Available online at www.sciencedirect.com
ScienceDirect
Journal of Financial Markets 17 (2014) 199–229
Journal of FINANCIAL MARKETS
www.elsevier.com/locate/finmar

Aggregate short selling, commonality, and stock market returns ☆

Andrew Lynch[a], Biljana Nikolic[b], Xuemin (Sterling) Yan[c,d,*], Han Yu[e]

[a]Binghamton University, School of Management, Binghamton, NY, USA
[b]University of San Diego, School of Business, San Diego, CA, USA
[c]University of Missouri, College of Business, Columbia, MO, USA
[d]Renmin University of China, School of Business, Beijing, China
[e]University of Kansas, School of Business, Lawrence, KS, USA

Received 23 January 2012; received in revised form 3 May 2013; accepted 3 May 2013
Available online 14 May 2013

Abstract

Using a comprehensive data set of short-sale transactions, we find strong evidence of commonality in daily shorting flows of individual stocks. More importantly, we find that aggregate shorting forecasts market returns. A one standard deviation increase in daily aggregate shorting is associated with a decrease in market excess return by up to 36 bps over the following 10 trading days (9% annualized). In addition, we find modest evidence that short sellers are informed about future aggregate earnings news, macroeconomic news, and investor sentiment. Overall, our results are consistent with short sellers possessing superior short-term market-wide information.

JEL classification: G10; G14

Keywords: Aggregate Short Selling; Market return; Commonality

☆We thank Grace Hao for helping us obtain a portion of the Regulation SHO data. We thank an anonymous referee, Azizjon Alimov, Chen Lin, Tarun Chordia (editor), Tao Li, Qianqiu Liu, Junbo Wang, Xueping Wu, Joe Zou, and seminar participants at the City University of Hong Kong, University of Kansas, and FMA 2011 Annual Meetings for helpful comments.
*Corresponding author at: University of Missouri, College of Business, Columbia, MO, USA. Tel.: +1 573 884 9708.
E-mail address: yanx@missouri.edu (X. Yan).

1386-4181/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.finmar.2013.05.001

1. Introduction

The role of short selling in financial markets has been a subject of intense debate among academics, practitioners, and regulators for decades. The issue is so contentious that over the course of a three-year period the SEC lifted price test restrictions on short sales, temporarily banned short selling in financial stocks, and then approved an alternative uptick rule.[1] Academic research, on the other hand, provides considerable evidence that short sellers are informed traders who contribute to market efficiency. For example, previous research finds that short sellers target overvalued stocks based on fundamental ratios (Dechow, Hutton, Meulbroek, and Sloan, 2001), enhance the informational efficiency of stock prices (Boehmer and Wu, forthcoming), anticipate negative news (Christophe, Ferri, and Angel, 2004; Karpoff and Lou, 2010), and predict future stock returns (Desai, Thiagarajan, and Balachandran, 2002; Boehmer, Jones, and Zhang, 2008; Diether, Lee, and Werner, 2009).[2]

However, nearly all of this research focuses on the information content of short selling at the individual stock level. In this paper, we extend the literature on individual stock short selling to the market level by investigating whether aggregate short selling contains information about future market returns. Markets aggregate information from economic agents and impound it into prices. This information can be firm-specific or market-wide. If short sellers trade only on firm-specific information, we should expect no significant relation between aggregate shorting and subsequent market returns because the idiosyncratic components of short sales and stock returns cancel out in the aggregate. Alternatively, if short sellers possess and trade on market-level information and this information is not immediately incorporated into prices, we should expect aggregate short selling to predict future market returns.

There are several reasons why we might expect short sellers to possess market-level information. Diamond and Verrecchia (1987) contend that short sellers are likely to be informed because investors should never initiate a short position for liquidity reasons. To the extent that the price of a stock is influenced by both firm-specific and aggregate news, short sellers might be informed at both the firm level and the market level. In fact, one could argue that there may be more potential for informed shorting at the market level. Lamont and Stein (2004), p. 29, for example, argue that "short-selling-based arbitrage would be more effective along the aggregate dimension than it is in the cross section." Similarly, Veldkamp and Wolfers (2007) contend that in equilibrium aggregate information will be produced and acquired more widely because information has a high fixed-cost of production and aggregate information, by definition, is relevant for all firms in the economy. Consistent with Veldkamp and Wolfers (2007), there is considerable evidence in the literature that various market participants including insiders, corporate managers, financial analysts, and mutual fund managers possess superior information about the aggregate market (Seyhun, 1988, 1992; Baker and Wurgler, 2000; Lakonishok and Lee, 2001; Howe, Unlu, and Yan, 2009; Bollen and Busse, 2001).[3] To the extent that short

[1]The SEC eliminated short-sale price tests (former Rule 10a-1) in July 2007, issued two emergency orders in July and September of 2008 to temporarily ban short selling in certain financial stocks, and then in February 2010 approved an alternative uptick rule (Rule 201), which imposes a restriction on the prices at which a stock can be shorted when the stock experiences a price decline of more than 10% from the close of the prior trading day.

[2]Numerous other studies including Asquith and Meulbroek (1995), Aitken, Frino, McCorry, and Swan (1998), Asquith, Pathak, and Ritter (2005), Cohen, Diether, and Malloy (2007), Boehmer, Jordan, and Huszar (2010), Christophe, Ferri, and Hsieh (2010) also find evidence of informed short selling.

[3]Albuquerque, De Francisco, and Marques (2008) and Beber, Brandt, and Kavajecz (2011) find significant evidence of market-wide information in equity order flows.

sellers are at least as sophisticated as the above market participants, it is plausible that short sellers might be informed at the market level.

We perform our tests using a comprehensive data set of short-sale transactions for a two and a half year sample period made available through Regulation SHO. Our tests yield three main results. First, we find strong evidence of commonality in daily shorting flows. Second, we find that daily aggregate shorting forecasts market returns over the following 5 to 20 trading days. Third, we show that this predictive ability is not due to liquidity provision, is not shared by regular sellers, and is not present in the monthly short interest data.

We begin our analysis by examining whether individual stock short selling contains common components. Intuitively, if short sellers trade on the basis of market-wide information, we should expect individual stock short selling to exhibit common components. Using daily shorting flows constructed from transaction-level data, we find strong evidence of commonality: short sales in individual stocks co-move significantly with both market- and industry-aggregated short sales.

To explore whether the commonality in short selling is driven by the proliferation of hedge funds and their arbitrage activities, we repeat our analysis for subsamples of stocks sorted by hedge fund ownership. We find that stocks with high hedge fund ownership exhibit significantly stronger commonality in short sales than stocks with low hedge fund ownership. Nevertheless, we continue to find significant evidence of commonality even among stocks with low hedge fund ownership, suggesting that hedge funds play only a partial role in explaining the commonality in short selling. We also find evidence that commonality in short selling is stronger in declining markets and among high-beta stocks. This evidence is consistent with the idea that, due to limits-to-arbitrage concerns, short sellers may be reluctant to short overvalued stocks when they anticipate positive market returns. Instead, they tend to time their price-correcting trades to coincide with falling markets.

We next investigate whether daily aggregate shorting flows predict future market returns. Our empirical design parallels that of Boehmer, Jones, and Zhang (2008). For each day, we construct our predictive variable based on aggregate shorting activity during the previous five trading days and forecast future market excess returns over the following 5, 10, or 20 trading days. We document a reliably negative relation between aggregate short selling and subsequent market returns during our sample period. The coefficient estimates on lagged aggregate shorting are negative in all regression specifications, and statistically significant in the majority of them. The result is economically meaningful; a one standard deviation increase in aggregate shorting is associated with a decrease in market excess return of 11–36 bps over the following 10 trading days (3–9% annualized). This predictive ability is not attributable to macroeconomic information readily available to the public, as the results remain significant even after controlling for several widely-used macroeconomic predictors of market returns (e.g., Fama and French, 1989; and Ferson and Harvey, 1999). Our result also persists after controlling for several market-level variables, such as market liquidity, market volatility, and aggregate order imbalance.

Our findings so far suggest that short sellers possess superior market-wide information. This superior information could be about the state of the economy or about market-wide mispricing. To explore the sources of this informational advantage, we perform several analyses. First, we examine whether aggregate short selling predicts future aggregate earnings news. We find modest evidence of an inverse relation between aggregate short selling and future aggregate earnings news. This finding suggests that short sellers are informed about future aggregate earnings. That is, when aggregate short selling is high, future aggregate earnings tends to be unexpectedly low. Second, we explore whether short sellers possess superior information about future economy-wide activities as measured by macroeconomic news announcements, such as the unemployment rate, consumer price index, and gross domestic product. We regress future macroeconomic news on lagged aggregate short selling and find that all regression coefficients on lagged aggregate short selling are negative, nearly half of which are statistically significant. This finding is consistent with the idea that short sellers possess superior information about the aggregate economy. Third, to provide evidence on whether the predictive ability of short sellers for market returns is related to market-wide mispricing, we regress future investor sentiment on lagged aggregate short selling. Our use of investor sentiment as a proxy for aggregate market mispricing is motivated by Brown and Cliff (2005) and Baker and Wurgler (2007), who show that high investor sentiment is associated with market overvaluation and low subsequent stock market returns. We find some evidence that high aggregate short selling predicts low investor sentiment. Taken as a whole, our findings from the above three analyses provide important insights into the informational sources of the predictive ability of aggregate short selling.

While our findings are consistent with the superior information hypothesis, an alternative explanation for the inverse relation between aggregate shorting and subsequent market returns is that short sellers act as liquidity providers when there is market-wide buying pressure (liquidity provision hypothesis). In this case, the predictive ability of aggregate short selling may simply reflect compensation for providing liquidity (e.g., Campbell, Grossman, and Wang, 1993; and Diether, Lee, and Werner, 2009).

To explore the liquidity provision hypothesis, we conduct two tests. First, we examine the relation between aggregate order imbalance and the level of aggregate short selling. The liquidity provision hypothesis predicts that elevated short selling will coincide with the buying pressure in the market. Using aggregate order imbalance constructed from the TAQ data, we find little support for this prediction. Second, we re-estimate our predictive regressions after excluding days when the aggregate order imbalance or market illiquidity are the highest. If liquidity provision is driving our results, we should find the predictive ability of aggregate shorting diminishes or even disappears once we exclude days with intense buying pressure or extreme market illiquidity. Contrary to this prediction, our results remain significant after the above exclusions. Thus our tests provide little support for the hypothesis that liquidity provision drives the predictive ability of aggregate short selling.

Until recently, much of the short selling research uses short interest data. To examine whether the predictive ability of daily aggregate shorting flows extend to the monthly short interest data, we next investigate the predictive ability of monthly aggregate short interest ratios. We fail to find any evidence that changes in aggregate short interest ratios predict future market returns during our sample period. This lack of evidence is not surprising because short sellers are short-term investors. Geczy, Musto, and Reed (2002) and Diether (2008) show that shorting contracts have a median duration of only three to seven trading days. As such, monthly short interest data are too coarse to fully capture the actions of short sellers. By contrast, daily shorting flows are much finer and permit a study of short-term trading strategies.

The primary contribution of our paper is to provide a comprehensive analysis of the information content of aggregate short selling using transaction-level data. In addition, we are the first to document evidence of commonality in short selling. Several papers have previously examined aggregate shorting using monthly short interest data. Seneca (1967) documents that higher aggregate short interest is associated with lower future S&P 500 Index levels. Hanna (1968), however, disputes Seneca's finding and reaches opposite conclusions using the same data. More recently, Lamont and Stein (2004) examine the behavior of aggregate short interest ratios and find no evidence that short sellers counter mispricing at the market level. In contrast to the above studies, we examine daily aggregate shorting flows constructed from transaction-level data and our results highlight the importance of using high-frequency data to detect informed short selling.

The rest of our paper proceeds as follows. In Section 2, we describe the data, variables, and descriptive statistics. In Section 3, we examine the commonality in daily shorting flows. In Section 4, we explore the predictive content of aggregate short selling. In Section 5, we provide concluding remarks.

2. Data, variables, and descriptive statistics

2.1. Data and sample

Publicly available short selling data have traditionally been in the form of shorting stock (i.e., open short positions or short interest). In this paper, we use high-frequency shorting flow data to investigate the informativeness of aggregate short selling.[4] Specifically, we obtain a comprehensive database of short-sale transactions made available through Regulation SHO. Regulation SHO was adopted by the SEC in June of 2004 to establish new rules regarding short sales and to evaluate the effectiveness of price test restrictions on short sales. At the same time, the SEC mandated that all self-regulatory organizations (SROs) make transaction-level short-sale data publicly available for the period from January 3, 2005 to July 6, 2007.[5] We collect short-sale transactions from all SROs including AMEX, Archipelago (ARCA), Boston, Chicago, NASD, NASDAQ, National (NSX), NYSE, and Philadelphia (PHLX).

The Regulation SHO data contain the ticker, date, and time of the short-sale transaction, number of shares shorted, and execution price. Because short sales for a given stock might occur at multiple venues, we aggregate the data across all exchanges for each stock on each day.[6] We use the TAQ master files to add CUSIPs to our data and then merge with the CRSP database to add PERMNOs. Our sample includes only common stocks (i.e., those securities with a sharecode of 10 or 11 in the CRSP database). To be included in our final sample, we also require that a stock have at least one short-sale transaction during our sample period.

For each stock on each day, we obtain the total number of trades, total share volume, quoted spread, and order imbalance from the TAQ database. We follow Chordia, Roll, and Subrahmanyam (2002) and purge the following trade and quote data: trades out of sequence, trades and quotes before the open or after the close, quotes not originated from the primary exchange, trades or quotes with negative prices, quotes with negative spread or negative depth, and quotes with spread greater than $4 or 20% of the mid-quote. We sign trades using the following algorithm: if a trade occurs above (below) the mid-point of the prevailing quote, it is classified as a buyer- (seller-) initiated trade. Trades that occur at the mid-point of the prevailing quote are unsigned.

[4]Both shorting flow and shorting stock capture the *quantity* of shorting. A number of recent studies (e.g., D'Avolio, 2002; Geczy, Musto, and Reed, 2002; Jones and Lamont, 2002; Ofek, Richardson, and Whitelaw, 2004; Cohen, Diether, and Malloy, 2007) examine the *price* of shorting (i.e., rebate rate or loan fee). The shorting price data, however, are not publicly available.

[5]In our analysis, we omit the three trading days in July 2007 to present a uniform sample of 30 months.

[6]Consistent with Diether, Lee, and Werner (2009), we find that a substantial percentage of short-sale transactions are executed away from the primary listing exchange. Based on the number of shares shorted, the NYSE accounts for 76.45% of the shorting activity in NYSE-listed stocks, while NASDAQ accounts for only 57.22% of shares shorted in NASDAQ stocks.

We obtain daily stock returns, share prices, shares outstanding, SIC codes, S&P 500 Index membership, and value-weighted market index returns from CRSP. We obtain book value of equity, short interest, and earnings announcement dates from Compustat. We obtain daily risk-free returns and Fama-French 48 industry classifications from Kenneth French's website. We obtain daily three-month T-bill rates, 10-year T-bond yields, and Aaa- and Baa-rated corporate bond yields from the website of Federal Reserve Bank of St. Louis. We obtain institutional ownership from Thomson 13F institutional holdings database. We obtain the sample of hedge funds from the CISDM database. We obtain investor sentiment data from Jeffery Wurgler's website. Finally, we obtain the following macroeconomics news announcements (along with professional economists' expectations) from Bloomberg: unemployment rate, non-farm payrolls, industrial production, retail sales, jobless claims, durable goods, consumer price index, and gross domestic product.

2.2. Aggregate shorting variables

To measure daily aggregate shorting activity, we first construct three shorting flow measures at the individual stock level. Specifically, for each stock on each day, we compute the percentage of share volume shorted (*sv*) as the ratio between the number of shares shorted and the total number of shares traded, the percentage of trades shorted (*st*) as the ratio between the number of short-sale trades and the total number of trades, and the percentage of shares outstanding shorted (*ss*) as the ratio between the number of shares shorted and the number of shares outstanding.

Next, we aggregate each of the above three measures in the following three ways: equal-weighted average across all stocks (*ew*), value-weighted average across all stocks (*vw*), and aggregated across all stocks (*agg*). This procedure results in a total of nine measures of daily aggregate shorting flows. For instance, *sv_ew* is the equal-weighted average of each stock's percentage of share volume shorted, *ss_vw* is the value-weighted average of each stock's percentage of shares outstanding shorted, and *st_agg* is the ratio between the aggregate number of short-sale trades across all stocks and the aggregate number of trades across all stocks.

We use this variety of measures in order to provide a comprehensive understanding of aggregate shorting and to make our results comparable to the individual stock shorting literature. First, our *sv* measure (percentage of share volume shorted) is analogous to the shorting measures used by Boehmer, Jones, and Zhang (2008) and Diether, Lee, and Werner (2009), and thus provides the closest comparison between our results and the individual stock shorting literature. Second, Jones, Kaul, and Lipson (1994) find that the number of trades conveys more information than total share volume or average trade size. As such, our *st* measure (percentage of trades shorted) may better capture informed shorting.[7] Third, a number of studies (e.g., Asquith, Pathak, and Ritter, 2005) scale short selling activities by total number of shares outstanding, which is similar to our *ss* measure. Finally, each of our weighting schemes has unique advantages. The equal-weighted measures assign more weight to small firms while the value-weighted measures give more weight to large firms. The aggregate measures are similar to value-weighted measures but are less likely to be influenced by extreme outliers.

[7] Boehmer, Jones, and Zhang (2008) also use the number of short-sale trades as a measure of shorting activity.

Table 1
Descriptive statistics.

We obtain data on short-sale transactions from NYSE, AMEX, NASDAQ, Archipelago (ARCA), Boston Stock Exchange, Chicago Stock Exchange, NASD, National Stock Exchange (NSX), and Philadelphia Stock Exchange (PHLX), which provide the data as part of the requirements under Regulation SHO. Our sample period is from January 3, 2005 to June 29, 2007. Our sample includes only common stocks (those with a sharecode of 10 or 11 in CRSP). We obtain daily stock returns, index returns, shares outstanding, share prices, and SIC codes from CRSP. We obtain number of trades and share volume from TAQ. For each stock on each trading day, we compute three short-sale measures, *sv*, *st*, and *ss*. *sv* is the number of shares shorted divided by total share volume. *st* is number of short-sale trades divided by total number of trades. *ss* is the number of shares shorted divided by number of shares outstanding. *sv_vw*, *st_vw*, and *ss_vw* are value-weighted average of *sv*, *st*, and *ss*, respectively, across all stocks. *sv_ew*, *st_ew*, and *ss_ew* are equal-weighted average of *sv*, *st*, and *ss*, respectively, across all stocks. *sv_agg*, *st_agg*, and *ss_agg* are ratios of the aggregate number of shares shorted to aggregate total share volume, aggregate number of short-sale trades to aggregate number of all trades, and aggregate number of shares shorted to aggregate shares outstanding, respectively. Aggregate order imbalance and market liquidity variables are computed from TAQ data. Aggregate order imbalance is defined as the difference between buyer-initiated share volume and seller-initiated share volume scaled by the total share volume. Aggregate quoted spread is value-weighted across all stocks. Daily market volatility is the sum of squared 5-minute S&P 500 Index returns for each day. Panel A presents univariate statistics for daily market-aggregate short-sale measures and other market variables. Panel B presents Pearson correlations. ***, **, * indicate statistical significance of 1%, 5%, and 10%, respectively.

Panel A: Univariate Statistics

	Mean	Median	Stdev	Maximum	Minimum
sv_vw	0.242	0.245	0.022	0.293	0.087
sv_ew	0.251	0.254	0.026	0.308	0.111
sv_agg	0.269	0.278	0.028	0.321	0.112
st_vw	0.262	0.270	0.030	0.315	0.108
st_ew	0.273	0.274	0.024	0.318	0.156
st_agg	0.314	0.328	0.036	0.367	0.169
ss_vw (%)	0.182	0.181	0.031	0.299	0.057
ss_ew (%)	0.203	0.201	0.030	0.318	0.060
ss_agg (%)	0.199	0.198	0.033	0.335	0.059
Value-weight CRSP Index Return—*vwret* (%)	0.050	0.095	0.671	2.386	−3.408
Total Number of Short-sale Trades (million)	2.603	2.516	0.635	5.277	0.768
Total Number of Shares Shorted (billion shares)	0.928	0.923	0.155	1.574	0.280
Total Number of Trades (million)	8.485	7.859	2.656	18.577	2.286
Total Share Volume (billion shares)	3.461	3.452	0.563	5.422	1.132
Aggregate Order Imbalance—*aggoi*	0.051	0.051	0.022	0.114	−0.023

Table 1 (*continued*)

Panel A: Univariate Statistics

	Mean	Median	Stdev	Maximum	Minimum
Aggregate Quoted Spread—*aggqp* (%)	0.054	0.036	0.005	0.053	0.026
Daily Market Volatility—*mktvol* (%)	0.540	0.503	0.171	1.461	0.210
Number of Stocks	4,565	4,564	33	4,681	4,394

Panel B: Correlations

	vwret	*sv_vw*	*sv_ew*	*sv_agg*	*st_vw*	*st_ew*	*st_agg*	*ss_vw*	*ss_ew*	*ss_agg*
sv_vw	0.068*									
sv_ew	−0.046	0.812***								
sv_agg	0.008	0.940***	0.861***							
st_vw	0.034	0.929***	0.783***	0.937***						
st_ew	0.053	0.706***	0.930***	0.757***	0.693***					
st_agg	0.013	0.819***	0.843***	0.915***	0.881***	0.846***				
ss_vw	−0.047	0.317***	0.341***	0.294***	0.274***	0.300***	0.233***			
ss_ew	−0.111**	0.380***	0.531***	0.394***	0.364***	0.494***	0.364***	0.882***		
ss_agg	−0.030	0.340***	0.390***	0.332***	0.309***	0.362***	0.288***	0.957***	0.914***	
aggoi	0.697***	0.078*	−0.087**	0.070*	0.151***	0.003	0.148***	−0.090**	−0.164***	−0.093**

2.3. Summary statistics

Panel A of Table 1 reports the summary statistics for our daily aggregate shorting measures and several market variables. The average values of *sv_vw*, *sv_ew*, and *sv_agg* indicate that short sales account for between 24.2% and 26.9% of the total share volume. The percentages based on the number of trades (i.e., *st_vw*, *st_ew*, and *st_agg*) are even higher, ranging from 26.2% to 31.4%. These results are consistent with those reported in Diether, Lee, and Werner (2009), who show that short selling accounts for 23.89% (31.33%) of the share volume on the NYSE (NASDAQ) in 2005. The percentage of shares outstanding shorted (*ss_vw*, *ss_ew*, and *ss_agg*) represents the short-sale turnover ratio, and has a mean value ranging from 0.182% to 0.203% on a daily basis, which corresponds to an annual turnover rate of 45–50%. Overall, our results show that short selling is prevalent during our sample period.[8]

Panel A of Table 1 also presents the summary statistics for market returns, aggregate trading activity, market volatility, and aggregate liquidity. Several results are noteworthy. The average value-weighted market return is approximately 5 bps per day, suggesting that our sample period is characterized by a generally rising market. Consistent with market returns, we find the average aggregate order imbalance is also positive. There are on average 8.5 million trades across all stocks each day, 2.6 million of which are short-sale trades. Similarly, the average daily total share volume is 3.46 billion shares, of which 928 million shares are involved in short selling.

Panel B of Table 1 presents the contemporaneous correlations among our variables. Not surprisingly, all aggregate shorting measures are significantly positively correlated with each other, with correlation coefficients ranging from 0.233 to 0.957. While some of these correlations are high, they are generally far below one, which suggests that our measures capture different aspects of the aggregate shorting activity. Correlations between aggregate shorting measures and contemporaneous market return are generally small in magnitude. As expected, we find a significant positive correlation between aggregate order imbalance and contemporaneous market return. The correlations between aggregate order imbalance and aggregate shorting measures are modest in magnitude and have mixed signs, which suggest that aggregate short selling is not merely picking up sell order imbalance.

2.4. Determinants of aggregate daily shorting flows

To better understand the behavior of aggregate daily shorting flows, we estimate the following time-series regression:

$$
\begin{aligned}
short_t = a_0 &+ \sum_{i=1}^{5} a_i short_{t-i} + \sum_{j=1}^{5} b_j vwret_{t-j} + \sum_{k=1}^{4} c_k day\_of\_week_t \\
&+ d_1 aggqp_t + d_2 mktvol_t + d_3 aggoi_t + e_t
\end{aligned}
\tag{1}
$$

The dependent variable ($short_t$) is one of our nine aggregate shorting flow measures (i.e., *sv_vw*, *sv_ew*, *sv_agg*, *st_vw*, *st_ew*, *st_agg*, *ss_vw*, *ss_ew*, and *ss_agg*). We include four sets of explanatory variables. First, Diether, Lee, and Werner (2009) establish that short sellers in individual stocks are contrarians. To examine whether this pattern extends to aggregate short selling, we include prior market returns in our analysis. Second, we include lagged values of aggregate shorting flows to explore whether aggregate shorting is persistent. Third, we include

[8] This finding is consistent with D'Avolio (2002), who shows that the vast majority of the stocks can be easily borrowed.

Table 2
Determinants of market-aggregated short-sales.

We obtain data on short-sale transactions from various exchanges under Regulation SHO. Our sample period is from January 3, 2005 to June 29, 2007. Our sample includes only common stocks (those with a sharecode of 10 or 11 in CRSP). Aggregate short sale measures are described in Table 1. Each market-aggregated short-sale measure is regressed on lagged market-aggregated short-sale measures ($Short_{t-1}$ through $Short_{t-5}$), lagged value-weight market returns ($vwret_{t-1}$ through $vwret_{t-5}$), day of the week indicator variables, and aggregate liquidity (*aggqp*), market volatility (*mktvol*), and aggregate order imbalance (*aggoi*). *aggqp* is the average quoted spread across all stocks. *mktvol* is the market volatility computed from 5-minute returns of the S&P 500 Index. *aggoi* is the aggregate order imbalance, defined as the difference between buyer-initiated share volume and seller-initiated share volume scaled by total share volume. Numbers in parentheses are t-statistics based on Newey-West standard errors with five lags.

	Dependent Variables								
	sv_vw	*sv_ew*	*sv_agg*	*st_vw*	*st_ew*	*st_agg*	*ss_vw*	*ss_ew*	*ss_agg*
Intercept	0.031 (2.90)	0.030 (3.01)	0.019 (2.15)	0.010 (1.41)	0.038 (2.88)	0.014 (1.62)	0.034 (2.73)	0.040 (3.17)	0.032 (2.31)
$Short_{t-1}$	0.252 (3.76)	0.257 (2.77)	0.264 (3.55)	0.303 (3.89)	0.316 (3.47)	0.345 (3.56)	0.440 (9.37)	0.451 (7.95)	0.429 (8.14)
$Short_{t-2}$	0.227 (4.28)	0.221 (4.42)	0.235 (4.29)	0.269 (4.55)	0.222 (3.89)	0.212 (3.34)	0.140 (2.92)	0.121 (2.44)	0.172 (3.51)
$Short_{t-3}$	0.172 (3.45)	0.202 (3.82)	0.188 (3.35)	0.180 (3.72)	0.167 (3.49)	0.170 (3.34)	0.145 (3.30)	0.144 (3.09)	0.158 (3.49)
$Short_{t-4}$	0.123 (4.09)	0.103 (2.81)	0.111 (3.45)	0.124 (3.11)	0.068 (1.77)	0.079 (1.97)	−0.007 (−0.16)	0.029 (0.64)	−0.014 (−0.30)
$Short_{t-5}$	0.014 (0.41)	0.039 (1.21)	0.041 (1.20)	0.018 (0.46)	0.044 (1.37)	0.042 (1.21)	−0.040 (−0.94)	−0.018 (−0.46)	−0.031 (−0.75)
$vwret_{t-1}$	−0.293 (−2.58)	−0.221 (−1.35)	−0.136 (−1.01)	−0.426 (−3.74)	−0.306 (−2.16)	−0.280 (−1.73)	−0.377 (−2.34)	−0.063 (−0.35)	−0.282 (−1.56)
$vwret_{t-2}$	−0.132 (−2.11)	−0.081 (−1.03)	−0.017 (−0.23)	−0.105 (−1.37)	−0.145 (−1.66)	−0.017 (−0.18)	−0.013 (−0.07)	−0.026 (−0.16)	−0.046 (−0.24)
$vwret_{t-3}$	−0.094 (−0.77)	0.071 (0.65)	0.002 (0.02)	−0.069 (−0.60)	0.001 (0.01)	0.055 (0.44)	−0.030 (−0.19)	0.113 (0.83)	0.007 (0.04)
$vwret_{t-4}$	0.043 (0.39)	0.028 (0.18)	0.027 (0.20)	0.040 (0.36)	−0.021 (−0.15)	0.023 (0.14)	0.092 (0.64)	0.099 (0.61)	0.089 (0.55)
$vwret_{t-5}$	0.045 (0.55)	0.137 (1.45)	0.083 (0.92)	0.074 (0.82)	0.125 (1.19)	0.140 (1.12)	0.046 (0.31)	0.109 (0.78)	0.093 (0.59)
Tue	−0.001 (−0.47)	−0.000 (−0.02)	0.001 (0.58)	0.000 (0.13)	0.001 (0.65)	0.004 (1.56)	0.022 (8.83)	0.017 (7.66)	0.026 (10.03)
Wed	0.001 (0.73)	−0.001 (−0.48)	0.004 (1.95)	0.003 (1.55)	−0.001 (−0.31)	0.004 (1.79)	0.027 (8.58)	0.017 (5.74)	0.029 (9.06)
Thu	0.002 (0.68)	0.001 (0.32)	0.004 (1.55)	0.002 (1.02)	0.001 (0.48)	0.006 (1.98)	0.023 (7.10)	0.018 (5.69)	0.026 (7.70)
Fri	0.001 (0.23)	0.002 (0.77)	0.002 (0.79)	0.002 (0.94)	0.004 (1.42)	0.005 (1.47)	0.006 (1.55)	0.001 (0.16)	0.007 (1.87)
aggqp	0.485 (3.12)	0.532 (3.18)	0.638 (3.39)	0.351 (1.85)	0.282 (1.95)	0.786 (2.94)	0.136 (0.70)	0.132 (0.71)	0.078 (0.39)
mktvol	0.323 (2.44)	0.103 (0.60)	0.237 (1.52)	0.291 (1.99)	0.051 (0.28)	0.151 (0.75)	1.790 (6.74)	1.405 (5.60)	1.612 (5.95)
aggoi	0.027 (1.29)	−0.086 (−2.94)	−0.019 (−0.82)	0.068 (2.97)	0.025 (0.88)	0.041 (1.47)	−0.098 (−2.39)	−0.148 (−3.92)	−0.081 (−1.95)
Adj. R^2	59.44%	63.11%	71.15%	78.99%	54.55%	74.66%	52.58%	52.16%	52.41%

day-of-the-week dummies to see if short selling exhibits any calendar effect. Finally, we include market liquidity (*aggqp*), market volatility (*mktvol*), and aggregate order imbalance (*aggoi*) to determine if short selling varies with these market-level variables. We include up to five lags for our lagged variables because there are five trading days each week. Accordingly, we estimate Newey-West standard errors with five lags to control for possible autocorrelations in regression residuals.

The results in Table 2 reveal that daily aggregate shorting flows are highly persistent. All nine shorting measures show significant evidence of positive autocorrelations. The aggregate shorting on day $t-1$ has the greatest impact on shorting on day t, with regression coefficients ranging from 0.252 to 0.451. The corresponding t-statistics range from 2.77 to 9.37.

Estimated coefficients on day-of-the-week dummies reveal an interesting pattern. We find that shorting measures scaled by shares outstanding (i.e., *ss_vw*, *ss_ew*, and *ss_agg*) are significantly higher on Tuesday, Wednesday, and Thursday than on Monday. Unreported tests indicate that this result is driven by the higher overall trading activity on Tuesday through Thursday. Consistent with this explanation, we find that the other shorting measures, which are scaled by the overall trading activity, generally do not display any significant day-of-the-week pattern.

The results in Table 2 also reveal that aggregate short selling is positively related to market illiquidity and market volatility. Six (five) of the nine coefficients on *aggqp* (*mktvol*) are positive and statistically significant at 10% level or better. These results suggest that aggregate short

Table 3
Short sellers' reactions to individual stock returns and market returns.

We obtain data on short-sale transactions from various exchanges under Regulation SHO. Our sample period is from January 3, 2005 to June 29, 2007. Our sample includes only common stocks (those with a sharecode of 10 or 11 in CRSP). Aggregate short sale measures are described in Table 1. We estimate a regression of firm-level short selling on past and current firm-specific returns and market returns. The dependent variable, $relss_t$, is the relative short sale defined as the number of shares shorted on each stock divided by the total share volume on that day. ret_t is contemporaneous stock return. $ret_{t-5,\ t-1}$ is lagged 5-day stock return. $mktret_t$ is contemporaneous market return. $mktret_{t-5,\ t-1}$ is lagged 5-day market return. $retex_t$ is contemporaneous stock return in excess of the market return. $retex_{t-5,\ t-1}$ is lagged 5-day stock return in excess of lagged 5-day market return. Numbers in parentheses are t-statistics based on clustered standard errors. ***, **, * indicate statistical significance of 1%, 5%, and 10%, respectively.

	Dependent Variable: $relss_t$					
	(1)	(2)	(3)	(4)	(5)	(7)
$ret_{t-5,\ t-1}$	0.014***	0.018***				
	(5.83)	(7.36)				
ret_t		0.522***				
		(40.54)				
$mktret_{t-5,\ t-1}$			−0.121***	−0.119***	−0.123***	−0.106***
			(−15.49)	(−15.28)	(−15.71)	(−13.77)
$mktret_t$				−0.169***		−0.264***
				(−8.16)		(−12.56)
$retex_{t-5,\ t-1}$					0.025***	0.029***
					(9.42)	(10.67)
$retex_t$						0.583***
						(39.48)
$relss_{t-5,\ t-1}$	0.146***	0.146***	0.146***	0.146***	0.146***	0.146***
	(169.94)	(170.40)	(170.35)	(170.36)	(169.46)	(169.89)
R^2	28.29%	28.81%	28.29%	28.30%	28.30%	28.90%

selling is higher when the market is more illiquid and more volatile. We also find that aggregate shorting is weakly negatively related to aggregate order imbalance, indicating that aggregate shorting tends to be low when there is market-wide buying pressure.

Finally, our results indicate a generally negative relation between aggregate short selling and past market returns. For example, all regression coefficients on day $t-1$ or day $t-2$ market returns are negative, with nearly half of them (7 out of 18) statistically significant at the 10% level. This result suggests that short sellers in the aggregate trade with, not against, the market. In contrast to our negative relation, Diether, Lee, and Werner (2009) find a significantly positive relation between shorting and past returns at the individual stock level.

To reconcile our finding with that of Diether, Lee, and Werner (2009), we modify their analysis by decomposing individual stock return (*ret*) into two components: the market return (*mktret*) and the firm-specific return (*retex*), where *retex* is the difference between *ret* and *mktret*. This analysis allows us to examine how individual stock short sales respond to firm-specific returns and market returns within a unified framework. Following Diether, Lee and Werner (2009), we use the relative short ratio (shares shorted divided by shares traded) as the dependent variable.

Table 3 reports the results for this analysis. To better compare with Diether, Lee, and Werner (2009), we include only firm-level returns in regressions (1) and (2). In regressions (3) and (4), we include only market returns. Finally, in regressions (5) and (6) we include both market returns and firm-specific returns. Overall, we find that short sellers respond positively to past individual stock returns and negatively to past market returns. Thus, we are able to reconcile our finding with that of Diether, Lee, and Werner (2009). We show that short sellers respond to firm-specific returns and market returns in fundamentally different ways: they are contrarians with respect to firm-specific returns and momentum traders with respect to market returns.

One potential explanation for the above differential responses is that firm-specific returns are more likely to reverse whereas market returns are more likely to continue. In particular, Jegadeesh (1990) and Lehmann (1990) have presented evidence of short-term reversal in stock returns, while Lo and MacKinlay (1988) have documented that market returns tend to be positively serially correlated at short horizons. More important, short sellers' differential responses to past market- and stock-level returns could be consistent with their attempt to reduce arbitrage risk (Shleifer and Vishny, 1997). Although informed short sellers look for overvalued firms to short, they may be reluctant to implement their trades until they anticipate negative market returns. Intuitively, low anticipated market returns not only increase expected shorting profits but also reduce arbitrage risk.[9]

3. Commonality in daily shorting flows

3.1. Potential reasons for the existence of commonality in short selling

Commonality in short sales could arise from a number of sources. The first source is market-wide information. If short sellers possess superior information about future market returns and trade on the basis of such information, then individual stock short sales will co-move.

The second source is based on trading activity and liquidity. Chordia, Roll, and Subrahmanyam (2000) and Hasbrouck and Seppi (2001) have shown that there is commonality

[9]We thank the referee for suggesting this explanation.

Table 4
Commonality in short sales.

We obtain data on short-sale transactions from various exchanges under Regulation SHO. Our sample period is from January 3, 2005 to June 29, 2007. Our sample includes only common stocks (those with a sharecode of 10 or 11 in CRSP). Aggregate short sale measures are described in Table 1.We obtain Fama and French 48-industry classifications and risk-free returns from Kenneth French's website. We estimate the following regression in Panel A: $\Delta short_{i,j,t}=a+b\ \Delta short_{mkt,t}+c\ \Delta short_{ind,j,t}+d\ vwret_t+e_{i,j,t}$. $short_i$ is *sv*, *st*, or *ss*. $short_{mkt}$ and $short_{ind}$ are the market- and corresponding industry-aggregated shorting measures. We exclude stock *i* in the calculation of market and industry aggregate short sales. In Panel B, we augment the regression in Panel A by controlling for lagged and contemporaneous returns of stock *i*. Numbers in parentheses are *t*-statistics following the Fama and MacBeth (1973) approach. ***, **, * indicate statistical significance of 1%, 5%, and 10%, respectively.

	Short-Sale Measures								
	sv_vw	*sv_ew*	*sv_agg*	*st_vw*	*st_ew*	*st_agg*	*ss_vw*	*ss_ew*	*ss_agg*
Market Only									
b	0.928***	0.951***	0.779***	0.568***	0.964***	0.572***	0.788***	0.885***	0.773***
t (*b*)	(53.70)	(60.78)	(50.44)	(36.90)	(64.40)	(45.80)	(25.92)	(23.59)	(25.41)
Adj. R^2	5.05%	6.37%	6.04%	6.08%	6.21%	6.78%	2.60%	2.74%	2.59%
Industry Only									
c	0.577***	0.746***	0.528***	0.493***	0.759***	0.546***	0.684***	0.561***	0.837***
t (*c*)	(55.12)	(59.86)	(52.71)	(48.21)	(57.35)	(49.61)	(26.50)	(27.33)	(27.32)
Adj. R^2	5.20%	5.58%	5.50%	6.51%	5.84%	7.51%	4.11%	2.78%	4.20%
Market and Industry									
b	0.648***	0.650***	0.562***	0.321***	0.560***	0.233***	0.256***	0.547***	0.141***
t (*b*)	(33.68)	(29.73)	(33.35)	(15.69)	(23.37)	(13.81)	(7.73)	(14.37)	(3.66)
c	0.257***	0.293***	0.208***	0.291***	0.395***	0.366***	0.594***	0.358***	0.781***
t (*c*)	(23.05)	(16.96)	(20.15)	(22.47)	(18.26)	(23.90)	(19.67)	(14.68)	(20.32)
Adj. R^2	7.16%	7.25%	7.78%	8.71%	7.90%	9.63%	5.23%	4.09%	5.29%

in trading volume, liquidity, and order imbalances. If short sales are systematically related to these variables, then short sales in individual stocks may exhibit commonality.

The third source is related to the arbitraging activities. Short sales in individual stocks may have a common component if short sales are motivated by arbitrage (e.g., index arbitrage, convertible bond arbitrage, and takeover arbitrage) and such arbitrage activities exhibit commonality across stocks. Moreover, commonality in short selling may arise from limits-to-arbitrage concerns (Shleifer and Vishny, 1997). In particular, the intensity of short selling should be higher (lower) when the anticipated market returns are negative (positive).

The fourth source is more mechanical and is related to how short sellers respond to *past* stock returns or market returns. Because individual stock returns have common components, short sales in individual stocks may exhibit commonality if short sellers react systematically to past individual stock returns (Diether, Lee, and Werner, 2009). Similarly, short sales in individual stocks may have a common component if short sellers trade systematically in response to past market returns.

We emphasize that the above-mentioned sources of commonality are neither exhaustive nor mutually exclusive. In our empirical analysis below, we first establish that short sales exhibit commonality and then explore the potential sources of the commonality in short sales.

3.2. Commonality results

Following Chordia, Roll, and Subrahmanyam (2000), we estimate the following regression to test for commonality in daily shorting flows:

$$\Delta short_{i,j,t} = a + b\Delta short\_mkt_t + c\Delta short\_ind_{j,t} + dvwret_t + e_{i,t} \tag{2}$$

The dependent variable is the daily percentage change of shorting in stock i that belongs to industry j. The individual stock shorting measure (*short*) is *sv*, *st*, or *ss*. *Short_mkt* and *short_ind* are market- and industry-aggregated shorting measures, respectively. We use Fama-French 48-industry classifications and construct industry-aggregated shorting measures similarly to market-aggregated shorting measures as described in Section 2.2. Following Chordia, Roll, and Subrahmanyam (2000), we exclude stock i (the stock in the dependent variable) when constructing market- and industry-aggregated shorting flow measures. In addition, we control for contemporaneous market returns to account for the possibility that individual stock shorting flows might respond to market movements. We estimate regression Eq. (2) stock by stock. To be included in the estimation, a stock must have short-sale transactions in at least 100 trading days during our sample period. Using an approach analogous to Fama and MacBeth (1973), we report the average regression coefficient (across stocks) and compute t-statistics based on the cross-sectional standard deviation of regression coefficients.

Table 4 presents the results. For brevity, we only report the coefficient estimates on market- and industry-aggregated shorting flows and do not report the intercept or the coefficient on market returns. Table 4 reveals strong evidence of commonality in short selling. Short sales in individual stocks co-move significantly with both market- and industry-aggregated short sales. When used alone, a 1% increase in market short selling corresponds to an average increase in the individual stock short selling of 0.56–0.96%, and a 1% increase in industry short selling corresponds to an average increase in the individual stock short selling of 0.49–0.84%, depending on the shorting measure used. When used together, the coefficients on market- and industry-aggregated shorting both remain highly statistically significant.

While the regression coefficients are highly significant, we recognize that the R-squares of these regressions are fairly modest, ranging from 4.09% to 9.63% when both market- and industry-aggregated shorting measures are included. These numbers are similar to those reported in Chordia, Roll, and Subrahmanyam (2000) in their analysis of commonality in trading activities,[10] and are in fact quite sensible. After all, it is reasonable to expect individual stock short selling to be primarily driven by firm-specific information.

3.3. Sources of commonality in short sales

Diether, Lee, and Werner (2009) present evidence that individual stock short selling is significantly related to contemporaneous and lagged stock returns. To the extent that individual stock returns are correlated across stocks, commonality in short sales might be driven by commonality in stock returns. To address this concern, we augment regression Eq. (2) by also controlling for current and lagged stock returns. Unreported results indicate that the coefficients on market- and industry-aggregated short sales continue to be positive and highly significant.

Because we control for market returns in our regressions, our finding is not due to short sellers' common response to market returns. Nor is our finding attributable to commonality in the overall trading activity (Chordia, Roll, and Subrahmanyam, 2000) because six of our nine shorting measures are scaled by total trading volume or total number of trades.

To explore whether commonality in short sales is influenced by firm-level determinants of short selling, we follow the methodology of Chordia, Roll, and Subrahmanyam (2000) and perform an analysis of commonality while controlling for size, institutional ownership, book-to-market, and liquidity. Unreported results indicate that individual stock short selling remains significantly related to industry-aggregated short selling after controlling for these firm characteristics.

To explore whether the commonality in short selling is driven by the proliferation of hedge funds and their arbitrage activities, we repeat our analysis for subsamples of stocks sorted by S&P 500 Index membership and hedge fund ownership. Table 5 reports the results. We fail to find a consistent difference between S&P 500 stocks and non-S&P 500 stocks across our nine aggregate shorting measures. In contrast, we find that stocks with high hedge fund ownership exhibit consistently stronger commonality in short sales than stocks with low hedge fund ownership. Nevertheless, we continue to find significant evidence of commonality even among stocks with low hedge fund ownership, suggesting that arbitrage activities play only a partial role in explaining the commonality in short selling.

Using the same portfolio sorting methodology, we also investigate whether the extent of commonality in short sales varies with institutional ownership (IO) and liquidity (proxied by quoted spread) in Table 5. The results are qualitatively similar to those based on hedge fund ownership. We find that high-IO (low-quoted-spread) firms exhibit significantly stronger commonality in short sales than do low-IO (high-quoted-spread) firms. However, significant commonality exists even among low-IO (high-quoted-spread) firms. These findings suggest that commonality in short sales is related to, but not solely driven by, institutions and liquidity.[11]

As noted earlier, informed short sellers may be more likely to short individual stocks when they anticipate negative market returns. If this is true, we would expect commonality in short selling to be stronger during (anticipated) falling markets and among high market beta stocks.[12] To explore this possibility, we divide our sample period into “UP” and “DOWN” markets based on whether the 10-day-ahead market excess return is above or below the sample median. We also

[10] Chordia, Roll, and Subrahmanyam (2000) report R-squares between 5.7% and 10% in their Table 7.

[11] We also perform an analysis similar to Kamara, Lou, and Sadka (2008) and find consistent results.

[12] High beta stocks tend to decline more in falling markets.

Table 5
Commonality in short sales—by firm characteristics and market states.

We obtain data on short-sale transactions from various exchanges under Regulation SHO. Our sample period is from January 3, 2005 to June 29, 2007. Our sample includes only common stocks (those with a sharecode of 10 or 11 in CRSP). Aggregate short sale measures are described in Table 1. S&P 500 Index membership is from CRSP. Institutional ownership is from Thomson 13 F institutional holdings database. We obtain hedge fund ownership by merging the 13 F institutional holdings database and the CISDM hedge fund database. Quoted spread is from the TAQ database. Beta is estimated from regressions of daily stock returns on market returns over the past 12 months. We divide all stocks into two equal-size portfolios based on hedge fund ownership, institutional ownership, quoted spread, and market beta, respectively. We classify all days into "UP" or "DOWN" market based on whether the 10-day-ahead market excess return is above or below the median. We estimate the following regression: $\Delta short_{i,j,t}=a+b\,\Delta short_{mkt,t}+c\ vwret_t+e_{i,j,t}$. We exclude stock i in the calculation of aggregate short sales. For brevity, the table reports only the regression coefficient on $\Delta short_{mkt}$. ***, **, * indicate statistical significance of 1%, 5%, and 10%, respectively.

	Short-Sale Measures								
	sv_vw	*sv_ew*	*sv_agg*	*st_vw*	*st_ew*	*st_agg*	*ss_vw*	*ss_ew*	*ss_agg*
S&P 500 stocks	0.929***	0.618***	0.713***	0.931***	0.732***	0.653***	1.128***	1.036***	1.041***
Non-S&P 500 Stocks	0.924***	0.981***	0.784***	0.530***	0.982***	0.563***	0.720***	1.057***	0.834***
Difference	0.005	−0.363***	−0.071**	0.401***	−0.250***	0.090***	0.308***	−0.021	0.207*
High Hedge Fund Ownership	0.958***	1.010***	0.795***	0.660***	1.073***	0.668***	1.007***	1.152***	1.014***
Low Hedge Fund Ownership	0.789***	0.950***	0.691***	0.492***	0.962***	0.572***	0.599***	0.705***	0.598***
Difference	0.169***	0.060*	0.104***	0.168***	0.111***	0.096***	0.408***	0.347***	0.416***
High Institutional Ownership	1.038***	1.069***	0.872***	0.726***	1.156***	0.870***	1.199***	1.324***	1.166***
Low Institutional Ownership	0.773***	0.898***	0.674***	0.413***	0.896***	0.476***	0.277***	0.365***	0.326***
Difference	0.265***	0.171***	0.198***	0.313***	0.260***	0.394***	0.922***	0.959***	0.840***
High Quoted Spread	0.842***	1.012***	0.721***	0.420***	0.978***	0.505****	0.365***	0.525***	0.393***
Low Quoted Spread	1.032***	1.056***	0.839***	0.733***	1.123***	0.725***	1.213***	1.335***	1.172***
Difference	−0.190***	−0.044	−0.118**	−0.313***	−0.145***	−0.220***	−0.848***	−0.810***	−0.779***
High Beta	0.880***	0.950***	0.797***	0.911***	1.033***	0.681***	0.981***	1.103***	0.981***
Low Beta	0.815***	0.712***	0.709***	0.529***	0.807***	0.522***	0.404***	0.632***	0.437***
Difference	0.065	0.238*	0.087*	0.382***	0.226***	0.159***	0.577***	0.471***	0.545***
UP market	0.892***	0.957***	0.751***	0.523***	0.974***	0.549***	0.751***	0.891***	0.753***
DOWN market	0.965***	0.939***	0.814***	0.624***	0.949***	0.607***	0.826***	0.882***	0.791***
Difference	−0.072**	0.018	−0.063**	−0.101***	0.025	−0.058***	−0.074*	0.009	−0.037

sort our sample stocks into high- and low-beta portfolios. The results in Table 5 indicate modest evidence that commonality in short selling is stronger in "DOWN" markets than "UP" markets. In addition, we find strong evidence that commonality in short selling is more pronounced among high-beta stocks.

Overall, we find strong evidence that individual stock short sales exhibit commonality and this commonality persists after controlling for market returns, individual stock returns, trading volume, hedge fund ownership, institutional ownership, firm-level liquidity, beta, and market states. We interpret this evidence as suggesting that short sellers base their trades, at least in part, on market-wide information.

4. The predictive content of aggregate short selling

4.1. Baseline results

If short sellers possess and trade on the basis of superior market-level information, we would expect aggregate shorting to predict future market returns. To test this prediction, we estimate the following time series regression:

$$VWRET_{t+1,t+d} = a + b \times short_{t-4,t} + c \times VWRET_{t-d+1,t} + e_t, \tag{3}$$

where $VWRET_{t+1,t+d}$ is the cumulative CRSP value-weighted return in excess of the risk-free return over the following 5, 10, or 20 trading days, $short_{t-4,t}$ is the mean aggregate shorting over the previous five days based on each of our nine shorting measures, and $VWRET_{t-d+1,t}$ is the lagged cumulative CRSP value-weighted excess return over the previous 5, 10, or 20 days. Our choice of these horizons is motivated by Boehmer, Jones, and Zhang (2008), and by the fact that short sellers are short-term traders.[13] For ease of exposition, all explanatory variables are standardized to have unit standard deviation. We estimate regression (3) at the daily frequency and compute standard errors using the Newey and West (1987) procedure with the number of lags equal to the number of trading days in the forecasting period.

Panel A of Table 6 reports the results of regression model (3). For brevity, we report only the coefficient on lagged aggregate shorting and the associated *t*-statistic. Consistent with the hypothesis that high aggregate shorting predicts low market returns, we find that all 27 coefficient estimates (nine shorting measures at three forecasting horizons) are negative, with 17 of them statistically significant at the 10% level or better. The statistical significance of our results is quite remarkable considering our relatively short sample period of two and a half years. Our results are also economically meaningful. For example, a one standard deviation increase in *sv_ew* predicts a decrease in market excess return of 28.1 bps over the following 10 trading days (7% annualized).

The results in Panel A of Table 6 also show that the magnitudes of the coefficients increase monotonically as the forecasting horizon increases from 5 to 20 days. Statistically speaking, the results are the strongest at the 10-day horizon, suggesting that much of the information contained in aggregate shorting flows is incorporated into prices within 10 days.

We would like to note that our findings do not mechanically follow from Boehmer, Jones, and Zhang (2008) and Diether, Lee, and Werner (2009), who document informed shorting at the firm level using daily shorting data. We test whether high aggregate shorting predicts low market returns in the time series. In contrast, the analyses in Boehmer, Jones, and Zhang (2008) and

[13]Our results are qualitatively identical if we skip a day when computing future market returns.

Table 6
Predictive ability of market-aggregated short sales.

We obtain data on short-sale transactions from various exchanges under Regulation SHO. Our sample period is from January 3, 2005 to June 29, 2007. Our sample includes only common stocks (those with a sharecode of 10 or 11 in CRSP). Aggregate short sale measures are described in Table 1.We obtain three-month T-bill rate, 10-year T-bond yields, Aaa and Baa bond yields from Federal Reserve Bank of St. Louis' website. Term spread is the difference between 10-year T-bond yield and three-month T-bill rate. Default spread is the difference between Baa yield and Aaa yield. *aggqp* is the average quoted spread across all stocks. *mktvol* is the market volatility computed from 5-minute returns of the S&P 500 Index. *aggoi* is the aggregate order imbalance. In Panel A, we estimate the following predictive regression: $\text{vwret}_{t+1,\ t+d}=a+b\ \text{short}_{t-4,\ t}+c\ vwret_{t-d+1,\ t}+e_t$, where short is *sv_vw*, *st_vw*, *ss_vw*, *sv_ew*, *st_ew*, *ss_ew*, *sv_agg*, *st_agg*, or *ss_agg*. In Panel B, we estimate the following predictive regression: $vwret_{t+1,\ t+d}=a+b\ short_{t-4,\ t}+c\ vwret_{t-d+1,\ t}+d\ tbill_t+f\ term_t+g\ default_t+h\ aggqp_t+i\ mktvol_t+j\ aggoi_t+e_t$. Market-aggregated shorting variable, $short_{t-4,\ t}$, is scaled by its time-series standard deviation. Numbers in parentheses are *t*-statistics based on Newey-West standard errors with 5, 10, and 20 lags. ***, **, * indicate statistical significance of 1%, 5%, and 10%, respectively.

Panel A: Baseline Regression

	Regression Coefficient (*t*-stat) on Aggregate Short-sale Measures								
	sv_vw	*sv_ew*	*sv_agg*	*st_vw*	*st_ew*	*st_agg*	*ss_vw*	*ss_ew*	*ss_agg*
5-day-ahead return									
b	−0.083	−0.123*	−0.096	−0.125*	−0.137*	−0.105	−0.029	−0.207***	−0.100
t (*b*)	(−1.05)	(−1.75)	(−1.33)	(−1.71)	(−1.85)	(−1.57)	(−0.40)	(−2.87)	(−1.36)
10-day-ahead return									
b	−0.190*	−0.281**	−0.187**	−0.245**	−0.289**	−0.202**	−0.111	−0.361***	−0.196
t (*b*)	(−1.83)	(−2.61)	(−1.99)	(−2.44)	(−2.59)	(−2.30)	(−0.94)	(−3.00)	(−1.64)
20-day-ahead return									
b	−0.257	−0.410**	−0.287*	−0.392**	−0.350**	−0.255*	−0.140	−0.533**	−0.262
t (*b*)	(−1.54)	(−2.44)	(−1.84)	(−2.28)	(−2.07)	(−1.76)	(−0.57)	(−2.23)	(−0.99)

Panel B: Controlling for Macroeconomic and Market Variables

	Regression Coefficient (*t*-stat) on Aggregate Short-sale Measures								
	sv_vw	*sv_ew*	*sv_agg*	*st_vw*	*st_ew*	*st_agg*	*ss_vw*	*ss_ew*	*ss_agg*
5-day-ahead return									
b	−0.288*	−0.328**	−0.376**	−0.386**	−0.300**	−0.445**	−0.051	−0.239**	−0.098
t (*b*)	(−1.91)	(−2.39)	(−2.39)	(−2.36)	(−2.21)	(−2.55)	(−0.48)	(−2.29)	(−0.94)

10-day-ahead return									
b	−0.400**	−0.542***	−0.478**	−0.508**	−0.480**	−0.575**	−0.105	−0.328**	−0.121
t (*b*)	(−2.17)	(−2.89)	(−2.37)	(−2.39)	(−2.57)	(−2.65)	(−0.81)	(−2.48)	(−0.98)
20-day-ahead return									
b	−0.377	−0.555**	−0.506*	−0.598**	−0.472*	−0.610**	−0.102	−0.357*	−0.066
t (*b*)	(−1.41)	(−2.16)	(−1.81)	(−1.98)	(−1.92)	(−2.00)	(−0.53)	(−1.78)	(−0.32)

Diether, Lee, and Werner (2009) are cross-sectional in nature (i.e., they test whether heavily shorted stocks outperform lightly shorted stocks in the cross-section). Indeed, to the extent that heavily shorted and lightly shorted stocks have similar market betas, the returns to the hedge portfolio in these papers are essentially insulated from market fluctuations.

4.2. Control for macroeconomic and market variables

It is possible that the documented predictive ability is attributable to macroeconomic variables that have been shown to predict market returns (e.g., Fama and French, 1989; and Ferson and Harvey, 1999). To test this possibility, we augment regression Eq. (3) with the following control variables: the three-month T-bill rate, the term spread between the 10-year T-bond yield and the three-month T-bill rate, and the default spread between the Baa yield and the Aaa yield.[14] In addition, we also control for market volatility, market liquidity (proxied by average quoted spread), and aggregate order imbalance to account for the possibility that aggregate shorting might predict market returns through these market-level variables.

The results in Panel B of Table 6 indicate that our findings are robust to the above-mentioned macroeconomic and market-level control variables. In fact, both the economic and statistical significance of our results are higher after controlling for these variables. All 27 regression coefficients on lagged aggregate shorting are negative, with 20 of them being statistically significant at the 10% percent or better (compared to 17 in the baseline regressions). These findings suggest that short sellers are not merely acting on publicly available macroeconomic or market-level information.

4.3. Sources of informational advantage

Short sellers might be able to predict market returns because they possess superior information about the state of the economy. Alternatively, short sellers might be informed about market-wide mispricing. There is growing evidence that stock prices can deviate from fundamentals.[15] If mispricing is market-wide, i.e., the entire market may be overvalued or undervalued, then aggregate short selling might predict future market returns. To explore the nature and sources of short sellers' informational advantage, we perform additional analyses.

4.3.1. Aggregate earnings news

Our first test examines whether aggregate short selling predicts future aggregate earnings news. We classify firm-level earnings news based on the cumulative abnormal returns around each earnings announcement. Specifically, we classify an earnings announcement as good news if the cumulative abnormal return over the three days around the announcement is greater than 2% and as bad news if the cumulative abnormal return is below −2%. We define the aggregate earnings news as the difference between the number of good news and number of bad news scaled by the total number of earnings announcements over a specific time period. We then regress future aggregate earnings news on lagged aggregate short selling. In addition to 20-day-

[14] Another widely used market return predictor is dividend yield. We do not control for dividend yield in our regression because dividend yields are not available at daily frequency. Unreported results indicate that using low-frequency dividend yield does not change our basic results.

[15] See Hirshleifer (2001), Barberis and Thaler (2003), Baker and Wurgler (2012), and the references therein.

Table 7
Predictive ability of market-aggregated short sales for future aggregate earnings news, macroeconomic news, and investor sentiment.

We obtain data on short-sale transactions from various exchanges under Regulation SHO. Our sample period is from January 3, 2005 to June 29, 2007. Our sample includes only common stocks (those with a sharecode of 10 or 11 in CRSP). Aggregate short sale measures are described in Table 1. We obtain the following macroeconomic announcement from Bloomberg: non-farm payrolls, unemployment rate, retail sales, consumer price index, industrial production, durable goods, jobless claims, and gross domestic product. We obtain monthly investor sentiment data from Jeff Wurgler's website. To construct the aggregate earnings news measure, we first compute the cumulative abnormal return over the three days around each earnings announcement. We then classify an earnings announcement as a good news if the abnormal return is greater than 2%, and as a bad news if the abnormal return is below -2%. We define aggregate earnings news as the difference between the number of good news and number of bad news scaled by the total number of earnings announcements over a specific time period. We code a macroeconomic announcement as 1 (i.e., good news) if the actual announcement exceeds the expectation, as 0 (i.e., neutral news) if the actual announcement is the same as the expectation, and as -1 (i.e., bad news) otherwise. We construct an aggregate macroeconomic news measure by summing all announcements over a certain time period and then scale by the total number of announcements. We estimate the following regression: $AggEarningsNews_{t+1,\ t+n}$ (or $MacroEconNews_{t+1,\ \mathrm{t+n}}$ or $Sentiment_{t+1}$)$=a+b\ short_{t-4,\ t}+c\ vwret_{t-n+1,\ t}+e_{i,t}$, where *short* is *sv_vw*, *st_vw*, *ss_vw*, *sv_ew*, *st_ew*, *ss_ew*, *sv_agg*, *st_agg*, or *ss_agg*. Numbers in parentheses are *t*-statistics based on Newey-West standard errors with 20 and 60 lags. ***, **, * indicate statistical significance of 1%, 5%, and 10%, respectively.

Panel A: Aggregate Earnings News

	Short-sale Measures								
	sv_vw	*sv_ew*	*sv_agg*	*st_vw*	*st_ew*	*st_agg*	*ss_vw*	*ss_ew*	*ss_agg*
20-day-ahead Aggregate Earnings News									
b	0.031	0.028	−0.136**	−0.055	−0.004	−0.090**	−0.062	0.012	−0.109**
t (*b*)	(0.73)	(0.44)	(−2.39)	(−1.46)	(−0.09)	(−1.98)	(−1.62)	(0.22)	(−2.15)
60-day-ahead Aggregate Earnings News									
b	−0.001	−0.013	−0.090***	−0.058*	−0.028	−0.073***	−0.070**	−0.027	−0.086***
t (*b*)	(−0.01)	(−0.24)	(−3.63)	(−1.93)	(−0.78)	(−3.19)	(−2.57)	(−0.60)	(−3.68)

Panel B: Macroeconomic News

	Short-sale Measures								
	sv_vw	*sv_ew*	*sv_agg*	*st_vw*	*st_ew*	*st_agg*	*ss_vw*	*ss_ew*	*ss_agg*
20-day-ahead Macroeconomic News									
b	−0.023	−0.042*	−0.020	−0.042*	−0.032	−0.035	−0.044	−0.061**	−0.035
t (*b*)	(−1.07)	(−1.74)	(−0.90)	(−1.90)	(−1.23)	(−1.55)	(−1.43)	(−2.07)	(−1.05)

Table 7 (*continued*)

60-day-ahead Macroeconomic News									
b	−0.027	−0.044*	−0.042**	−0.049***	−0.038	−0.040*	−0.007	−0.038*	−0.023
t (b)	(−1.62)	(−1.87)	(−2.06)	(−2.90)	(−1.45)	(−1.95)	(−0.32)	(−1.90)	(−0.99)

Panel C: Investor Sentiment

	Short-sale Measures								
	sv_vw	*sv_ew*	*sv_agg*	*st_vw*	*st_ew*	*st_agg*	*ss_vw*	*ss_ew*	*ss_agg*
Baseline Regression									
b	−0.150***	−0.138***	−0.171***	−0.171***	−0.118***	−0.165***	0.000	−0.027	−0.009
t (b)	(−5.88)	(−5.78)	(−8.72)	(−8.17)	(−4.19)	(−6.70)	(0.01)	(−1.15)	(−0.37)
Controlling for Lagged Sentiment									
b	−0.001	−0.001*	−0.003*	−0.002	−0.001	−0.001	0.001	0.001	0.001
t (b)	(−1.43)	(−1.69)	(−1.72)	(−1.53)	(−1.07)	(−0.93)	(0.70)	(0.45)	(1.08)

ahead aggregate earnings news, we also predict 60-day-ahead aggregate earnings news because firm-level earnings are reported on a quarterly basis.

Panel A of Table 7 present the results for this analysis. We find modest evidence of an inverse relation between aggregate short selling and future aggregate earnings news. Most of the regression coefficients on lagged aggregate short selling (15 out of 18) are negative, and 8 of them are statistically significant at the 10% level or better. This finding suggests that short sellers are informed about future market earnings. That is, when aggregate short selling is high, future aggregate earnings tends to be unexpectedly low.

4.3.2. Macroeconomic news announcements

We next explore whether short sellers possess superior information about future economy-wide activities as measured by the following macroeconomic news announcements: unemployment rate, non-farm payrolls, industrial production, retail sales, jobless claims, durable goods, consumer price index, and gross domestic product. We code an announcement as 1 (i.e., good news) if the actual announcement exceeds the expectation of professional economists, as −1 (i.e., bad news) if the actual announcement falls short of the expectation, and as 0 (i.e., neutral news) otherwise. We construct an aggregate macroeconomic news measure by summing all announcements and then scale this sum by the total number of announcements. We then regress the 20- or 60-day-ahead macroeconomic news on lagged aggregate short selling. Panel B of Table 7 presents the results. We find that all regression coefficients on lagged aggregate short selling are negative, and nearly half of them are statistically significant. This finding is consistent with the idea that short sellers possess superior information about the aggregate economy.

4.3.3. Investor sentiment

Previous studies find strong evidence that investor sentiment impacts market valuation and subsequent returns. Brown and Cliff (2005) and Baker and Wurgler (2007) show that high investor sentiment is associated with market overvaluation and low subsequent market returns. To provide evidence on whether the predictive ability of short sellers for market returns is related to market-wide mispricing, we regress one-month-ahead investor sentiment on lagged aggregate short selling. We estimate two specifications: one controls for lagged investor sentiment and the other does not. The results are reported in Panel C of Table 7. When not controlling for lagged investor sentiment, we find significant evidence of an inverse relation between lagged aggregate short selling and future investor sentiment. After controlling for lagged sentiment, this relation is significantly weakened but remains negative. Overall, we find some evidence that high aggregate short selling predicts low investor sentiment.

4.4. Liquidity provision hypothesis

Informed trading is not the only explanation consistent with our findings. Short sellers as a group may act as liquidity providers when there is market-wide buying pressure. Chakrabarty, Moulton and Shkilko (2012), for example, present evidence that the majority of short sales are initiated by the buyers, which suggests that short sellers are more likely to provide liquidity than demand liquidity. If short sellers are compensated for providing liquidity, then we should expect an inverse relation between aggregate shorting and subsequent market returns.

We explore the liquidity provision hypothesis by performing two tests. In the first test, we examine the relation between aggregate order imbalance and the level of aggregate short selling. The liquidity provision hypothesis predicts that elevated short selling should coincide with

Table 8
Market-aggregated short sales and aggregate order imbalance.

We obtain data on short sale transactions from various exchanges under Regulation SHO. Our sample period is from January 3, 2005 to June 29, 2007. Our sample includes only common stocks (those with a sharecode of 10 or 11 in CRSP). Aggregate short sale measures are described in Table 1. Aggregate order imbalance is defined as the difference between buyer-initiated share volume and seller-initiated share volume scaled by the total share volume, and is computed from TAQ data. Numbers in parentheses are *t*-statistics. ***, **, * indicate statistical significance of 1%, 5%, and 10%, respectively.

	Aggregate Short-sale Measures								
Aggregate Order Imbalance Quintile	*sv_vw*	*sv_ew*	*sv_agg*	*st_vw*	*st_ew*	*st_agg*	*ss_vw*	*ss_ew*	*ss_agg*
q1- low	0.242	0.256	0.269	0.259	0.274	0.309	0.187	0.212	0.204
q2	0.242	0.255	0.271	0.263	0.274	0.316	0.185	0.208	0.202
q3	0.242	0.250	0.269	0.261	0.272	0.313	0.180	0.200	0.196
q4	0.236	0.242	0.262	0.256	0.267	0.307	0.176	0.195	0.191
q5- high	0.248	0.251	0.276	0.274	0.276	0.326	0.183	0.201	0.199
q5-q1	0.006**	−0.005*	0.007**	0.015***	0.002	0.017***	−0.004	−0.011***	−0.005
	(2.56)	(−1.67)	(2.31)	(4.12)	(0.49)	(4.03)	(−1.03)	(−2.90)	(−1.25)

Table 9
Liquidity provision and the predictive ability of market-aggregated short sales.

We obtain data on short-sale transactions from various exchanges under Regulation SHO. Our sample period is from January 3, 2005 to June 29, 2007. Our sample includes only common stocks (those with a sharecode of 10 or 11 in CRSP). Aggregate short sale measures are described in Table 1. Aggregate order imbalance and market liquidity are computed from TAQ data. Aggregate order imbalance is defined as the difference between buyer-initiated share volume and seller-initiated share volume scaled by the total share volume. Average quoted spread is value-weighted across all stocks. We estimate the following predictive regression for the sample that excludes the top quintiles for aggregate order imbalance or quoted spread over the past five days: $vwret_{t+1,\ t+10} = a + b\ short_{t-4,\ t} + c\ vwret_{t-9,\ t} + e_t$, where short is *sv_vw*, *st_vw*, *ss_vw*, *sv_ew*, *st_ew*, *ss_ew*, *sv_agg*, *st_agg*, or *ss_agg*. Market-aggregated shorting variable, $short_{t-4,\ t}$, is scaled by its time series standard deviation. Numbers in parentheses are *t*-statistics based on Newey-West standard errors with 10 lags. ***, **, * indicate statistical significance of 1%, 5%, and 10%, respectively.

	Regression Coefficient (*t*-stat) on Aggregate Short-sale Measures								
Sorting Variable	*sv_vw*	*sv_ew*	*sv_agg*	*st_vw*	*st_ew*	*st_agg*	*ss_vw*	*ss_ew*	*ss_agg*
Order Imbalance									
b	−0.208	−0.280**	−0.201	−0.282**	−0.305**	−0.228*	−0.201*	−0.409***	−0.289**
t (*b*)	(−1.46)	(−2.21)	(−1.49)	(−2.04)	(−2.49)	(−1.84)	(−1.67)	(−3.66)	(−2.36)
Quoted Spread									
b	−0.306*	−0.381***	−0.316**	−0.365**	−0.373***	−0.325**	−0.207*	−0.408***	−0.252*
t (*b*)	(−1.87)	(−2.71)	(−2.10)	(−2.31)	(−2.73)	(−2.35)	(−1.67)	(−3.57)	(−1.86)

Table 10
Predictive ability of market-aggregated short sales – controlling for total sell volume.

We obtain data on short-sale transactions from various exchanges under Regulation SHO. Our sample period is from January 3, 2005 to June 29, 2007. Our sample includes only common stocks (those with a sharecode of 10 or 11 in CRSP). Aggregate short sale measures are described in Table 1.We estimate the following predictive regression: $vwret_{t+1,\ t+d}=a+b\ short_{t-4,\ t}+c\ sellvolume_{t-4,\ t}+d\ vwret_{t-d+1,\ t}+e_t$. Market-aggregated shorting variable, $short_{t-4,\ t}$, is scaled by its time-series standard deviation. *sellvolume* is total sell volume scaled by total trading volume. Numbers in parentheses are *t*-statistics based on Newey-West standard errors with 5, 10, and 20 lags. ***, **, * indicate statistical significance of 1%, 5%, and 10%, respectively.

	Regression Coefficient on Aggregate Short Sells (*b*) and Total Sells (*c*)								
	sv_vw	*sv_ew*	*sv_agg*	*st_vw*	*st_ew*	*st_agg*	*ss_vw*	*ss_ew*	*ss_agg*
5-day-ahead return									
b	−0.082	−0.135*	−0.089	−0.115	−0.148**	−0.093	−0.052	−0.244***	−0.125*
t (*b*)	(−1.03)	(−1.94)	(−1.22)	(−1.53)	(−2.03)	(−1.33)	(−0.71)	(−3.32)	(−1.70)
c	0.117	0.139	0.106	0.093	0.135	0.092	0.137	0.209**	0.161
t (*c*)	(1.10)	(1.32)	(0.98)	(0.85)	(1.29)	(0.83)	(1.27)	(1.96)	(1.50)
10-day-ahead return									
b	−0.187*	−0.300***	−0.175*	−0.229**	−0.309***	−0.185*	−0.143	−0.412***	−0.229**
t (*b*)	(−1.76)	(−2.82)	(−1.78)	(−2.10)	(−2.82)	(−1.88)	(−1.28)	(−3.43)	(−1.99)
c	0.174	0.225	0.158	0.123	0.217	0.130	0.226	0.321**	0.250
t (*c*)	(1.05)	(1.43)	(0.94)	(0.71)	(1.37)	(0.73)	(1.46)	(2.12)	(1.61)
20-day-ahead return									
b	−0.259	−0.407**	−0.296*	−0.415**	−0.346**	−0.278*	−0.134	−0.555**	−0.262
t (*b*)	(−1.52)	(−2.53)	(−1.80)	(−2.21)	(−2.15)	(−1.68)	(−0.59)	(−2.46)	(−1.06)
c	−0.094	−0.028	−0.124	−0.179	−0.050	−0.157	−0.038	0.111	0.001
t (*c*)	(−0.33)	(−0.10)	(−0.42)	(−0.59)	(−0.18)	(−0.50)	(−0.16)	(0.48)	(0.00)

market-wide buying pressure. To test this prediction, we first construct daily aggregate order imbalance as the difference between buyer-initiated and seller-initiated share volume across all stocks, scaled by aggregate share volume. Next, we sort all trading days into quintiles based on aggregate order imbalance and then compute the average aggregate shorting flows for each quintile. Finally, we compare the difference between the two extreme quintiles of aggregate order imbalance.

Results in Table 8 do not show a consistent relation between the level of short selling and aggregate order imbalance. The relation is positive for five shorting measures and negative for the remaining four measures. Thus, we find little evidence that short selling is positively related to buy-order imbalance at the market level.

In our second test, we re-estimate our predictive regression Eq. (3) by excluding days when the aggregate order imbalance or market illiquidity is the highest. The intuition behind this test is that if liquidity provision is driving our results, we should find that the predictive ability of aggregate shorting dissipates or disappears once we exclude days when the buying pressure is the greatest. The rationale for using market liquidity as a sorting variable is that liquidity provision should be more valuable, and hence should receive greater compensation when market liquidity is poorer. We use the average quoted spread across all stocks as our measure of market liquidity and our results are robust to alternative measures such as effective spread and quoted depth.

Table 9 presents the regression results. Overall, they are similar to the full sample results reported in Table 6. For example, the coefficient on lagged aggregate shorting is negative and statistically significant in 7 out of 9 regressions after excluding days in the top order imbalance quintile. The results are even stronger when we exclude the top market illiquidity quintiles. Thus, our results are robust to the exclusion of days when the impact of liquidity provision is expected to be the greatest, suggesting that the predictive ability of aggregate shorting flows is not driven by liquidity provision.

Overall, our tests provide little support for the liquidity provision hypothesis. There is little evidence that the level of aggregate shorting is positively related to the aggregate order imbalance. More importantly, the predictability of aggregate shorting flows remains significant after controlling for aggregate buying pressure or market liquidity.

4.5. *Short sellers versus regular sellers*

Following Boehmer, Jones, and Zhang (2008), we next investigate whether short sellers are more or less informed than regular sellers. Diamond and Verrecchia (1987) contend that investors do not short for liquidity reasons and hence short sellers should be more informed than regular sellers. To test this prediction, we first construct a measure of total selling as the total seller-initiated volume scaled by the total trading volume from TAQ data. We then re-estimate our predictive regressions by including both short selling and total selling.

Table 10 reports the regression coefficients on short sells as well as total sells. We find that lagged aggregate short selling remains a negative predictor of future market returns after controlling for total sells. All 27 regression coefficients are negative and 18 of them are statistically significant. In contrast, the regression coefficients on total selling are mostly positive and insignificant. Taken together, our results suggest that the market-level information possessed by short sellers is not shared by other regular sellers.

4.6. The predictive ability of aggregate short interest ratios

Until recently, much of the short selling research uses short interest data. To examine if our findings based on the daily shorting flow data extend to the monthly short interest data, we next investigate whether changes in aggregate short interest ratios predict future market returns during our sample period. We construct two aggregate short interest ratio variables: the first measure (*SIS*) scales the aggregate short interest by aggregate shares outstanding while the second measure (*SIV*) scales the aggregate short interest by aggregate trading volume. We then regress one-month-ahead market returns on changes in aggregate short interest ratios.

We report the results for this analysis in Table 11. We fail to find any evidence that changes in aggregate short interest ratios predict future market returns. We note that the lack of significant result is not driven by our small sample size. In two of the six regression specifications, the *t*-statistics are merely −0.01 and −0.11. In the other four specifications, the regression coefficients are of the wrong sign.

Short sellers are short-term investors. For example, Geczy, Musto, and Reed (2002) find that the median duration of a position in the equity lending market is just three days. Diether (2008), using a proprietary data set, shows that shorting contracts have a median duration of only seven trading days. The implication of this short investment horizon is that monthly short interest data may be too coarse to fully capture the actions of short sellers. In particular, changes in monthly short interest do not reflect short transactions initiated and closed within the month (round-trip trades). Furthermore, monthly short interest cannot capture the exact timing of short selling activities. By contrast, daily shorting flows are much finer and permit a study of short-term trading strategies.

Table 11
Predictive ability of aggregate short interest ratios.

We obtain short interest data from Compustat. Market returns, share volume, and shares outstanding are from CRSP. Our sample includes only common stocks (those with a sharecode of 10 or 11 in CRSP). *SIV* is the ratio between aggregate short interest and aggregate monthly trading volume. *SIS* is the ratio between aggregate short interest and aggregate shares outstanding. We obtain T-bill rate (*TBILL*), default premium (*DEF*), and term premium (*TERM*) from Amit Goyal's website. *VWRET* is the value-weighted CRSP index returnin excess of the risk-free return. We regress 20-day-ahead market excess returns on changes in aggregate short interest ratios and control variables. Numbers in parentheses are *t*-statistics. ***, **, * indicate statistical significance of 1%, 5%, and 10%, respectively.

	Dependent Variable: 20-day-ahead Market Return ($VWRET_{t,t+20}$)					
	(1)	(2)	(3)	(4)	(5)	(6)
ΔSIS_t	−0.074	−0.565	2.331			
	(−0.01)	(−0.11)	(0.44)			
ΔSIV_t				0.250	0.250	0.326
				(1.00)	(0.99)	(1.39)
$VWRET_{t-20,t}$		−0.070	−0.105		−0.064	−0.151
		(−0.31)	(−0.48)		(−0.30)	(−0.74)
$TBILL_t$			−0.017			−0.017
			(−0.84)			(−0.94)
DEF_t			0.174**			0.174**
			(2.41)			(2.51)
$TERM_t$			−0.006			−0.009
			(−0.21)			(−0.31)
R^2	0.00%	0.35%	24.13%	3.48%	3.81%	29.24%

Overall, we find no evidence that the low-frequency aggregate short interest ratios have any predictive power for future market returns during our sample period. This result is in stark contrast to our finding based on daily aggregate short selling and highlights the importance of using high-frequency data to detect informed shorting.

5. Conclusions

Markets aggregate information from economic agents and impound it into prices. This information can be firm-specific or market-wide. In this paper, we study whether an important group of market participants, short sellers, possess and trade on market-wide information. We provide a comprehensive analysis of the information content of aggregate short selling by using the Regulation SHO database of short-sale transactions. We extend the short selling literature, which has thus far focused on individual stock short sales, to the market level and provide fresh insights into the informativeness of aggregate short selling.

We document several important results. First, we find strong evidence of commonality in daily individual stock shorting flows. This result suggests that short sales are not entirely firm-specific and that short sellers also act on market-wide information. Second, and more important, we find that short sellers are proficient at anticipating short-run market movements. This predictability is not attributable to publicly available macroeconomic information or due to compensation for liquidity provision. Third, we explore the sources of short sellers' superior market-level information and find evidence that aggregate shorting predicts future aggregate earnings news, macroeconomic news announcements, and investor sentiment. Fourth, we find no evidence that the low-frequency aggregate short interest ratios have any predictive power for future market returns during our sample period. This finding highlights the importance of using high-frequency data to detect informed shorting.

We emphasize that because daily shorting flow data are not publicly available, one cannot devise a trading strategy based on the return predictability we document in this paper. In addition, all returns reported in this study are gross returns; they do not account for various transaction costs including the cost of short selling. The objective of our study is not to show that investors can realize abnormal returns based on aggregate shorting flows. Rather, we are interested in the question of whether or not short sellers are informed about market returns. From this perspective, the documented return predictability can be interpreted as evidence of superior information possessed by short sellers.

Overall, we present evidence that the information set that short sellers use to make trades includes both firm-level information and market-level information. Our finding has important implications to academics, practitioners, and regulators. In particular, our results suggest that regulations designed to constrain short selling will likely adversely affect price efficiency not only at the individual stock level, but also at the market level.

References

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

Albuquerque, R., De Francisco, R., Marques, L., 2008. Marketwide private information in stocks: forecasting currency returns. Journal of Finance 63, 2297–2343.

Asquith, P., Meulbroek, L., 1995. An empirical investigation of short interest. Working paper, MIT.

Asquith, P., Pathak, P., Ritter, J.R., 2005. Short interest, institutional ownership, and stock returns. Journal of Financial Economics 78, 243–276.
Baker, M., Wurgler, J., 2000. The equity share in new issues and aggregate stock returns. Journal of Finance 55, 2219–2257.
Baker, M., Wurgler, J., 2007. Investor sentiment in the stock market. Journal of Economic Perspectives 21, 129–151.
Baker, M., Wurgler, J., 2012. Behavioral corporate finance: an updated survey. In: Constantinides, G.M., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance. Elsevier Science, North-Holland, Amsterdam and Boston.
Barberis, N., Thaler, R., 2003. A survey of behavioral finance. In: Constantinides, G.M., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance. Elsevier Science, North-Holland, Amsterdam and Boston.
Beber, A., Brandt, M., Kavajecz, K., 2011. What does equity sector order flow tell us about the economy? Review of Financial Studies 24, 3688–3730.
Boehmer, E., Jones, C.M., Zhang, X., 2008. Which shorts are informed? Journal of Finance 63, 491–527.
Boehmer, E., Jordan, B., Huszar, Z.R., 2010. The good news in short interest. Journal of Financial Economics 96, 80–97.
Boehmer, E., Wu, J.J.. Short selling and the informational efficiency of prices. Review of Financial Studies, forthcoming.
Bollen, N.P., Busse, J.A., 2001. On the timing ability of mutual fund managers. Journal of Finance 56, 1075–1094.
Brown, G.W., Cliff, M.T., 2005. Investor sentiment and asset valuation. Journal of Business 78, 405–440.
Campbell, J.Y., Grossman, S.J., Wang, J., 1993. Trading volume and serial correlation in stock returns. Quarterly Journal of Economics 108, 905–939.
Chakrabarty, B., Moulton, P.C., Shkilko, A., 2012. Short sales, long sales, and the Lee-Ready trade classification algorithm revisited. Journal of Financial Markets 15, 467–491.
Chordia, T., Roll, R., Subrahmanyam, A., 2000. Commonality in liquidity. Journal of Financial Economics 56, 3–28.
Chordia, T., Roll, R., Subrahmanyam, A., 2002. Order imbalance, liquidity and market returns. Journal of Financial Economics 65, 111–130.
Christophe, S.E., Ferri, M.G., Angel, J., 2004. Short-selling prior to earnings announcements. Journal of Finance 59, 1845–1875.
Christophe, S.E., Ferri, M.G., Hsieh, J., 2010. Informed trading before analyst downgrades: evidence from short sellers. Journal of Financial Economics 95, 85–106.
Cohen, L., Diether, K.B., Malloy, C.J., 2007. Supply and demand shifts in the shorting market. Journal of Finance 62, 2061–2096.
D'Avolio, G., 2002. The market for borrowing stock. Journal of Financial Economics 66, 271–306.
Dechow, P., Hutton, A., Meulbroek, L., Sloan, R.G., 2001. Short-sellers, fundamental analysis, and stock returns. Journal of Financial Economics 61, 77–106.
Desai, H., Ramesh, K., Thiagarajan, S.R., Balachandran, B.V., 2002. An Investigation of the informational role of short interest in the Nasdaq market. Journal of Finance 57, 2263–2287.
Diamond, D.W., Verrecchia, R.E., 1987. Constraints on short-selling and asset price adjustment to private information. Journal of Financial Economics 18, 277–311.
Diether, K.B., 2008. Short selling, timing and profitability. Working paper, Dartmouth College.
Diether, K.B., Lee, K., Werner, I.M., 2009. Short-sale strategies and return predictability. Review of Financial Studies 22, 575–607.
Fama, E.F., French, K.R., 1989. Business conditions and expected stock returns. Journal of Financial Economics 25, 23–50.
Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: empirical tests. Journal of Political Economy 81, 607–636.
Ferson, W.E., Harvey, C.R., 1999. Conditioning variables and the cross section of stock returns. Journal of Finance 54, 1325–1360.
Geczy, C.C., Musto, D.K., Reed, A.V., 2002. Stocks are special too: an analysis of the equity lending market. Journal of Financial Economics 66, 241–269.
Hanna, M., 1968. Short interest: bullish or bearish?—Comment. Journal of Finance 23, 520–523.
Hasbrouck, J., Seppi, D.J., 2001. Common factors in prices, order flows, and liquidity. Journal of Financial Economics 59, 383–411.
Hirshleifer, D., 2001. Investor psychology and asset pricing. Journal of Finance 56, 1533–1597.
Howe, J., Unlu, E., Yan, X.S., 2009. The predictive content of aggregate analyst recommendations. Journal of Accounting Research 47, 799–821.
Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. Journal of Finance 45, 881–898.
Jones, C.M., Kaul, G., Lipson, M.L., 1994. Transactions, volume, and volatility. Review of Financial Studies, 631–651.

Jones, C.M., Lamont, O.A., 2002. Short-sale constraints and stock returns. Journal of Financial Economics 66, 207–239.
Kamara, A., Lou, X., Sadka, R., 2008. The divergence of liquidity commonality in the cross-section of stocks. Journal of Financial Economics 89, 444–466.
Karpoff, J.M., Lou, X., 2010. Short sellers and financial misconduct. Journal of Finance 65, 1879–1913.
Lakonishok, J., Lee, I., 2001. Are insider trades informative? Review of Financial Studies 14, 79–111.
Lamont, O.A., Stein, J.C., 2004. Aggregate short interest and market valuations. American Economic Review 94, 29–32.
Lehmann, B.N., 1990. Fads, martingales, and market efficiency. Quarterly Journal of Economics 105, 1–28.
Lo, A.W., MacKinlay, A.C., 1988. Stock market prices do not follow a random walk: evidence from a simple specification test. Review of Financial Studies 1, 41–66.
Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.
Ofek, E., Richardson, M., Whitelaw, R.F., 2004. Limited arbitrage and short sales restrictions: evidence from the options markets. Journal of Financial Economics 74, 305–342.
Seneca, J, 1967. Short interest: bearish or bullish? Journal of Finance 22, 67–70.
Seyhun, H.N., 1988. The information content of aggregate insider trading. Journal of Business 61, 1–24.
Seyhun, H.N., 1992. Why does aggregate insider trading predict future stock returns? Quarterly Journal of Economics 107, 1303–1331.
Shleifer, A., Vishny, R., 1997. The limits of arbitrage. Journal of Finance 52, 35–53.
Veldkamp, L., Wolfers, J., 2007. Aggregate shock or aggregate information? Costly information and business cycle comovement. Journal of Monetary Economics 54, 37–55.