.18). The difference in coefficients is 0.036 (with a t-value of 6.65). The results from Panel B of Table 4 support H2b, which predicts that the relationship between short-selling activity and short-term IO is stronger than that between short selling-activity and long-term IO.

To further check the relationship between IO (or IO_LONG, IO_SHORT) and SIR, I rank the stocks into 10 deciles by IO (or IO_LONG, IO_SHORT independently) in each month and compute the average SIR in each decile. I then compute the time-series average of the SIR in each decile. The results are plotted in Figure 4. There is a positive relationship between SIR and IO, which is consistent with previous studies that show that the IO level is a proxy for short-sale constraints (Chen et al., 2002). Figure 4 shows a steeper curve among the Nasdaq stocks, which is consistent with the Nasdaq IO being lower than that of the NYSE/Amex. Therefore, short-selling demand is more sensitive to IO in the Nasdaq than in the NYSE/Amex.

[Figure 4 here]

Figure 4 also shows that the IO_SHORT curve is steeper than the IO_LONG curve. On the NYSE/Amex, the IO_LONG curve is quite flat, indicating that short selling is insensitive to the IO_LONG level. However, short selling is sensitive to the IO_SHORT level, and SIR varies from 2% in decile 0 to 5% in decile 9. On the Nasdaq, the SIR level of decile 9 is even lower than those of decile 7 and 8 if sorted by IO_LONG. The IO_LONG curve is less steep compared with the IO_SHORT curve. The IO_SHORT curve is more sensitive among the high IO_SHORT deciles. In conclusion, Figure 4 further confirms the results in Table 4, which shows that IO_SHORT has a stronger relationship with short-selling activity than IO_LONG.

The results in Table 4 and Figure 4 are different from those of previous studies. For example, D'Avolio (2002) reports that custody banks are the largest lenders in the U.S. market, representing large institutional owners such as pension funds, public retirement funds, mutual funds and endowments. Passive indexers are the most extensive participants in the custodian's lending program. D'Avolio's (2002) findings suggest that most of the shares supplied for loans come from institutions with stable, low-turnover portfolios (Kolasinski, Reed and Ringgenberg, 2013). However, Table 4 shows that the relationships of short-term IO are stronger than those of long-term IO. The difference could be a result of the different sample periods used. D'Avolio focuses on the 18-month period from April 2000 to September 2001. This paper uses a longer period from 1988 to 2011. Figure 2 in a study by Hanson and Sunderam (2013) shows that the SIR of large cap stocks is higher than that of small cap stocks around 2000 based on the size of the NYSE decile. This pattern has changed in the last decade, as the SIR level of small-median cap stocks is now higher than that of large cap stocks. My results are consistent with those of Hanson and Sunderam (2013), reflecting the changes in supply share sources and the rising role of short-term IO.

4.3 Liquidity and short selling

To test H3 on the relationship between liquidity and short-selling activity, I perform the following univariate regression:

$$SIR_{it} = \alpha + \beta SPREAD_{it} (or\ ILLIQ_{it}) + \varepsilon \quad (10)$$

The results are reported in Table 5. The t-values are clustered by the firm-plus-year dummy (Peterson, 2009). Because the tick size changed in 1997 and 2001, I split the sample period into sub-periods covering 1988-1996, 1997-2000 and 2001-2011. Apart from the coefficients, the corresponding t-values and the R^2 from Equation (10), I also report the average value of SPREAD and ILLIQ in the different periods.

Panel A of Table 5 shows that the average spread of the NYSE/Amex stocks increases from 2.68% to 4% from 1988-1996 to 1997-2000, and then decreases sharply to 0.8% in 2001-2011. The coefficient of SPREAD from Equation (10) shows a strong negative relationship between SPREAD and SIR, which is consistent with H3. The difference in coefficients is also tested formally. The differences between 1988-1996 and 2001-2011 and between 1997-2000 and 2001-2011 are significant at the 1% level, and the difference between 1988-1996 and 1997-2000 is significant at the 10% level.

[Table 5 here]

Panel B of Table 5 shows that the average spread of the Nasdaq stocks decreases from 7.05% to 4.09% from 1988-1996 to 1997-2000, and further decreases to 1.82% in 2001-2011. The coefficient of SPREAD from Equation (10) shows that the relationship between SPREAD and SIR is significantly negative at the 1% level, which is also consistent with H3. The differences in coefficients among the three periods are also statistically significant at the 1% level.

I also investigate the relationship between ILLIQ and SIR in the three sub-periods. Table 5 shows that the average ILLIQ does not change much for the NYSE/Amex stocks, although it does increase from 1988-1996 to 1997-2000 and then decrease in 2001-2011. Panel B shows that the average ILLIQ decreases from 1988-1996 through to 1997-2000 and 2001-2011. The regression results show a statistically negative relationship between ILLIQ and SIR in all of the sub-periods on both the NYSE/Amex and Nasdaq. These results are consistent with H3.

The differences among the sub-periods are also consistent. The differences between the three sub-periods are statistically significant across the NYSE/Amex and Nasdaq, except the difference between 1997-2000 and 2001-2011 on the NYSE/Amex.

In summary, the results in Table 5 support the hypothesis that there is a positive relationship between liquidity and short-selling activity. This relationship is stronger in later periods, especially on the SPREAD measure of liquidity.

5. Multivariate regression

In Section 4, I investigate the univariate relationships between short-selling activity and hedge fund assets, IO and liquidity one by one. The results support all of the hypotheses. In this section, I perform a multivariate regression with control variables. The t-values reported are also clustered by the firm-plus-year dummy.

$$
\begin{aligned}
SIR_{it} = {} & \alpha + \beta_1 LNHFA_t + \beta_2 IO\_SHORT_{it} + \beta_3 SPREAD_{it} + \beta_4 ILLIQ_{it} \\
& \beta_5 LNCAP_{it} + \beta_6 BM_{it} \\
& \beta_7 PR1Y_{it} + \beta_8 D\_DEBT_{it} + \beta_9 D\_OPT_{it} + \beta_{10} VOLAT_{it} + \beta_{11} D\_SP500_{it}
\end{aligned}
$$

$$\beta_{12}MKT_t + \beta_{13}LNVIX_t + \varepsilon \tag{11}$$

5.1 Whole period

Table 6 reports five regression models for the NYSE/Amex and Nasdaq stocks. Model 1 combines LNHFA, IO_SHORT, SPREAD and ILLIQ. Model 2 adds LNCAP and BM as control variables. Model 3 adds PR1Y, D_DEBT, D_OPT, VOLAT and D_SP500 as control variables. Model 4 adds MKT and LNVIX as control variables. Model 5 includes all of the variables.

Model 1 shows that ILLIQ plays a less important role in determining short-selling activity on the NYSE/Amex, where general liquidity is high. The sign of the ILLIQ coefficient also changes from negative in Table 5 to positive in Model 1 of Table 6. However, Model 1 shows that ILLIQ remains statistically negative on the Nasdaq. For the coefficients of LNHFA, IO_SHORT and ILLIQ are consistent with the univariate regression results in Section 4.

[Table 6 here]

Model 2 shows that once the firm characteristics are controlled, ILLIQ become statistically negative with the SIR on the NYSE/Amex. The sign of the SPREAD coefficient on the NYSE/Amex changes from positive to negative and becomes insignificant on the Nasdaq. This means that there are higher correlations between market capitalization and liquidity, and that the effect of liquidity is captured by firm size. The sign of LNCAP is negative on the NYSE/Amex, indicating that the information asymmetric effect dominates the borrowing cost effect. However, the positive coefficient of LNCAP on the Nasdaq means that the borrowing cost effect dominates the information asymmetric effect. The positive BM coefficient on the NYSE/Amex indicates that short sellers prefer to short value stocks, although the same is not true for the Nasdaq.

Model 3 controls for other possible short-selling reasons. The negative coefficient of PR1Y indicates that short sellers follow momentum rather than contrarian strategies and prefer to short stocks with poor past returns. The significant coefficients of D_DEBT and D_OPT indicate that active hedging and arbitrage demands exist for short selling, which is consistent with the literature that has claimed that more investors will short if the difference of investors' opinion is large. Short sellers attempt to avoid the S&P 500 Index constituent stocks, as they are highly transparent in the market and less likely to be overvaluated.

Model 4 controls for market condition. The positive coefficient of MKT shows that short sellers play a contrarian role when the overall market performance is better and choose to short stocks. The positive coefficient of LNVIX shows that short sellers like to short during periods of high volatility.

Model 5 combines all of the variables. The coefficient signs are generally consistent with those of the other models. The signs of MKT and LNVIX change from positive to negative once the firm level variables are controlled, indicating that short sellers rely on firm-level information rather than market-level information.[7]

5.2 Sub-periods

I split the sample into two roughly equal sub-periods, i.e., 1988-2000 and 2001-2011. The first period covers the Internet bubble burst on the Nasdaq, and the second period covers the 2007-2009 financial crisis. Figure 1 shows that short selling increases significantly in the second period. I run Model 5 from Table 6 in the two sub-periods and compute the difference between the two.

Table 7 reports the results. Consistent with the univariate regression from Tables 3 to 5, there is a significant difference in coefficients between the two sub-periods aside from the

[7] The high correlation between the current SIR and SIR_{t-1} also bears consideration. In an unreported table, I add SIR_{t-1} as an additional control variable, and the results are similar to Model 5 of Table 6. These results are available on request.

difference in ILLIQ on the Nasdaq. For example, short-term IO plays a more important role in determining short-selling activity. The coefficients of IO_SHORT are 0.0283 (with a t-value of 9.16) and 0.0839 (with a t-value of 12.42) in 1988-2000 and 2001-2011 on NYSE/Amex, respectively.

[Table 7 here]

There are also significant changes from the first period to the second period among the control variables. For example, the short sellers in the NYSE/Amex prefer to short high book-to-market ratio stocks in the 2001-2011 period but not in the 1988-2000 period. In contrast, the short sellers in the Nasdaq prefer to short low book-to-market ratio stocks in the 1988-2000 period but not in the 2001-2011 period. The sign and difference of the BM coefficients reflect the effects of the Internet bubble burst on the Nasdaq and the recent financial crisis mainly on the NYSE/Amex stocks. The corresponding BM coefficients are 0.001 (with a t-value of 1.04) and 0.0066 (with a v-value of 4.42) in the first and second sub-periods on the NYSE/Amex, respectively, and -0.0021 (with a t-value of 6.00) and 0.0000 (with a t-value of 0.03) in the first and second sub-periods on the Nasdaq, respectively.

Arbitrage and hedging between the listed stock options and underlying stocks also increase significantly in the second period. The short sellers attempt to avoid shorting S&P 500 Index constituent stocks more often in the second period.

In an unreported table, I also investigate whether a significant difference exists between the financial crisis (2007-2009) and non-financial crisis (2001-2011, excluding 2007-2009) periods. Both LNHFA and IO_SHORT are more strongly related to SIR during the crisis period, but not in terms of the liquidity variables. For example, the LNHFA coefficients are 0.1005 (with a t-value of 14.7) and 0.0112 (with a t-value of 15.94) in the crisis and non-crisis periods, respectively. The results suggest that hedge funds are much more active during a crisis period, as they offer the use of long-short strategies that are unavailable for most mutual funds (Chen, Desai and Krishnamurthy, 2012). These results are available on request.

6. Expected and unexpected components of short interest

Section 4.2 shows that IO_SHORT has a stronger relationship with SIR. Short selling could result from daily risk management practices such as hedging that contain less information and are well expected, or from unexpected private information that has not yet been incorporated into a stock price. Therefore, I predict that unexpected short selling should have a stronger relationship with IO_SHORT than with IO_LONG.

Similar to a study by Chan, Kot and Ni (2012), I further decompose SIR into expected and unexpected components as follows:

$$E\_SIR_{it} = \frac{\sum_{t-1}^{t-12} SIR_{it}}{12} \quad (12)$$

$$U\_SIR_{it} = SIR_{it} - E\_SIR_{it} \quad (13)$$

where E_SIR is the average SIR over the past 12 months and U_SIR is the difference between SIR and E_SIR. I run the following two regressions to investigate the relationships between E_SIR/U_SIR and IO_LONG/ IO_SHORT:

$$E\_SIR_{it} = \alpha + \beta_1 IO\_LONG_{it} + \beta_2 IO\_SHORT_{it} + \varepsilon \quad (14)$$

$$U\_SIR_{it} = \alpha + \beta_1 IO\_LONG_{it} + \beta_2 IO\_SHORT_{it} + \varepsilon \quad (15)$$

Table 8 reports the results. For E_SIR on the NYSE/Amex in Panel A, the IO_LONG coefficient is negative in 1988-2000 and insignificant in 2001-2011. IO_SHORT is significantly positive in two sub-periods. For E_SIR on the Nasdaq in Panel A, the IO_LONG coefficients are significantly positive, as are those of IO_SHORT in two sub-periods. The difference in the coefficients of IO_SHORT and IO_LONG is significant on the NYSE/Amex and Nasdaq and in two sub-periods. The results suggest that IO_SHORT remains the main factor in determining expected short-selling activity.

[Table 8 here]

Panel B of Table 8 reports the results for U_SIR. Consistent with the prediction that the relationship between IO_SHORT and U_SIR is stronger than that between IO_LONG and U-SIR, the difference between IO_SHORT and IO_LONG is significant on the NYSE/Amex and Nasdaq and in two sub-periods. I also find that the IO_LONG coefficient is significantly negative on the Nasdaq in both sub-periods once IO_SHORT is controlled for.[8]

In the robustness checks in Model 5 of Table 6, I re-run the regressions using E_SIR and U_SIR as the dependent variables. Table 9 reports the results. The coefficients of LNHFA, IO_SHORT, SPREAD and ILLIQ are consistent with the previous results when E_SIR and U_SIR are used as dependent variables. Among the control variables, the arbitrage and hedging demand from convertible bonds and stock options contain less private information, and consequently have less of a relationship with unexpected short selling. The signs of D_DEBT and D_OPT become negative when U_SIR is used as a dependent variable. The signs of PR1Y and D_SP500 are similarly different when E_SIR and U_SIR are used as dependent variables.

[Table 9 here]

7. Conclusions

Short selling plays an important role in the financial market, as it enhances the market's quality and liquidity. Short selling has increased sharply over the past two decades. It has attracted much attention from academics for different reasons, such as the predictive ability of future returns from short interest, whether the abnormal returns of predictive ability are a result of

[8] The univariate relationship between IO_LONG and U_SIR is significantly positive.

public or private information and the effect of short sales on market quality among others. However, the increase in short selling itself has received limited attention. This paper attempts to understand the increase based on two main aspects: the growth of the hedge fund industry and short selling demand from hedge funds, and how the increased IO has caused borrowing costs to lower and attracted more investors to short stocks.

Using monthly short interest on the NYSE/Amex and Nasdaq from 1988 to 2011, hedge fund assets data and institutional ownership data, I find strong evidence to support my hypotheses. In particular, I find that short-selling activity is more related to the holds of short-term institutions rather than those of long-term institutions. This finding persists for both the expected and unexpected components of short interest. A sub-period analysis shows that the relationships of hedge fund assets, institutional ownership and liquidity with short selling are stronger in 2001-2011 than in 1988-2000. My findings help further our understanding of why short selling has increased over the past two decades, and could be applied to other markets.

References

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effect, *Journal of Financial Markets* 5, 31–56.

Arnold, Tom, Alexander W. Butler, Timothy F. Crack, and Yan Zhang, 2005, The information content of short interest: A natural experiment, *Journal of Business* 78, 1307–1335.

Asquith, Paul, Parag A. Pathak, and Jay Ritter, 2005, Short interest, institutional ownership, and stock returns, *Journal of Financial Economics* 78, 243–276.

Blau, Benjamin M., Bonnie F. Van Ness, and Robert A. Van Ness, 2009, Short selling and the weekend effect for NYSE securities, *Financial Management* 38, 603–630.

Boehme, Rodney D., Bartley R. Danielsen, and Sorin M. Sorescu, 2006, Short sales constraints, differences of opinion, and overvaluation, *Journal of Financial and Quantitative Analysis* 41, 455–487.

Boehmer, Ekkehart, Charles M. Jones, and Xiaoyan Zhang, 2008, Which shorts are informed?, *Journal of Finance* 63, 491–528.

Boehmer, Ekkehart, Charles M. Jones, and Xiaoyan Zhang, 2012, What do short sellers know?, Working paper, Columbia University.

Boehmer, Ekkehart, Charles M. Jones, and Xiaoyan Zhang, 2013, Shackling short sellers: The 2008 shorting ban, *Review of Financial Studies*, In press.

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

Chan, Kalok, Hung Wan Kot, and Sophie X. Ni, 2012, Does option trading affect the return predictability of short selling activity?, Working paper, HKUST.

Chen, Honghui, Memang Desai, and Srinivasan Krishnamurthy, 2012, A first look at mutual funds that use short sales, *Journal of Financial and Quantitative Analysis*, forthcoming.

Chen, Joseph, Harrison Hong, and Jeremy C. Stein, 2002, Breadth of ownership and stock returns, *Journal of Financial Economics* 66, 171–205.

Christophe, Stephen E., Michael G. Ferri, and James J. Angel, 2009, Short-selling and the weekend effect in Nasdaq stock returns, *Financial Review* 44, 31–57.

Christophe, Stephen E., Michael G. Ferri, and James J. Angel, 2004, Short-selling prior to earnings announcement, *Journal of Finance* 59, 1845–1875.

Cohen, Lauren, Karl B. Diether, and Christopher J. Malloy, 2007, Supply and demand shifts in the shorting market, *Journal of Finance* 62, 2061–2096.

Corwin, Shane, and Paul Schultz, 2012, A simple way to estimate bid-ask spreads from daily high and low prices, *Journal of Finance* 67, 719–759.

D'Avolio, Gene, 2002, The market for borrowing stock, *Journal of Financial Economics* 66, 271–

306.

Dechow, Patricia M., Amy P. Hutton, Lisa Meulbroek, and Richard G. Sloan, 2001, Short-sellers, fundamental analysis, and stock returns, *Journal of Financial Economics* 61, 77–106.

Desai, Hemang, K. Ramesh, S. Ramu Thiagarajan, and Bala V. Balachandran, 2002, An investigation of the informational role of short interest in the Nasdaq market, *Journal of Finance* 57, 2263–2287.

Diether, Karl B., Kuan-Hui Lee, and Ingrid M. Werner, 2009a, Short-sale strategies and return predictability, *Review of Financial Studies* 22, 575–607.
.
Diether, Karl B., Kuan-Hui Lee, and Ingrid M. Werner, 2009b, It's SHO time! Short-sale price-tests and market quality, *Journal of Finance* 64, 37–73.

Edwards, Amy K., and Kathleen Weiss Hanley, 2010, Short selling in initial public offering, *Journal of Financial Economics* 98, 21–39.

Engelbert, Joseph E., Adam V. Reed, and Matthew C. Ringgenberg, 2012, How are shorts informed? Short sellers, news, and information processing, *Journal of Financial Economics*, in press.

Geczy, Christopher C., David K. Musto, and Adam V. Reed, 2002, Stocks are special too: An analysis of the equity lending market, *Journal of Financial Economics* 66, 241–269.

Goldman Sachs, 2009, Hedge fund trend monitor (August 24).

Grundy, Bruce D., Bryan Lim, and Patrick Vervijmeren, 2012, Do option markets undo restrictions on short sales? Evidence from the 2008 short-sale ban, *Journal of Financial Economics*, in press.

Hanson, Samuel G., and Adi Sunderam, 2013, The growth and limits of arbitrage: Evidence from short interest, Working paper, Harvard Business School.

Henry, Tyler R., and Jennifer L. Koski, 2010, Short selling around seasoned equity offerings, *Review of Financial Studies* 23, 4389–4418.

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

Kolasinksi, Adam, Adam V. Reed, and Matthew C. Ringgenberg, 2013, A multiple lender approach to understanding supply and search in the equity lending market, *Journal of Finance* 68, 559–595.

Massoud, Nadia, Debarshi Nandy, Anthony Saunders, and Keke Song, 2011, Do hedge funds trade on private information? Evidence from syndicated lending and short-selling?, *Journal of Financial Economics* 99, 477–499.

Nagel, Stefan, 2005, Short sales, institutional investors, and the cross-section of stock returns, *Journal of Financial Economics* 78, 277–309.

Petersen, Mitchell A., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of Financial Studies* 22, 435–480.

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

Seneca, Joseph J., 1976, Short interest: Bearish or bullish?, *Journal of Finance* 22, 67–70.

Yan, Xuemin (Sterling), and Zhe Zhang, 2009, Institutional investors and equity returns: Are short-term institutions better informed?, *Review of Financial Studies* 22, 893–924.

Figure 1 Short interest from 1998 to 2011

This figure reports the aggregate short interest (million shares) and average short interest ratio (SIR, defined as the short interest scaled by the number of outstanding shares) on the NYSE/Amex and Nasdaq from 1988 to 2011. The short-interest data were obtained from the exchanges and Compustat. The number of outstanding shares was obtained from CRSP.

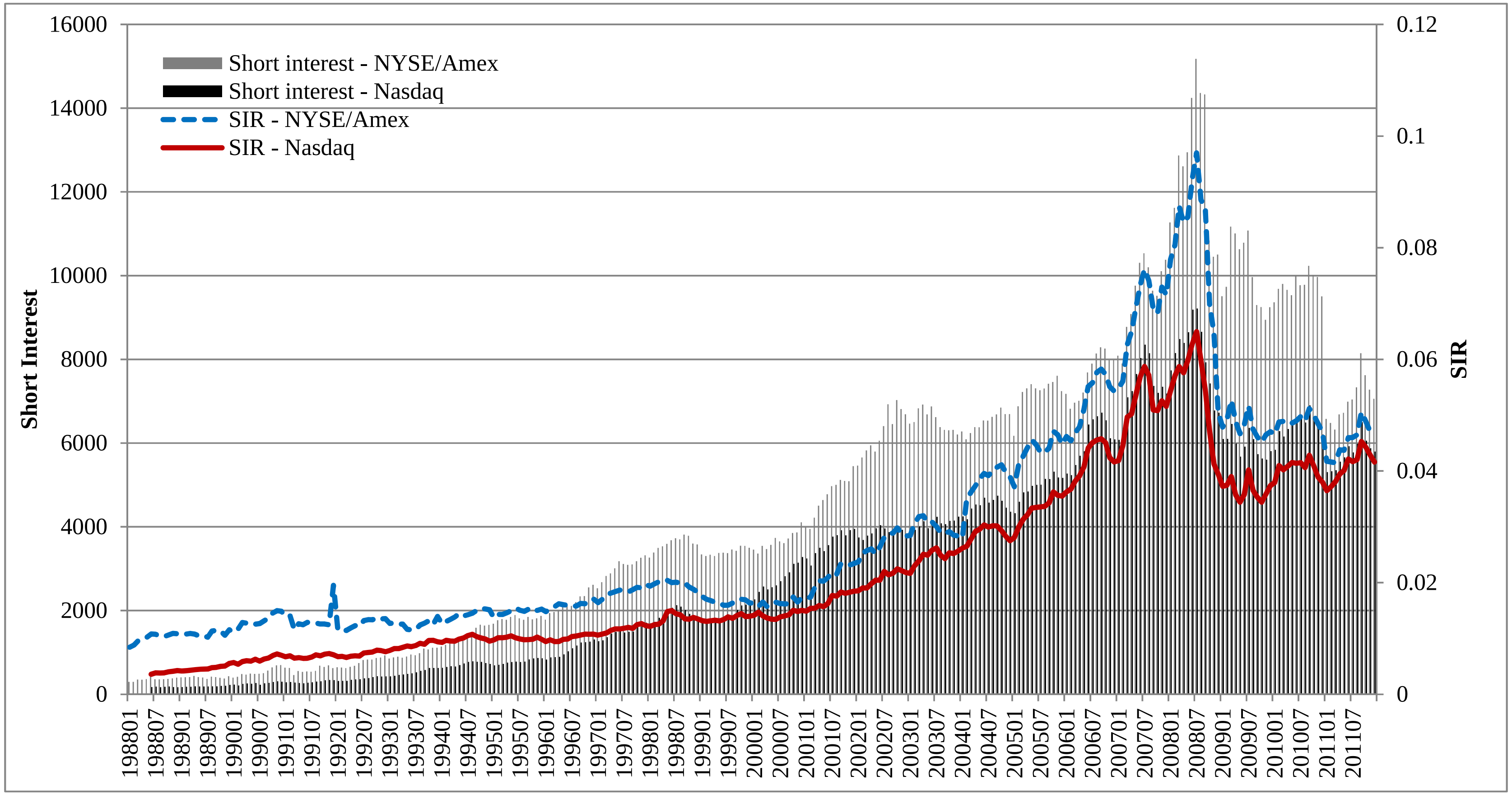


Figure 2 Hedge fund number and total assets from 1988 to 2011

This figure reports the number of funds and asset values (million USD) from 1988 to 2011. The investment region is restricted to North America; the fund asset currency to USD; and the main strategy to equity hedge, even-driven, macro and relative value. The hedge fund data were obtained from HFR hedge fund database and dead hedge fund database.

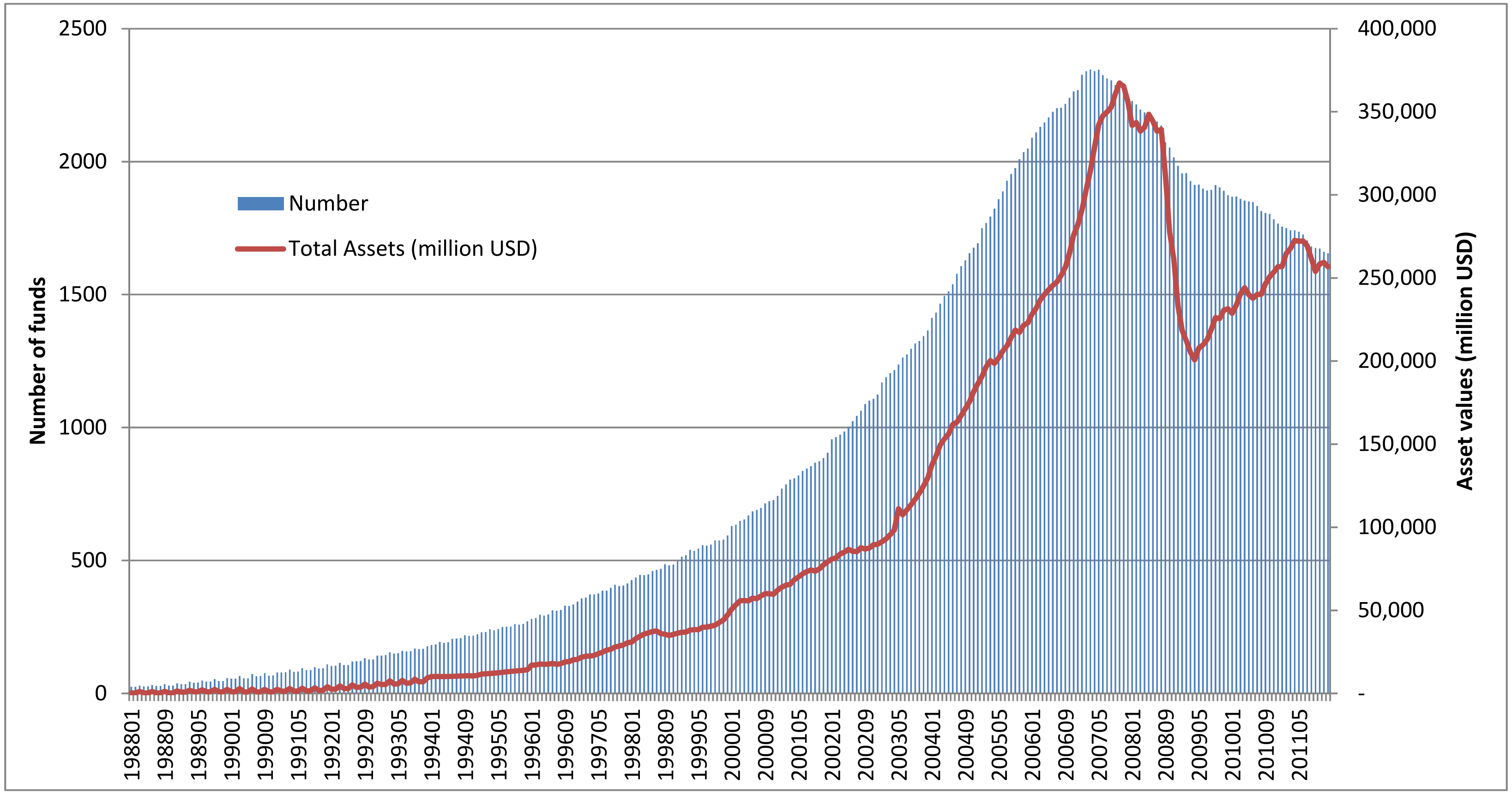


Figure 3 Institutional ownership from 1988 to 2011

Institutional ownership (IO) is defined as the shares owned by institutions scaled by the number of outstanding shares. Following Yan and Zhang (2009), IO_LONG represents the IO holds of long-term institutions and IO_SHORT represents the IO holds of short-term institutions. The IO data were obtained from Thomson Financial's institutional database. The stock prices and number of outstanding shares were obtained from CRSP.

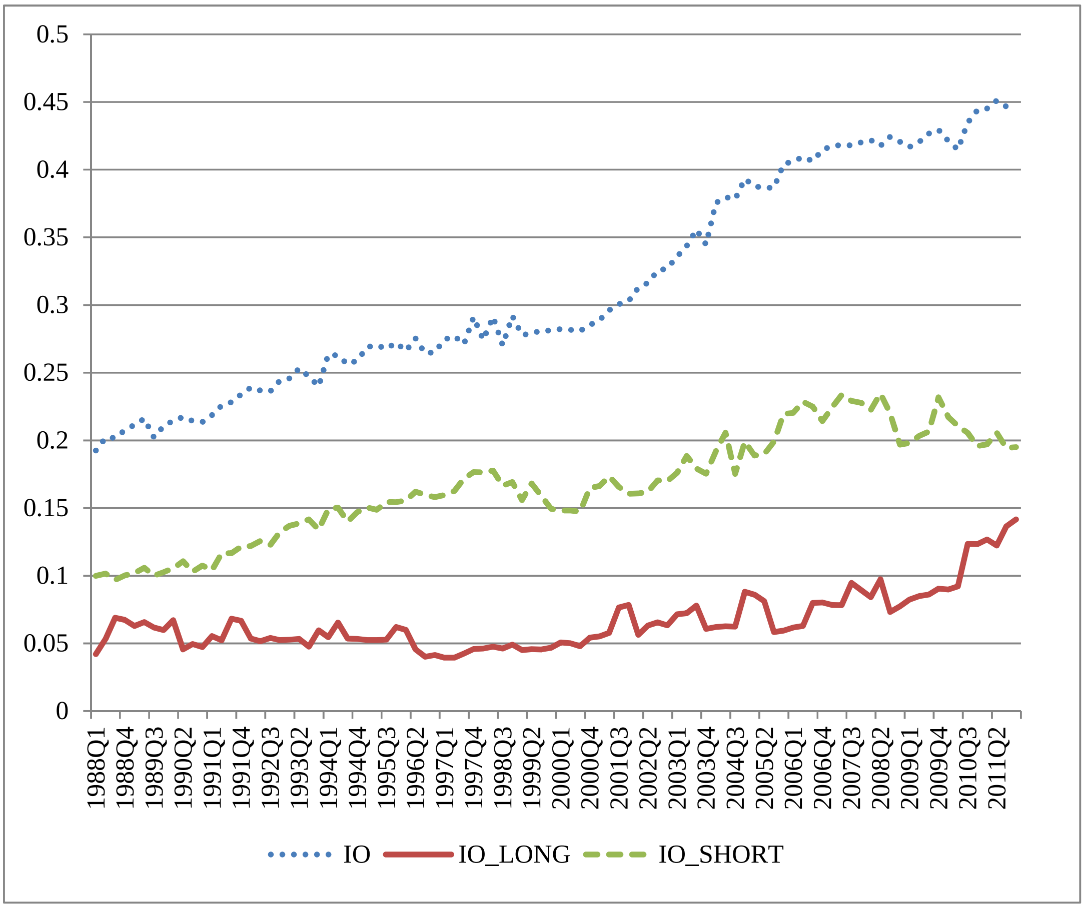
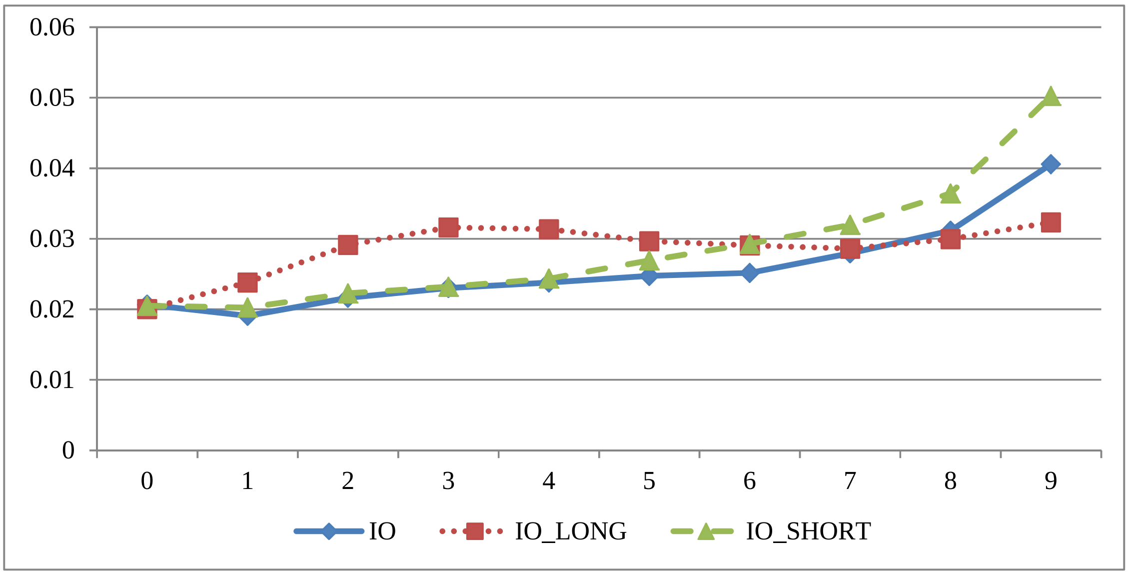


Figure 4 Institutional ownership decile and short interest ratio

This figure reports the relationship between the IO deciles (X-axis) and SIR (Y-axis) from 1988 to 2011. In each month I sort the stocks by IO (or IO_LONG, IO_SHORT individually) into 10 deciles and compute the average SIR in each decile. SIR is defined as the short interest scaled by the number of outstanding shares. IO is defined as the shares owned by institutions scaled by the number of outstanding shares. Following Yan and Zhang (2009), IO_LONG represents the IO holds of long-term institutions and IO_SHORT represents the IO holds of short-term institutions. The short-interest data were obtained from the exchanges and Compustat. The IO data were obtained from Thomson Financial's institutional database. The stock prices and number of outstanding shares were obtained from CRSP.

Panel A: NYSE/Amex



Panel B: Nasdaq

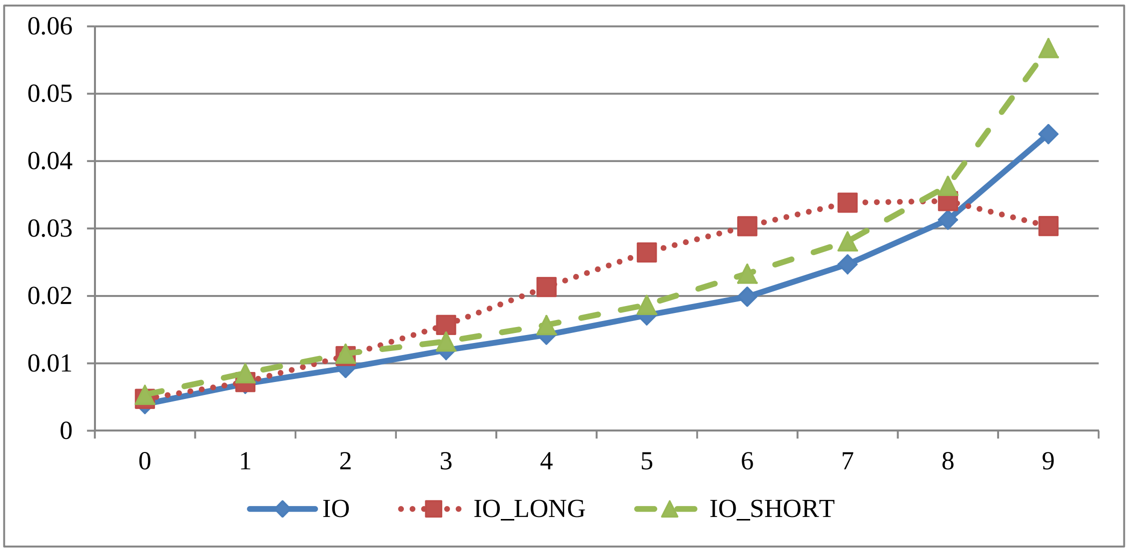


Table 1 Summary statistics

SIR is the short interest ratio, defined as the short interest scaled by the number of outstanding shares. HFA is the total assets of the hedge funds in millions of dollars. IO is the institutional ownership scaled by the number of outstanding shares. IO_LONG and IO_SHORT are the IO holds of the long- and short-term institutions, classified by Yan and Zhang (2009). SPREAD is the average daily closing bid-ask spread within 1 month. ILLIQ is Ahihud's illiquidity measure. CAP is the market capitalization in millions of dollars. BM is the industry-adjusted book-to-market ratio. PR1Y is the past 1-year stock return. D_DEBT is a dummy variable representing the availability of convertible debt (1 for yes, 0 for no). D_OPT is a dummy variable signifying the availability of stock options (1 for yes, 0 for no). VOLAT is the standard deviation of the daily stock returns over 3 months. D_SP500 is a dummy variable signifying inclusion in the S&P 500 Index (1 for yes, 0 for no). MKT represents the CRSP value-weighted stock returns. VIX represents the close of the CBOE VIX index. The sample period covers 1988.1-2011.12 for the NYSE/Amex stocks and 1988.6-2011.12 for the Nasdaq stocks. The stock return data were obtained from CRSP. The accounting data were obtained from Compustat. The short-interest data were obtained from the exchanges and Compustat. The hedge fund data were obtained from Hedge Fund Research Inc. The IO data were obtained from Thomson Financial's institutional database. The options availability was obtained from the Berkeley Options Database (1988-1995) and Option Metrics Database (1996-2011). p10 and p90 denote the 10th and 90th percentiles, respectively.

	p10	Median	p90	Mean	S.D.
Panel A: NYSE/Amex					
SIR	0	0.01	0.08	0.03	0.05
HFA	1,471	40,453	264,449	99,790	110,473
IO	0.07	0.46	0.83	0.45	0.28
IO_LONG	0.01	0.08	0.2	0.1	0.09
IO_SHORT	0.02	0.2	0.46	0.22	0.17
SPREAD	0	0.01	0.05	0.03	0.06
ILLIQ	0	0.01	2.53	1.33	4.23
CAP	21	503	7,163	3,860	15,538
BM	-0.65	-0.16	0.61	0.01	1.88
PR1Y	-0.43	0.07	0.66	0.13	0.63
D_DEBT	0	0	1	0.24	0.43
D_OPT	0	0	1	0.38	0.49
VOLAT	0.01	0.02	0.05	0.03	0.03
D_SP500	0	0	1	0.19	0.39
MKT	-0.05	0.01	0.06	0.01	0.04
VIX	12	19	31	20	8

	p10	Median	p90	Mean	S.D.
Panel B: Nasdaq					
SIR	0	0	0.06	0.02	0.04
HFA	1,292	35,787	258,414	86,293	104,871
IO	0.02	0.21	0.7	0.29	0.26
IO_LONG	0	0.04	0.13	0.06	0.07
IO_SHORT	0	0.11	0.42	0.16	0.17
SPREAD	0	0.03	0.1	0.04	0.06
ILLIQ	0	0.28	13.91	3.29	6.27
CAP	7	68	744	604	6,367
BM	-0.62	-0.15	0.64	0	1.6
PR1Y	-0.58	0	0.85	0.14	0.95
D_DEBT	0	0	1	0.16	0.37
D_OPT	0	0	1	0.18	0.38
VOLAT	0.02	0.04	0.08	0.05	0.04
D_SP500	0	0	0	0.02	0.12
MKT	-0.05	0.02	0.06	0.01	0.04
VIX	12	19	31	20	8

Table 2 Correlations

SIR is the short interest ratio, defined as the short interest scaled by the number of outstanding shares. LNHFA is the natural logarithm of the total hedge fund assets. IO_SHORT is the IO hold of the short-term institutions classified by Yan and Zhang (2009), where IO is the institutional ownership scaled by the number of outstanding shares. SPREAD is the average daily closing bid-ask spread within 1 month. ILLIQ is Ahihud's illiquidity measure. LNCAP is the natural logarithm of market capitalization. BM is the industry-adjusted book-to-market ratio. PR1Y is the past 1-year stock return. D_DEBT is a dummy variable representing the availability of convertible debt (1 for yes, 0 for no). D_OPT is a dummy variable signifying the availability of stock options (1 for yes, 0 for no). VOLAT is the standard deviation of the daily stock returns over 3 months. D_SP500 is a dummy variable signifying inclusion in the S&P 500 Index (1 for yes, 0 for no). MKT represents the CRSP value-weighted stock returns. LNVIX represents the natural logarithm of the close of the CBOE VIX index. The sample period covers 1988-2011. The stock return data were obtained from CRSP. The accounting data were obtained from Compustat. The short interest data were obtained from the exchanges and Compustat. The hedge fund data were obtained from Hedge Fund Research Inc. The IO data were obtained from Thomson Financial's institutional database. The options availability was obtained from the Berkeley Options Database (1988-1995) and Option Metrics Database (1996-2011).

Panel A: NYSE/Amex

	SIR	LNHFA	IO_SHORT	SPREAD	ILLIQ	LNCAP	BM	PR1Y	D_DEBT	D_OPT	VOLAT	D_SP500	MKT
LNHFA	0.30												
IO_SHORT	0.28	0.28											
SPREAD	-0.14	-0.21	-0.29										
ILLIQ	-0.08	-0.13	-0.29	0.59									
LNCAP	0.06	0.25	0.50	-0.47	-0.54								
BM	0.06	-0.01	-0.09	0.11	0.13	-0.17							
PR1Y	-0.02	0.03	0.11	-0.15	-0.16	0.13	0.04						
D_DEBT	0.08	-0.11	0.07	0.02	-0.02	0.04	0.00	-0.01					
D_OPT	0.26	0.42	0.46	-0.27	-0.24	0.59	-0.07	0.03	0.01				
VOLAT	0.08	0.01	-0.18	0.45	0.48	-0.44	0.10	-0.14	0.05	-0.13			
D_SP500	-0.07	0.01	0.18	-0.15	-0.15	0.58	-0.08	0.00	0.06	0.33	-0.18		
MKT	-0.04	-0.08	-0.02	0.00	-0.01	0.00	0.00	-0.02	0.01	-0.03	-0.02	0.00	
LNVIX	0.04	0.09	0.04	0.07	0.04	-0.01	0.00	-0.13	-0.04	0.13	0.20	0.00	-0.21

Panel B: Nasdaq

	SIR	LNHFA	IO_SHORT	SPREAD	ILLIQ	LNCAP	BM	PR1Y	D_DEBT	D_OPT	VOLAT	D_SP500	MKT
LNHFA	0.31												
IO_SHORT	0.46	0.24											
SPREAD	-0.26	-0.35	-0.36										
ILLIQ	-0.22	-0.20	-0.29	0.71									
LNCAP	0.39	0.42	0.57	-0.61	-0.57								
BM	-0.07	0.01	-0.08	0.10	0.13	-0.15							
PR1Y	0.03	0.02	0.12	-0.16	-0.17	0.20	0.03						
D_DEBT	0.08	-0.02	0.03	0.00	-0.01	0.02	-0.02	-0.02					
D_OPT	0.49	0.37	0.49	-0.29	-0.24	0.54	-0.08	0.03	0.03				
VOLAT	-0.06	-0.06	-0.19	0.56	0.41	-0.38	0.06	-0.10	0.04	-0.10			
D_SP500	0.05	0.07	0.12	-0.08	-0.07	0.31	-0.03	0.00	0.02	0.20	-0.07		
MKT	-0.03	-0.08	-0.01	0.02	0.00	-0.02	0.00	-0.01	0.00	-0.03	0.00	-0.01	
LNVIX	0.04	0.09	-0.01	-0.08	0.02	-0.01	0.00	-0.10	0.00	0.10	0.12	0.01	-0.20

Table 3 Hedge funds and short selling

This table reports the coefficients of the following univariate regressions:

$$LNSI_t = \alpha + \beta LNHFA_t + \varepsilon \quad (6)$$

$$SIR_t = \alpha + \beta LNHFA_t + \varepsilon \quad (7)$$

LNSI is the natural logarithm of the dollar value of the short interest of the common stocks. SIR is the average SIR value of the common stocks. LNHFA is the natural logarithm of the total hedge fund assets. The sample period covers 1988-2011. The hedge fund data were obtained from the HFR hedge fund and dead hedge fund databases. The short interest was obtained from the exchanges and Compustat.
*, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	Equation (6)		Equation (7)	
	NYSE/Amex	Nasdaq	NYSE/Amex	Nasdaq
Panel A: 1988-2000				
Beta	0.1076***	0.1467***	0.0008**	0.0004***
t-value	3.66	3.86	2.14	3.31
R^2	0.99	0.99	0.91	0.98
Panel B: 2001-2011				
Beta	1.0475***	1.0714***	0.0419***	0.0283***
t-value	7.38	7.85	4.17	5.46
R^2	0.92	0.92	0.95	0.95
Panel C: diff. in two periods				
Diff	0.9399***	0.9247***	0.0411***	0.0279***
t-value	6.49	6.52	4.09	5.38

Table 4 Institutional ownership and short selling

This table reports the coefficients of the following regressions:

$$SIR_{it} = \alpha + \beta IO_{it}(or\ IO\_LONG, IO\_SHORT) + \varepsilon \quad (8)$$

$$SIR_{it} = \alpha + \beta_1 IO\_LONG_{it} + \beta_2 IO\_SHORT_{it} + \varepsilon \quad (9)$$

SIR is defined as the short interest scaled by the number of outstanding shares. IO is the institutional ownership scaled by the number of outstanding shares. IO_LONG and IO_SHORT are the IO holds of the long- and short-term institutions, classified by Yan and Zhang (2009). The sample period covers 1988-2011. The short interest was obtained from the exchanges and Compustat. The institutional ownership data were obtained from Thomson Financial's institutional database. The standard errors are clustered by the firm-plus-year dummy.

*, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	NYSE/Amex			Nasdaq		
	IO	IO_LONG	IO_SHORT	IO	IO_LONG	IO_SHORT
Panel A: Equation (8)						
1988-2000						
Beta	0.0124***	0.0072***	0.0272***	0.0292***	0.0408***	0.0472***
T-value	9.91	1.98	13.14	23.87	9.61	25.05
R^2	0.01	0.01	0.02	0.06	0.02	0.06
2001-2011						
Beta	0.0403***	0.1024***	0.0977***	0.0659***	0.1686***	0.1293***
T-value	24.22	11.28	16.01	43.31	17.22	29.52
R^2	0.07	0.05	0.11	0.18	0.10	0.23
Difference in sub-periods						
Beta	0.0279***	0.0952***	0.0704***	0.0367***	0.1278***	0.0821***
T-value	14.90	11.26	11.35	19.83	12.47	17.60
Panel B: Equation (9)						
Beta		0.0202***	0.0615***		0.0526***	0.0886***
T-value		4.04	18.04		10.59	36.18
Difference between IO_LONG and IO_SHORT						
Beta		0.0412***			0.0360***	
T-value		6.40			6.65	

Table 5 Liquidity and short selling

This table reports the coefficients of the following univariate regression:

$$SIR_{it} = \alpha + \beta SPREAD_{it}(or\ ILLIQ_{it}) + \varepsilon \tag{10}$$

SIR is the short interest ratio, defined as the short interest scaled by the number of outstanding shares. SPREAD is the average daily closing bid-ask spread within one month. ILLIQ is Ahihud's illiquidity measure. Mean is the average level of SPREAD (ILLIQ) in different periods. Coefficient, T-values and R^2 are obtained from Equation (10). The sample period covers 1988-2011. The short-interest data were obtained from the exchanges and Compustat. The stock prices, trading volume, number of outstanding shares and close bid-ask spread were obtained from CRSP. The standard errors are clustered by the firm-plus-year dummy.

*, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	SPREAD				ILLIQ			
Period	Mean	Coefficient	T-value	R^2	Mean	Coefficient	T-value	R^2
Panel A: NYSE/Amex								
(1) 1988-1996	0.0268	-0.0254***	-2.62	0.00	0.4984	-0.0005***	-5.02	0.00
(2) 1997-2000	0.0400	-0.0422***	-11.16	0.00	0.6021	-0.0009***	-13.03	0.00
(3) 2001-2011	0.0080	-0.1083***	-6.54	0.02	0.4176	-0.0011***	-4.36	0.02
Difference of (3) & (1)		-0.0829***	-4.38			-0.0006**	-2.02	
Difference of (3) & (2)		-0.0660***	-4.01			-0.0002	-0.76	
Difference of (2) & (1)		-0.0168*	-1.85			-0.0004***	-3.66	
Panel B: Nasdaq								
(1) 1988-1996	0.0705	-0.0510***	-22.55	0.02	4.4090	-0.0006***	-29.48	0.03
(2) 1997-2000	0.0409	-0.1357***	-23.58	0.04	2.5531	-0.0009***	-36.31	0.03
(3) 2001-2011	0.0182	-0.4501***	-31.74	0.09	2.1013	-0.0021***	-51.54	0.08
Difference of (3) & (1)		-0.3990***	-27.91			-0.0015***	-34.06	
Difference of (3) & (2)		-0.3143***	-21.42			-0.0012***	-26.81	
Difference of (2) & (1)		-0.0847***	-14.87			-0.0003***	-12.11	

Table 6 Multivariate regression

This table reports the coefficients of the multivariate regressions. The dependent variable is SIR. All of the variables and data sources are described in Table 2. The sample period covers 1988-2011. The standard error is clustered by the firm-plus-year dummy.
*, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	NYSE/Amex					Nasdaq				
	1	2	3	4	5	1	2	3	4	5
LNHFA	0.0193***	0.0204***	0.0272***	0.0196***	0.0261***	0.0017***	0.0020***	0.0027***	0.0019***	0.0026***
	19.85	18.87	21.69	20.56	22.11	18.16	16.51	18.40	20.57	18.79
IO_SHORT	0.0611***	0.0771***	0.0541***	0.0611***	0.0626***	0.0946***	0.0879***	0.0710***	0.0945***	0.0652***
	16.16	19.92	11.84	16.16	15.35	36.87	27.78	24.11	36.86	19.86
SPREAD	0.0286***	-0.0183**	-0.0518***	0.0285***	-0.0711***	-0.0295***	0.0013	-0.0909***	-0.0299***	-0.0769***
	3.21	-1.98	-4.57	3.20	-6.21	-7.75	0.33	-15.59	-7.81	-13.59
ILLIQ	-0.0005	-0.0013***	-0.0006	-0.0005	-0.0014***	-0.0005***	-0.0003***	-0.0003***	-0.0005***	-0.0001***
	-1.51	-4.91	-1.51	-1.51	-5.16	-21.29	-9.25	-11.34	-21.29	-4.20
LNCAP		-0.0034***			-0.0033***		0.0028***			0.0021***
		-11.02			-7.41		9.97			6.67
BM		0.0054***			0.0058***		-0.0006			-0.0003
		4.19			4.46		-1.00			-0.62
PR1Y			-0.0032***		-0.0042***			-0.0004*		-0.0007***
			-6.67		-7.80			-1.90		-3.00
D_DEBT			0.0109***		0.0109***			0.0084***		0.0083***
			9.98		10.17			11.25		11.15
D_OPT			0.0132***		0.0177***			0.0288***		0.0265***
			9.14		11.29			29.51		26.50
VOLAT			0.3704***		0.3154***			0.1642***		0.1793***
			12.18		10.33			15.04		14.68
D_SP500			-0.0172***		-0.0093***			-0.0194***		-0.0253***
			-14.47		-7.48			-10.23		-12.11
MKT				0.0015**	-0.0069***				0.0024***	-0.0036***
				2.15	-5.98				5.74	-5.49
LNVIX				0.0009**	-0.0040***				0.0014***	-0.0014***
				2.52	-6.89				8.02	-5.23
Adj-R^2	0.17	0.21	0.22	0.17	0.24	0.27	0.28	0.34	0.27	0.34
Obs	384,629	354,671	331,043	384,629	330,739	820,519	705,246	588,040	820,519	586,511

Table 7 Sub-period analysis

This table reports the coefficients from Model 5 of Table 6 in the sub-periods. The dependent variable is SIR. All of the variables and data sources are described in Table 2. The sample period covers 1988-2011. The standard error is clustered by the firm-plus-year dummy.

*, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	NYSE/Amex			Nasdaq		
	1988–2000	2001–2011	Diff	1988–2000	2001–2011	Diff
LNHFA	0.0057***	0.0511***	0.0454***	0.0005***	0.0303***	0.0299***
	7.97	19.69	16.91	5.51	19.97	19.60
IO_SHORT	0.0283***	0.0839***	0.0556***	0.0260***	0.0957***	0.0697***
	9.16	12.42	7.75	9.90	16.40	11.20
SPREAD	-0.0574***	-0.1097***	-0.0523**	-0.0948***	-0.1540***	-0.0592***
	-6.81	-4.18	-1.98	-17.81	-11.74	-4.26
ILLIQ	-0.0008***	-0.0017***	-0.0010**	0.0000	0.0001**	0.0000
	-5.21	-4.27	-2.31	1.57	1.95	0.87
LNCAP	0.0002	-0.0070***	-0.0072***	0.0031***	0.0009*	-0.0022***
	0.77	-9.40	-9.41	11.34	1.68	-3.73
BM	0.0010	0.0066***	0.0056***	-0.0021***	0.0000	0.0021***
	1.04	4.42	3.07	-6.00	0.03	4.64
PR1Y	-0.0014***	-0.0068***	-0.0055***	-0.0007***	0.0001	0.0008*
	-2.67	-7.86	-5.35	-3.93	0.21	1.77
D_DEBT	0.0132***	0.0082***	-0.0050**	0.0078***	0.0069***	-0.0009
	11.85	4.79	-2.55	10.92	5.16	-0.62
D_OPT	0.0062***	0.0285***	0.0223***	0.0112***	0.0319***	0.0207***
	5.11	10.12	7.60	9.16	22.94	11.73
VOLAT	0.3492***	0.2510***	-0.0982*	0.1749***	0.1997***	0.0248
	11.61	5.67	-1.88	19.34	8.15	0.96
D_SP500	-0.0051***	-0.0101***	-0.0050**	-0.0133***	-0.0301***	-0.0168***
	-4.63	-5.16	-2.34	-6.82	-10.33	-5.01
MKT	-0.0084***	-0.0045***	0.0039*	-0.0078***	-0.0016	0.0062***
	-6.80	-2.62	1.83	-10.27	-1.56	4.91
LNVIX	-0.0017***	-0.0017**	0.0000	-0.0003	0.0016***	0.0019***
	-4.04	-2.03	-0.04	-1.17	4.22	4.22
R^2	0.12	0.22		0.16	0.36	

Table 8 Institutional ownership and (un)expected components of short interest
This table reports the coefficients of the following two regressions:

$$E\_SIR_{it} = \alpha + \beta_1 IO\_LONG_{it} + \beta_2 IO\_SHORT_{it} + \varepsilon \quad (14)$$

$$U\_SIR_{it} = \alpha + \beta_1 IO\_LONG_{it} + \beta_2 IO\_SHORT_{it} + \varepsilon \quad (15)$$

E_SIR is the expected component of the SIR, which is defined as the average SIR over months t-1 to t-12. U_SIR is the unexpected component of the SIR, which is defined as SIR minus E_SIR. IO_LONG and IO_SHORT are the IO holds of the long- and short-term institutions, classified by Yan and Zhang (2009). The sample period covers 1988-2011. The short interest was obtained from the exchanges and Compustat. The institutional ownership data were obtained from Thomson Financial's institutional database. The standard errors are clustered by the firm-plus-year dummy.
*, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

		IO_LONG	IO_SHORT	Difference
Panel A: Equation (14) of E_SIR				
NYSE/Amex	1988-2000	-0.0156***	0.0252***	0.0408***
		-4.62	10.37	9.59
	2001-2011	0.0131	0.0677***	0.0546***
		1.44	12.17	5.71
Nasdaq	1988-2000	0.0177***	0.0405***	0.0228***
		4.92	21.21	5.56
	2001-2011	0.0740***	0.1146***	0.0405***
		8.35	30.83	4.46
Panel B: Equation (15) of U_SIR				
NYSE/Amex	1988-2000	0.0014**	0.0050***	0.0036***
		1.98	8.25	3.58
	2001-2011	0.0024	0.0149***	0.0125***
		1.49	12.67	7.00
Nasdaq	1988-2000	-0.0048***	0.0116***	0.0164***
		-4.76	19.49	12.55
	2001-2011	-0.0074***	0.0199***	0.0273***
		-4.78	24.04	13.19

Table 9 Robustness checks

This table provides the robustness checks for Model 5 of Table 6 for the different dependent variables E_SIR and U_SIR. E_SIR is the average SIR over the past 12 months. U_SIR is SIR minus E_SIR. The other variables and data sources are described in Table 2. The sample period covers 1988-2011. The standard errors are clustered by the firm-plus-year dummy.

*, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	NYSE/Amex		Nasdaq	
	E_SIR	U_SIR	E_SIR	U_SIR
LNHFA	0.0115***	0.0140***	0.0016***	0.0022***
	14.11	16.84	12.78	13.98
IO_SHORT	0.0495***	0.0135***	0.0486***	0.0171***
	13.49	15.85	16.65	22.51
SPREAD	-0.0779***	-0.0096**	-0.0620***	-0.0159***
	-6.28	-2.23	-11.74	-9.43
ILLIQ	-0.0013***	0.0000	-0.0003***	0.0001***
	-4.43	-0.13	-9.37	16.08
LNCAP	-0.0029***	-0.0007***	0.0019***	0.0002***
	-6.75	-8.00	6.26	3.30
BM	0.0059***	-0.0004	-0.0001	-0.0001***
	4.24	-1.58	-0.30	-3.00
PR1Y	-0.0049***	0.0005**	-0.0031***	0.0023***
	-9.53	1.65	-10.41	13.22
D_DEBT	0.0125***	-0.0017***	0.0099***	-0.0015***
	11.65	-7.23	12.98	-8.33
D_OPT	0.0181***	-0.0011***	0.0284***	-0.0021***
	12.18	-4.14	28.54	-9.58
VOLAT	0.3477***	-0.0009	0.1461***	0.0360***
	10.77	-0.10	13.41	10.89
D_SP500	-0.0094***	0.0008***	-0.0231***	-0.0025***
	-7.93	3.51	-11.14	-6.33
MKT	-0.0076***	0.0025***	-0.0041***	0.0008
	-7.40	3.12	-7.93	1.44
LNVIX	0.0003	-0.0039***	0.0022***	-0.0035***
	0.50	-11.04	9.70	-13.24
R-square	0.27	0.04	0.35	0.05