Days to Cover and Stock Returns*

Harrison Hong† Frank Weikai Li‡ Sophie X. Ni §

José A. Scheinkman ¶ Philip Yan ‖

First Draft: April 2014

This Draft: December 2015

Abstract

A crowded trade problem emerges when speculators' positions are large relative to the liquidity of the asset, thereby making exit difficult. We study this problem, which has been a point of concern in the Dodd Frank Financial Reforms regarding systemic risk, through the lens of short-selling. We show in a simple model that days to cover (DTC), the ratio of short interest to trading volume, measures the costliness of exiting crowded trades. We find that arbitrageurs are worried about the crowding problem as short-sellers avoid illiquid stocks and require a significant premium to enter into such positions. A strategy shorting high DTC stocks and buying low DTC stocks generates a 1.2% monthly return. We show that there is a comparably large days-to-cover effect on the long positions of levered hedge funds.

*We thank Jeremy Stein, Charles Lee and seminar participants at European Finance Association, Seoul National University, Hong Kong University of Science and Technology, Baruch College, and Princeton University for helpful comments.

†Princeton University and NBER.

‡Hong Kong University of Science and Technology.

§Hong Kong University of Science and Technology.

¶Columbia University, Princeton University, and NBER.

‖Goldman Sachs.

1 Introduction

One of the most dramatic changes in financial markets in the last twenty years has been the rise of sophisticated investors. Hedge funds, perhaps the leading example of such investors, attracted negligible assets under management before the late nineties. The hedge fund sector now stands at 3.13 trillion dollars in 2015, which is roughly 10% of the size of the worldwide mutual fund sector (Mutual Fund Fact Book (2015)). These sophisticated investors, or arbitrageurs as we will loosely also refer to them, are practiced in both the use of short-selling and leverage, thereby magnifying their impact on financial markets.

A number of papers have pointed to this trend and its consequences for market efficiency (French (2008); Stein (2009)). One potentially important consequence of having so many sophisticated investors is the crowded trade problem. If too many arbitrageurs are on the same side of the trade, their coordinated exit from speculative and potentially levered strategies, which are typically absent from retail investors' portfolios, might be destabilizing for financial markets. This problem is a likely motivation behind increased disclosure requirements, as part of the Dodd Frank Financial Reforms following the Financial Crisis of 2008, to address systemic risk. Sophisticated investors are now required to disclose how much assets they have in certain strategies and at a much higher frequency.

The literature up to this point has focused on how crowded trades emerge naturally in price-untethered quantitative strategies such as momentum (Stein 2009; Lou and Polk (2013)). If investors are buying on price trends, it is likely that these investors end up with some probability in over-priced positions. But this problem is likely to emerge more generically whenever the aggregate position of arbitrageurs is large relative to the liquidity of the asset or trade, which would then make exit difficult. Indeed, such crowdedness across a number of trades are increasingly cited as a problem by the expanding hedge fund sector (see, e.g., "Crowded Trades Collapse," *Wall Street Journal*, December 3, 2015).

In this paper, we examine the link between crowded trades and liquidity in the context of short-selling. Short-selling is a good setting to study this question for a few reasons. First, shorting is by and large executed by sophisticated investors. Second, the ratio of shares shorted to shares outstanding (SR) has also been shown to predict negative returns consistent with informed arbitrageurs trading against mispricing.[1] Third, these trades are short-term in nature and implicitly levered in that the loan of the shares might be recalled at any moment, thereby triggering forced short-covering (i.e. the buying back of shares to pay back the loans) and hence potentially large price movements.

We show that days-to-cover (DTC), which divides short ratio by average daily share turnover—is a natural statistic for measuring the crowdedness of short trades. Consider as an example two stocks X and Y. Stock X with a short ratio of 5% and average daily turnover of 1% has a DTC of 5 days. Stock Y with a short ratio of 5% and average daily turnover of 5% has a DTC of 1 day. DTC is widely monitored by short-sellers and thought of as a risk management tool similar in spirit to the crowded trade interpretation we propose. Short-sellers report that they prefer stocks where they are able to close or cover their positions quickly without having to account for a big part of the daily market volume (i.e. Y compared to X). DTC is a measure of the ease with which they can achieve this goal.[2] Yet, as far as we know, it has not been previously studied.

We develop a simple model with A competitive arbitrageurs who face an over-priced stock but incur trading costs in establishing their short positions. The trading cost is increasing, all else equal, in the aggregate position of arbitrageurs. This is a reduced-form way of capturing the one-sidedness of arbitrage trades relative to liquidity in the market. We then solve for a symmetric equilibrium among arbitrageurs and prove that DTC captures the costliness of entering crowded positions.

Our model generates two key predictions to the extent that such costs are a concern. First, the short interest ratio of a stock should decline in the illiquidity of the stock. The reason is that stocks with higher price impact costs are more at risk for a crowded trade problem. So arbitrageurs will in equilibrium short less of these stocks, all else equal. Second, to the extent DTC is a serious concern for arbitrageurs, we should see high DTC stocks under-performing significantly to compensate the arbitrageurs for entering these positions in

[1] See, e.g., Figlewski (1981), Dechow, Hutton, Meulbroek, and Sloan (2001), and Asquith, Pathak, and Ritter (2005).

[2] We can equivalently, as practitioners do, calculate DTC using shares shorted to average shares traded daily; but for our empirical analysis below, it is more convenient to scale everything first by shares outstanding. The two ways of calculating are identical (absent significant changes in shares outstanding within a given month) and yield identical results.

the first place. Our insights also apply to long positions of hedge funds, which are typically levered; we will examine this levered long-side at end of the paper.

Our sample period runs from 1988 to 2012. Our baseline DTC measure divides the short-interest ratio measured in a given month by the average daily turnover during the same month.[3] In our sample, the mean DTC for a cross-section of stocks is 5.5 with a standard deviation of 8 days. In the time series, DTC has risen from a low of 3 days to 7 days during the recent sample, consistent with worries from the practitioners about crowded trades being an increasing problem.

We then test our first prediction, which is that short ratio (SR) of a stock should be correlated with its liquidity. We consider a variety of liquidity measures and find that the turnover proxy has the most explanatory power for SR. To address the causality of the relationship between SR and turnover, we instrument for turnover using the decimilization reforms of the early 2000s on the US stock exchanges. Our theory is agnostic on which liquidity measure is the best. But the fact that turnover is much stronger compared to bid-ask spread is consistent with worries of forced short covering. Such trades are bound to be large and turnover is likely a better measure of the true trading costs compared to bid-ask spread or other price impact measures which implicitly assume that investors can trade slowly.

Consistent with our second prediction, we find that arbitrageurs are compensated for entering high DTC positions. We can also consider an alternative form of DTC whereby we regress SR on all liquidity measures (including turnover) to create a Residual SR measure. Though Residual SR is also significant in predicting returns, consistent with a crowded trade effect, it turns out that DTC generates a much larger spread than Residual SR, especially for value-weighted portfolios. So we will focus on DTC in our discussions, though one could use the analogous Residual SR variable.

A strategy long low DTC decile stocks and short high DTC decile stocks yields 1.19% per month with a t-statistic of 6.67. A value-weighted DTC strategy also yields a statistically significant .67% per month. Since big stocks are more liquid, the DTC premium is smaller

[3] Our results are similar when we scale by average daily turnover calculated using data from prior months instead since liquidity tends to be persistent.

when considering larger stocks. We also show that DTC is statistically significant across a variety of return benchmarks, such as Daniel, Grinblatt, Titman, and Wermers (1997) adjusted returns and the Carhart (1997) four-factor and Carhart (1997) plus Pastor and Stambaugh (2003) five-factor alphas. In other words, DTC's predictive power is not simply picking up a liquidity effect, which one might worry is the case since DTC scales SR by daily turnover.

This excess return predictability, in our model, reflects two forces. The first is the usual limits of arbitrage assumption that there is limited capital in short-selling (see, e.g. Hanson and Sunderam (2014)). The second, and new to our model, is the explicit use of trading costs motivated by the risk of covering shorts. Consistent with this effect, during the Short-selling Ban of US financial stocks in 2008, we find that shorting high DTC stocks indeed experienced significant losses compared to other stocks as short-sellers were forced to cover their shorts. The cumulative drawdown for the value-weighted DTC strategy is -64% in the August, September and October of 2008. More generally, forced covering of shorts can come at the wrong time or be destabilizing and such effects are exacerbated for high DTC stocks.

In multiple regressions, we then show that the DTC effect is distinct from an SR effect. SR does not account for crowded trade issues, which are likely to vary across stocks depending on the liquidity of the stock. Second, we show that our DTC variable is not mechanically related to other turnover-based measures such as illiquidity (Amihud (2002)), which divides the absolute value of returns by dollar trading volume. These two variables, though both have turnover in the denominator, are not strongly correlated. The same is true if we use the inverse of turnover. The reason, as we showed above, is that SR is strongly correlated with turnover.

One might still worry that our DTC effect is somehow a redux of the well-known result that pricing anomalies such as the market-to-book effect are stronger in low turnover or less liquid stock since the limits of arbitrage is stronger in these stocks. To the extent SR is associated with high market-to-book stocks, perhaps DTC is inadvertently capturing this old phenomenon. We show that this is both theoretically and empirically not the case.

Our model also generates the usual prediction that stocks with higher lending fees are more overpriced.[4] But the lending fee effect is distinct from our DTC effect. Over a short sample since 2003, we also have lending fees by stock, which has also been shown to be a proxy for over-pricing. We find that DTC is as powerful a predictor as lending fees. Our lending fee data covers only the recent period.[5]

While we have focused on the short side, our analysis also applies theoretically to the long side. We show that this is the case using hedge fund holdings. Hedge funds are typically levered with a leverage ratio of 2 to 1, which is comparable to the implicit leverage of short-selling. The leverage hence makes forced covering of hedge fund long positions also plausible in contrast to the positions of unlevered mutual funds. We find a comparably large crowded trade effect on the long side of hedge fund trades as for short-selling.

Our paper proceeds as follows. Section 2 develops our model and outlines our predictions. Section 3 describes the various data we used in the analysis and presents summary statistics. Section 4 examines our first prediction concerning the relationship between short ratio and stock liquidity. Section 5 presents the second prediction regarding the DTC results. Section 6 looks at the whether there is also a crowded trade problem on the long side of levered hedge funds. Section 7 concludes.

2 Model

There are three dates $t = 0, 1, 2$. The asset yields a payoff at $t = 2$. There are two types of agents. A fraction γ of the agents are arbitrageurs and a fraction $1 - \gamma$ of the agents are optimists. Optimists believe that the random payoff has a mean μ_o. They start with one unit of the asset per-capita. Arbitrageurs have no endowment but believe that the payoff has mean $\mu_a < \mu_o$. All agents are risk-neutral.

Before trading at $t = 1$, half the optimists learn that they will value each dollar at $t = 2$ as $1 < \delta < 2$ dollars. The other half at the same time learn that they will value each dollar at $t = 2$ as $\delta' = 2 - \delta$ dollars. In equilibrium, the optimists that receive the positive time

[4]See, e.g., D'avolio (2002), Jones and Lamont (2002), Cohen, Diether, and Malloy (2007), and Beneish, Lee, and Nichols (2013).

[5]Following the literature, we also take short interest divided by institutional ownership, SIO, to be a measure of lending fees (see, e.g., Nagel (2005) and Drechsler and Song (2014)). Our effects are again robust to controlling for the SIO proxy of lending fees.

preference shock δ, more preference for the future, buy shares, whereas the ones that receive a negative time preference shock $2-\delta$, less preference for the future, would like to sell shares.

As such, there are two sources of trading in our model. The first is the differences in beliefs between the optimists and the arbitrageurs. The other is liquidity needs among the optimists which we generate using this preference shock.

To solve our model, we start with the portfolio maximization problem of the optimists. Optimists that receive the positive preference shock δ will choose a net demand n_o^+ at $t = 1$ that solves:

$$\max_n \left\{(1+n)\delta\mu_o - np_0 - \frac{c_o}{2}n^2\right\} \tag{1}$$

where p_0 is the price at $t = 0$ and c_o is the (perceived) trading cost parameter of optimists.[6] Notice that n denotes the net demand and $(1+n)$ is the investor's endowments of the shares plus his position. Thus the net demand by optimists that receive the positive time preference shock is:

$$n_o^+ = \frac{\delta\mu_o - p_0}{c_o} \tag{2}$$

Optimists that receive the negative preference shock $2 - \delta$ face a similar maximization problem and will choose a net demand n_o^- that solves:

$$\max_n \left\{(1+n)(2-\delta)\mu_o - np_0 - \frac{c_o}{2}n^2\right\}. \tag{3}$$

It follows then that the net demand by optimists that receive the negative time preference shock is:

$$n_o^- = \frac{(2-\delta)\mu_o - p_0}{c_o} \tag{4}$$

We will assume that there are A risk neutral arbitrageurs, which together represent a fraction γ of the total number of traders. We will make parameter choices that will imply that, in equilibrium, arbitrageurs are short while optimists are long. The cost of trading faced by an individual arbitrageur ℓ that acquires n_a^ℓ depends on the total amount traded

[6]It is not crucial that the optimists who we view as retail investors and creating the mispricing perceive the trading cost correctly. Indeed, studies on retail investor trading such as Barber and Odean (2000) indicate that retail investors under-estimate these trading costs.

by arbitrageurs and is given by

$$\frac{c}{2} n_a^\ell \sum_j n_a^j.$$

The motivation for the presence of this externality is as follows. As in the literature on trading costs (see, e.g., Vayanos (1998)), we think of the quadratic trading cost function as a reduced form for price impact, which are important in many markets and vary across stocks. Under a price impact interpretation, the trading costs can be identified with shocks that force shorts to prematurely close their positions. With probability $\frac{1}{2}$, the short-sellers receive shocks and have to close their short position and buy back the shares. Since the optimists still want to hold onto their shares, this will lead to price impact. Market makers can then provide liquidity in this situation but will set an ask that depends on the aggregate amount of shares that arbitrageurs wish to sell. Hence arbitrageurs have to trade away from the average price by an amount that depends on the aggregate amount of shorts. Here c can then be interpreted as inventory considerations in Grossman and Miller (1988).

An individual arbitrageur ℓ would take the trading of all other arbitrageurs as given and maximize:

$$\max_{n_a^\ell} \left\{ n_a^\ell(\mu_a + f - p_0) - \frac{c}{2} n_a^\ell \sum_j n_a^j \right\}, \tag{5}$$

where f is the fee to shorting. We assume that the lending fee is exogenous to start and that the fee is collected by a broker. In the Appendix, we endogenize the lending fee and show that the main conclusions remain. If the arbitrageurs short n, then nf is the total short-fees paid by the arbitrageurs. Recall that they have no initial endowment of shares to begin with. In an interior symmetric equilibrium, the demand by any arbitrageur must satisfy:

$$n_a^\ell = \frac{\mu_a + f - p_0}{c + \frac{c}{2}(A-1)}. \tag{6}$$

Hence the aggregate demand by the arbitrageur sector n_a satisfies

$$n_a = \frac{\mu_a + f - p_0}{\frac{c}{A} + \frac{c}{2}\frac{(A-1)}{A}}. \tag{7}$$

In the sequel we fix A and write

$$c_a = \frac{c}{A} + \frac{c}{2}\frac{(A-1)}{A}. \tag{8}$$

Thus the aggregate demand by the arbitrageur sector satisfies[7] :

$$n_a = \frac{\mu_a + f - p_0}{c_a}. \tag{9}$$

We will focus on the equilibrium in which the optimists that receive a negative preference shock sell some of their shares and the arbitrageurs short. To this end, we require the following set of parameter restrictions:

$$\delta\mu_o > p_0 > \max\{(2-\delta)\mu_o; \mu_a + f\} \tag{10}$$

and

$$p_0 \leq (2-\delta)\mu_o + c. \tag{11}$$

The first set of three parameter restrictions in Equation (10) essentially says that the optimists with the positive preference shock δ buy but the optimists with the negative preference shock sell while the arbitrageurs short-sell. The second parameter restriction in Equation (11) says that the optimists with the negative preference shock do not short-sell.

Adding up the three types we get:

$$\frac{\gamma}{c_a}[\mu_a + f - p_0] + \frac{1-\gamma}{2c_o}[(2-\delta)\mu_0 - p_0] + \frac{1-\gamma}{2c_o}[\delta\mu_0 - p_0] = 0 \tag{12}$$

or

$$p_0 = \frac{\frac{1-\gamma}{c_o}\mu_o + \frac{\gamma}{c_a}(\mu_a + f)}{\frac{\gamma}{c_a} + \frac{1-\gamma}{c_o}} = \mu_a + \frac{\frac{1-\gamma}{c_o}(\mu_o - \mu_a) + \frac{\gamma}{c_a}f}{\frac{\gamma}{c_a} + \frac{1-\gamma}{c_o}} \tag{13}$$

We can think of the first term μ_a as the fundamental value associated with the expectation

[7] We may think as c_a as the cost perceived by arbitrageurs. The larger the number of arbitrageurs A, the smaller is the cost that arbitrageurs perceive, since each arbitrageur ignores the impact of their actions on the other arbitrageurs.

of the risk-neutral arbitrageurs. The second two terms, both of which are positive, reflect then the overpricing due to costly short-selling γf and costly trading $(1-\gamma)(\mu_o - \mu_a)$.

This then leads us to our first proposition that is the basis of our first prediction.

Proposition 1. *Short interest is given by*

$$\gamma|n_a| = \gamma\frac{p_0 - f - \mu_a}{c_a} = \frac{\frac{\gamma(1-\gamma)}{c_a c_o}[\mu_o - \mu_a - f]}{\frac{\gamma}{c_a} + \frac{1-\gamma}{c_o}}. \tag{14}$$

The equilibrium short ratio then satisfies:

$$SR = \frac{\gamma}{1-\gamma}|n_a| = \frac{\frac{\gamma}{c_a c_o}[\mu_o - \mu_a - f]}{\frac{\gamma}{c_a} + \frac{1-\gamma}{c_o}} = \frac{\gamma[\mu_o - \mu_a - f]}{\gamma c_o + (1-\gamma)c_a}. \tag{15}$$

Since only the optimists that receive the positive preference shock are buys, volume is given by

$$V = \frac{1-\gamma}{2}\frac{\delta\mu_o - p_0}{c_o}. \tag{16}$$

So

$$V = \frac{1-\gamma}{2c_o}\frac{(\delta\mu_o - \mu_a - f)\frac{\gamma}{c_a} + (\delta - 1)\mu_o\frac{1-\gamma}{c_o}}{\frac{\gamma}{c_a} + \frac{1-\gamma}{c_o}}. \tag{17}$$

Furthermore,

$$\frac{\partial V}{\partial c_a} < 0 \tag{18}$$

and

$$\frac{\partial SR}{\partial c_a} < 0. \tag{19}$$

Thus, SR is positively correlated with V.

Proposition 1 points out the problematic nature of the short interest ratio (SR) as a measure of the crowded trades. High SR might simply reflect low trading costs. Ideally, a crowded trade measure captures both the number of arbitrageurs in the trade as well as the liquidity on the other side. If c_a and c_o increase in the same proportions, the capital gains or compensation that arbitrageurs expect for entering the trade, $p_0 - \mu_a$, does not change, but SR decreases by that same proportion. If the cost of trading varies across assets, SR is not a good proxy for the compensation for crowded trades.

We next show that DTC, which divides short-interest by the number of shares traded, is a more robust measure of crowded trades.

Proposition 2. *Days to cover is given by*

$$DTC := \frac{\gamma |n_a|}{V}. \tag{20}$$

DTC is a better measure of crowded trades than SR since the elasticity of DTC with respect to c_a is smaller than the elasticity of SR with respect to c_a:

$$0 > e_{c_a}(DTC) = e_{c_a}(SR) - e_{c_a}(V) > e_{c_a}(SR). \tag{21}$$

The logic of Proposition 2 then implies that we can sort on DTC in the data and see the extent to which short-sellers are compensated to enter these positions.

Finally, we explain in Proposition 3 below why our DTC effect is not simply a redux of a standard result, whereby anomalies are stronger in smaller or less liquid stocks that are more difficult to arbitrage. To be more concrete, consider the well-known fact that the market-to-book effect is stronger in low turnover stocks. In our model, we can measure the market-to-book effect as $p_o - \mu_a$ or the degree of over-pricing. And c_a captures the cost of arbitrage, which one can associate with turnover. To the extent SR is associated with high market-to-book stocks, perhaps DTC is simply capturing this old result. We show that this is theoretically not the case.

We need to a few calculations to understand why. First, consider a Taylor expansion of $p_0 - \mu_a$ around $\gamma = 0$ to a first order:

$$p_0 - \mu_a = \mu_o - \mu_a + \left[\frac{c_o(f - \mu_o + \mu_a)}{c_a}\right]\gamma. \tag{22}$$

Hence for small γ, variations of $p_o - \mu_a$ are dominated by variations in μ_o and not c_a. On the other hand, an expansion of SR gives

$$SR = \left[\frac{\mu_o - \mu_a - f}{c_a}\right]\gamma. \tag{23}$$

The lack of a zero-th order γ term indicates that even for small γ, variations of SR depend on both μ_o and c_a. In other words, a double sort on SR and turnover is not equivalent to a double sort on market-to-book and turnover. Whereas a sort on price essentially picks up variation in sentiment μ_o, a sort on SR depends on both μ_o and cost c_a.

Notice that the following inequalities that are easily derived from our equilibrium. First,

$$\frac{\partial(p_o - \mu_a)}{\partial \mu_o} = \frac{(1-\gamma)c_a}{\gamma c_o + (1-\gamma)c_a} > 0, \tag{24}$$

(i.e. the more optimistic the sentiment, the more the over-pricing). Second,

$$\frac{\partial(p_o - \mu_a)}{\partial c_a} = \frac{(1-\gamma)\gamma c_o[\mu_o - \mu_a - f]}{(\gamma c_o + (1-\gamma)c_a)^2} > 0, \tag{25}$$

(i.e. the higher the arbitrage cost, the more the over-pricing). Moreover, sentiment μ_o has a larger effect on over-pricing among high arbitrage cost stocks,

$$\frac{\partial^2(p_o - \mu_a)}{\partial \mu_o \partial c_a} = \frac{(1-\gamma)\gamma c_o}{(\gamma c_o + (1-\gamma)c_a)^2} > 0. \tag{26}$$

This latter positive cross-partial derivative is the essence of the statement that anomalies like market-to-book effect are stronger in harder to arbitrage or low turnover stocks.

However, the SR effect is not necessarily stronger in harder to arbitrage or low turnover stocks to the extent SR varies also because of c_a as it is easy to show that

$$\frac{\partial^2(p_o - \mu_a)}{\partial c_a^2} < 0. \tag{27}$$

Proposition 3. *The DTC effect is not a redux of a standard anomalies effect, whereby the SR effect (i.e. similar to the market-to-book effect) is stronger in harder to arbitrage (i.e. low turnover) stocks.*

3 Data, Variables and Summary Statistics

We obtain monthly short interest data from the NYSE, Amex and Nasdaq exchanges from 1988 to 2008. The exception is for Amex from 2005-2008, which are from Compustat. Short interest data from 2009 to 2012 are obtained from Compustat.[8] We use the short interest data that is reported in a given month, typically the mid-point. We start our sample in 1988 since there is little shorting earlier than this date. To form short interest ratio (SR), we normalize short interest by total shares outstanding from CRSP.

In addition to data on the level of short interest, we also construct two variables proxy for stocks' loan fees from the Markit equity lending database. The first variable, Fee1, is the simple average fees of stock borrowing transactions from hedge funds in a given security, which is the difference between the risk-free rate and the rebate rate. Fee1 is only available for a stock to the extent that the stock is being shorted by a hedge fund. The second variable, Fee2, which covers all stocks, is a score from 1 to 10 created by Markit using their proprietary information meant to capture the cost of borrowing the stock. Here 1 is the cheapest to short and 10 the most difficult. The first fee variable is available since November of 2006 while the second fee variable is available since October of 2003.

In the second part of our empirical analysis, we also utilize hedge fund holdings data provided by Jiang (2015).[9] For each stock in the sample, we compute its quarterly hedge fund holdings (HFH) as the sum of shares held by all hedge funds reported at each quarter divided by the total number of shares outstanding. If the stock is not held by even a single hedge fund in that quarter, its HFH is set to zero.

Data on monthly stock returns and daily trading volume are obtained from CRSP. We require stocks to be listed on NYSE, AMEX and NASDAQ and common stocks (i.e. share type code equals to 10 or 11). We remove stocks with month end price less than \$3. Turnover is calculated as the daily ratio of the number of total shares traded to the number of total shares outstanding. The daily turnover ratio is averaged within a month to get a monthly

[8]The NYSE-AMEX data is available on Compustat starting in 1976. The NASDAQ data is only available starting in 2003. There are two versions: unadjusted and adjusted for stock splits. The exchange data we are using is unadjusted.

[9]The detailed method to extract hedge fund holdings data could be found in his paper as well as in Griffin and Xu (2009).

variable. Since the dealer nature of the NASDAQ market makes the turnover on it difficult to compare with the turnover observed on NYSE and AMEX, we follow Gao and Ritter (2010) by adjusting trading volume for NASDAQ stocks.[10]

We use standard control variables in our empirical analysis part. Following Fama and French (1992), market beta (Beta) of an individual stock is estimated by running a time-series regression of monthly stock excess return on market excess return over the prior 60 months if available (minimum of 24 months). Size (LnME) is defined as natural logarithm of market capitalization at the end of June in each year. Book value equals the value of common stockholders' equity, plus deferred taxes and investment tax credit, minus the book value of preferred stock. Book-to-market (LnBM) ratio equals to the most recent fiscal year-end report of book value divided by market capitalization at the end of calendar year t-1. Momentum (Mom) is defined as the cumulative holding-period return from month t-12 and t-2. We follow the literature by skipping the most recent month return when constructing Momentum variable. The short term reversal measure (REV) is the prior month's return. Institutional ownership (IO) is the sum of shares held by institutions from 13F filings in each quarter divided by total shares outstanding.

Idiosyncratic volatility (IVOL) is the standard deviation of the residuals from the regression of daily stock excess return on Fama-French three factor returns within a month (Ang et al., 2006). Following Diether, Malloy, and Scherbina (2002), analyst earnings forecast dispersion (DISP) is the standard deviation of annual earnings-per-share forecasts scaled by the absolute value of the average outstanding forecast. Firm-level variables are obtained from Compustat annual files. Analyst forecast data is from I/B/E/S. Data on institutional holdings is from Thompson Reuters Financial.

In addition to trading volume, we create several commonly used measures of stock liquidity based on market microstructure literature. Our liquidity measures include the Amihud (2002) illiquidity measure, the FHT measure which is backed out from the frequency of zero returns (Fong, Holden and Trzcinka 2014), the Pastor and Stambaugh (2003) liquidity mea-

[10]Specifically, we divide NASDAQ volume by 2.0, 1.8, 1.6, and 1 for the periods prior to February 2001, between February 2001 and December 2001, between January 2002 and December 2003, and January 2004 and later years, respectively.

sure and the percentage quoted spread using daily close price (Chung and Zhang 2014). The details of construction of these liquidity measure is in the Appendix B.

In empirical analysis section, we compute monthly characteristic adjusted return by subtracting the stock's raw return by the return of the benchmark group to which the stock belongs to (see, e.g., Daniel, Grinblatt, Titman, and Wermers (1997)). The 5*5*5 benchmark groups are formed at the end of June of each year based on size, book-to-market ratio, and past one year return. The monthly benchmark returns and stock assignments are obtained from Russ Wermer's website [11].

3.1 Days-to-Cover: Scale Short Ratio by Share Turnover

Based on the analysis in our hypothesis development section, we consider Days-to-cover (DTC) measure, which scales SR by share turnover, as a superior measure of arbitrageur opinion about a stock's over-pricing. Specifically, DTC is defined as

$$DTC = \frac{SR}{Average\,Daily\,Turnover} \tag{28}$$

Recall SR is typically measured at the middle of a given month. The average of daily turnover is taken with respect to the same month's daily share turnover.[12] Below, we consider averaging turnover over prior months and find similar results. The DTC measure has an intuitive interpretation as roughly how many days of average share volume it would take for all short sellers to cover their short positions. Figure 1 plots the market average short ratio, turnover and DTC measure in time series. There is an increasing trend for both SR and turnover. Short ratio is negligible in the beginning of our sample period, at around 1%, and then steadily rises and peaks in 2008 at around 5.5%. It has subsequently fallen back to around 4% in the most recent years of our sample.

Share turnover also improved dramatically during this period, as indicated by the upward trend of market mean turnover ratio. To put daily turnover on the same scale, we multiplied by 1000. In 2008, mean daily turnover is 8 in units of 1000, which translates into 0.8%

[11] http://terpconnect.umd.edu/ wermers/ftpsite/Dgtw/coverpage.htm

[12] We can also calculate DTC using shares shorted to average shares traded daily but for exposition it is more convenient to scale everything first by shares outstanding.

per day or roughly 200% per year. In the beginning of the sample, the daily turnover is a fraction of this, at around .20% per day or 50% per year.

However, SR increases more than turnover ratio, so mean days-to-cover (DTC) also increased a lot, from around 3 day to about 7 days. The rise of SR due to hedge fund or arbitrageur activity is well known. One can interpret that the latter part of our sample might be especially relevant in terms of evaluating our model since this is when there are significant levels of short-selling or arbitrageur activity.

3.2 Summary Statistics

To dig deeper into these numbers, Table 1 presents the summary statistics for the variables used in our analysis. Panel A reports the time series average of the cross-sectional mean and standard deviation of the variables for the full sample and by market capitalization quintiles. We start with SR. It has a mean of 2.26%. Note that there is more shorting in Size Quintiles 2-4 than in Quintiles 1 and 5, consistent with Hanson and Sunderam (2014). The mean turnover in our sample is .46% per day. The mean DTC in our sample is 5.45 days with a standard deviation of 8.26. Notice that the standard deviation of DTC is quite large and this will play a key role in our analysis below. The mean institutional ownership in our sample is 42%. The remaining summary statistics are well known and do not require additional discussion.

Panel B of Table 1 reports these statistics for our sub-sample where we also have lending fee data. More specifically, we use the sub-sample where Fee2 is available which is starting from October 2003. Fee2 is Markit's internal rating system for whether a stock is difficult to borrow for shorting, where 1 is cheapest and 10 is the most difficult. We also report summary statistics for Fee1 which is the simple average of fees of the stock borrowing transactions among hedge funds, which is available only in more recent sample starting in November 2006.

Notice that both SR and DTC have higher means, 4.32% and 6.85 respectively, in this sub-sample since it is more recent. A similar comment holds for institutional ownership, which is 58.71% in this sample. Importantly, notice that Fee2 has a mean of 1.39 and a standard deviation of 3.56. Importantly, notice that this number is fairly consistent across Size Quintiles. If we look at Fee1, we see that the mean is 48 basis points (annualized) with a standard deviation of 91 basis points. We treat the fees as given and use them as control variables.

To take this point a bit further, Panel C of Table 1 reports the pairwise correlations among our variables where they overlap. The correlation between DTC and SR is high at .83 but far from perfectly correlated, as what be predicted by our model since trading costs vary across stocks (i.e. turnover varies) and shorts are influenced by the same underlying unobservable costs (i.e. SR covaries with turnover as we have already established). Indeed, there is different information captured in these two variables which we will exploit in our asset pricing exercises below.

The other notable correlation is that DTC is not very correlated with IO, Fee1 and Fee2. They are .3, .17 and .04, respectively. This points to the fact that the issue we are dealing with regarding the influence of heterogeneous trading costs on shorting is an independent issue from the lending market frictions that have been emphasized in the literature.

4 Short Interest and Stock Liquidity

4.1 The Relation between Short Interest and Stock Liquidity

The first prediction of our model is that arbitrageurs' aggregate short position should be positively correlated with stock's liquidity. To test this, we run the following regression to examine the relation between short ratio (SR) and stock liquidity:

$$SR_{i,t} = a + bLIQ_{i,t} + cLnMEi,t + dLnBMi,t + eMOMi,t + fIOi,t + gIVOLi,t + \epsilon_{i,t} \quad (29)$$

where LIQ represents one of our five liquidity measures: Turnover, Amihud, FHT, Pastor and Stambaugh and Daily Percent Quoted Spread.

Table 2 reports the regression results. In column (1) to (7) we use Fama-Macbeth (1973) regression method. All variables are standardized to have mean 0 and standard deviation of 1, so the coefficients on independent variables are directly comparable to each other. In column (1), we only include turnover as the explanatory variables in the regression. As we can see, the coefficient on turnover is 0.47 and highly significant. Because our variables are standardized, we could interpret the coefficient as that one standard deviation shock to turnover is associated with 0.47 standard deviation of short ratio. The average R-square from this univariate regression is 22.4%, indicating turnover alone explain a significant portion of cross-sectional variation of short interest ratio.

In column (2), we add control variables along with turnover. The coefficient on these control variables are all significant and the sign is consistent with previous studies on the determinants of SR (Dechow et al., 2001; Hirshleifer, Teoh and Yu, 2011). Short interest is higher among small stocks, growth stocks, loser stocks, stocks with high institutional ownership and low idiosyncratic volatility. However, the average R-square increases from 22.4% to only 28.2% in this multiple regression and the magnitude of the coefficient on these control variables are much smaller than the coefficient on turnover. This indicates that among all the stock attributes, short sellers care most about the stocks' liquidity and how easy they could exit their position, which is consistent with our model.

In columns (3) to (6), we replace turnover with other liquidity measures and find the results mostly support our prediction that short interest is positively correlated with stocks' liquidity. In column (3), we find the coefficient on Amihud illiquidity is -.0511 (t=-9.19). Since a higher Amihud measures lower liquidity, this indicates that short sellers reduce their position when the stock is more illiquid. Similarly, we find negative coefficient when liquidity is measured by the FHT (which is based on frequency of zero return trading days) and daily closing quoted spread measure, as higher value of these two measures indicate less liquidity. The only exception is Pastor-Stambaugh (2003) liquidity measure in column (5), which should be a positive coefficient. As pointed out by Pastor and Stambaugh (2003), their meausre of liquidity is quite noisy so that they caution against using it as a measure of liquidity at individual stock level.

In column (7), we add all these liquidity measures in the regression along with other control variables. As we can see, the coefficient on turnover barely changes and is also the most important determinant of short interest ratio based on the coefficient magnitude. The average R-square in this regression is 30.9%, which represents less than 40% increase relative to the regression with turnover in column (1).

In column (8), we run panel regression with firm-fixed effect. The firm-fixed effect regression helps alleviate the concern that some unobservables correlated with both turnover and short interest may be present and bias the coefficient. The coefficient on turnover is 0.315, similar to the Fama-Macbeth regression result.

4.2 Instrumental Variable Regression Using 2001 Shift to Decimalization

The results in Table 2 show that short interest is strongly associated with stock's liquidity level, especially trading volume. However, we are still concerned that some time-varying unobservables may drive the strong correlation between turnover and SR. In this section, we try to establish a causal effect of liquidity on short interest by exploiting a large exogenous shock to stock liquidity during our sample period. Prior to 2001, the minimum tick size for quotes and trades on the three major U.S. exchanges was $1/16. Over the period of August 28, 2000 to January 29, 2001, NYSE and Amex reduced the minimum tick size to pennies and terminated the system of fractional pricing. NASDAQ decimalized shortly thereafter over the period of March 12, 2001 to April 9, 2001. Prior studies show significant increases in liquidity as a result of decimalization, especially among actively traded stocks (Bessembinder (2003), Furfine (2003)). Decimalization appears to be a good candidate to generate exogenous variation in liquidity since it directly affects liquidity. But it is unlikely to directly affect short selling, and changes in liquidity surrounding decimalization exhibit variation in the cross-section of stocks.

We use decimalization as an instrument for share turnover and examine how exogenous change in liquidity affect short interest in a 2SLS regression. Our test uses the fact that decimalization is first implemented for NYSE/Amex listed stocks and subsequently for NASDAQ listed stocks. In the first stage, we run the following regression:

$$Turnover_{i,t} = a + bNyseamex + cPost_t + dNyseamex * Post_t + eXi, t + \epsilon_{i,t} \tag{30}$$

Here Nyseamex is a dummy equal to one for NYSE/Amex listed stocks and zero for NASDAQ listed stocks. Post is a dummy equal to 1 for the period of February and March of 2001 and 0 for the period from March 2000 to January 2001. We restrict our sample period to March 2000 to March 2001 in this test. Notice we are taking advantage of the staggered reforms to compare NYSE/AMEX stocks under-going reforms to NASDAQ stocks that have yet to undergo reforms.

The coefficient on NYSE/Amex dummy measures the average difference in turnover between NYSE/Amex listed stocks and NASDAQ listed stocks in the pre-decimalization period. The coefficient on Post dummy measures the change in turnover in the post-decimalization period compared to pre-decimalization period for NASDAQ stocks. The coefficient on d therefore measures the change in stock turnover on NYSE/Amex listed stocks in the post-decimalization period relative to the change in turnover on NASDAQ-listed stocks in a period when only NYSE/Amex stocks have gone through the decimalization process.

The results from this first-stage regression is reported in the left column of Table 3. The coefficient on Nyseamex dummy is significantly negative, indicating stocks listed on NYSE/Amex have less share turnover. More importantly, we find the coefficient on the interaction term is significantly positive with a coefficient of 0.074 (t=4.01), indicating share turnover increase significantly for NYSE/Amex listed stocks during the post-decimalization period compared to NASDAQ stocks.

In the second stage, we use the predicted turnover from first stage IV regression to explain short interest. The result is reported in the right column of table 3. As we can see, the coefficient on turnover is 1.074 and highly significant. The result supports our prediction that short selling activity react strongly to liquidity changes.

4.3 Correlation between Short Ratio and Turnover Over Time

Now that we have established the causal relationship between turnover and SR, we next look at how this relationship has changed over time. Panel A of Figure 2 depicts the cross-sectional correlation between SR and turnover over time. That is, for every month, we calculate the cross-sectional correlation between SR and turnover. We then plot these cross-sectional correlations over time. Note that the correlation between SR and turnover is positive and has increased over time. In the beginning of our sample in 1988, the correlation is around 0.3 to 0.4. But since the 2000s, this correlation is around 0.5 to 0.6. These results are using contemporaneous values of short ratio and turnover but the same thing holds if we used lagged values of turnover.

We next show that these correlations are not driven by omitted valuation factors. To control for other factors that can potentially confound the cross sectional correlations, we compute the partial correlation between short ratio and share turnover controlling for size, book-to-market, past 12 months cumulative returns and institutional ownership. Specifically, every month t we run the regressions

$$SR_{it} = \alpha_t^{SR} + \beta_{it}^{SR} X_{it} + u_{it}^{SR} \tag{31}$$

$$Turnover_{it} = \alpha_t^{T} + \beta_{it}^{T} X_{it} + u_{it}^{T} \tag{32}$$

where X_{it} is our collection of control variables, and $Turnover_{it}$ is our trading cost proxy share turnover. The cross-sectional partial correlation between short ratio and turnover in time t is given by

$$\rho_t = Corr(\hat{u}_{it}^{SR}, \hat{u}_{it}^{T}) \tag{33}$$

Panel B of Figure 2 plots the partial correlation of SR with turnover. We see that the observations made in the univariate correlation case earlier remain true after controlling for other variables. The overall magnitude of the correlation is smaller, dropping in the post-2000 sample from around .5 to .6 around .4 to .5. The partial correlations are nonetheless still positive. Moreover, we see a similar pattern overtime with the correlation becoming stronger in the recent period. These findings are consistent with worries on the part of the hedge fund sector in the crowded trade problem over time and wanting to avoid illiquid stocks as a result.

5 Compensation for Crowded Trades

To the extent that sophisticated investors worry about the crowded trade problem, our second prediction is that they should get compensated for being in these positions. Hence, DTC should strongly and negatively predict subsequent stock returns. In this section, we test this second prediction of our model using both portfolio sorts and multiple regressions.

5.1 Portfolio Sorts

In this section, we show that stocks sorted on DTC generate significant return spreads. We conduct the decile portfolio sorts as follows. At the end of each month, we sort stocks into deciles based on DTC. We then compute the average return of each decile portfolio over the next month, both equal-weighted and value-weighted. This gives us a time series of monthly returns for each decile. We use these time series to compute the average return of each decile over the entire sample. As we are most interested in the return spread between the two extreme portfolios, we only report the return to a long-short portfolio, i.e., a zero investment portfolio that goes long the stocks in the lowest DTC decile and shorts the stocks in the highest DTC decile.[13] We report the average return (and associated t-statistics) of this long-short portfolio in the left columns, the characteristics-adjusted return spread, computed in the way described by Daniel, Grinblatt, Titman, and Wermers (1997) and denoted DGTW in the second columns, the Cahart (1997) 4-factor alphas in the third column and 5-factor adjusted alpha (the return adjusted by the Fama-French three factors, the momentum factor, and the Pastor and Stambaugh (2003) liquidity risk factor) in the right columns.

The result is reported in Table 4. In Panel A, the equal-weighted return to the long-short portfolio sorted on DTC is a monthly 1.19% per month, with a t-stat of 6.67. The Sharpe ratio is 1.33. For DGTW returns, the numbers are very similar. The four-factor

[13]The mean DTC for the decile 1 portfolio is 0.126, while the mean DTC is 25 for the decile 10 portfolio.

and five-factor alphas are 1.3% per month with t-statistics of around 8. In Panel B, we see that the value-weighted results are weaker but are nonetheless statistically and economically significant across the board. For excess returns, it is .67% with a t-statistic of 2.24. For DGTW returns, it is .59% with a t-statistic of 2.56. The figures are around .70% for the four- and five-factor alphas and both figures are also statistically significant. So regardless of the metric, high DTC stocks underperform low DTC stocks significantly.

In Table 5, we examine the robustness of our portfolio sorts on a variety of samples. As a baseline, we report the full sample results. The DGTW-adjusted return on the long-short DTC strategy is more pronounced when returns are equally-weighted. This is to be expected, as small stock are more illiquid. In the remaining sub-samples, we use equal-weighted portfolios.

We next check whether our results hold not only in the full sample, but also in each of two sub-periods: one that starts in January 1988 and ends in December 1999, and another that starts in January 2000 and ends in December 2012. We choose 2000 as the breakpoint as hedge fund activity is more significant after 2000. Notice that DTC strategy generates a monthly return of 0.70% in the first half of the sample and 1.10% in the second half of the sample. The higher return to DTC strategy in the most recent sample is consistent with our model that arbitrageurs become more concerned about the crowdedness of their short positions over time, which manifested as higher compensation for taking short position on illiquid stocks.

The third row of Table 5 shows that our results hold for stocks listed on both NYSE-Amex and NASDAQ stock exchanges. This assures us that our results are robust with respect to the different ways trading volume is counted for NYSE-Amex and NASDAQ listed stocks. DTC strategy generates a monthly DGTW-adjusted return of 0.96% in the NYSE-Amex sample and 0.76% in the Nasdaq sample.

The fourth row shows that sorting stocks into either 5 portfolios or 20 portfolios does not change our results. Using 5 portfolios, high DTC stocks underperform low DTC stocks by a monthly 0.78% (t=6.04). Using 20 portfolios, the return spread is 1.26% (t=7.44).

In Figure 3, we take a more detailed look and report by decile portfolios the equal-weighted DGTW adjusted portfolio returns. One can see that the pattern for DTC is fairly monotonic, consistent with our model. The lowest DTC decile portfolio has a excess return of nearly .45% and the highest DTC decile portfolio has a excess return of -0.50%. So the spread across 10 to 1 is almost 1% and monotonic across deciles. Boehmer, Huszar and Jordan (2009) document that stocks with low short interest have positive abnormal returns in the future. The result in our paper is consistent with their findings that low DTC stocks also have positive DGTW-adjusted excess return. In untabulated results, we show that this is mainly due to the part of low DTC stocks that also heavily held by hedge funds. Specifically, we show that low DTC stocks with low hedge fund holdings do not have abnormal positive returns.

The fifth row shows the results when we remove micro-cap stocks. First we remove stock in the bottom 10% of market capitalization using the NYSE cutoff. The return spread is a monthly 0.84% for DTC strategy in this sample. When we drop the bottom 20%, the return spread is 0.79%. Both are highly significant, although the magnitude decreases compared to full sample results. DTC is more valuable when considering all stocks than when considering just big stocks. The reason is that dropping the smaller stocks reduces variation in trading costs in our sample, which means adjusting for trading costs becomes less valuable. Nonetheless, we still see even among fairly large stocks in the universe that we see high DTC stocks underperform low DTC stocks significantly.

The sixth row of the Table 5 shows that we obtain similar results if we exclude stocks whose price at the sorting month is less than $5. Again, DTC strategy generates monthly excess return of 0.59%. Across almost all the specifications in Table 5, high DTC stocks underperform low DTC stocks significantly.

To get a better sense of how the DTC strategy performs from an investor's perspective, we compute the cumulative returns to DTC spread portfolio. Figure 4 Panel A shows that the equal-weighted DTC strategy generate significant cumulative returns in our sample period. One dollar invested in the long-short portfolio sorted on DTC at the beginning of 1988 will grow to 30 dollars at the end of our sample period.[14] Panel B of Figure 4 plots

[14]We have to be careful in interpreting these graphs since they are dependent on how the strategies did initially. The mean monthly differences are more robust to when one starts the sample.

the cumulative returns for the value-weighted portfolio. The DTC strategy also works for value-weighted returns, although the magnitudes are less dramatic for the reasons outlined above. Interestingly, there is a significant draw-down in the strategy of DTC during the short-selling ban of 2008. The noticeable drawdown in the DTC strategy occurs mostly in the months of short-selling ban in August and September of 2008. The ban means that the short-sellers had to cover their short positions and high DTC stocks means it is harder to cover, resulting in greater losses. We view it as comforting that DTC is indeed capturing accurately the marginal cost of shorts, i.e. stocks that are more difficult to buy back.

Figure 5 plots the annual equal-weighted returns to DTC spread portfolio. This figure also highlights the low volatility of returns to DTC strategy. As we can see, DTC strategy delivers positive annual returns in 24 out of 25 years.

5.2 Is DTC a Redux of More Price Anomalies in Less Liquid Stocks?

One might think that sorting on DTC is equivalent to the stronger abnormal return sorted on SR among low turnover stock. That is, we know almost every anomaly produce larger abnormal profits within the group of stocks subject to greater limits-to-arbitrage. For example, the value and momentum effect is stronger among stocks with less liquidity (Ali, Hwang and Trombley 2003; Sadka 2006). Table 6 shows that our DTC measure is distinct from a double sort on SR and turnover.

To produce this table, in every month we sort all the stocks into quintiles based on short interest ratio (SR). We independently sort stocks into quintiles based on turnover (Turn). The monthly excess returns of the 25 portfolios are reported in Panel A. If DTC is equivalent to a double sort, we would expect the return spread to a long-short portfolio based on SR to be largest among low turnover stocks. However, we see the return spread is actually lower among low turnover stocks. The equal-weighted return to the high SR minus low SR portfolio generates monthly excess return of -0.91% in the lowest turnover quintile, while this number is -1.76% among highest turnover quintile. For the value-weighted portfolios, we observe similar pattern. SR portfolio generates monthly excess return of -0.35% in lowest turnover stocks and -1.21% among highest turnover stocks. Panel B of Table 6 reports the three-factor alphas of the 25 portfolios. Here we see similar patterns. SR strategy generate smaller and less significant alphas among low turnover stocks than among high turnover stocks. The result in this table clearly shows that our DTC effect is not the same as a double sort on SR and turnover.

5.3 Fama-Macbeth Regressions

We now test our main hypothesis using the Fama-Macbeth (1973) regression methodology. One advantage of this methodology is that it allows us to examine the predictive power of DTC while controlling for known predictors of cross-sectional stock returns. This is important because, as shown in Table 1, DTC is correlated with some of these predictors. We conduct the Fama-Macbeth regressions in the usual way. Each month, starting in February 1988 and ending in December 2012, we run a cross-sectional regression of stock returns on DTC and a set of control variables known to predict returns, including market beta (Beta), the natural logarithm of the book-to-market ratio (LnBM), the natural logarithm of the market value of equity (LnME), and past returns for the prior month (Rev) and for the prior 12-month period excluding month t-1 (Mom).

Table 7 reports the time-series averages of the coefficients on the independent variables. Column (1) and (2) shows that both DTC and SR strongly and negatively predict subsequent stock returns in the cross-section when entering into the regression alone, even after controlling for other known predictors of returns. The statistical significance, however, is much larger for DTC (t=-9.15) than for SR (t=-5.71). To compare the explanatory power of the DTC and SR, we focus on t statistics. The average coefficient estimates in a Fama and MacBeth (1973) regression can be interpreted as monthly returns on long-short trading strategies that trade on that part of the variation in each regressor that is orthogonal to every other regressor. The t-values associated with the Fama and MacBeth slopes are, therefore, proportional to the Sharpe ratios of the self-financing strategies. They equal the annualized Sharpe ratios times $\sqrt{T}$, where T represents the number of years in the sample.

In column (3), we run a horse race between DTC and SR by including both in the

Fama-Macbeth regression. The coefficient on SR is cut by half and is significant only at 5% level, while the coefficient on DTC is largely unchanged and remains highly significant. To get a sense of the magnitude, the coefficient of -0.042 on SR implies that a one-standard deviation spread in SR generates a differential in expected returns of 0.16%. The coefficient of -0.0003 on DTC, however, implies that a one-standard deviation spread in DTC generates a differential in expected returns of 0.25%, which is 50% higher than SR.[15]

In column (4) and (5), we use residual short ratio (RSR) as the return predictor. A higher residual short ratio means arbitrageurs' willingness to short the stocks given its low liquidity. RSR1 is the residual from the cross-sectional regression of short ratio on turnover. We see that the return predictability of SR becomes stronger after adjusting for liquidity. The coefficient on RSR1 is -0.119 with a t-stat of -8.05. The Sharpe ratio implied by the t-statistic for RSR1 is 60% larger than the Sharpe ratio for the raw SR. RSR2 is the residual from the cross-sectional regression of short interest ratio on all trading cost proxies, including turnover, Amihud illiquidity, FHT measure, Pastor-Stambaugh (2003) liquidity measure and the daily percent quoted spread. The coefficient on RSR2 further improves to -0.130 (t=-8.74).

5.4 The Effect of DTC Controlling for Lending Fees

In Table 8, we consider how our results change when we add in lending fees as a control. We have three measures of lending fees. The first is Fee1, which again is the average of the fee observed in hedge fund borrowing transactions. The results are in columns (1)-(3). In columns (4)-(6), we add Fee2, Markit's estimate of the lending fee, as a covariate. In columns (7)-(9), we take short interest scaled by institutional ownership, SIO, to be a measure of the lending fee. The motivation for SIO is to proxy for the size of the lending fee by taking the ratio of demand for shorts to a proxy for the supply of shorts in the form of institutional ownership (Asquith, Pathak, and Ritter (2005), Nagel (2005), Drechsler and Song (2014)).

In columns (1)-(3), we see that DTC in this sample is marginally significant with a t-

[15]Note that our long-short DTC portfolios from the previous section make a much more extreme comparison than the one-standard deviation move considered here, thereby yielding more dramatic differences in portfolio returns (see Figure 3).

statistic of -1.71. But Fee1 is also not significant. Notice that Fee1 is a very short sample starting in November 2006. So the lack of statistical significance is not surprising. SR is insignificant in the Fee1 sample. Moreover, the coefficient on SR is -.03, which is still economically much smaller than before. In column (3), we find that DTC is a more significant predictor than SR. Indeed, in this specification, DTC attracts a coefficient of -.0003 and a t-statistic of -2.48. In other words, DTC is very robust to different sub-periods and controlling for Fee1.

In columns (4)-(6), we control for Fee2. In this larger sample, we see that DTC attracts a coefficient of -.0003 with a t-statistic of -2.86 in column (4). The coefficient on Fee2 is -.0025 with a t-statistic of -3.40. This is consistent with the literature that lending fees are a significant predictor of poor returns consistent with binding short-sales constraints and over-pricing. The economic and statistical significance of DTC is comparable to Fee2. In column (5), we see that the coefficient in SR is still small, at around -.0251 but now has a t-statistic of -1.56.

In columns (7)-(9), our effects are again robust to controlling for lending fees. In column (7), the coefficient on DTC is -.0002 with a t-statistic of -4.05. The coefficient on SIO is -.0142 with a t-statistic of -5.34. So DTC is comparable in economic significance to SIO in this sample. In column (8), we see that the coefficient on SR is insignificant. A similar conclusion holds from column (9).

5.5 Robustness

In Table 9, we consider a number of alternative explanations for the power of DTC to forecast returns. Our primarily concern is that the literature has found that lagged turnover measured at different horizons forecasts stock returns.

First, the DTC measure used in our previous analysis is short ratio (SR) scaled by the average of this month's daily turnover ratio. We show that our results are robust to the horizon length over which we average daily turnover. For DTC2, we average daily turnover in the previous month. For DTC3, we use the past 6-months of data to calculate daily turnover. For DTC4, we use the past one-year of daily observations to calculate average daily turnover.

Second, our DTC results could be driven by the high volume return premium as documented in Gervais, Kaniel, and Mingelgrin (2001). They find stocks experiencing abnormal recent increases in trading volume this week have high average returns in subsequent weeks. Stocks experiencing abnormal increases in trading volume would have low DTC measure by construction, and that could contribute to the strong return predictability of DTC.

In column (1) and (2), we run a horse race between DTC and DTC2 with 1/Turnover and 1/Turnover2, where Turnover and Turnover2 are the average daily turnover measured at the same horizon as the ones used to construct DTC and DTC2.[16] The coefficient on 1/Turnover and 1/Turnover2 indeed comes in with the expected sign. Low turnover stocks do worst the next month, consistent with Gervais, Kaniel, and Mingelgrin (2001). But our DTC effect and this low turnover effect are different because SR is highly correlated with turnover and so our DTC variable is not very correlated (just .09) with 1/Turnover.

Alternatively, we worry that our results might be related to Amihud (2002), where share turnover is in the denominator and is typically measured using longer horizons of data going back as far as one year. So we control for the Amihud (2002) illiquidity factor in column (3) and (4) and find that our results are largely unchanged. The coefficient on DTC3 is -.0004 and is statistically significant with a t-statistic of -7.20. The coefficient on the illiquidity measure is not strong in this sample. A similar conclusion holds for DTC4.

In column (5) and (6), we consider a more simplified version of Amihud (2002) which is the inverse of turnover, 1/Turnover3 and 1/Turnover4, from Lou and Shu (2014). Again, we find that DTC3 and DTC4 remains economically and statistically significant. The coefficient is virtually unchanged. In this specification, the inverse of turnover weakly forecast stock returns. Yet, the power of DTC3 and DTC4 remain the same.

In column (7), we show that DTC captures the marginal cost associated with trading as opposed to fundamental risks as captured say idiosyncratic volatility (IVOL) of Ang, Hodrick, Xing, and Zhang (2006). Another reason why we are interested in idiosyncratic volatility is that Stambaugh, Yu, and Yuan (2015) find that IVOL captures potential overpricing due to short-selling costs. We find again that the coefficient on DTC is unchanged

[16]We get the same results if we used Turnover and Turnover2 in the specification instead.

when adding in IVOL to our regression. To the extent DTC again is capturing over-valuation, it might be driven by disagreement and binding short-sales constraints effects as measured by Diether, Malloy, and Scherbina (2002)'s analyst forecast dispersion. In column (8), we show that DTC is not capturing the same effect as this over-valuation factor.

6 Evidence from Hedge Fund Holdings

Our model predicts that arbitrageurs will require extra compensation to enter into crowded positions also on the long-side. We simply need to modify the model by replacing the optimistic retail investors with pessimistic retail investors. The arbitrageurs would then want to take a long position. Assuming similarly that the cost of trading increases in the aggregate long positions of the arbitrageurs, we would obtain similar results on the long-side. So it is an empirical question the extent to which crowded trades on the long-side also matter.

6.1 Hedge Fund Holdings and Stock Liquidity

We first examine whether hedge fund's long-side position is also correlated with stock liquidity. To test this, we regress hedge fund holding (HFH) on various liquidity measures and a set of control variables including size, book-to-market, past 1-year return, institutional ownership and idiosyncratic volatility, similar to the short ratio regression. The results from this regression is reported in Table 10. In column (1), we include turnover as the only explanatory variable. The coefficient on turnover is 0.198 (t=13.75), suggesting that hedge fund's long-side position is also strongly associated with trading volume. However, both the coefficient magnitude and average R-square (4.4%) are much smaller compared to the regression when short ratio is the dependent variable. This means that while hedge funds are also concerned about stock's liquidity when taking long positions, the concern is much less compared to their short positions. When we include other control variables in column (2), the coefficient on turnover is reduced by half to 0.104, but nonetheless still highly significant. The results also hold when we use other proxies for stock liquidity, as reported in column

(3) to (6). Finally, hedge fund holding is strongly correlated with trading volume when we control for firm-fixed effect in panel regression in column (8).

6.2 Return Predictability of Turnover-adjusted Hedge Fund Holdings

Our model predicts that arbitrageurs require additional compensation on stocks with abnormally large arbitrage positions. We have shown that days-to-cover, which measures how crowdedness the stock is, predict return negatively. When this intuition is applied to arbitrageur's long-side positions, we expect stocks with abnormal large hedge fund holding relative to its level of liquidity should outperform other stocks in the future. Ideally, we would like to create a measure similar to days-to-cover and examine its return predictability. However, simply dividing hedge fund holding by average daily turnover has many problems. First, a large portion of stocks have zero hedge fund holding, especially during the early sample period when hedge fund assets are negligible. This will result in a large proportion of stocks with DTC of zero. Secondly, hedge fund holding is not as strongly correlated with turnover as we see in short interest ratio, so putting turnover in the denominator will lead to some stocks with extremely high DTC measures.

Instead, we use the residual hedge fund holding (RHFH) as measure of the our long-side version of DTC. The residual hedge fund holding is obtained by taking the residual from cross-sectional regression of quarterly hedge fund holding on average turnover at each quarter. High residual hedge fund holding therefore measures hedge funds' willingness to take large long positions despite its lower level of liquidity.

We first use portfolio sorts to examine the return predictability of residual hedge fund holding (RHFH). At the end of each quarter, we sort all the stocks into decile portfolios based on residual hedge fund holding (RHFH) and a long-short portfolio is formed by buying the highest RHFH portfolio and shorting the lowest RHFH portfolio. Portfolios are rebalanced at each quarter. We report the excess returns, characteristics-adjusted abnormal returns calculated following Daniel, Grinblatt, Titman and Wermers (1997), denoted as DGTW, four-factor alpha (following Carhart (1997)) and five-factor adjusted alpha (Carhart 4-factor augmented by Pastor and Stambaugh's (2003) liquidity factor) in table 11.

As we can see, the long-short portfolio based on residual hedge fund holding generates significant positive returns for both equal- and value-weighted portfolios. The average equal-weighted monthly return to this long-short portfolio is 0.65% with a t-stat of 3.52. The Sharpe ratio is 0.77. A DGTW adjustment leads to a lower return of 0.46% but still significantly different from zero. The four-factor and five-factor alphas are around 0.70% and highly significant. In panel B, we see the value-weighted results are still significant in this case. For excess return, it is 0.74% with a t-stat of 3.07. For DGTW adjusted returns, it is 0.58% with a t-stat of 2.72. The alphas from four-factor and five-factor model are also significantly positive at around .6%.

We also use the Fama-Macbeth (1973) regression to examine the return predictability of residual hedge fund holding. This approach allows us to control for the usual firm characteristics. The result is reported in Table 12. We first examine whether the raw hedge fund holding measure could predict next quarter's stock returns in column (1). The coefficient on hedge fund holding (HFH) is 0.013 but not significant (t=1.47). This is consistent with the existing evidence on the weak return predictability of hedge funds' holdings[17]. In column (2), we use the residual hedge fund holding to predict returns. In contrast to the previous result, the coefficient on residual hedge fund holding (RHFH) is 0.024 and significant at 1% level. The strong return predictability of residual hedge fund holding is in sharp contrast to the raw hedge fund holding measure and shows hedge fund's positions are informative when the position is large relative to stock liquidity. In column (3) and (4), we add days-to-cover (DTC) in the multiple regression. The coefficient on residual hedge fund holding is still significantly positive, and the coefficient on DTC is significantly negative. This result demonstrates that arbitrageurs are also compensated for entering crowded positions on the long side.[18]

[17]Griffin and Xu (2009) document that hedge fund ownership does not predict next quarter stock returns once they control for momentum.

[18]In untabulated results, we also adjust the institutional ownership for stock turnover, and we do not find the evidence that residual institutional ownership adjusted for liquidity predict next quarter's returns. This is consistent with the vast majority of institutions not being levered.

7 Conclusion

The recent Dodd Frank Financial Reforms require increased disclosure by institutional investors. These reforms are in part motivated by worries of a crowded trade problem, whereby aggregate speculators' positions are large relative to the liquidity of the asset. The exit from crowded trades can be destabilizing as there is little liquidity on the other side of the trade. These reforms are consonant with similar worries on the part of practitioners as the sophisticated investors are becoming an increasingly large part of the market.

We study this problem in the context of short-sellers. We develop a simple model to analyze days to cover (DTC), a widely used statistic by short-sellers to monitor crowded trades. We find that arbitrageurs are worried about the crowding problem as they systematically avoid illiquid stocks, all else equal, and require a significant premium to enter into crowded positions. We also show that there is a crowded trade problem on the long-side.

8 Appendix

8.1 Appendix A

Endogenizing the lending fee f

Now suppose that in addition to arbitrageurs and optimists there is a third type of traders, index funds, with portfolio choices that are insensitive to prices or payoff forecasts. We suppose that on aggregate these new type of agents hold ν shares for each agent of the other two types. Since we normalize the number of arbitrageurs plus optimists to 1, index funds hold a total of ν shares, and the total number of shares is $\nu + 1 - \gamma$.

Index funds are the only suppliers of borrowed shares. We assume that the market clearing fee is given by

$$f = f_0 + f_1 x, \tag{34}$$

where x is the amount of shares borrowed, provided $x < \nu$.

For simplicity, we set $A = 1$ and $c = c_o$. As before the demand by each arbitrageurs is given by

$$n_a = \frac{\mu_a + f - p_0}{c} < 0. \tag{35}$$

Hence the total demand for shorting by arbitrageurs is:

$$\gamma |n_a| = \gamma \frac{p_0 - \mu_a - f}{c}$$

or using equation (34), we get

$$\gamma |n_a| = \frac{\gamma (p_0 - \mu_a - f_0)}{\gamma f_1 + c}$$

Notice that this coincides with results on the text whenever $f_1 = 0$

Since index funds do not adjust their portfolio, the market clearing price for a given f stays the same that is:

$$p_0 = (1-\gamma)\mu_o + \gamma \mu_a + \gamma f = \mu_a + (1-\gamma)(\mu_o - \mu_a) + \gamma f. \tag{36}$$

f is now endogenous and in equilibrium

$$p_0 - \mu_a = \frac{\gamma f_1 + c}{\gamma f_1 + c - \gamma^2 f_1} \left[(1-\gamma)(\mu_o - \mu_a) + \gamma f_0 - \frac{\gamma^2 f_0 f_1}{\gamma f_1 + c} \right] \tag{37}$$

However for γ small (*i.e.* $\gamma^2 \sim 0$), $p_0 - \mu_a$ is invariant to c. More precisely,

$$p_0 - \mu_a - (1-\gamma)(\mu_o - \mu_a) - \gamma f_0 = \frac{\gamma^2 f_1}{\gamma f_1 + c - \gamma^2 f_1} [(1-\gamma)(\mu_o - \mu_a) + (\gamma - 1) f_0]$$

Hence,

$$p_0 - \mu_a = (1-\gamma)(\mu_o - \mu_a) + \gamma f_0 + \gamma^2 f_1 \Gamma(c, f_0, f_1, \gamma),$$

with Γ bounded.

On the other hand, short interest satisfies:

$$SR = \frac{\gamma|n_a|}{1-\gamma+\nu} = \frac{\gamma(p_0-\mu_a-f_0)}{(\gamma f_1+c)(1-\gamma+\nu)} = \frac{(1-\gamma)\gamma(\mu_o-\mu_a-f_0)+\gamma^3 f_1\Gamma(c,f_0,f_1,\gamma)}{(\gamma f_1+c)(1-\gamma+\nu)}$$

Thus if γ^2 is small enough, SR decrease with c.

Furthermore, since f does not enter the optimization problem of the optimists, we still obtain that turnover V equals

$$\frac{1-\gamma}{2(1-\gamma+\nu)} \times \frac{(\delta\mu_o - p_o)}{c}$$

Thus

$$DTC = \frac{2c[\gamma(\mu_o-\mu_a-f_0)+\gamma^3 f\Gamma]}{(\gamma f_1+c)[(\delta-1)\mu_o+\gamma(\mu_o-\mu_a-f_0)-\gamma^2 f_1\Gamma]}$$

Notice that for any fixed $c > 0$ $\frac{\partial\Gamma}{\partial\gamma}$ is bounded above.

Thus taking a first order expansion of DTC around $\gamma = 0$ we obtain, as before,

$$DTC \sim 2\frac{\gamma(\mu_o-\mu_a-f_0)}{(\delta-1)\mu_o}$$

8.2 Appendix B

Construction of liquidity measures:

1. $Amihud = Average(\frac{|r_t|}{Volume_t})$, where r_t is the stock return on day t and $Volume_t$ is the dollar trading volume on day t. The higher the Amihud illiqudity measure, the more illiquid the stock is.

2. $Daily\ Percent\ Quoted\ Spread = Average(\frac{ClosingAsk_t - ClosingBid_t}{(ClosingAsk_t + ClosingBid_t)/2})$. A higher quoted percent spread means higher trading cost.

3. The Pastor and Stambaugh liquidity measure is the coefficient Γ from the regression: $r^e_{t+1} = \theta + \phi r_t + \Gamma sign(r^e_t)(Volume_t) + \xi_t$, where r^e_t is the stock's excess return above the CRSP value-weighted market return on day t, θ is the intercept, ϕ and Γ are regression coefficients, and ξ_t is the error term. The more negative the coefficient Γ, the less liquidity the stock is.

4. The FHT measure is the trading cost backed out from the frequency of zero return trading days: $FHT = 2\sigma N^{-1}(\frac{1+z}{2})$, where z is the empirical observed frequency of zero returns, σ is the stock return volatility and $N^{-1}()$ is the inverse function of the cumulative normal distribution. A higher FHT measures larger trading cost.

References

Ali, A., L.-S. Hwang, and M. A. Trombley, 2003, "Arbitrage risk and the book-to-market anomaly," *Journal of Financial Economics*, 69(2), 355–373.

Amihud, Y., 2002, "Illiquidity and stock returns: cross-section and time-series effects," *Journal of financial markets*, 5(1), 31–56.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang, 2006, "The cross-section of volatility and expected returns," *The Journal of Finance*, 61(1), 259–299.

Asquith, P., P. A. Pathak, and J. R. Ritter, 2005, "Short interest, institutional ownership, and stock returns," *Journal of Financial Economics*, 78(2), 243–276.

Barber, B. M., and T. Odean, 2000, "Trading is hazardous to your wealth: The common stock investment performance of individual investors," *Journal of Finance*, pp. 773–806.

Beneish, M. D., C. Lee, and C. Nichols, 2013, "In short supply: Equity overvaluation and short selling," *Rock Center for Corporate Governance at Stanford University Working Paper*, (165).

Bessembinder, H., 2003, "Trade execution costs and market quality after decimalization," *Journal of Financial and Quantitative Analysis*, 38(04), 747–777.

Boehmer, E., Z. R. Huszar, and B. D. Jordan, 2010, "The good news in short interest," *Journal of Financial Economics*, 96(1), 80–97.

Carhart, M. M., 1997, "On persistence in mutual fund performance," *The Journal of finance*, 52(1), 57–82.

Chung, K. H., and H. Zhang, 2014, "A simple approximation of intraday spreads using daily data," *Journal of Financial Markets*, 17, 94–120.

Cohen, L., K. B. Diether, and C. J. Malloy, 2007, "Supply and demand shifts in the shorting market," *The Journal of Finance*, 62(5), 2061–2096.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers, 1997, "Measuring mutual fund performance with characteristic-based benchmarks," *The Journal of finance*, 52(3), 1035–1058.

D'avolio, G., 2002, "The market for borrowing stock," *Journal of Financial Economics*, 66(2), 271–306.

Dechow, P. M., A. P. Hutton, L. Meulbroek, and R. G. Sloan, 2001, "Short-sellers, fundamental analysis, and stock returns," *Journal of Financial Economics*, 61(1), 77–106.

Diether, K. B., C. J. Malloy, and A. Scherbina, 2002, "Differences of opinion and the cross section of stock returns," *The Journal of Finance*, 57(5), 2113–2141.

Drechsler, I., and Q. F. Song, 2014, "The Shorting Premium and Asset Pricing Anomalies," *Available at SSRN.*

Fama, E. F., and K. R. French, 1992, "The cross-section of expected stock returns," *the Journal of Finance*, 47(2), 427–465.

———, 1993, "Common risk factors in the returns on stocks and bonds," *Journal of financial economics*, 33(1), 3–56.

Fama, E. F., and J. D. MacBeth, 1973, "Risk, return, and equilibrium: Empirical tests," *The Journal of Political Economy*, pp. 607–636.

Figlewski, S., 1981, "The informational effects of restrictions on short sales: some empirical evidence," *Journal of Financial and Quantitative Analysis*, 16(04), 463–476.

Fong, K. Y., C. W. Holden, and C. Trzcinka, 2014, "What are the best liquidity proxies for global research?," *Available at SSRN 1558447.*

French, K. R., 2008, "Presidential address: The cost of active investing," *The Journal of Finance*, 63(4), 1537–1573.

Furfine, C., 2003, "Decimalization and market liquidity," *Economic Perspectives*, 27(4), 2.

Gao, X., and J. R. Ritter, 2010, "The marketing of seasoned equity offerings," *Journal of Financial Economics*, 97(1), 33–52.

Gervais, S., R. Kaniel, and D. H. Mingelgrin, 2001, "The high-volume return premium," *The Journal of Finance*, 56(3), 877–919.

Griffin, J. M., and J. Xu, 2009, "How smart are the smart guys? A unique view from hedge fund stock holdings," *Review of Financial Studies*, 22(7), 2531–2570.

Grossman, S. J., and M. H. Miller, 1988, "Liquidity and market structure," *the Journal of Finance*, 43(3), 617–633.

Hanson, S. G., and A. Sunderam, 2014, "The growth and limits of arbitrage: Evidence from short interest," *Review of Financial Studies*, 27(4), 1238–1286.

Hirshleifer, D., S. H. Teoh, and J. J. Yu, 2011, "Short arbitrage, return asymmetry, and the accrual anomaly," *Review of Financial Studies*, 24(7), 2429–2461.

Hong, H., J. D. Kubik, and T. Fishman, 2012, "Do arbitrageurs amplify economic shocks?," *Journal of Financial Economics*, 103(3), 454–470.

Jegadeesh, N., and S. Titman, 1993, "Returns to buying winners and selling losers: Implications for stock market efficiency," *The Journal of Finance*, 48(1), 65–91.

Jiang, W., 2014, "Leveraged Speculators and Asset Prices," *Available at SSRN 2525986*.

Jones, C. M., and O. A. Lamont, 2002, "Short-sale constraints and stock returns," *Journal of Financial Economics*, 66(2), 207–239.

Lou, D., and C. Polk, 2013, *Comomentum: Inferring arbitrage activity from return correlations*. Paul Woolley Centre for the Study of Capital Market Dysfunctionality; Financial Markets Group.

Lou, X., and T. Shu, 2014, "Price Impact or Trading Volume: Why is the Amihud (2002) Illiquidity Measure Priced?," *Available at SSRN 2291942*.

Nagel, S., 2005, "Short sales, institutional investors and the cross-section of stock returns," *Journal of Financial Economics*, 78(2), 277–309.

Newey, W. K., K. D. West, et al., 1987, "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3), 703–08.

Pastor, L., and R. F. Stambaugh, 2003, "LIQUIDITY RISK AND EXPECTED STOCK RETURNS," *Journal of Political Economy*, 111(3).

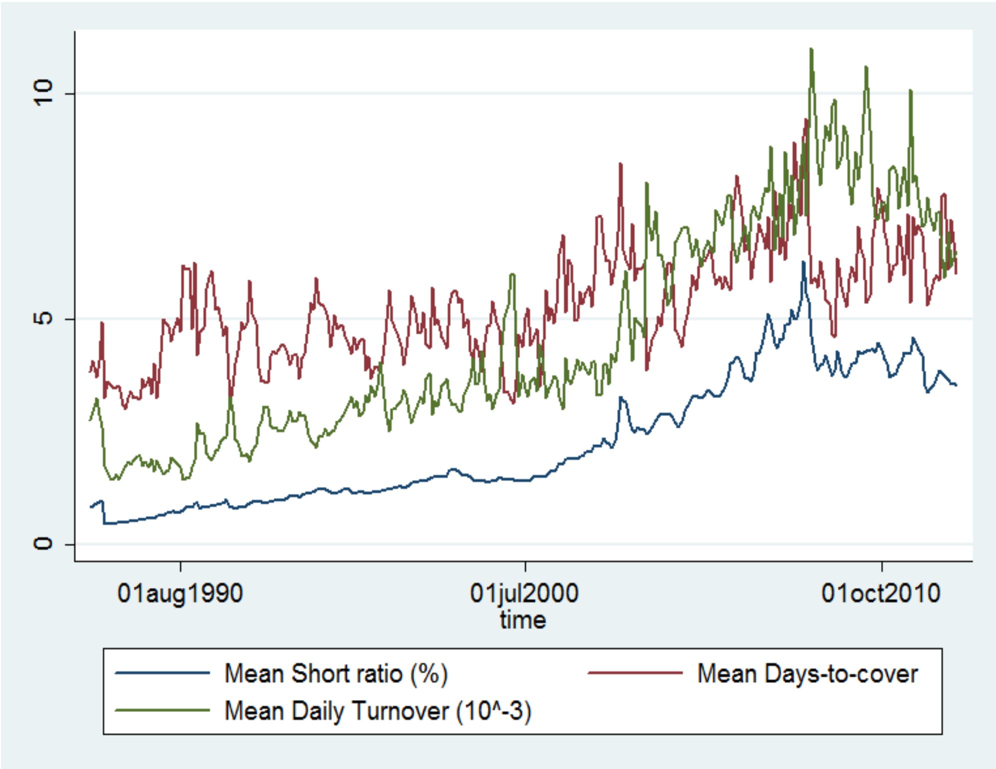
Sadka, R., 2006, "Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk," *Journal of Financial Economics*, 80(2), 309–349.

Stambaugh, R. F., J. Yu, and Y. Yuan, 2015, "Arbitrage asymmetry and the idiosyncratic volatility puzzle," *The Journal of Finance*.

Stein, J. C., 2009, "Presidential address: Sophisticated investors and market efficiency," *The Journal of Finance*, 64(4), 1517–1548.

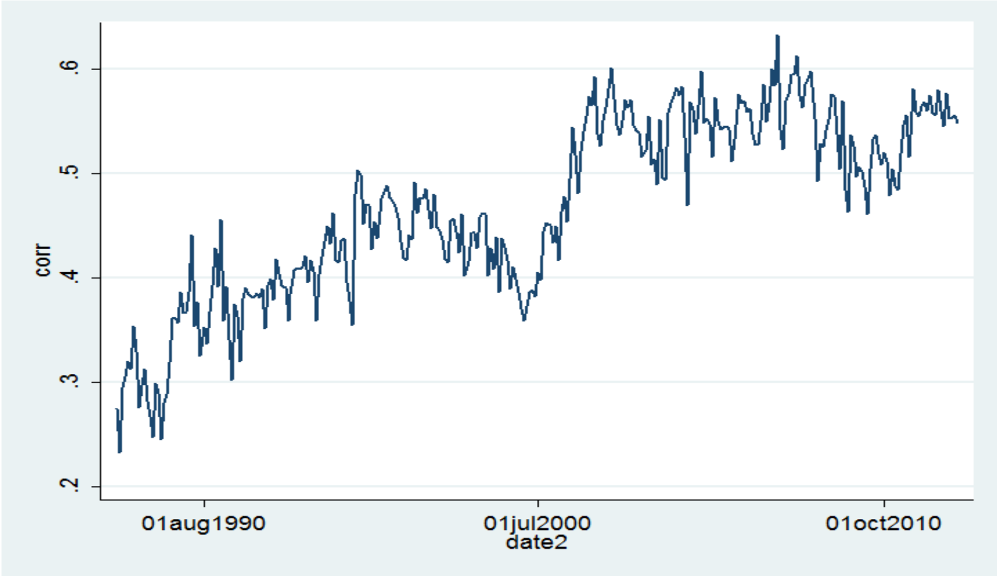
Vayanos, D., 1998, "Transaction costs and asset prices: A dynamic equilibrium model," *Review of financial studies*, 11(1), 1–58.

Figure 1: **Time Series of Mean SR and DTC.**

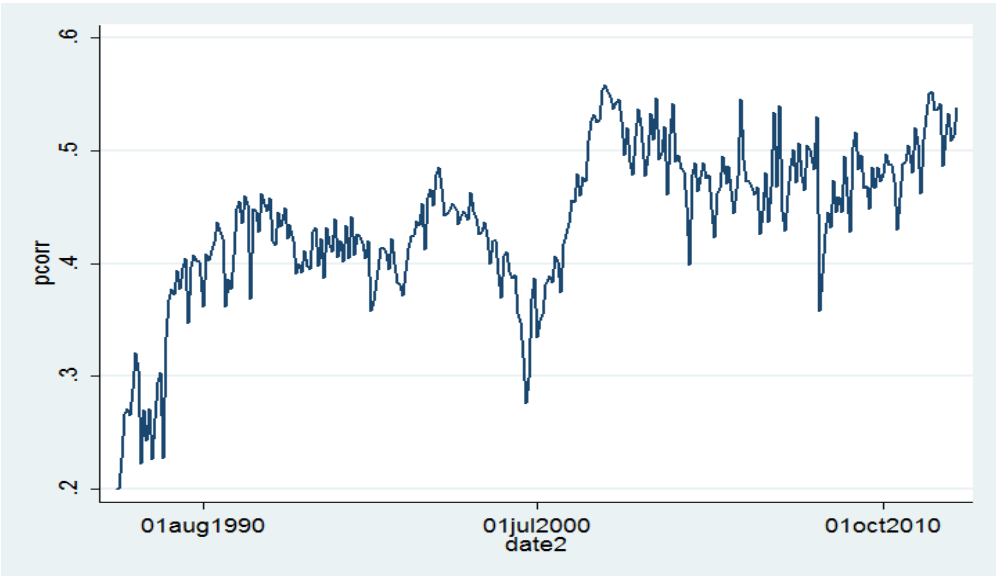


This figure plots the time-series of average short interest (SR) ratio (in percentage), days-to-cover (DTC) and daily turnover (scaled by 1000) from January 1988 to December 2012 based on all NYSE, Amex, and NASDAQ stocks.

Figure 2: **Correlations between Short ratio (SR) and Turnover, 1988:01 to 2012:12**



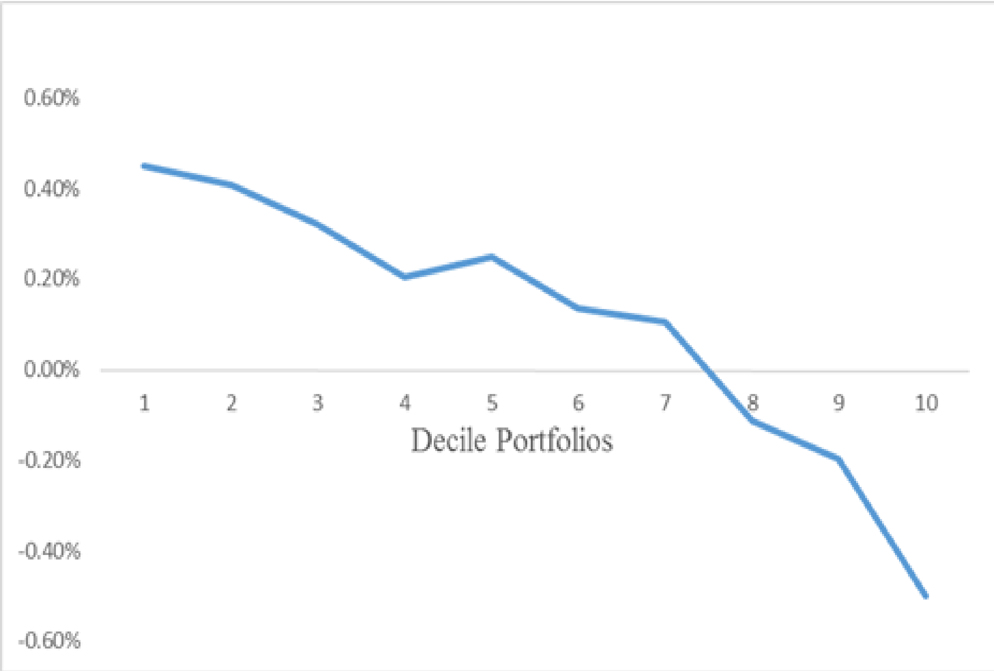
(a) Correlation of SR and Turnover



(b) Partial Correlation of SR and Turnover

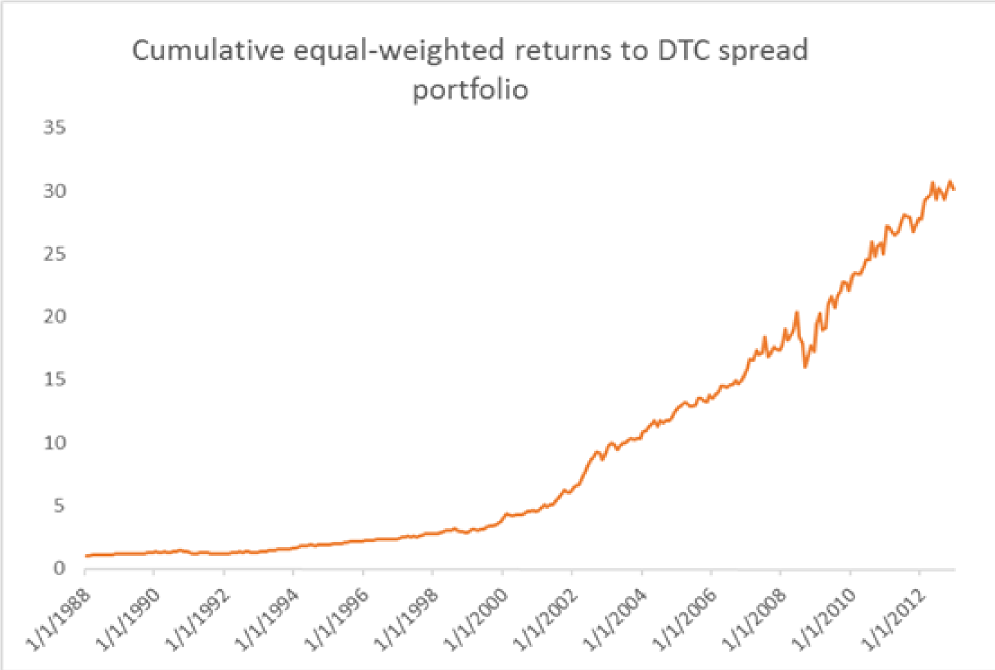
This figure plots the cross-sectional correlation (Panel A) and partial correlation (Panel B) between short interest ratio (SR) and share turnover. The univariate correlation coefficient is computed in cross section every month and is plotted over time. The partial correlation between short interest ratio (SR) and share turnover is computed after controlling for size, book-to-market, past 12 months cumulative returns and institutional ownership. The sample period runs from January of 1988 to December of 2012.

Figure 3: **Decile Portfolio Performance**

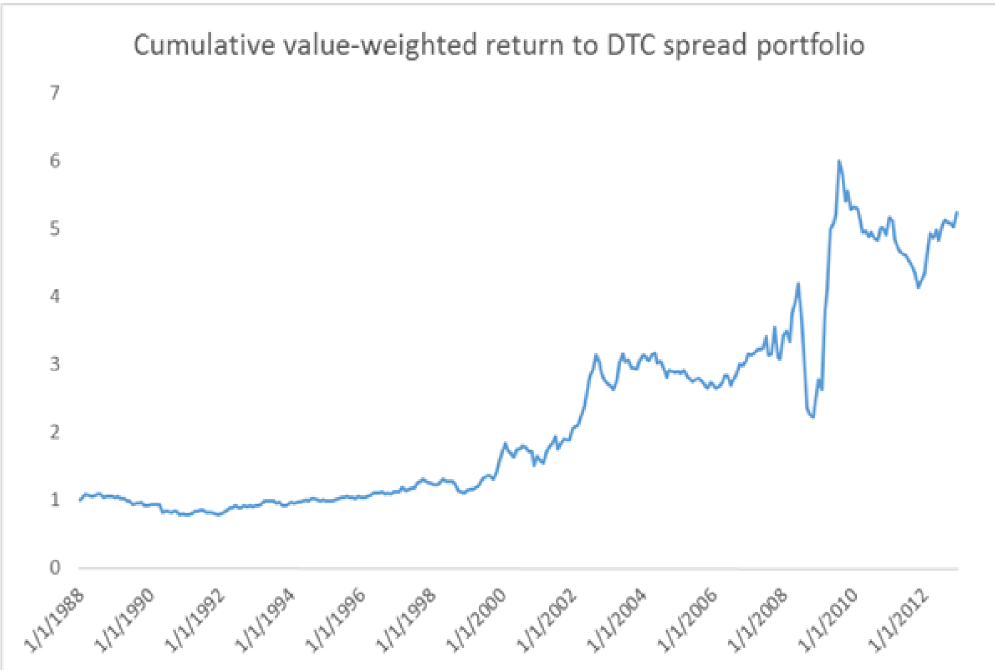


This figure shows the average DGTW-adjusted returns for decile portfolios sorted on days-to-cover (DTC). Returns are equally weighted within each portfolio. The y-axis is monthly returns and x-axis is the decile portfolio from low DTC to high DTC.

Figure 4: **Cumulative Returns to DTC Spread Portfolio**



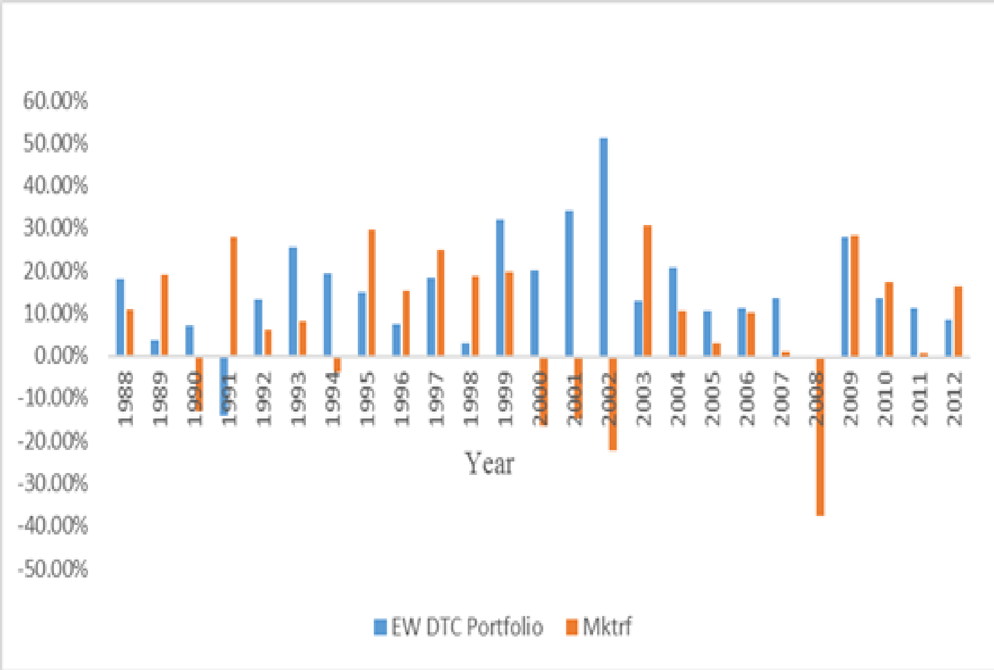
(a) Cumulative return (Equal weighted)



(b) Cumulative return (Value weighted)

This figure plots the cumulative equal-weighted returns (Panel A) and value-weighted returns (Panel B) to the bottom-minus-top decile portfolio formed on days-to-cover (DTC). The sample period runs from January of 1988 to December of 2012.

Figure 5: **Annual Returns to Equal-Weighted DTC Spread Portfolio.**



This figure shows the annual equal-weighted returns to the bottom-minus-top decile portfolio formed on days-to-cover (DTC) and market excess return. The sample period runs from January of 1988 to December of 2012.

Table 1: **Descriptive Statistics**

This table presents the summary statistics of the sample, including the mean and standard deviation for each variable, and the correlations among them. Days-to-cover (DTC) is short interest ratio (SR) over daily turnover. Short interest ratio (SR) is the shares shorted over total shares outstanding. Share turnover (Turnover) is the daily trading volume over total shares outstanding averaged within a given month. Hedge fund holdings (HFH) is the sum of shares held by all hedge funds at the end of each quarter divided by the total number of shares outstanding. If the stock is not held by even a single hedge fund in that quarter, its HFH is set to zero. Institutional ownership (IO) is the sum of shares held by institutions from 13F filings in each quarter divided by shares outstanding. Market beta (Beta) is calculated from past five years' monthly return, following Fama and French (1992). Size (LnME) is the natural log of firm market capitalization at the end of the June of each year. Book-to-market (LnBM) is the natural log of book-to-market ratio. The cases with negative book value are deleted. Momentum (MOM) is defined as the cumulative return from month t-12 to t-2. The short term reversal measure (REV) is the lagged monthly return. IVOL is the idiosyncratic volatility, calculated following Ang et al. (2006). Dispersion (DISP) is the analysts' earnings forecast dispersion measure, following Diether, Malloy and Scherbina (2002). Panel A reports the summary statistics for the full sample and by NYSE size quintile. Panel B reports these statistics for our sub-sample where we also have the lending fee data. Panel C reports the pairwise correlations among our variables where they overlap. Fee1 is the simple average fees of stock borrowing transactions from hedge funds in a given security, which is the difference between the risk-free rate and the rebate rate. Fee2 is a score from 1 to 10 created by Markit using their proprietary information meant to capture the cost of borrowing the stock. Here 1 is the cheapest to short and 10 the most difficult. Fee1 is available since November of 2006 and fee2 is available since October of 2003. The overall sample period runs from January of 1988 to December 2012 except for hedge fund holdings when the data starts from 1992.

Panel A: Summary Statistics-Full Sample

Variables		**All Firms**	**Quintile 1**	**Quintile 2**	**Quintile 3**	**Quintile 4**	**Quintile 5**
Days to Cover (DTC)	Mean	5.45	5.08	7.16	6.27	5.32	3.94
	Std.	8.26	9.03	9.06	7.24	5.99	3.92
Short Ratio (SR)	Mean	2.26%	1.59%	3.51%	3.48%	2.92%	1.87%
	Std.	3.65%	3.32%	4.59%	4.16%	3.23%	1.97%
Turnover	Mean	0.46%	0.32%	0.55%	0.66%	0.71%	0.60%
	Std.	0.61%	0.51%	0.63%	0.71%	0.72%	0.59%
Hedge Fund Holdings (HFH)	Mean	3.32%	3.39%	3.50%	3.62%	3.54%	2.06%
	Std.	4.72%	5.41%	4.52%	4.25%	3.89%	2.37%
Institutional Ownership (IO)	Mean	42.04%	27.36%	52.85%	59.52%	63.43%	63.02%
	Std.	28.99%	24.85%	26.29%	24.56%	21.94%	17.70%
Market Beta (Beta)	Mean	1.30	1.40	1.32	1.22	1.14	1.00
	Std.	1.06	1.21	0.96	0.85	0.74	0.66
Size (LnME)	Mean	5.51	3.96	5.97	6.88	7.80	9.39
	Std.	2.14	1.25	0.62	0.56	0.55	1.00
Book-to-market (LnBM)	Mean	-0.61	-0.44	-0.70	-0.78	-0.85	-1.01
	Std.	0.89	0.94	0.78	0.75	0.75	0.80
Reversal (REV)	Mean	1.21%	1.22%	1.22%	1.25%	1.22%	1.07%
	Std.	17.47%	20.80%	14.41%	12.99%	11.54%	9.86%
Momentum (MOM)	Mean	13.92%	11.21%	17.56%	18.12%	16.84%	15.04%
	Std.	73.22%	84.79%	63.57%	62.10%	53.18%	41.35%
Amihud Illiquidity (Amihud)	Mean	1.11	2.51	0.16	0.02	0.01	0.00
	Std.	7.18	10.54	0.47	0.09	0.04	0.00
FHT	Mean	0.60%	1.00%	0.52%	0.33%	0.22%	0.14%
	Std.	0.70%	0.86%	0.49%	0.35%	0.25%	0.17%
Pastor-Stambaugh (PS)	Mean	2.69E-06	6.56E-06	3.04E-07	-1.65E-07	-8.96E-08	-1.25E-08
	Std.	0.02%	0.03%	0.00%	0.00%	0.00%	0.00%
Daily percent quoted spread (QS)	Mean	1.92%	2.81%	1.41%	0.99%	0.77%	0.56%
	Std.	1.68%	1.91%	0.80%	0.59%	0.46%	0.30%
Idiosyncratic Volatility (IVOL)	Mean	2.59%	3.25%	2.50%	2.14%	1.86%	1.59%
	Std.	1.56%	1.72%	1.34%	1.19%	1.01%	0.79%
Analyst Forecast Dispersion (DISP)	Mean	0.18	0.31	0.19	0.14	0.12	0.08
	Std.	1.01	1.31	0.83	0.63	0.63	0.33
# of Obs.		906377	483318	142346	102408	90635	87670

Table 1 Continued

Panel B: Summary Statistics - Sample with Lending Fee data							
Variables		**All Firms**	**Quintile 1**	**Quintile 2**	**Quintile 3**	**Quintile 4**	**Quintile 5**
Days to Cover (DTC)	Mean	6.85	7.59	9.03	6.63	4.46	2.98
	Std.	7.37	8.75	6.61	5.20	3.73	2.28
Short Ratio (SR)	Mean	4.32%	3.47%	7.06%	6.10%	4.36%	2.46%
	Std.	4.71%	4.60%	5.27%	4.87%	3.72%	2.17%
Turnover	Mean	0.80%	0.54%	0.99%	1.15%	1.19%	0.96%
	Std.	0.82%	0.70%	0.83%	0.92%	0.90%	0.72%
Hedge Fund Holding (HFH)	Mean	5.41%	5.66%	5.59%	5.84%	5.91%	3.36%
	Std.	6.51%	7.31%	6.19%	6.10%	5.95%	3.60%
Institutional Ownership (IO)	Mean	58.71%	43.90%	72.73%	75.57%	76.65%	71.83%
	Std.	28.34%	27.51%	22.83%	20.92%	17.64%	14.37%
Market Beta (Beta)	Mean	1.31	1.44	1.31	1.24	1.16	0.97
	Std.	0.99	1.10	0.96	0.87	0.79	0.67
Size (LnME)	Mean	6.38	4.85	6.67	7.50	8.33	9.93
	Std.	1.95	1.01	0.34	0.30	0.35	0.86
Book-to-market (LnBM)	Mean	-0.64	-0.48	-0.69	-0.75	-0.85	-0.99
	Std.	0.83	0.87	0.73	0.73	0.73	0.75
Reversal (REV)	Mean	1.10%	1.09%	1.27%	1.20%	1.06%	0.82%
	Std.	14.04%	16.42%	12.64%	11.25%	10.32%	8.96%
Momentum (MOM)	Mean	12.75%	10.39%	15.66%	16.31%	16.36%	12.19%
	Std.	66.89%	82.48%	52.39%	47.27%	43.00%	32.66%
Amihud Illiquidity (Amihud)	Mean	0.97	2.15	0.01	0.00	0.00	0.00
	Std.	11.80	17.50	0.03	0.00	0.00	0.00
FHT	Mean	0.09%	0.15%	0.06%	0.04%	0.03%	0.02%
	Std.	0.21%	0.27%	0.14%	0.11%	0.08%	0.06%
Pastor-Stambaugh (PS)	Mean	-9.00E-07	-1.71E-06	-8.94E-08	-7.39E-08	-4.10E-08	-5.03E-09
	Std.	0.01%	0.02%	0.00%	0.00%	0.00%	0.00%
Daily percent quoted spread (QS)	Mean	0.45%	0.86%	0.16%	0.11%	0.09%	0.06%
	Std.	0.84%	1.11%	0.11%	0.09%	0.06%	0.04%
Idiosyncratic Volatility (IVOL)	Mean	2.15%	2.67%	2.03%	1.77%	1.61%	1.38%
	Std.	1.28%	1.42%	1.09%	1.00%	0.85%	0.70%
Analyst Forecast Dispersion (DISP)	Mean	0.16	0.27	0.16	0.11	0.11	0.06
	Std.	0.84	1.07	0.74	0.51	0.54	0.28
Fee1 (basis point)	Mean	48.10	73.04	48.20	39.96	33.11	28.74
	Std.	91.44	128.92	89.10	75.83	53.89	34.29
Fee2 (score)	Mean	1.39	1.63	1.20	1.12	1.07	1.13
	Std.	3.56	1.42	0.85	0.64	0.43	0.17
# of Obs.		279891	142429	43499	32639	30392	30932

Table 1 Continued

Panel C: Correlations

	SR	Turnover	DTC	IO	Beta	LnME	LnBM	Rev	Mom	Fee1	Fee2
SR	1.00										
Turnover	0.69	1.00									
DTC	0.83	0.21	1.00								
IO	0.57	0.59	0.33	1.00							
Beta	0.13	0.22	0.01	0.02	1.00						
LnME	0.53	0.49	0.36	0.71	-0.11	1.00					
LnBM	-0.29	-0.27	-0.19	-0.17	-0.11	-0.30	1.00				
Rev	0.01	0.07	-0.02	0.04	-0.03	0.05	0.02	1.00			
Mom	0.02	0.09	-0.04	0.11	-0.05	0.14	0.03	0.03	1.00		
Fee1	0.18	0.01	0.17	-0.19	0.10	-0.27	-0.02	-0.07	-0.06	1.00	
Fee2	-0.03	-0.12	0.04	-0.39	0.08	-0.34	-0.03	-0.07	-0.18	0.42	1.00

Table 2: **Regression of Short Interest Ratio (SR) on Trading Cost Measures**

This table reports results from regression of monthly short interest ratio (SR) on various trading costs measures. Turnover is the monthly average of daily turnover ratio. Amihud is the Amihud (2002) illiquidity measure. FHT is the transaction costs estimated from the frequency of zero returns (Fong, Holden and Trzcinka 2014). PS is the Pastor and Stambugh (2003) liquidity measure. QS is the percentage quoted spread using daily close price (Chung and Zhang 2014). Size (LnME) is the natural log of firm market capitalization at the end of the June of each year. Book-to-market (LnBM) is the natural log of book-to-market ratio. The cases with negative book value are deleted. Momentum (MOM) is defined as the cumulative return from month t-12 to t-2. IO is the institutional ownership ratio. IVOL is the idiosyncratic volatility following Ang et al. (2005). We use Fama-Macbeth regression method in column (1) to (7). In column (8), we use panel regression controlling firm-fixed effect. All variables are standardized to have mean 0 and standard deviation of 1. For Fama-Macbeth regression, all the t-statistics are Newey and West (1987) adjusted to control for heteroskedasticity and autocorrelation. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Turnover	0.4676***	0.4512***					0.4436***	0.3150***
	(51.59)	(59.28)					(67.29)	(301.00)
Amihud			-0.0511***				0.0287***	0.0123***
			(-9.19)				(11.74)	(14.87)
FHT				-0.0966***			-0.0309***	-0.0182***
				(-27.24)			(-10.37)	(-17.44)
PS					-0.0036***		0.0007	0.0002
					(-3.26)		(0.68)	(0.31)
QS						-0.2726***	-0.0826***	-0.0088***
						(-24.27)	(-8.22)	(-6.13)
LnME		-0.0647***	0.0135	0.0023	0.0156	-0.0456***	-0.0798***	0.0315***
		(-4.57)	(1.40)	(0.24)	(1.62)	(-3.52)	(-4.19)	(12.18)
LnBM		-0.0984***	-0.1443***	-0.1445***	-0.1468***	-0.1416***	-0.1077***	-0.0658***
		(-28.34)	(-31.35)	(-33.20)	(-31.71)	(-23.13)	(-21.96)	(-55.43)
Mom		-0.0647***	0.0132**	0.0054	0.0145**	-0.0112*	-0.0738***	-0.0603***
		(-13.31)	(2.32)	(0.93)	(2.54)	(-1.87)	(-12.45)	(-78.85)
IO		0.1503***	0.3090***	0.3013***	0.3100***	0.2820***	0.1462***	0.3171***
		(10.03)	(17.96)	(17.74)	(17.89)	(20.88)	(11.32)	(181.23)
IVOL		-0.0877***	0.1017***	0.1309***	0.0925***	0.2108***	-0.0525***	-0.0478***
		(-17.43)	(21.68)	(27.01)	(20.01)	(25.17)	(-8.46)	(-39.10)
Constant	0.0007	-0.0029*	-0.0034*	-0.0037**	-0.0031*	-0.0095***	-0.0004	-0.0038***
	(1.33)	(-1.84)	(-1.96)	(-2.21)	(-1.84)	(-3.23)	(-0.16)	(-5.54)
Method	FM	FM	FM	FM	FM	FM	FM	firm-fixed
Ave.R-sq	0.224	0.282	0.145	0.149	0.143	0.188	0.309	0.244
N.of Obs.	1249818	1046916	1046916	1045834	1046916	938855	938847	968041

Table 3: **Instrumental Variable Regression Using 2001 Shift to Decimalization**

This table reports instrumental variable regression using 2001 Decimalization as an exogenous shock to liquidity. The left column reports first-stage regression results and the right column reports second-stage regression results. Nyseamex is a dummy equal to 1 for stocks listed in Nyse/Amex exchanges and 0 for Nasdaq-listed stocks. Post is a dummy equal to 1 for the period of February and March of 2001 and 0 for the period from March 2000 to January 2001. Size (LnME) is the natural log of firm market capitalization at the end of the June of each year. Book-to-market (LnBM) is the natural log of book-to-market ratio. The cases with negative book value are deleted. Momentum (MOM) is defined as the cumulative return from month t-12 to t-2. IO is the institutional ownership ratio. IVOL is the idiosyncratic volatility following Ang et al. (2005). Control variables are Standardized to have mean 0 and standard deviation of 1. . ***, **, and * stands for significance level of 1%, 5% and 10%, respectively. The sample period runs from March 2000 to March 2001.

	Fisrt Stage		Second Stage
Nyseamex	-0.0822***	Turnover	1.0739***
	(-10.04)		(8.76)
Post	0.0011	LnME	-0.1595***
	(0.10)		(-5.03)
Nyseamex*Post	0.0744***	LnBM	-0.0425***
	(4.01)		(-2.66)
LnME	0.2670***	Mom	-0.2636***
	(53.28)		(-9.32)
LnBM	-0.1152***	IO	-0.0787**
	(-30.13)		(-2.32)
Mom	0.2237***	IVOL	-0.3410***
	(68.61)		(-6.55)
IO	0.2765***	Constant	-0.0181***
	(64.77)		(-4.38)
IVOL	0.4177***	Ave.R-sq	0.14
	(109.37)	N.of Obs.	60272
Constant	0.0261***		
	(5.75)		
Ave.R-sq	0.384		
N.of Obs.	60272		

Table 4: **Returns to portfolio strategies based on Days-to-Cover (DTC)**

This table provides portfolio returns and alphas, sorted on Days-to-Cover (DTC). At the end of each month, all the stocks are sorted into deciles based on days-to-cover and a long-short portfolio is formed by buying the lowest decile and shorting the highest decile portfolio. Portfolio returns are computed over the next month. We report the excess returns, characteristics-adjusted abnormal returns calculated following Daniel, Grinblatt, Titman and Wermers (1997), denoted as DGTW, four-factor alpha (following Carhart (1997)) and five-factor adjusted alpha (Carhart 4-factor augmented by Pastor and Stambaugh's (2003) liquidity factor). Sharpe ratios are annualized. The sample runs from January 1988 to December 2012. Panel A reports the results for the equal-weighted returns and panel B reports value-weighted returns.

Panel A: Equal-weighted Portfolio Returns

	Excess Return	DGTW Adjusted Return	Four-factor Alpha	Five-factor Alpha
Mean	1.19%	0.95%	1.35%	1.31%
t-stat	(6.67)	(5.93)	(8.32)	(8.04)
Std.Dev.	3.09%	2.77%		
Sharpe Ratio	1.33	1.19		
Skewness	-0.30	-0.30		
Kurtosis	1.70	2.46		
No. of obs	300	300		

Panel B: Value-weighted Portfolio Returns

	Excess Return	DGTW Adjusted Return	Four-factor Alpha	Five-factor Alpha
Mean	0.67%	0.59%	0.72%	0.79%
t-stat	(2.24)	(2.56)	(2.39)	(2.60)
Std.Dev.	5.16%	3.97%		
Sharpe Ratio	0.45	0.51		
Skewness	1.96	1.50		
Kurtosis	20.93	16.40		
No. of obs	300	300		

Table 5: **Robustness of Portfolio Sorts on DTC**

This table reports portfolio sorting results for various robustness checks. The first set of results look at two sub-periods: one from 1988 to 1999 and one from 2000 to 2012. In the second set of results, we separately examine stocks listed on NYSE-Amex versus stocks listed on NASDAQ. In the third set of results, we sorted stocks based on DTC into five and twenty portfolios instead of deciles. In the fourth set of results, we drop all the micro-cap stocks whose market capitalization are in the bottom decile/quintile of all the NYSE stocks. Lastly, we exclude the stocks with beginning period of price lower than \$5. The sample period runs from January of 1988 to December 2012. Reported are the equal-weighted DGTW-adjusted abnormal returns of a long-short portfolio that long in stocks of lowest DTC and short in stocks of highest DTC.

Full Sample	Equal-weight	0.95%
	t-stat	(5.93)
	Value-weight	0.59%
	t-stat	(2.56)
Subperiod	1988/01-1999/12	0.79%
	t-stat	(4.38)
	2000/01-2012/12	1.10%
	t-stat	(4.25)
Stock Exchanges	NYSE-Amex	0.96%
	t-stat	(6.49)
	Nasdaq	0.76%
	t-stat	(3.00)
Alternative sorts	5 portfolios	0.78%
	t-stat	(6.04)
	20 portfolios	1.26%
	t-stat	(7.44)
Remove micro-cap stocks	Bottom 10% of NYSE size cutoff	0.84%
	t-stat	(6.14)
	Bottom 20% of NYSE size cutoff	0.79%
	t-stat	(6.71)
Alternative price filter	price¿=5\$	0.59%
	t-stat	(3.55)

Table 6: **Returns to Portfolios Double Sorted on Share Turnover (Turn) and Short Interest Ratio (SR)**

This table reports monthly portfolio returns (in percentage) independently sorted on share turnover (Turn) and short interest ratio (SR). At the end of each month, all the stocks are sorted into quintiles based on turnover and independently sorted into quintiles based on short interest ratio. We report the monthly excess returns in panel A and Fama-French 3-factor alphas in panel B. The sample runs from January 1988 to December 2012. The left panel reports the results for the equal-weighted returns and the right panel reports value-weighted returns.

Panel A: Mean Excess Return

	Equal-weighted return							**Value-weighted return**					
	Low SR	SR2	SR3	SR4	High SR	High - Low		Low SR	SR2	SR3	SR4	High SR	High - Low
Low Turn	0.92	0.56	0.33	0.24	0.01	-0.91***	Low Turn	0.75	0.67	0.61	0.69	0.39	-0.35
						(-3.36)							(-1.36)
Turn2	1.58	1.10	0.75	0.51	0.21	-1.37***	Turn2	1.26	0.85	0.87	0.89	0.60	-0.67***
						(-7.54)							(-2.70)
Turn3	1.78	1.48	1.08	0.83	0.39	-1.39***	Turn3	1.15	0.94	0.94	1.00	0.71	-0.45*
						(-6.62)							(-1.95)
Turn4	2.10	1.80	1.33	1.11	0.54	-1.56***	Turn4	1.42	0.87	0.89	0.98	0.72	-0.70**
						(-6.23)							(-2.29)
High Turn	2.24	1.80	1.49	1.24	0.48	-1.76***	High Turn	2.00	1.68	0.93	1.10	0.79	-1.21***
						(-3.70)							(-2.74)
5-1	1.32**	1.24***	1.16***	1.00***	0.47	-0.85*	5-1	1.25**	1.01***	0.31	0.42	0.40	-0.85*
						(-1.81)							(-1.78)

Panel B: FF 3-factor Alpha

	Equal-weighted return							**Value-weighted return**					
	Low SR	SR2	SR3	SR4	High SR	High - Low		Low SR	SR2	SR3	SR4	High SR	High - Low
Low Turn	0.09	-0.39	-0.67	-0.82	-1.18	-1.27***	Low Turn	-0.09	-0.19	-0.35	-0.35	-0.75	-0.66***
						(-4.96)							(-2.66)
Turn2	0.63	0.07	-0.29	-0.55	-0.93	-1.56***	Turn2	0.52	0.10	0.03	-0.03	-0.43	-0.95***
						(-8.54)							(-3.93)
Turn3	0.87	0.48	0.01	-0.25	-0.75	-1.62***	Turn3	0.33	0.16	0.07	0.08	-0.30	-0.63***
						(-8.46)							(-2.71)
Turn4	1.11	0.78	0.28	0.00	-0.62	-1.72***	Turn4	0.48	-0.08	0.01	0.01	-0.32	-0.80***
						(-7.03)							(-2.66)
High Turn	1.21	0.71	0.38	0.12	-0.72	-1.93***	High Turn	0.97	0.59	-0.14	0.10	-0.28	-1.25***
						(-4.03)							(-2.63)
5-1	1.11**	1.10***	1.05***	0.94***	0.46*	-0.66	5-1	1.06**	0.78***	0.21	0.46*	0.47	-0.59
						(-1.45)							(-1.21)

Table 7: **Fama-MacBeth Regressions of Monthly Returns on DTC**

This table reports results from Fama-Macbeth (1973) regression of monthly stock returns on DTC and residual short ratio (RSR). Days-to-cover (DTC) is short interest ratio (SR) over daily turnover. Short interest ratio (SR) is the shares shorted over total shares outstanding. Residual short ratio (RSR) is the residual from the regression of SR on liquidity measures in each month. We use turnover to calcualte RSR1 and use all five liquidity proxies to calculate RSR2. Market beta (Beta) is calculated from past five years' monthly return, following Fama and French (1992). Size (LnME) is the natural log of firm market capitalization at the end of the June of each year. Book-to-market (LnBM) is the natural log of book-to-market ratio. The cases with negative book value are deleted. Momentum (MOM) is defined as the cumulative return from month t-12 to t-2. The short term reversal measure (REV) is the lagged monthly return. All the t-statistics are Newey and West (1987) adjusted to control for heteroskedasticity and autocorrelation. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)
DTC	-0.0004***		-0.0003***		
	(-9.15)		(-5.29)		
SR		-0.0912***	-0.0417**		
		(-5.71)	(-1.97)		
RSR1				-0.1194***	
				(-8.05)	
RSR2					-0.1301***
					(-8.74)
Beta	0.0017	0.0020	0.0018	0.0018	0.0018
	(1.17)	(1.41)	(1.29)	(1.28)	(1.23)
LnME	-0.0003	-0.0001	-0.0002	-0.0002	-0.0004
	(-0.50)	(-0.24)	(-0.45)	(-0.42)	(-0.74)
LnBM	0.0024***	0.0023***	0.0023***	0.0024***	0.0023***
	(2.75)	(2.69)	(2.70)	(2.70)	(2.68)
Rev	-0.0341***	-0.0337***	-0.0340***	-0.0344***	-0.0346***
	(-8.17)	(-8.14)	(-8.18)	(-8.30)	(-8.38)
Mom	0.0039**	0.0044**	0.0040**	0.0040**	0.0040**
	(2.13)	(2.39)	(2.16)	(2.21)	(2.16)
Constant	0.0111**	0.0090**	0.0109**	0.0085*	0.0093**
	(2.52)	(2.04)	(2.42)	(1.92)	(2.11)
Ave.R-sq	0.047	0.049	0.051	0.048	0.047
N.of Obs.	877047	877047	877047	877047	877047

Table 8: **Fama-MacBeth Regressions of Monthly Returns on DTC - Controlling for Stock Lending Fees**

This table reports results from Fama-Macbeth (1973) regression of monthly stock returns on DTC, after controlling short interest ratio (SR) and stock lending fees. Short interest ratio (SR) is the shares shorted over total shares outstanding. Market beta (Beta) is calculated from past five years' monthly return, following Fama and French (1992). Size (LnME) is the natural log of firm market capitalization at the end of the June of each year. Book-to-market (LnBM) is the natural log of book-to-market ratio. The cases with negative book value are deleted. Momentum (MOM) is defined as the cumulative return from month t-12 to t-2. The short term reversal measure (REV) is the lagged monthly return. Institutional ownership (IO) is the sum of shares held by institutions from 13F filings in each quarter divided by shares outstanding. Fee1 is the simple average fees of stock borrowing transactions from hedge funds in a given security, which is the difference between the risk-free rate and the rebate rate. Fee2 is a score from 1 to 10 created by Markit using their proprietary information meant to capture the cost of borrowing the stock. Here 1 is the cheapest to short and 10 the most difficult. Fee1 is available since November of 2006 while fee2 is available since October of 2003. SIO is short interest ratio (SR) divided by institutional ownership. All the t-statistics are Newey and West (1987) adjusted to control for heteroskedasticity and autocorrelation. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
DTC	-0.0003*		-0.0003**	-0.0003***		-0.0003**	-0.0002***		-0.0002***
	(-1.71)		(-2.48)	(-2.86)		(-2.47)	(-4.05)		(-3.34)
SR		-0.0338	-0.0047		-0.0251	-0.0013		-0.0263	0.0048
		(-1.01)	(-0.14)		(-1.56)	(-0.08)		(-1.12)	(0.16)
Beta	0.0022	0.0023	0.0022	0.0024	0.0027	0.0024	0.0011	0.0012	0.0011
	(0.77)	(0.78)	(0.76)	(1.02)	(1.13)	(1.05)	(0.80)	(0.88)	(0.80)
LnME	-0.0005	-0.0005	-0.0006	-0.0012**	-0.0011**	-0.0012**	-0.0010**	-0.0010**	-0.0011**
	(-0.77)	(-0.78)	(-0.87)	(-2.08)	(-2.03)	(-2.14)	(-2.31)	(-2.27)	(-2.34)
LnBM	-0.0012	-0.0011	-0.0012	0.0008	0.0006	0.0007	0.0021***	0.0020***	0.0021***
	(-0.97)	(-0.88)	(-0.95)	(0.77)	(0.62)	(0.69)	(2.82)	(2.74)	(2.83)
Rev	-0.0238*	-0.0258*	-0.0260*	-0.0360***	-0.0358***	-0.0357***	-0.0507***	-0.0508***	-0.0507***
	(-1.71)	(-1.81)	(-1.82)	(-4.94)	(-4.98)	(-4.92)	(-12.44)	(-12.48)	(-12.47)
Mom	-0.0047	-0.0046	-0.0047	-0.0029	-0.0027	-0.0028	0.0036**	0.0038**	0.0037**
	(-0.71)	(-0.71)	(-0.72)	(-0.70)	(-0.65)	(-0.69)	(2.06)	(2.13)	(2.05)
Fee1	0.0000	0.0000	0.0000						
	(0.37)	(0.43)	(0.24)						
Fee2				-0.0025***	-0.0027***	-0.0025***			
				(-3.40)	(-3.81)	(-3.62)			
SIO							-0.0142***	-0.0186***	-0.0157***
							(-5.34)	(-7.25)	(-5.64)
Ave.R-sq	0.086	0.088	0.090	0.050	0.050	0.053	0.055	0.056	0.058
N.of Obs.	102980	102980	102980	289026	289026	289026	987678	987678	987678

Table 9: **Fama-MacBeth Regressions of Monthly Returns on DTC - Robustness**

This table shows results from Fama-Macbeth (1973) regression for various robustness checks. Days-to-cover (DTC) is short interest ratio (SR) over daily turnover. DTC2 is short interest ratio (SR) over average daily turnover measured over prior month. DTC3 is short interest ratio (SR) over average daily turnover measured over past six months. DTC4 is short interest ratio (SR) over average daily turnover measured over past twelve months. Market beta (Beta) is calculated from past five years' monthly return, following Fama and French (1992). Size (LnME) is the natural log of firm market capitalization at the end of the June of each year. Book-to-market (LnBM) is the natural log of book-to-market ratio. The cases with negative book value are deleted. Momentum (MOM) is defined as the cumulative return from month t-12 to t-2. The short term reversal (REV) is the lagged monthly return. Amihud is the Amihud (2002)'s illiquidity measure. 1/Turnover is one divided by daily turnover ratio. 1/Turnover2 is one divided by average daily turnover measured over prior month. 1/Turnover3 is one divided by average daily turnover measured over past six months. 1/Turnover4 is one divided by average daily turnover measured over past twelve months. IVOL is the idiosyncratic volatility, calculated following Ang et al. (2006). Dispersion (DISP) is the analysts' earnings forecast dispersion measure, following Diether, Malloy and Scherbina (2002). All the t-statistics are Newey and West (1987) adjusted to control for heteroskedasticity and autocorrelation. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DTC	-0.0004***						-0.0004***	-0.0006***
	(-8.96)						(-9.87)	(-4.12)
1/Turnover	-0.0000***							
	(-3.17)							
DTC2		-0.0003***						
		(-6.38)						
1/Turnover2		-0.0000						
		(-1.42)						
DTC3			-0.0004***		-0.0003***			
			(-7.20)		(-6.90)			
1/Turnover3					-0.0000*			
					(-1.71)			
DTC4				-0.0004***		-0.0003***		
				(-6.28)		(-6.00)		
1/Turnover4						-0.0000		
						(-1.03)		
Amihud			0.0000	0.0000				
			(0.23)	(0.24)				
LnME	-0.0005	-0.0003	-0.0003	-0.0003	-0.0003	-0.0002	-0.0008*	-0.0010*
	(-0.87)	(-0.59)	(-0.56)	(-0.54)	(-0.55)	(-0.41)	(-1.93)	(-1.80)
LnBM	0.0029***	0.0029***	0.0027***	0.0027***	0.0028***	0.0028***	0.0023***	0.0014
	(3.50)	(3.45)	(3.08)	(3.09)	(3.40)	(3.34)	(2.83)	(1.44)
Rev	-0.0331***	-0.0321***	-0.0315***	-0.0314***	-0.0322***	-0.0321***	-0.0315***	-0.0278***
	(-8.61)	(-8.32)	(-7.97)	(-7.94)	(-8.33)	(-8.31)	(-7.67)	(-5.74)
Mom	0.0037**	0.0038**	0.0036**	0.0038**	0.0038**	0.0040**	0.0033*	0.0005
	(2.07)	(2.15)	(2.05)	(2.15)	(2.18)	(2.28)	(1.91)	(0.12)
IVOL							-0.1188***	
							(-3.10)	
DISP								0.0017
								(0.80)
Constant	0.0131***	0.0108**	0.0108**	0.0106**	0.0112**	0.0103**	0.0156***	0.0155***
	(2.91)	(2.38)	(2.48)	(2.43)	(2.37)	(2.14)	(4.70)	(3.15)
Ave.R-sq	0.047	0.047	0.047	0.047	0.048	0.048	0.050	0.071
N.of Obs.	1041563	1041563	974842	974823	1041563	1041540	1003934	627762

Table 10: **Regression of Hedge Fund Holding (HFH) on Trading Cost Measures**

This table reports results from Fama-Macbeth (1973) regression of quarterly hedge fund holding (HFH) on various trading costs measures. Turnover is the monthly average of daily turnover ratio. Amihud is the Amihud (2002) illiquidity measure. FHT is the transaction costs estimated from the frequency of zero returns (Fong, Holden and Trzcinka 2014). PS is the Pastor and Stambugh (2003) liquidity measure. QS is the percentage quoted spread using daily close price (Chung and Zhang 2014). Size (LnME) is the natural log of firm market capitalization at the end of the June of each year. Book-to-market (LnBM) is the natural log of book-to-market ratio. The cases with negative book value are deleted. Momentum (MOM) is defined as the cumulative return from month t-12 to t-2. IO is the institutional ownership ratio. IVOL is the idiosyncratic volatility following Ang et al. (2005). We use Fama-Macbeth regression method in column (1) to (7). In column (8), we use panel regression controlling for firm-fixed effect. All variables are standardized to have mean 0 and standard deviation of 1. For Fama-Macbeth regression, all the t-statistics are Newey and West (1987) adjusted to control for heteroskedasticity and autocorrelation. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Turnover	0.1983***	0.1043***					0.0920***	0.0641***
	(13.75)	(9.20)					(8.32)	(31.76)
Amihud			-0.0669***				-0.0370***	-0.0126***
			(-16.85)				(-13.21)	(-6.51)
FHT				-0.0202*			0.0227**	0.0125***
				(-1.85)			(2.25)	(5.52)
PS					-0.0075***		-0.0023	0.0004
					(-3.77)		(-1.19)	(0.28)
QS						-0.0973***	-0.0315***	-0.0235***
						(-10.62)	(-2.72)	(-7.82)
LnME		-0.2775***	-0.2569***	-0.2547***	-0.2518***	-0.2750***	-0.2808***	-0.2979***
		(-31.55)	(-27.28)	(-28.53)	(-26.24)	(-27.74)	(-27.32)	(-61.75)
LnBM		-0.0273***	-0.0319***	-0.0379***	-0.0396***	-0.0292***	-0.0215***	-0.0029
		(-5.06)	(-5.90)	(-7.00)	(-7.49)	(-5.31)	(-3.95)	(-1.35)
Mom		-0.0085*	0.0070	0.0087	0.0102*	0.0003	-0.0071*	-0.0064***
		(-1.90)	(1.44)	(1.65)	(1.95)	(0.07)	(-1.76)	(-4.64)
IO		0.4361***	0.4716***	0.4751***	0.4763***	0.4571***	0.4325***	0.4079***
		(30.72)	(34.69)	(30.57)	(32.69)	(38.58)	(33.32)	(131.18)
IVOL		-0.0001	0.0746***	0.0605***	0.0475***	0.0870***	0.0201*	-0.0294***
		(-0.01)	(6.48)	(5.24)	(3.75)	(6.50)	(1.93)	(-11.33)
Constant	-0.0011	-0.0329***	-0.0315***	-0.0326***	-0.0323***	-0.0313***	-0.0301***	-0.0317***
	(-1.47)	(-13.68)	(-13.33)	(-13.86)	(-13.71)	(-13.18)	(-11.10)	(-26.55)
Method	FM	FM	FM	FM	FM	FM	FM	firm-fixed
Ave.R-sq	0.044	0.158	0.152	0.150	0.149	0.153	0.162	0.066
N.of Obs.	376684	323087	323087	323074	323087	318528	318528	318528

Table 11: **Returns to Portfolio Strategies based on Residual Hedge Fund Holding (RHFH)**

This table provides portfolio returns and alphas, sorted on residual hedge fund holding (RHFH). At the end of each quarter, all the stocks are sorted into 10 portfolios based on residual hedge fund holding and a long-short portfolio is formed by buying the highest RHFH portfolio and shorting the lowest RHFH portfolio. Portfolios are rebalanced at each quarter. We report the excess returns, characteristics-adjusted abnormal returns calculated following Daniel, Grinblatt, Titman and Wermers (1997), denoted as DGTW, four-factor alpha (following Carhart (1997)) and five-factor adjusted alpha (Carhart 4-factor augmented by Pastor and Stambaugh's (2003) liquidity factor). Sharpe ratios are annualized. Residual hedge fund holding is the residual of the cross-sectional regression of hedge fund holding on average turnover at each quarter. The sample runs from January 1992 to December 2012. Panel A reports the results for the equal-weighted returns and panel B reports value-weighted returns.

Panel A: Equal-weighted Portfolio Returns

	Excess Return	DGTW Adjusted Return	Four-factor Alpha	Five-factor Alpha
Mean	0.65%	0.46%	0.70%	0.64%
t-stat	3.52	2.95	4.49	4.09
Std.Dev.	2.94%	2.48%		
Sharpe Ratio	0.77	0.64		
Skewness	-0.29	-0.21		
Kurtosis	1.78	2.20		
No. of obs	252	252	252	252

Panel B: Value-weighted Portfolio Returns

	Excess Return	DGTW Adjusted Return	Four-factor Alpha	Five-factor Alpha
Mean	0.58%	0.40%	0.62%	0.53%
t-stat	2.72	2.42	3.37	2.91
Std.Dev.	3.37%	2.66%		
Sharpe Ratio	0.59	0.53		
Skewness	-0.16	0.04		
Kurtosis	1.16	1.87		
No. of obs	252	252	252	252

Table 12: **Fama-MacBeth Regressions of Monthly Returns on Hedge Fund Holding (HFH) and Residual Hedge Fund Holding (RHFH)**

This table reports results from Fama-Macbeth (1973) regression of monthly excess return on lagged hedge fund holding (HFH) and residual hedge fund holding (RHFH). Hedge fund holding (HFH) is the number of shares held by hedge fund over total shares outstanding at quarter end. Residual hedge fund holdings (RHFH) is the residual from cross-sectional regression of hedge fund holdings on turnover within each quarter. Days-to-cover (DTC) is short interest ratio (SR) over daily turnover. Size (LnME) is the natural log of firm market capitalization at the end of June of each year. Book-to-market (LnBM) is the natural log of book-to-market ratio. The cases with negative book value are deleted. Short-term reversal (Rev) is the return from prior month. Momentum (MOM) is defined as the cumulative return from month t-12 to t-2. All the t-statistics are Newey and West (1987) adjusted to control for heteroskedasticity and autocorrelation. ***, **, and * stands for significance level of 1%, 5% and 10%, respectively.

	(1)	(2)	(3)	(4)
HFH	0.0134		0.0137	
	(1.47)		(1.51)	
RHFH		0.0236***		0.0233***
		(3.35)		(3.27)
DTC			-0.0001***	-0.0001***
			(-6.43)	(-6.44)
LnME	-0.0012	-0.0012	-0.0011	-0.0011
	(-1.57)	(-1.58)	(-1.49)	(-1.47)
LnBM	0.0037***	0.0036***	0.0036***	0.0036***
	(3.03)	(3.02)	(3.00)	(2.95)
Rev	-0.0405***	-0.0402***	-0.0405***	-0.0403***
	(-7.37)	(-7.41)	(-7.40)	(-7.36)
Mom	-0.0005	-0.0004	-0.0007	-0.0007
	(-0.16)	(-0.13)	(-0.20)	(-0.21)
Constant	0.0179***	0.0189***	0.0183***	0.0189***
	(2.79)	(2.94)	(2.85)	(2.90)
Ave.R-sq	0.035	0.034	0.035	0.035
N.of Obs.	996629	996629	992182	992182