Are retail traders compensated for providing liquidity?[☆]

Jean-Noel Barrot[a,d,*], Ron Kaniel[b,c,d], David Sraer[e,f]

[a]Massachusetts Institute of Technology
[b]Simon School of Business, University of Rochester
[c]Interdisciplinary Center Herzliya
[d]Centre for Economic Policy Research
[e]University of California, Berkeley
[f]National Bureau of Economic Research

Abstract

This paper examines the extent to which individual investors provide liquidity to the stock market and whether they are compensated for doing so. We show that the ability of aggregate retail order imbalances, contrarian in nature, to predict short-term future returns is significantly enhanced during times of market stress, when market liquidity provisions decline. While a weekly rebalanced portfolio long in stocks purchased and short in stocks sold by retail investors delivers 19% annualized excess returns over a four-factor model from 2002 to 2010, it delivers up to 40% annualized returns in periods of high uncertainty. Despite this high aggregate performance, individual investors do not reap the rewards from liquidity provision because they experience a negative return on the day of their trade and they reverse their trades long after the excess returns from liquidity provision are dissipated. During the financial crisis, French active retail stock traders stepped up to the plate, increased stock holdings, and provided liquidity. In contrast, mutual fund investors fled from delegation by selling their mutual funds.

Keywords: Liquidity, Retail investors, Crisis.
JEL classification: G01, G11, G14.

*Corresponding author. Tel: 617-715-4820; Fax: 617-258-6855. Email address: jnbarrot@mit.edu.

1. Introduction

What is the contribution of individual investors to the formation of prices and liquidity in financial markets? A long-standing literature has considered them as noise traders, in the sense of Black (1986) and Shleifer and Summers (1990), who push prices away from fundamentals and destabilize markets. In contrast to this literature, recent empirical evidence suggests that individual investors' trades provide liquidity to meet the demand for immediacy of other market participants (Kaniel, Saar, and Titman, 2008; Kaniel, Liu, Saar, and Titman, 2012; Kelley and Tetlock, 2013). While retail investors are possibly less sophisticated than their institutional counterparts, they also face lower agency costs and liquidity constraints than institutional investors, such as mutual funds (Chevalier and Ellison, 1999; Coval and Stafford, 2007). Retail traders could thus have some ability to act as market makers, especially when institutional liquidity dries up, as was the case during the 2008-2009 financial crisis.

This paper examines the extent to which individual investors provide liquidity to the stock market and whether or not they are compensated for doing so. We use a unique data set obtained from a leading European online broker in personal investing and online trading. This data set allows us to track the trades of a large sample of individuals from January 2002 to December 2010. The data cover the recent financial crisis, when the liquidity-provision capacity of traditional market makers was plausibly reduced (Nagel, 2012). We uncover a series of new findings.

First, individuals provide liquidity especially at times when conventional liquidity providers are constrained. We begin by showing that in our sample, consistent with recent literature, aggregate retail buy-sell imbalances are contrarian and positively predict the cross section of stock returns at a horizon of a few weeks. A one standard deviation increase in daily order imbalances is associated with an increase in returns of about 15 additional basis points over the following three trading weeks (a 4% increase in annualized returns). We then test whether this increase in returns earned by retail investors corresponds to compensation for liquidity provision. To do so, we first construct a weekly rebalanced portfolio that goes long in stocks purchased and short in stocks sold by retail investors (the retail portfolio). We then compare the returns on this portfolio with time series variation in the supply of liquidity provided by institutional investors. Guided by prior work showing that intermediaries are especially constrained in their ability to provide liquidity in times of high uncertainty, we split our sample into periods of high and low VIX [Chicago Board Options Exchange (CBOE) volatility index], when the VIX is higher or lower than 20, its 2002-2010 median. We contrast the returns on the retail portfolio in these two subsamples and find robust evidence that they increase sharply in times of high uncertainty. While the retail portfolio earns 19% annualized

1

excess returns over a four-factor model from 2002 to 2010, it earns up to 40% annualized returns when traded over the weeks when the VIX is above its sample median. We also provide suggestive evidence that, during times of high uncertainty, retail investors do in fact step up to provide more liquidity and that their risk-bearing capacity seems to increase. These results indicate that retail traders do provide liquidity to the stock market, especially when institutional liquidity dries up.

Second, we exploit the unique panel feature of our data set to show that retail investors fail to reap the returns from liquidity provision. We provide two explanations for this result. The first has to do with the price at which retail orders are executed on the day of trading. To benefit from the predictable short-term returns that follow a day of intense imbalances, individual investors need to avoid being picked off on day 0. To understand why, suppose that institutions holding stock S are hit with liquidity shocks and need to fire sell their shares of S. The price of S plummets on day 0 and recovers in the short-term thereafter. Individuals buying stock S at its lowest on day 0 fully benefit from the price reversal in the subsequent days. However, those who purchase S before it reaches its lowest price experience a negative intra-day return on day 0, which can more than offset the gain from price reversal. Our analysis of order-level data indicates that, in our sample, retail investors get picked off on day 0. The average retail trade experiences large and negative returns on this day, so much so that returns on day 0 more than offset the rewards from liquidity provision that could arise subsequently.

The second reason for the low performance of individual investors in our sample has to do with the speed at which they reverse their trades. Individuals cannot benefit from liquidity provision unless they reverse their trades quickly enough thereafter, before the benefits are dissipated. This is exactly what retail investors in our sample fail to do. The average holding period among retail investors in our sample is above three hundred days, and most of the returns from liquidity provision are dissipated, on average, after 20 days. Thus, surprisingly, low trading frequency – specifically, slow reversal of trades – is one of the reasons that individual investors in our sample under-perform. f retail traders were to close their positions earlier, they could be demanding liquidity themselves, which might adversely affect their returns. Nonetheless, while Odean (1998) or Barber and Odean (2000) argue that overtrading is associated with high transactions costs and is, therefore, responsible for the low performance of retail traders, our findings suggest that retail traders could paradoxically capture a larger liquidity premium by trading more quickly.

Finally, we take advantage of the richness of our data and uncover substantial cross-sectional heterogeneity in the returns to liquidity provision. We first sort trades based on the experience of the individual placing them. We find that highly experienced individuals are much less

prone to the picking-off effect. In addition, they flip their trades much more quickly.[1] These two components explain a significant share of their outperformance relative to less experienced traders. We also sort trades based on the average speed at which the individuals who place them usually reverse their trades. We find that fast traders are less prone to the picking-off effect and, thus, experience higher returns relative to slower traders.

Our findings are in contrast to the view expressed by part of the financial press arguing that the financial crisis led to a massive exodus of small retail investors from the stock market. Using our data, we show that, during the financial crisis, French retail investors on aggregate fled from delegation by selling their mutual funds, yet, at the same time, active retail stock traders stepped up to the plate, increased stock holdings, and provided liquidity.

This paper adds to the ongoing debate on the contribution of retail trades to stock market efficiency. A number of papers find that individual trades positively predict short-term returns. A first body of work has interpreted this as evidence of noise trading pushing prices away from fundamentals. Barber, Odean, and Zhu (2009) find that stocks that individual investors are buying (selling) during one week have positive (negative) abnormal returns on that week and in the subsequent two weeks. These returns then reverse over the next several months. Although Barber, Odean, and Zhu (2009) interpret their results as evidence of noise trading, their findings are also consistent with individual investors providing liquidity to institutional investors. Another body of work has associated the short-term predictability of retail trades with liquidity provision. Kaniel, Saar, and Titman (2008) identify individual investor trades using the NYSE Consolidated Audit Trail Data files, which contain detailed information on all orders executed on the exchange, including a field that identifies whether the order comes from an individual investor. They show that the top decile of stocks heavily bought by individuals outperform those heavily sold by individuals, a result again consistent with retail traders providing liquidity to institutions that require immediacy. Kaniel, Liu, Saar, and Titman (2012) also find evidence that stocks purchased by individual investors prior to earnings announcement outperform those they sell and that compensation for risk-averse liquidity provision accounts for approximately half of this over-performance. Finally, Kelley and Tetlock (2013) argue that retail traders provide liquidity to the market and benefit from the reversal of transitory price movements. Our contribution to this body of work is twofold. First, we show that the predictability of individual trades increases

---

[1]Some notable contributions to the recent and growing literature on learning dynamics in finance are List (2003), Mahani and Bernhardt (2007), Agarwal, Driscoll, Gabaix, and Laibson (2008), Choi, Laibson, Madrian, and Metrick (2009), Kaustia, Alho, and Puttonen (2008), Greenwood and Nagel (2009), Nicolosi, Peng, and Zhu (2009), Seru, Shumway, and Stoffman (2009), Chiang, Hirshleifer, Qian, and Sherman (2011), and Linnainmaa (2011).

when the rewards to liquidity provision are high, consistent with the idea that they provide liquidity. Second, we are able to evaluate whether retail investors benefit from supplying liquidity, which has been an open question since Kaniel, Saar, and Titman (2008). The data used in either Kaniel, Saar, and Titman (2008) or Kelley and Tetlock (2013) were not suited to answer this question, as those data sets do not make it possible to track individual orders. Utilizing our trade-level data, we demonstrate that individuals fail to benefit from their liquidity provision role.

Because we focus on returns at the order level, our results also relate to the literature on individual investors' performance.[2] The average household trades in excess of what liquidity and hedging motives would command and loses money in the process (Odean, 1998; Barber and Odean, 2000; Barber, Lee, Liu, and Odean, 2006; Grinblatt and Keloharju, 2000). This is generally attributed to behavioral biases such as overconfidence or gambling (Statman, Thorley, and Vorkink, 2006; Glaser and Weber, 2007; Grinblatt and Keloharju, 2009; French, 2008). A small group of retail traders, however, manage to generate absolute performance (Barber, Lee, Liu, and Odean, 2014), with some persistence (Coval, Hirshleifer, and Shumway, 2005). Linnainmaa (2010) explores the source of performance in more detail. He uncovers losses on limit orders and gains on market orders in Finland.[3] We add to this body of literature by relating this picking-off effect to the short-term returns from liquidity provision. In particular, we show that individual investors' returns are low because they get picked off and they fail to reverse their trades soon enough. In other words, retail investors do not trade fast enough to collect the benefits from their liquidity provision. Another contribution, particularly relative to Linnainmaa (2010), is to show that more experienced retail investors tend to be less picked-off.

The rest of the paper is organized as follows. We present the data in Section 2. Section 3 presents the results, and Section 4 concludes.

2. Data

In this section, we present the data, the construction of the main variables, and we compare our sample with other samples used in the literature.

2.1. Data source

We consider information about a large sample of French retail investors trading between January 2002 and December 2010, provided by a leading European broker in personal

---

[2]For an extensive review of individual investors' behavior and performance, see Barber and Odean (2011).

[3]In the same spirit, Stoffman (forthcoming) finds that prices move in the direction of institutional trading: When institutions sell to households, prices tend to fall.

investing and online trading. In the past 12 years, this broker accounted for an average 15% of online brokers' stock trades on Euronext Paris, which collectively represented 14% of all trades in the market.[4] This sample is thus fairly representative of the behavior of individual investors directly investing in the French stock market. These data are also used in Foucault, Sraer, and Thesmar (2011), who study the effect of retail investors on the volatility of stock returns.

There are 91,647 investors making approximately 4.6 million trades in 730 stocks in our sample. The data contain executed orders.[5] For each order, we track the trading exchange identifier (International Securities Identification Number, ISIN), the trading date, the quantity, and the amount traded in euros.[6] Given that we do not have the exact timing of the execution within the day, we aggregate trades by individual × stock × day. The average trade size in our sample is 7,741 euros. We obtain daily stock returns from EUROFIDAI.[7]

Our sample stands out in a number of ways. First, it includes information at the order level, which allows us to perform detailed analyses of retail trading. Second, it spans a long time period that includes episodes of market stress, such as the recent financial crisis. This makes it possible to contrast the behavior of a large number of retail investors at different times, when the returns to liquidity provision vary. Recent work, including Adrian, Etula, and Muir (2012), Ang, Gorovyy, and Van Inwegen (2011), and Ben-David, Franzoni, and Moussawi (2012) suggest that intermediaries are especially constrained for high levels of the VIX index of implied volatilities of Standard & Poor's (S&P) index options. Nagel (2012) uses the returns to short-term reversal strategies as proxies for the returns to liquidity provision. He finds that they are almost perfectly correlated with the level of the VIX. To contrast the predictability of retail imbalances with the liquidity condition on the market, we construct three proxies for the level of uncertainty on the market, which we expect to be associated with higher rewards for liquidity provisions. We first use the CBOE S&P 500 implied volatility index on day $t-5$, which we normalize to a daily volatility measure by dividing it by $\sqrt{250}$.[8] As an alternative proxy, we construct the variable *Idiosyncratic volatility* as the cross-sectional standard deviation of individual stock returns on day $t-$

---

[4]This is according to Acsel, the association of French online brokers (see http://www.associationeconomienumerique.fr/), which collects monthly data on online trading.

[5]Throughout the paper, we use the terms "order" or "trade" interchangeably.

[6]The data also specify the type of order used (limit or market order). At the retail trader account level, data include monthly fees associated with the account as well as limited demographical information, namely, their gender, age, and geographical location.

[7]EUROFIDAI is a research institute funded by the CNRS (French National Center for Scientific Research), whose mission is to develop European stock exchange databases for academic research.

[8]We check and find that all results are unchanged when we use the VSTOXX, the European volatility index, instead of the VIX.

5. Finally, we define Crisis as a dummy variable equal to one in the seven months from September 2008 to April 2009.

In some of the analyses, we adjust returns and cumulative returns for systematic risk. To do so, we estimate the exposure of each of the 730 stocks in the sample to systematic risk factors (market, small-minus-big, high-minus-low, and momentum) over the sample period (2002-2010) at the weekly level. All risk factors are obtained from EUROFIDAI.[9] We run the following ordinary least squares (OLS) model for each stock in the sample:

$$Ret_t - r_f = a + b(Mkt_t - Rf_t) + cSMB_t + d.HML_t + eMOM_t + \epsilon_{it}, \tag{1}$$

where Ret_t is a given stock's return in week t, and $Mkt_t - Rf_t$, SMB_t, HML_t, and MOM_t are, respectively, the returns of the market factor, small-minus-big, high-minus-low, and momentum. The estimated coefficients $\hat{b}$, $\hat{c}$, $\hat{d}$, and $\hat{e}$ are then used to define the risk-adjusted return on any given stock i in any given period t, $AdjRet_t$, as the difference between the realized return $Ret[t]$ and its predicted value:

$$AdjRet_{it} = Ret_{it} - (\hat{b}_i(Mkt_t - Rf_t) + \hat{c}_iSMB_t + \hat{d}_iHML_t + \hat{e}_iMOM_t + r_f) \tag{2}$$

2.2. Comparison with the US

While French and US financial institutions are fairly different, the behavior of retail investors in our sample appears similar to what has been reported in previous studies using US data. In terms of trading activity, 49% of the retail investors in our sample make at least one trade per year. Kumar and Lee (2006) report that, in their sample of retail investors at a major US discount brokerage house, the corresponding statistic is 45%. The average trade size in our sample is 7,741 euros, which is about $10,000 dollars. Kelley and Tetlock (2013) report an average trade size of $11,566 using their US-based brokerage account data. Finally, according to press reports quoting Euronext insiders, retail investors represent 6% of Euronext volumes, which is similar to the share of individual investors in total volume in the US.[10]

Investors in our sample also exhibit similar biases to what has been shown using US data. For instance, Boolell-Gunesh, Broihanne, and Merli (2008) use the same data set as ours (although only up to 2008) and report that individual investors in this sample are also prone to the disposition effect.

---

[9]The procedure used by EUROFIDAI to compute factor returns for the French market is similar to the one used by Kenneth French for the US market.

[10]See L'Agefi (2012).

3. Results

In this section, we present and discuss the results of our analyses at the stock, at the portfolio, and the order level.

3.1. Evidence on liquidity provision

This paper studies the relation between retail orders and future short-term returns, both at the stock-day level and the order level. We start by aggregating individual trades at the stock-day level. Our main measure of imbalances, $Imb[0]$, is similar to the one used in Kelley and Tetlock (2013). It is computed daily as the number of shares bought by retail investors minus the number of shares sold by retail investors divided by shares bought plus shares sold.[11] As in Kelley and Tetlock (2013), we exclude from the sample stock $\times$ days with less than five orders, leaving 217,511 stock-days. We control for the size of firms with the log of their market capitalization ($LogME$). We denote as $Ret[x, y]$ the holding period return between day x and day y.

Panel A of Table 1 provides summary statistics. The average and median of $Imb[0]$ are very close to zero. The average aggregate volume traded in a stock on a given day is just over 165,000 euros, which represents an average of 2.4% of the total daily volume for these stocks.[12] We first estimate whether retail order imbalance dynamics are consistent with liquidity provision, i.e., if they seem to respond to past liquidity shocks. We want to measure the sensitivity of retail imbalances to past returns, controlling for market conditions and stock invariant characteristics. We do so by running the following linear regression:

$$Imb[0]_{it} = c + \alpha_0 Ret]-1, 0]_{it} + \alpha_1 Ret[-5, -1]_{it} + \alpha_2 Ret[-26, -6]_{it} + \alpha_3 LogME_{it} + \pi_t + \eta_i + \epsilon_{it}, \tag{3}$$

where $Imb[0]_{it}$ is the imbalance of stock i on day t, and $Ret[-5, -1]_{it}$ and $Ret[-26, -6]_{it}$ are the cumulative returns over the past week and the prior month on stock i, respectively. π_t and η_i are, respectively, day and stock fixed effects. $Ret]-1, 0]_{it}$ is the date-t return of stock i and is included in the regression to account for the possibility that retail orders are positively correlated with date-t stock returns, which are typically followed by short-run reversal (Nagel, 2012).[13] Standard errors are clustered at the stock level, allowing the ϵ_{it} to be correlated within a stock but not across stocks within a given day. Results are presented

---

[11] We check below that our results are robust when we use alternative measures.

[12] Our coverage is comparable to Kelley and Tetlock (2013), in which retail trades account for 2.3% of total listed (NYSE, Amex, and NASDAQ) volume, over a period of five years.

[13] We use this notation to later distinguish stock returns over the day ($Ret]-1, 0]$) from the return from order execution until the end of the day ($Ret[0]$).

in Table 2. In Column 1, the specification includes only day fixed effects. We add stock fixed effects in Column 2. While imbalances can correlate with some omitted variable, such as the differences in households' participation rate in various stocks, the inclusion of stock fixed effects ensures that those are not driving the results. Consistent with evidence in Kaniel, Saar, and Titman (2008) and Kelley and Tetlock (2013), we find that retail imbalances react strongly to past returns. The estimates are highly statistically significant, and the coefficients are economically large. A one standard deviation decrease in the past week's returns, $Ret[-5, -1]_{it}$, leads to an increase of about 7 percentage points in $Imb[0]$, which represents 12% of the sample standard deviation of $Imb[0]$. The estimate is unaffected by the inclusion of stock fixed effects, suggesting that time-invariant stock-level characteristics are not responsible for the cross-sectional correlation between retail imbalances and past returns. To make sure that we are not picking up a different functional form of past returns, we estimate a more flexible model allowing for a different impact of positive and negative stock returns on date-t imbalances. The results are presented in Columns 3 and 4 of Table 2. We still find a significant and negative effect of past-week returns on current imbalances. A one standard deviation increase in the positive past-week returns leads to a 4 percentage point increase in imbalances. The effect is larger following negative returns. A one standard deviation decrease in negative past-week returns leads to a 5.2 percentage point increase in date-t imbalances. Overall, the pattern of buying and selling in reversal strategies is robust and resembles the trading of a liquidity provider who takes opposite positions to the rest of the market.

[Insert Tables 1 and 2 near here]

We then turn to the analysis of the returns to liquidity provision. Our regression model for predicting cumulative holding period returns from day x to y is

$$Ret[x, y]_{it} = \beta_0 \quad + \quad \beta_1 Imb[0]_{it} + \beta_2 Ret] - 1, 0]_{it} + \beta_3 Ret[-5, -1]_{it} + \beta_4 Ret[-26, -6]_{it} \quad (4)$$
$$+ \quad \beta_5 LogME_{it} + \pi_t + \eta_i + \epsilon_{it}$$

Standard errors are clustered at the stock level.[14] The coefficient of interest is β_1, which measures the sensitivity of future returns to current imbalances from retail orders.

---

[14]These standard errors are correct if residuals are uncorrelated within a given day. In Appendix B, Table B1, we explore the sensitivity of our results to this clustering assumption. We cluster standard errors at the day level and also at the day and stock level, and we find slightly smaller standard errors. Clustering at the day (day and stock) level implies standard errors are correct under the assumption that there is no source of correlation of residuals for Number of observations across stocks (days and stocks). We report the most conservative standard errors in the main table.

We start by estimating Eq. (4) using as a dependent variable the cumulative returns from day $x = 1$ to day y, where y takes values from one to one hundred. We plot in Fig. 1 the one hundred β_1 coefficients estimated, along with their 95% confidence intervals. The graph shows that stocks heavily purchased by retail investors outperform those that are heavily sold by retail investors by a significant 18 basis points over the first 16 days. This outperformance then gradually dissipates over the subsequent 85 days. We obtain identical results when we perform the same analysis with risk-adjusted cumulative returns. This finding is comparable to the results presented in Linnainmaa (2010) on Finnish retail investor data and is also consistent with the patterns shown by Gutierrez and Kelley (2008) on US data.

[Insert Fig.1. near here]

We formalize this result in a regression setting in Table 3. Columns 1 and 2 estimate Eq. (4) including day fixed effects, and Columns 3 and 4 also include stock fixed effects. The main finding is that retail imbalances positively predict cumulative returns from day 1 to day 16, even after controlling for date-t returns and including stock fixed effects. Quantitatively, a one standard deviation increase in $Imb[0]$ leads to a 15 basis point increase in cumulative returns over the following three trading weeks. These estimates are comparable in magnitudes to those obtained by Kelley and Tetlock (2013). Columns 2 and 4 show that the effect is short-lived, as it is nearly fully reversed after one hundred days. As we already noted in Table 2, the estimates are virtually unaffected by the inclusion of stock fixed effects. Taken together, Table 2 and Table 3 suggest that individual investors provide liquidity by placing contrarian trades and receive a significant compensation in the form of high returns over the next three weeks. We perform a number of additional tests to ensure that these results are robust to the proxy we use for imbalances created by retail investors' trades. We show in Table 4 that the results hold when we standardize $Imb[0]$ (by subtracting its within-stock mean and scaling it by its within-stock standard deviation), when we use terciles of $Imb[0]$, or when we define imbalances on stock i as the ratio of buy minus sell orders by retail investors normalized by the market-wide volume of the stock. In Appendix B, Table B2, we also show that our results are robust to splitting $Ret] - 1, 0]$, $Ret[-5, -1]$, and $Ret[-26, -6]$ into positive and negative returns.

[Insert Table 3 and 4 near here]

The finding that retail imbalances predict future stock returns after conditioning on contemporaneous and past returns is consistent with Kelley and Tetlock (2013) and suggests that incremental information is embedded in retail trades that cannot be mechanically exploited by looking at past daily, weekly, and monthly returns. While we cannot directly

9

test for the exact information structure of retail investors in our sample, one possibility is that they have some information about whether past order flow is motivated by liquidity needs and that they act on this information by placing their orders accordingly.

3.2. Retail investors' performance and the demand for liquidity

An alternative to the natural liquidity provision interpretation of the results in Subsection 3.1 is that the predictive power of retail imbalances results from noisy correlated trading pressure from individuals. To disentangle these two alternative hypotheses, we look at how the sensitivity of retail imbalances to past returns, as well as the predictive power of these imbalances on future returns, vary with general liquidity conditions in the market. If the predictability of individual imbalances increases when liquidity dries up elsewhere in financial markets, and the returns to liquidity provision go up, then it is likely that these trades are providing liquidity, instead of pushing prices away from fundamentals. Our first step in answering this question thus consists in investigating whether the liquidity provision behavior shown in Table 2 is stronger when the returns to liquidity provision increase.

To do so, we estimate an augmented version of Eq. (3), in which we regress retail imbalances on the interaction of past returns with our proxies for liquidity conditions on the market: the level of the VIX, the cross-sectional idiosyncratic volatility, and the *Crisis* dummy. However, when the variance of asset returns is larger, absorbing demand shocks is riskier for retail traders so that they naturally require a larger swing in prices for a given demand shock. In Appendix A, we formalize this intuition with a simple trading model in the spirit of Grossman and Miller (1988). Therefore, we consider how retail imbalances react to risk-adjusted returns instead of raw returns. We estimate Eq. (3) but use standardized measures of past returns as our control variables. To standardize returns, we first subtract the average past-20-day return of the stock and then divide this difference by the past-20 day standard deviation of returns.

The results presented in Table 5 support the hypothesis that, in our sample, the risk-bearing capacity of retail investors increases during periods of high uncertainty. The coefficients on the interaction between standardized past returns and *VIX*, *Idiosyncratic volatility*, and the *Crisis* dummy are significantly negative, suggesting that the sensitivity of imbalances to past standardized returns increases during episodes of market stress, including the height of the financial crisis. Quantitatively, we find that when the VIX is at its 25th percentile value, a one standard deviation increase in past-week standardized returns leads to a 9.6 percentage point decrease in date-t retail imbalances. However, when the VIX is at its 75th percentile value, a similar increase in returns leads to a 10.4 percentage point decrease in imbalance.

[Insert Table 6]

10

The increase in risk-bearing capacity could be coming from a decline in retail investors risk aversion or from an increase in their relative weight in the stock market. We have no way to fully disentangle these two explanations, particularly because one limitation of our data is that we do not observe institutional trades. As a first path, however, we can look at the aggregate share of individual investors in total volume and their aggregate portfolio turnover in the time series. In Fig. 2, we show that the share of volume traded by retail investors and the rate at which they turnover their portfolio went up sharply in 2008 and 2009, at times in which the aggregate uncertainty measured by the VIX or the average idiosyncratic volatility is high. This complements Fig. 3, which shows that, on aggregate, retail investors were net sellers of equity mutual funds and net buyers of equity in the financial crisis of 2008. This suggests that part of the increased negative elasticity of retail imbalances to past returns could be coming from an increase in the relative contribution of retail investors.

[Insert Fig.2. and Fig.3. near here]

We then investigate how the returns following large retail imbalances vary with our two proxies for the returns to liquidity provision. We start with a graphical analysis, based on Eq. (4). We split the sample into low and high VIX days, in which high VIX days occur when the VIX is higher than 20, its 2002-2010 median. In each of these subsamples, we run separate regressions of the cumulative returns from day $x = 1$ to day y, where y takes values one to one hundred on day 0 retail imbalance, controlling for contemporaneous, past weekly, and past monthly returns, as well as the log of the market capitalization of the stock. We plot the estimated coefficients for $Imb[0]$ at each horizon (one to one hundred), along with their 95% confidence intervals in Fig. 4. Panel A (Panel B) corresponds to high VIX (low VIX) days. Two interesting facts emerge from Fig. 4. First, on high VIX days, retail imbalances are followed by a much larger price increase over the subsequent couple of weeks than on low VIX days. Stocks heavily purchased by individuals reach 24 basis points in cumulative returns over the subsequent 16 days, and they reach about 14 basis points in cumulative returns on low VIX days. Second, the reversal is much more pronounced on low VIX days than it is on high VIX days. A potential interpretation of these results is that, when uncertainty is high, retail trades provide liquidity and are significantly compensated for this. Conversely, in times of low uncertainty, retail trades are more likely to be picked off by informed traders and eventually generate negative cumulative returns.

[Insert Fig.4. near here]

We confirm these results in formal regression tests. We interact all terms in Eq. (4) with our three proxies for market liquidity: the level of the VIX (Columns 1 and 2), the

cross-sectional standard deviation of individual stock returns (Columns 3 and 4), and the *Crisis* dummy equal to one in the seven months from September 2008 to April 2009. The results are presented in Table 6. Columns 1, 3, and 5 include day fixed effects, and Columns 2, 4, and 6 also include stock fixed effects. Consistent with the intuition obtained from Fig. 4, the two-week returns following high retail imbalances are significantly higher when the VIX is high, when the cross-sectional standard deviation of stock returns is high, or during the crisis period. Short-term returns to retail investors' liquidity provision almost triple during the financial crisis. While the sensitivity of three-week cumulative returns to retail imbalances is 0.0016 when computed over months outside the financial crisis, this sensitivity increases to 0.0063 during the financial crisis.

[Insert Table 6 near here]

Altogether, these results provide evidence that the outperformance of stocks heavily purchased by individuals over those heavily sold by individuals amounts to compensation for liquidity provision. One objection to our interpretation is that, during high VIX periods, limits to arbitrage increase so that correlated trading by individual investors would be less likely to be arbitraged away by constrained sophisticated arbitrageurs. As a result, the increased predictability during high VIX periods would simply be the result of noisy price pressure from individuals and would be unrelated to the returns to liquidity provision. However, we find in unreported tests that the autocorrelation of our measure of retail imbalances is significantly lower on high VIX days, making this alternative interpretation much less compelling than our hypothesis based on liquidity provision.

These results also emphasize the value of our data set. Because we track individuals for a long period of time, we are able to reconcile apparently contradictory findings from the literature with respect to the reversal of short-term returns. While Hvidkjaer (2008) and Barber, Odean, and Zhu (2009) find evidence of reversal following short-term returns, Kaniel, Saar, and Titman (2008) and Kelley and Tetlock (2013) find none. Our results suggest that different samples and time periods could explain these differences.

3.3. Portfolio analysis

In this subsection, we show that portfolios mimicking trades of individual investors generate significant positive abnormal returns at the weekly horizon. Doing so allows us to check that the effects we shown in Subsection 3.2 are economically meaningful and that they are not driven by the smallest trades observed in our sample of individual investors. The portfolio we build aggregates trades across individuals and assets and, thus, gives a greater weight to stocks heavily traded. In addition, we want to make sure that the results

12

are not driven by a particular feature of $Imb[0]$, our measure of retail imbalances. In the analysis that follows, we pool all stocks into a long portfolio and a short portfolio, and we therefore abstract from any measure of imbalances. Finally, we want to analyze how the returns of the short-term reversal strategies of individual investors load on systematic risk factors.

We proceed as follows. Each week over the sample period spanning 2002 to 2010, we aggregate individual trades at the stock level. We sort stocks based on their net retail aggregate position into two subsets of stocks purchased or sold. We form a long and a short portfolio by value-weighting the stocks in each of these two subsets. In each of these two portfolios, returns are value-weighted based on the dollar values of weekly aggregate retail traders' volumes in each stock. Hence, a stock enters in the long and the short portfolio with a weight that reflects the size of the aggregate individual imbalance in that stock in a given week. We rebalance each of the two portfolios at the end of each week. We consider the returns on the long-short portfolio (the retail portfolio). More precisely, we regress the returns on the long-short portfolio on a model similar to that of Eq. (2). The coefficient of interest here is a, which is an estimate of the weekly returns on the retail portfolio, adjusted for exposure to systematic risk. We introduce the risk factors one by one in the model to assess their effects on the estimate of excess returns. The results of the main specification are presented in Panel A of Table 7. The unadjusted return is 26 weekly basis points, which amounts to 15% annualized returns. A market model (Column 2) generates an alpha of 23 basis points per week, which amounts to an annualized risk-adjusted return of 13%. Moving to a Fama and French risk model (Column 4) increases the estimates to 33 basis points, an annualized return of 19%. Introducing the momentum factor (Column 5) in the model does not affect the estimate of alpha.

[Insert Table 7 near here]

If these excess returns represent compensation for liquidity provision, then we should find, consistent with the analysis presented above, that these excess returns increase in times of high uncertainty, i.e., when the rewards to liquidity provision increase. We thus split the sample based on the level of the VIX in the last day of the portfolio formation week. We then regress the excess returns on the zero cost retail portfolio on the same risk factors as those used in Eq. (2), but using only those weeks when the level of the VIX is above its sample median. The results are presented in Panel B of Table 7. The unadjusted returns are 52 basis points weekly, 30% in annualized terms. Adjusting for the exposure to the market increases the estimates to 57 basis points, which amounts to an annualized risk-adjusted performance of 34%. Introducing the three other risk factors pushes the returns even further

13

to an annualized return of nearly 40%. In other words, irrespective of the particular risk adjustment, the excess returns earned by the retail portfolio are two times greater in high VIX weeks than on average. These returns strengthen the conclusion that retail trades provide liquidity to the markets.

Finally, we check whether the effects are stronger if we restrict the sample to stock-weeks when retail imbalances are extreme. We define an imbalance as extreme if the stock-week lies in the top or the bottom terciles of the distribution of the stock's retail imbalances between 2002 and 2010. We then compute the long and short portfolios as described above. Given that the returns to liquidity provision are higher for large retail imbalances, we expect the abnormal returns on this portfolio to be higher than what we obtain using the whole sample. We run the same model and present the results in Panel C of Table 7. Unsurprisingly, the results are larger than those obtained in Panel A. A market model delivers a return of 34 basis points weekly, which amounts to an annualized return of 19%. A four-factor model increases the alpha to 47 basis points weekly, or about 28% annually.

A natural question to ask is how much of these risk-adjusted returns would be left after taking into account transaction cost. Unfortunately, we have no way to directly observe these transactions costs. As a first path, we rely on De Groot, Huij, and Zhou (2012), who provide single-trip transaction cost information for the one thousand largest European stocks each year between 1995 and 2009, by volume deciles, which they obtain from Nomura securities. In Table 3 of their paper, transaction cost estimates range between 19 (top decile) and 76 (bottom decile) on average. In Panel A (C) of Table 7, we find that, over our sample period, a weekly rebalanced portfolio long in stocks purchased and short in stocks sold by retail traders yields a weekly alpha over a four-factor model of 33 (47) basis points. We compute that, on average, 60% of the positions of our aggregate portfolio are closed at the end of the week.[15] Hence, an omniscient investor exploiting retail investor trades would break-even (earn 15 basis points a week if restricting to extreme imbalance stock-weeks) if, on average, individuals in our sample trade stocks that lie in or above the ninth volume decile of the one thousand largest European stocks, for which single-trip transaction costs are 20 basis points.[16] For extreme imbalance stock-weeks, there seems to be a small potential gain down

---

[15]This result is not directly comparable with the order-level holding period that we document in Subsection 3.4. The main reason for this is aggregation. Suppose trader A purchases one share of stock X in week 1, trader B sells one share in week 2, and traders A and B reverse their respective positions in weeks 52 and 53. In aggregate, the position initially opened by trader A would be closed by week 2, a one-week holding period. At the order level, however, the average holding period would be 52 weeks. Another reason why the two are not directly comparable is that aggregate positions overweight larger orders, which are typically reversed sooner (see Table 11).

[16]Given that 60% of the positions are closed weekly, we approximate the transaction cost of a stock in the ninth volume decile as $1.6 \times 20 = 32$ basis points.

14

to the seventh decile, where single-trip transaction costs are 26 basis points. Panel B shows that, in high VIX weeks, the alpha is 64 basis points, implying a higher potential gain. However, if a large enough amount of money was invested in a strategy replicating retail trades, then the liquidity premium associated with this strategy would probably be gradually dissipated.

3.4. Order-level analysis

The results aggregated at the stock level seem at odds with the results commonly found in the literature that individual investors lose money on average (Odean, 1998; Barber and Odean, 2000; Barber, Lee, Lin, and Odean, 2006; Grinblatt and Keloharju, 2000), either because they trade too much or because they pick losing stocks. In this subsection, we attempt to reconcile these results with our findings. We first notice that the fact that the retail portfolio earns positive and significant excess returns does not necessarily mean that retail traders earn significant trading profits. In fact, for individual investors to collect the returns from liquidity provision, it needs to be the case that the return on the day of the trade (day 0) should not be lower than the subsequent excess returns and that their trades are reversed before the rewards from liquidity provision are dissipated. We exploit the richness of our data to explore the behavior of individual investors along these two dimensions.

For each of the 4.6 million trades in the sample, we construct the following variables. We define the return on day 0, $Ret[0]$, as the difference between the closing price at the end of the day when the trade was placed and the price at which the order was executed during the day. We define days to reversal as the number of days between the date of a trade and the earliest date at which the trade was at least partially reversed.[17] We measure the holding period return as the cumulative return from the time of execution to the close of the earliest date at which the order is partially reversed.[18] For ease of comparison across holding periods, we also compute the internal rate of return of each trade. For most of the analysis, we adjust both the holding period returns and the internal rate of returns for exposure to systematic risk, following the procedure described in Section 2. In addition, because we are interested in the heterogeneity in experience across individual investors, we define the cumulative number of orders for a trader i as the total number of orders placed prior to placing a given order.

Summary statistics for the sample at the order level are presented in Panel B of Table 1. The average order is worth 7,741 euros. There are slightly more purchases than sales in the

---

[17]We would obtain similar results by considering the number of days until the position is fully reversed. However, our measure is more conservative for the purpose of this study.

[18]For simplicity, we cap the holding period to five hundred days. When an order is never reversed in the sample, we cap its holding period to the earliest of the last day of trading in the sample and the last quotation day of the stock if it delisted.

sample, but purchases are slightly smaller (7,186 euros) than sales (8,342 euros). Turning to our variables of interest, the average holding period is 310 days and the median is 40. This is much longer than the average time at which the returns to liquidity provision are dissipated on average. As is apparent from Fig. 1, the cumulative returns following retail order imbalances peak at around 16 days, are only 10 basis points after 30 days, and are 0 after 80 days. In addition, the average return on day 0, $Ret[0]$, is -90 basis points.[19] This is much larger (in absolute value) than the average estimated rewards from liquidity provision. Hence, the average trade in the sample does not reap the returns to liquidity provision because it is picked off on day 0 and it is not reversed quickly enough.

Table 1 indicates that the average holding period return is -2.7% and the average internal return is 4 basis points per day. On a risk-adjusted basis, these numbers are, respectively, -90 basis points and 3 basis points.[20] Table 8 presents the correlations between these variables. The number of days to reversal is negatively related to $Ret[0]$ and the internal rate of return, which are positively correlated. Quickly reversed trades are picked off less, on average. We decompose the holding period return in Table 9 and present the results in both unadjusted and risk-adjusted terms. The loss on day 0, $Ret[0]$, accounts for approximately one third of the negative holding period returns, while the rest comes from $Ret[17, R]$, the returns from day 16 to the reversal of the trade. On a risk-adjusted basis, the results are very similar. One of the striking implications of these results is that individuals seem to be losing money because they do not reverse their trades soon enough, i.e., because they do not trade enough.

[Insert Tables 8 and 9 near here]

3.5. Heterogeneous effects

In this subsection, we explore the heterogeneity in the behavior of individual investors. We show that certain individual characteristics seem to be associated with a better ability to capture the returns to liquidity provision. The first characteristic we consider is a trader's experience. We sort trades in our sample based on the experience of the trader, measured by the total number of prior orders placed. We expect that experienced individual investors should be less picked-off and should reverse their trades quicker. We thus simply compute the average of all our return variables across all deciles of experience. The results are presented

---

[19]Fong, Gallagher, and Lee (2014) find negative day 0 returns for retail investor trades in Australia.

[20]The discrepancy between holding period returns and internal rates of returns comes from the fact that losing positions tend to be held longer, so that their holding period returns are larger in absolute value than the holding period returns of winning positions that are reversed quickly. The internal rate of return rescales returns to the daily horizon, therefore correcting this bias. The adjusted return on day 0 is computed as the difference between $Ret[0]$ and the stock return predicted by the four-factor model presented in Eq. (2).

in Panel A of Table 10. The average cumulative number of orders placed in the first decile of experience is 12 versus 3,323 in the highest decile of experience. Experienced traders flip their trades much faster than inexperienced ones. Interestingly, experienced traders also get less picked-off. The difference between the first and the tenth deciles of experience in $Ret[0]$ is 100 basis points, which is more than half their difference in holding period returns in risk-adjusted terms. Experienced traders have slightly better returns between day 1 and day 16. Their risk-adjusted holding period return and internal rate of return are, respectively, 20 and 10 basis points larger than those of inexperienced investors.

[Insert Table 10 near here]

The second characteristic we consider is a trader's average holding period. We sort orders in the sample based on the propensity of each trader to quickly reverse its trades. We expect individuals with shorter holding periods to be more capable of seizing the returns from liquidity provision. To check whether this is the case, we compute the average holding period of the trader over the sample.[21] We then compute the average of all our return variables across the ten deciles of the distribution of holding periods. The results are presented in Panel B of Table 10. The average cumulative number of orders placed in the first decile of speed to reversals is 33, versus 1,065 in the highest decile of reversals. Traders who usually reverse their positions faster also get less picked-off. The difference between decile 1 and 10 in the average date-0 return, $Ret[0]$, is 35 basis points. The risk-adjusted holding period returns and internal rates of returns of traders quickly reversing their trades are, respectively, 4.5% and 10 basis points higher than those of traders in the bottom decile, and the internal rate of return is monotonically decreasing in the time to reversal.

One way to evaluate whether experienced traders, and traders who reverse their trades faster, are smarter in how they supply liquidity is to see whether, although they get picked-off less, they still buy stocks with large day-t negative returns. Columns 6 and 11 of Table 10 show that these traders select stocks that experience larger swings in returns (by about 40 basis point). Overall, Table 10 suggests that experienced traders, relative to inexperienced ones, seem to provide more liquidity, to receive a higher compensation for doing so, and yet to be less picked-off on the day they trade. They thus appear to be smarter in the way they provide liquidity.

In principle, this result could emanate from two different channels. First, it could be that the worst-performing types (low experience, long holding periods) exit the sample

---

[21]While we exclude the order itself from the computation of the average holding period of the trader, it could still be the case that traders that typically have losing positions tend to keep them for longer.

more frequently. Second, it could be that individual investors experience some form of learning-by-doing. To get a quantitative sense of these two channels, we run simple regressions of $Ret[0]$ and of the log of the number of days to reversal on vectors of time-varying trader characteristics (including the log of the cumulative number of past orders and its square, the log of the size of the account, and the log of the monthly volume traded) and day and stock × day fixed effects. Most importantly, we add individual fixed effects. If there is no learning-by-doing, trading ability is fixed within individual. Therefore, individual fixed effects should capture trading ability, and our proxies for experience, the log of the cumulative number of past orders, should not predict $Ret[0]$ or the log of the number of days to reversal.

We present the results in Table 11. The results confirm that experience is strongly positively related to $Ret[0]$ and negatively related to the number of days to reversal. Traders with a larger number of past orders are less picked-off and reverse their trades faster. However, the coefficients decrease substantially when we introduce individual fixed effects. This suggests that an important part of the learning occurs via the attrition of our sample, i.e., the survival of the traders that are less picked-off and of those that reverse their trades more quickly. The results are similar when we proxy for experience with the cumulative volume traded by a given retail investor, as evidenced in Table 12. Interestingly, larger orders get picked-off less and start to get reversed faster. Younger traders tend to reverse trades earlier.

[Insert Tables 11 and 12 near here]

3.6. Discussion

The results presented in this paper indicate that individual investors provide liquidity to stock markets and that some of them are compensated for doing so, especially in periods of high uncertainty such as the 2008-2009 financial crisis, when institutional liquidity providers were most constrained. This seems at odds with the common view according to which retail investors flee to liquidity during times of financial market stress and thereby amplify the initial stress. Financial newspapers, both in Europe and in the US, reported a massive exodus of small retail investors from the stock market following the financial crisis, with potentially worrisome consequences. According to the Wall Street Journal, in the US, "from 2007 through 2009, [retail investors] withdrew money [from mutual funds that invest in US stocks] for three consecutive years," which "marked the first three-year period of withdrawals since 1979-1981."[22]

---

[22]See Wall Street Journal (2010).

Using our data, we find that, on aggregate, individual investors decreased their exposure to mutual funds. However, they also significantly increased their exposure to equities. As evidenced from Fig. 3, the net outflows of individual investors in our sample from equity mutual funds reached 150 million euros from mid-2007 to the first quarter of 2009. In the meantime, inflows into stocks amount to approximately 100 million euros over the same period. The results presented in this paper offer a new perspective on this somewhat surprising finding: In the aggregate, individual investors acted as liquidity providers for the rest of the market.

Finally, to shed some light on where the positive net flows of retail investors into stocks during the crisis period came from, we attempt to decompose the result presented in Fig. 3. We compute aggregate cumulative imbalances across categories of investors. To allocate investors to various categories, we examine their behavior in the precrisis period, which runs from January 2004 to December 2007. Our first category relates to trading frequency. Each quarter, we sort investors into terciles of number of trades realized in the quarter. A retail investors is in the active (inactive) category if she falls in the top (bottom) tercile in more than two-thirds of the quarters. Our second category relates to the type of securities mostly traded by the investors. Each quarter, we sort investors based on the ratio of stock volume they trade to the volume of stock and mutual funds they trade. Investors are allocated to the stock (mutual funds) category if they are above (below) the median ratio in more than two-thirds of the quarters. Importantly, because information on mutual fund trading is available starting only in January 2006, this split between mutual fund and stock traders has to be performed only on the period from January 2006 to December 2007.

Fig. 5 shows the results of the decomposition. Panel A of the figure shows that most of the positive flows into stocks was due to active and stock traders, and Panel B, in contrast, indicates that almost all of the negative flows out of equity mutual funds came from inactive and mutual fund traders. This dichotomy between active stock traders and inactive mutual fund investors is striking, and it emphasizes that not all retail investors are created equal. A first order heterogeneity in behavior exists within the group, which has not been fully appreciated in the literature. The flight from the market during the crisis discussed in the financial press seems to have been a flight from delegation, at least in France. Passive investors who get exposure to markets mostly through mutual funds panicked and fled. At the same time, active stock traders stepped up to the plate to provide liquidity helping to assuage market stress. Not withstanding retail traders' important liquidity-provision role in times of stress, our analysis has shown that they typically fail to capture the returns to liquidity provision because they get picked off and do not reverse their trades fast enough. Experience, another important dimension of heterogeneity, improves traders ability to capture returns

to liquidity provision.

[Insert Fig. 5. near here]

4. Conclusion

This paper examines the extent to which individual investors provide liquidity to the stock market and whether or not they are compensated for doing so. We start by confirming with our data that aggregate retail buy-sell imbalances are contrarian and positively predict the cross section of stock returns at a horizon of three weeks, and we proceed to show that rewards to liquidity provision increase sharply during the financial crisis of 2008-2009, or more generally in times of high uncertainty. We then show that individuals' liquidity-provision role is enhanced when uncertainty is high.

Our data enable us to look into a question that has been outstanding since Kaniel, Saar, and Titman (2008) first suggested that individuals supply liquidity as a group: Do individual retail investors benefit from this service? Our answer is no. We find that individual investors fail to reap the benefits from liquidity provision for two reasons. First, they get picked-off on the day of trading. Second, they do not reverse their trades quickly enough so that when they close their trades, the returns to liquidity provision are dissipated. Thus, procrastination – rather than frequent trading – leads to under-performance, at least for investors in our sample.

We also take advantage of the richness of our data to uncover substantial heterogeneity in the returns to liquidity provision across individuals. We show that experienced traders, who are smarter in how they supply liquidity, get less picked-off and reverse their trade much faster than less experienced ones. Overall, these two components explain a significant share of the outperformance of experienced traders relative to less experienced traders.

Finally, our data suggest that the view expressed by part of the financial press that argued that the financial crisis led to a massive exodus of small retail investors from the stock market is incomplete. Instead, at least in France, during the financial crisis, retail investors on aggregate fled from delegation by selling their mutual funds, yet, at the same time, active retail stock traders stepped up to the plate, increased stock holdings, and provided liquidity. A more detailed analysis of the source of this heterogeneity in retail investors trading behavior, specifically the dichotomous trading behavior observed for stock versus mutual funds, is an interesting topic for future research.

The similarity in French and US retail trading suggests that our results likely generalize to the US, as well as probably to other developed markets. More generally, the results indicate that when market liquidity dries up due to the increased demand of institutional

investors, or to the constraints on professional market makers, individual investors are the ones absorbing the shocks. This highlights an important role of individual investors for the efficiency of financial markets.

References

Adrian, T., Etula, E., Muir, T., 2012. Financial intermediaries and the cross section of asset returns. Staff Reports 464, Federal Reserve Bank of New York, NY.

Agarwal, S., Driscoll, J. C., Gabaix, X., Laibson, D., 2008. Learning in the credit card market. Unpublished working paper. National Bureau of Economic Research, Cambridge, MA.

Ang, A., Gorovyy, S., Van Inwegen, G. B., 2011. Hedge fund leverage. Journal of Financial Economics 102, 102-126.

Barber, B., Lee, Y.-T., Liu, Y.-J., Odean, T., 2014. The cross section of speculator skill: evidence from daytrading. Journal of Financial Markets 18, 1-24.

Barber, B. M., Lee, Y.-T., Liu, Y.-J., Odean, T., 2006. Just how much do individual investors lose by trading? Review of Financial Studies 22, 609-632.

Barber, B. M., Odean, T., 2000. Trading is hazardous to your wealth: the common stock investment performance of individual investors. Journal of Finance 55, 773-806.

Barber, B. M., Odean, T., 2002. Online investors: do the slow die first? Review of Financial Studies 15, 455-487.

Barber, B. M., Odean, T., 2011. The behavior of individual investors. Handbook of the Economics of Finance 2, 1533-1570.

Barber, B. M., Odean, T., Zhu, N., 2009. Do retail trades move markets? Review of Financial Studies 22, 151-186.

Ben-David, I., Franzoni, F., Moussawi, R., 2012. Hedge fund stock trading in the financial crisis of 2007-2009. Review of Financial Studies 25, 1-54.

Black, F., 1986. Noise. Journal of Finance 41, 529-543.

Boolell-Gunesh, M.-H., Broihanne, S., Merli, M., 2008. Are French individual investors reluctant to realize their losses? Working papers of LaRGE Research Center 2008-09. Universite de Strasbourg, Laboratoire de Recherche en Gestion et Economie (LaRGE), Strasbourg, France.

Chevalier, J., Ellison, G., 1999. Career concerns of mutual fund managers. Quarterly Journal of Economics 114, 389-432.

Chiang, Y. M., Hirshleifer, D., Qian, Y., Sherman, A. E., 2011. Do investors learn from

experience? Evidence from frequent IPO investors. Review of Financial Studies 24, 1560-1589.

Choi, J. J., Laibson, D., Madrian, B. C., Metrick, A., 2009. Reinforcement learning and savings behavior. Journal of Finance 64, 2515-2534.

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

Coval, J. D., Hirshleifer, D. A., Shumway, T., 2005. Can individual investors beat the market? Social Science Research 1001, 1-37.

De Groot, W., Huij, J., Zhou, W., 2012. Another look at trading costs and short-term reversal profits. Journal of Banking and Finance 36, 371-382.

Foucault, T., Sraer, D., Thesmar, D. J., 2011. Individual investors and volatility. Journal of Finance 66, 1369-1406.

Fong, K. Y. L., Gallagher, D. R., Lee A. D., 2014 Individual Investors and Broker Types. Journal of Financial and Quantitative Analysis 49, 431-451

French, K. R., 2008. Presidential address: the cost of active investing. Journal of Finance 63, 1537-1573.

Glaser, M., Weber, M., 2007. Overconfidence and trading volume. Geneva Risk and Insurance Review 32, 1-36.

Greenwood, R., Nagel, S., 2009. Inexperienced investors and bubbles. Journal of Financial Economics 93, 239-258.

Grinblatt, M., Keloharju, M., 2000. The investment behavior and performance of various investor types: a study of Finlands unique data set. Journal of Financial Economics 55, 43-67.

Grinblatt, M., Keloharju, M., 2009. Sensation seeking, overconfidence, and trading activity. Journal of Finance 64, 549-578.

Grossman, S. J., Miller, M. H., 1988. Liquidity and market structure. Journal of Finance 43, 617-637.

Gutierrez, R. C., Kelley, E. K., 2008. The long-lasting momentum in weekly returns. Journal of Finance 63, 415-447.

Hvidkjaer, S., 2008. Small trades and the cross section of stock returns. Review of Financial Studies 21, 1123-1151.

Kaniel, R., Liu, S., Saar, G., Titman, S., 2012. Individual investor trading and return pat-

terns around earnings announcements. Journal of Finance 67, 639-680.

Kaniel, R., Saar, G., Titman, S., 2008. Individual investor trading and stock returns. Journal of Finance 63, 273-310.

Kaustia, M., Alho, E., Puttonen, V., 2008. How much does expertise reduce behavioral biases? The case of anchoring effects in stock return estimates. Financial Management 37, 391-412.

Kelley, E. K., Tetlock, P. C., 2013. How wise are crowds? Insights from retail orders and stock returns. Journal of Finance 68, 1229-1265.

Kumar, A., Lee, C. M., 2006. Retail investor sentiment and return co-movements. Journal of Finance 61, 2451-2486.

L'Agefi, 2012. NYSE Euronext prépare une bourse dédiée aux particuliers. Philippe de Baillencourt, November 22.

Linnainmaa, J. T., 2010. Do limit orders alter inferences about investor performance and behavior? Journal of Finance 65, 1473-1506.

Linnainmaa, J. T., 2011. Why do (some) households trade so much? Review of Financial Studies 24, 1630-1666.

List, J. A., 2003. Does market experience eliminate market anomalies? Quarterly Journal of Economics 118, 41-71.

Mahani, R., Bernhardt, D., 2007. Financial speculators under-performance: learning, self-selection, and endogenous liquidity. Journal of Finance 62, 1313-1340.

Nagel, S., 2012. Evaporating liquidity. Review of Financial Studies 25, 2005-2039.

Nicolosi, G., Peng, L., Zhu, N., 2009. Do individual investors learn from their trading experience? Journal of Financial Markets 12, 317-336.

Odean, T., 1998. Do investors trade too much? SSRN Electronic Journal 89, 1279-1298.

Seru, A., Shumway, T., Stoffman, N., 2009. Learning by trading. Review of Financial Studies 23, 705-739.

Shleifer, A., Summers, L. H., 1990. The noise trader approach to finance. Journal of Economic Perspectives 4, 19-33.

Statman, M., Thorley, S., Vorkink, K., 2006. Investor overconfidence and trading volume. Review of Financial Studies 19, 1531-1565.

Stoffman, N., forthcoming. Who trades with whom? Individuals, institutions, and returns.

Journal of Financial Markets.

Wall Street Journal, 2010. Small investors flee stocks, changing market dynamics. Jim Browning, July 12.
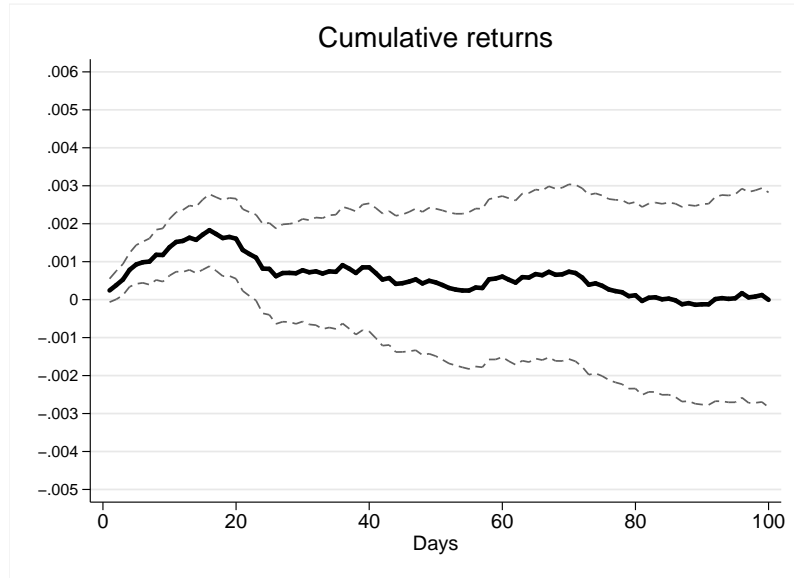
Fig. 1. Predicting returns using retail order imbalances. This graph plots the coefficient on aggregate retail imbalances, $Imb[0]$, in stock $\times$ day-level regressions in which the dependent variable is $R[1, x]$, the cumulative return from day 1 to day x (from one to one hundred) following the trading day. Controls are contemporaneous, past weekly, and past monthly returns, as well as market equity [see Eq. (4)]. $Imb[0]$ is defined as shares bought minus shares sold divided by shares bought plus shares sold. There are 730 distinct stocks traded between 2002 and 2010.

Panel A: Portfolio turnover
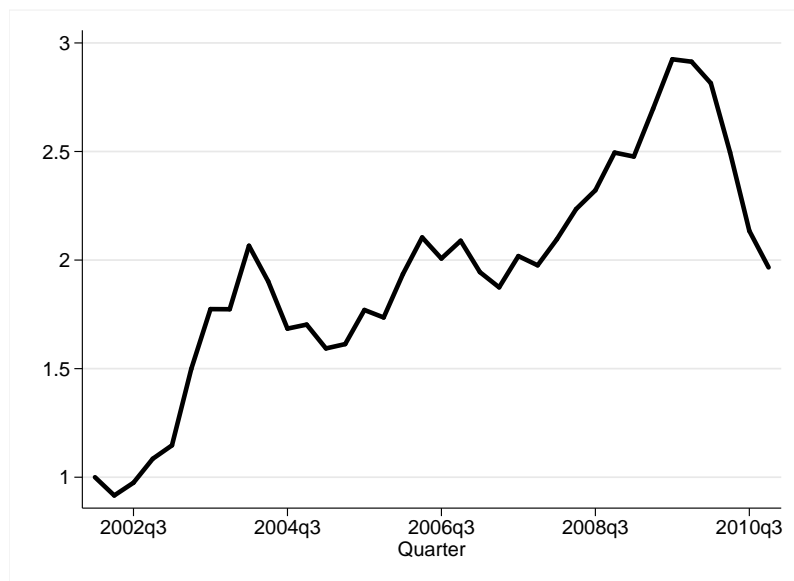
Panel B: Share of total volume
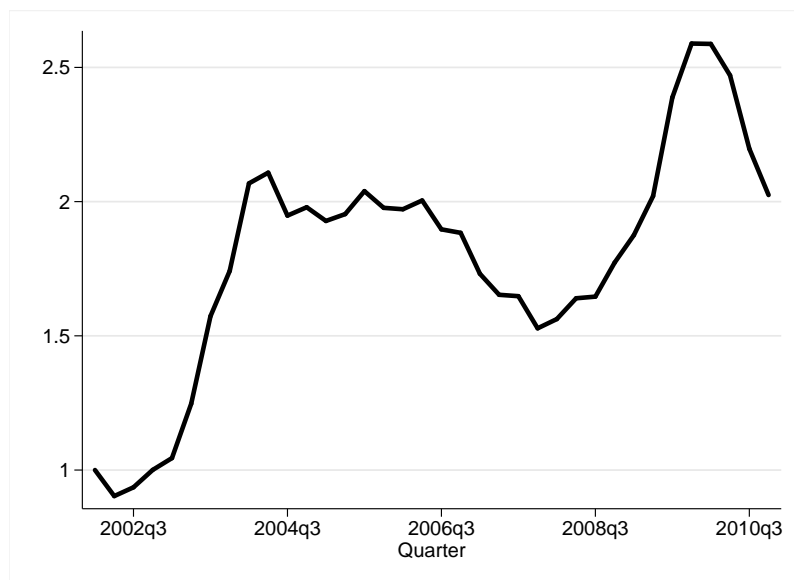
Fig. 2. Aggregate retail investor volumes. This graph plots the aggregate quarterly portfolio turnover, computed as the ratio of total volume to total portfolio holdings at the beginning of the quarter (Panel A) and the aggregate share of market-wide volume traded by retail investors in our sample (Panel B). We show the rolling average of both variables over the current and previous three quarters. Both variables are normalized to one as of the first quarter of 2002.
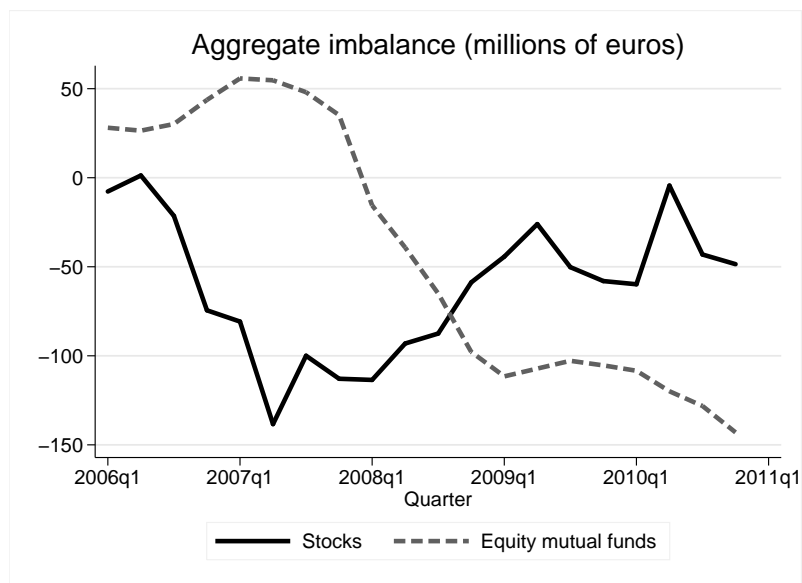
Fig. 3. Equity investment by individual investors during the crisis. This graph plots the cumulative aggregate net flows into stocks and equity mutual funds from 2006 to 2010, in millions of euros. The sample includes the trades of the 81,946 investors in our sample who traded during this period.

Panel A: High VIX days
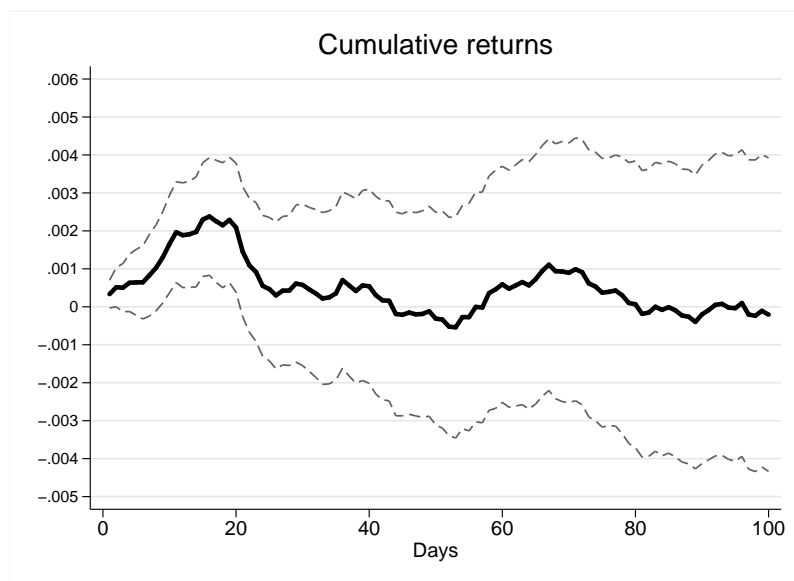
Panel B: Low VIX days
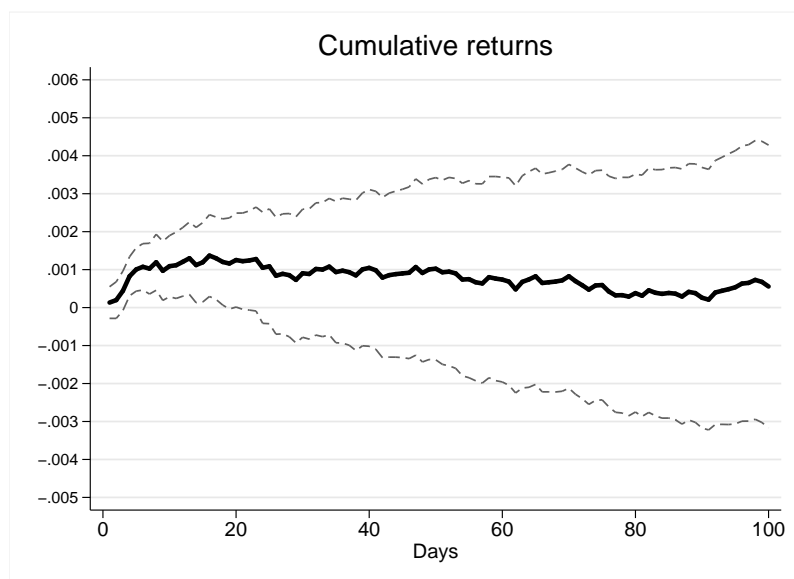
Fig. 4. Predicting returns using retail order imbalances, high versus low VIX days. This graph plots the coefficient on aggregate retail imbalances, $Imb[0]$, in stock $\times$ day-level regressions in which the dependent variable is $R[1, x]$, the cumulative return from day 1 to day x (from 1 to 100) following the trading day. Controls are contemporaneous, past weekly, and past monthly returns, as well as market equity. $Imb[0]$ is measured using shares bought minus shares sold divided by shares bought plus shares sold. There are 730 distinct stocks traded between 2002 and 2010. Panel A and B present the estimates on the subsample of days when the VIX (Chicago Board Options Exchange Volatility Index) is, respectively, above and below its 2002-2010 median.

Panel A: Stocks
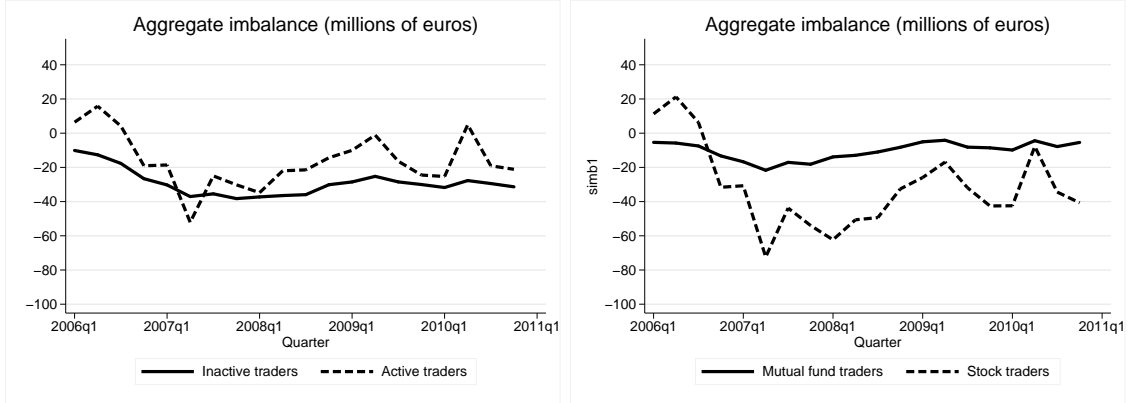
Panel B: Equity mutual funds
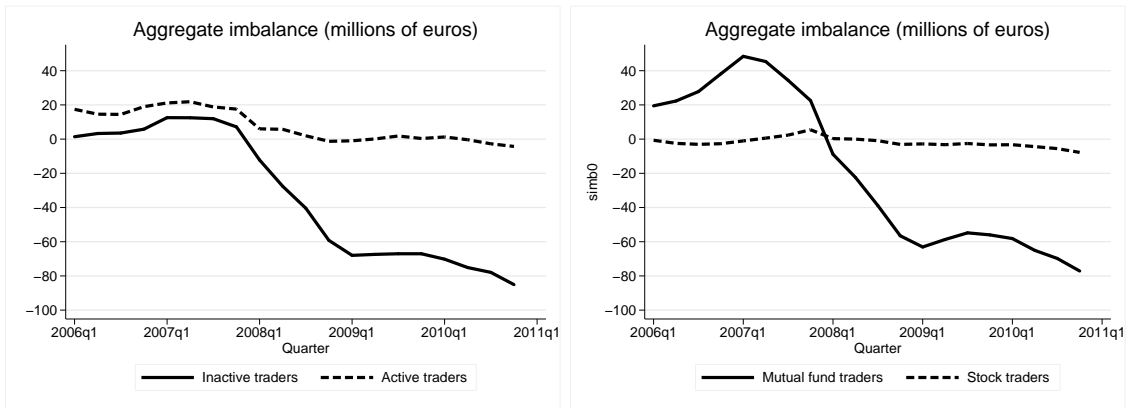
Fig. 5. Decomposed aggregate retail investor imbalances. This figure presents the decomposition of aggregate cumulative imbalances across different types of investors around the financial crisis. To allocate investors to various categories, we examine their behavior in the precrisis period, which runs from January 2004 to December 2007. Our first category relates to the trading frequency. Each quarter, we sort investors into terciles of number of trades realized in the quarter. A retail investor is in the active (inactive) category if she falls in the top (bottom) tercile in more than two thirds of the quarters. Our second category relates to the type of securities mostly traded by the investors. Each quarter, we sort investors based on the ratio of stock volume they trade to the volume of stock and mutual funds they trade. Investors are allocated to the stock (mutual funds) category if they are above (below) the median ratio in more than two-thirds of the quarters. Importantly, because information on mutual fund trading is available starting only in January 2006, this split between mutual fund and stock traders has to be performed only on the period from January 2006 to December 2007.

Table 1

Summary statistics.

   This table presents summary statistics for the stock × day-level sample (Panel A) and the order-level sample (Panel B). There are a total of 91,647 traders placing approximately 4.6 million orders in 730 stocks from 2002 to 2010 in our sample, which leaves 217,511 stock-days. $Imb[0]$ is our measure of retail imbalances, defined for a stock-day as the number of shares purchased minus sold over the number of shares purchased plus sold. $LogME$ is the log of the market capitalization of the stock. Retail volume is the absolute value, in euros, of trades in the stock originating from traders in our sample. Share of retail volume is the ratio of the number of shares of the stock traded in our sample divided by the market-wide number of shares traded. $Ret[x,y]$ is the cumulative holding period return from day x to day y. Standardized $Ret[x,y]$ is the standardized cumulative period return from day x to day y. To standardize returns, we first subtract the average past-20 day return of the stock and then divide this difference by the past-20 day standard deviation of daily returns. $Ret]-1,0]$ is the stock return on day 0. Days to reversal is the number of days from the day the order was placed until the earliest date at which the order is at least partially reversed. VIX, the Chicago Board Options Exchange Standard & Poor's (S&P) 500 implied volatility index, is normalized to a daily volatility measure by dividing it by $\sqrt{250}$. We use the value of the VIX on day $t-5$. Idiosyncratic volatility is measured as the cross-sectional standard deviation of individual stock returns on day $t-5$. *Crisis* is a dummy taking the value of one in the seven months from September 2008 to April 2009. Ret is the order holding period return. $Ret[0]$ is the percentage change from the execution price to the closing price on the day the order is placed. Variables with an *Adj* prefix are the corresponding risk-adjusted (four-factor model) quantities. IRR variables are the corresponding internal rate of return quantities.

Variable	Number of observations	Mean	Standard deviation	Median
Panel A: Stock×day-level statistics				
$Imb[0]$	217,511	-0.0034	0.5381	0.0014
Log market equity	217,511	21.61	2.07	21.82
Retail volume	217,511	165160	362246	54939
Share of retail volume	217,480	0.0241	0.0388	0.0073
$Ret]-1,0]$	217,511	0.0016	0.0308	0.0000
$Ret[-26,-6]$	217,511	0.0116	0.1239	0.0108
$Ret[-5,-1]$	217,511	0.0050	0.0646	0.0031
$Ret[1,16]$	217,511	0.0035	0.1001	0.0044
$Ret[17,100]$	217,511	0.0198	0.2336	0.0218
Max($Ret]-1,0]$,0)	217,511	0.0114	0.0207	0.0000
Min($Ret]-1,0]$,0)	217,511	-0.0098	0.0173	0.0000
Max($Ret[-5,-1]$,0)	217,511	0.0248	0.0437	0.0031
Min($Ret[-5,-1]$,0)	217,511	-0.0198	0.0358	0.0000
Max($Ret[-26,-6]$,0)	217,511	0.0493	0.0837	0.0108
Min($Ret[-26,-6]$,0)	217,511	-0.0377	0.0682	0.0000
Standardized $Ret]-1,0]$	217,470	0.0239	1.1233	-0.0370
Standardized $Ret[-5,-1]$	217,470	0.0888	1.0905	0.0606
Standardized $Ret[-26,-6]$	212,040	0.1098	1.0635	0.1601
Days to reversal	217,511	223.81	381.98	62.00
VIX	217,511	1.2962	0.6327	1.1485
Idiosyncratic volatility	217,511	2.5609	0.5196	2.4493
Crisis	217,511	0.0618	0.2408	0.0000

Panel B: Order-level statistics

Unadjusted stock returns

Ret	4,603,607	-0.0269	0.2974	0.0045
Ret[0]	4,639,850	-0.0090	0.0251	-0.0066
IRR	4,534,305	0.0004	0.0101	0.0001
IRR[0]	4,639,850	-0.0090	0.0251	-0.0066

Adjusted stock returns

AdjRet	4,637,552	-0.0093	0.2339	0.0034
AdjRet[0]	4,637,552	-0.0066	0.0269	-0.0051
AdjIRR	4,622,152	0.0003	0.0084	0.0001
AdjIRR[0]	4,637,552	-0.0066	0.0269	-0.0051

Individual trader characteristic

Log cumulative number of orders	4,639,850	5.3528	1.6332	5.4889
Squared log cumulative number of orders	4,639,850	31.3194	16.7155	30.1284
Log days to reversal	4,639,101	3.8015	2.3198	3.6889
Log monthly volume traded	4,639,850	2.9508	2.1494	2.9247
Log account size	4,639,850	2.5069	1.7928	2.7193
Log number of months since inception	4,637,671	4.1298	0.9851	4.3820

Order characteristics

Purchase	4,639,850	0.5196	0.4996	1.0000
Order size (euros)	4,639,850	7741.3	36543.9	2493.6
Purchase order size (euros)	2,410,808	7185.6	33002.7	2358.4
Sale order size (euros)	2,229,042	8342.2	40014.0	2670.0
Log order size	4,639,850	7.9	1.3	7.8
Days to reversal	4,639,850	309.7	518.1	40.0
Log days to reversal	4,639,101	3.8	2.3	3.7

Table 2
Liquidity provision.

This table presents the results of stock×day-level ordinary least squares regressions of retail order imbalances on past returns, controls, and day and stock fixed effects. $Imb[0]$ is our measure of retail imbalances, defined for a stock-day as the number of shares purchased minus sold over the number of shares purchased plus sold. $LogME$ is the log of the market capitalization of the stock. $Ret[x,y]$ is the cumulative holding period return from day x to day y. $Ret]-1,0]$ is the stock return on day 0. Standard errors are corrected for clustering at the stock level and are presented in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	Dependent variable: $Imb[0]$			
Variable	(1)	(2)	(3)	(4)
$Ret]-1,0]$	-5.22***	-5.37***		
	(0.23)	(0.24)		
$Ret[-5,-1]$	-1.01***	-1.08***		
	(0.05)	(0.05)		
$Ret[-26,-6]$	-0.26***	-0.28***		
	(0.02)	(0.02)		
Max($Ret]1,0]$,0)			-4.18***	-4.48***
			(0.21)	(0.22)
Min($Ret]1,0]]$,0)			-6.69***	-6.60***
			(0.29)	(0.28)
Max($Ret[-5,-1]$,0)			-0.79***	-0.85***
			(0.05)	(0.06)
Min($Ret[-5,-1]$,0)			-1.39***	-1.46***
			(0.07)	(0.07)
Max($Ret[-26,-6]$,0)			-0.18***	-0.16***
			(0.02)	(0.02)
Min($Ret[-26,-6]$,0)			-0.40***	-0.48***
			(0.03)	(0.03)
$LogME$	-0.02***	-0.02***	-0.02***	-0.00
	(0.00)	(0.01)	(0.00)	(0.01)
Day fixed effects	Yes	Yes	Yes	Yes
Stock fixed effects	No	Yes	No	Yes
Number of observations	217,511	217,511	217,511	217,511
R^2	0.140	0.157	0.143	0.159

Table 3

Returns to liquidity provision.

    This table presents the results of stock×day-level ordinary least squares regressions of future returns on retail order imbalances, past returns, controls, and day and stock fixed effects. There are 730 distinct stocks traded between 2002 and 2010. $Imb[0]$ is our measure of retail imbalances, defined for a stock-day as the number of shares purchased minus sold over the number of shares purchased plus sold. $LogME$ is the log of the market capitalization of the stock. $Ret[x, y]$ is the cumulative holding period return from day x to day y. $Ret]-1, 0]$ is the stock return on day 0. Standard errors are corrected for clustering at the stock level and are presented in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Variable	$Ret[1, 16]$	$Ret[17, 100]$	$Ret[1, 16]$	$Ret[17, 100]$
$Imb[0]$	0.0018***	-0.0021*	0.0019***	-0.0012
	(0.0005)	(0.0012)	(0.0005)	(0.0011)
$Ret]-1, 0]$	-0.0611***	0.0499**	-0.0598***	0.0643***
	(0.0127)	(0.0238)	(0.0129)	(0.0197)
$Ret[-5, -1]$	-0.0390***	0.0416**	-0.0429***	0.0393**
	(0.0081)	(0.0185)	(0.0079)	(0.0178)
$Ret[-26, -6]$	-0.0004	0.0482***	-0.0063	0.0337*
	(0.0070)	(0.0182)	(0.0065)	(0.0183)
$LogME$	0.0011***	0.0041**	-0.0252***	-0.1203***
	(0.0004)	(0.0016)	(0.0026)	(0.0129)
Day fixed effects	Yes	Yes	Yes	Yes
Stock fixed effects	No	No	Yes	Yes
Number of observations	217,511	217,511	217,511	217,511
R^2	0.362	0.406	0.389	0.496

Table 4
Returns to liquidity provision, alternative proxies.

This table presents the results of stock×day-level ordinary least squares regressions of future returns on retail order imbalances, past returns, controls, and day and stock fixed effects. There are 730 distinct stocks traded between 2002 and 2010. $Imb[0]$ is our measure of retail imbalances, defined for a stock-day as the number of shares purchased minus sold over the number of shares purchased plus sold. $LogME$ is the log of the market capitalization of the stock. $Ret[x, y]$ is the cumulative holding period return from day x to day y. $Ret]-1, 0]$ is the stock return on day 0. Standard errors are corrected for clustering at the stock level and are presented in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Variable	$Ret[1, 16]$	$Ret[17, 100]$	$Ret[1, 16]$	$Ret[17, 100]$
Panel A: Imb[0], normalized				
$Imb[0]$	0.0013***	0.0009	0.0009***	-0.0006
	(0.0003)	(0.0007)	(0.0003)	(0.0006)
$Ret]-1, 0]$	-0.0582***	0.0694***	-0.0605***	0.0645***
	(0.0128)	(0.0237)	(0.0129)	(0.0196)
$Ret[-5, -1]$	-0.0385***	0.0455**	-0.0431***	0.0394**
	(0.0081)	(0.0184)	(0.0078)	(0.0177)
$Ret[-26, -6]$	-0.0004	0.0491***	-0.0064	0.0337*
	(0.0070)	(0.0182)	(0.0065)	(0.0183)
$LogME$	0.0011***	0.0042***	-0.0252***	-0.1203***
	(0.0004)	(0.0016)	(0.0026)	(0.0129)
Day fixed effects	Yes	Yes	Yes	Yes
Stock fixed effects	No	No	Yes	Yes
Number of observations	217,483	217,483	217,483	217,483
R^2	0.362	0.406	0.389	0.496
Panel B: Buy-sell over market wide volume, normalized				
$Imb[0]$	0.0007***	0.0006	0.0004*	-0.0005
	(0.0002)	(0.0005)	(0.0002)	(0.0005)
$Ret]-1, 0]$	-0.0662***	0.0653***	-0.0666***	0.0676***
	(0.0122)	(0.0221)	(0.0123)	(0.0182)
$Ret[-5, -1]$	-0.0403***	0.0444**	-0.0445***	0.0399**
	(0.0081)	(0.0182)	(0.0078)	(0.0175)
$Ret[-26, -6]$	-0.0008	0.0489***	-0.0067	0.0340*
	(0.0070)	(0.0181)	(0.0065)	(0.0183)
$LogME$	0.0011***	0.0042***	-0.0252***	-0.1202***
	(0.0004)	(0.0016)	(0.0026)	(0.0129)
Day fixed effects	Yes	Yes	Yes	Yes
Stock fixed effects	No	No	Yes	Yes
Number of observations	217,458	217,458	217,458	217,458
R^2	0.362	0.406	0.388	0.496
Panel C: Tertiles of within-stock Imb[0]				
Medium $Imb[0]$	0.0012**	-0.0007	0.0010*	-0.0019
	(0.0005)	(0.0014)	(0.0005)	(0.0014)
Large $Imb[0]$	0.0026***	0.0020	0.0019***	-0.0013
	(0.0006)	(0.0014)	(0.0006)	(0.0013)
$Ret]-1, 0]$	-0.0604***	0.0696***	-0.0622***	0.0665***
	(0.0127)	(0.0234)	(0.0128)	(0.0193)
$Ret[-5, -1]$	-0.0390***	0.0454**	-0.0435***	0.0399**
	(0.0081)	(0.0184)	(0.0078)	(0.0177)
$Ret[-26, -6]$	-0.0005	0.0492***	-0.0065	0.0338*
	(0.0070)	(0.0182)	(0.0065)	(0.0183)
$LogME$	0.0011***	0.0042***	-0.0252***	-0.1202***
	(0.0004)	(0.0016)	(0.0026)	(0.0129)
Day fixed effects	Yes	Yes	Yes	Yes
Stock fixed effects	No	No	Yes	Yes
Number of observations	217,511	217,511	217,511	217,511
R^2	0.362	0.406	0.389	0.496

Table 5

Liquidity provision and the crisis: standardized returns.

This table presents the results of stock×day-level ordinary least squares regressions of retail order imbalances on past returns, controls, and day and stock fixed effects. There are 730 stocks traded from 2002 to 2010 in our sample. $Imb[0]$ is our measure of retail imbalances, defined for a stock-day as the number of shares purchased minus sold over the number of shares purchased plus sold. $LogME$ is the log of the market capitalization of the stock. $Ret[x,y]$ is the cumulative holding period return from day x to day y. $Ret]-1,0]$ is the stock return on day 0. We standardize returns by substracting the average past-20-day return of the stock and then dividing this difference by the past-20-day standard deviation of returns. $Crisis$ is a dummy taking the value of one in the seven months from September 2008 to April 2009. VIX, the Chicago Board Options Exchange Standard & Poor's (S&P) 500 implied volatility index, is normalized to a daily volatility measure by dividing it by $\sqrt{250}$. We use the value of the VIX on day $t-5$. Idiosyncratic volatility is measured as the cross-sectional standard deviation of individual stock returns on day $t-5$. Standard errors are corrected for clustering at the stock level and are presented in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	Dependent variable: $Imb[0]$					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
VIX × Standardized $Ret]-1,0]$	-0.03***	-0.03***				
	(0.00)	(0.00)				
VIX × Standardized $Ret[-5,-1]$	-0.02***	-0.01***				
	(0.00)	(0.00)				
VIX × Standardized $Ret[-26,-6]$	-0.01***	-0.01**				
	(0.00)	(0.00)				
VIX × $LogME$	0.01***	0.00				
	(0.00)	(0.00)				
Idiosyncratic volatility × Standardized $Ret]-1,0]$			-0.03***	-0.02***		
			(0.00)	(0.00)		
Idiosyncratic volatility × Standardized $Ret[-5,-1]$			-0.02***	-0.02***		
			(0.00)	(0.00)		
Idiosyncratic volatility × Standardized $Ret[-26,-6]$			-0.01***	-0.01***		
			(0.00)	(0.00)		
Idiosyncratic volatility × $LogME$			0.01***	0.00		
			(0.00)	(0.00)		
$Crisis$ × Standardized $Ret]-1,0]$					-0.04***	-0.04***
					(0.01)	(0.01)
$Crisis$ × Standardized $Ret[-5,-1]$					0.01	0.01
					(0.01)	(0.01)
$Crisis$ × Standardized $Ret[-26,-6]$					0.00	0.01
					(0.01)	(0.01)
$Crisis$ × $LogME$					0.01***	0.01***
					(0.00)	(0.00)
Standardized $Ret]-1,0]$	-0.14***	-0.15***	-0.11***	-0.12***	-0.17***	-0.18***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Standardized $Ret[-5,-1]$	-0.07***	-0.08***	-0.04***	-0.05***	-0.09***	-0.09***
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)
Standardized $Ret[-26,-6]$	-0.05***	-0.05***	-0.02**	-0.03***	-0.06***	-0.06***
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)
$LogME$	-0.03***	-0.02***	-0.04***	-0.02***	-0.02***	-0.02***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Day fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock fixed effects	No	Yes	No	Yes	No	Yes
Number of observations	212,077	212,077	212,077	212,077	212,077	212,077
R^2	0.180	0.196	0.179	0.196	0.179	0.196

Table 6

Returns to liquidity provision and the crisis.

This table presents the results of stock×day-level ordinary least squares regressions of future returns on retail order imbalances, past returns, controls, and day and stock fixed effects. There are 730 stocks traded from 2002 to 2010 in our sample. $Imb[0]$ is our measure of retail imbalances, defined for a stock-day as the number of shares purchased minus sold over the number of shares purchased plus sold. $LogME$ is the log of the market capitalization of the stock. $Ret[x, y]$ is the cumulative holding period return from day x to day y. $Ret]-1, 0]$ is the stock return on day 0. $Crisis$ is a dummy taking the value of one in the seven months from September 2008 to April 2009. VIX, the Chicago Board Options Exchange Standard & Poor's (S&P) 500 implied volatility index, is normalized to a daily volatility measure by dividing it by $\sqrt{250}$. We use the value of the VIX on day $t - 5$. Idiosyncratic volatility is measured as the cross-sectional standard deviation of individual stock returns on day $t - 5$. Standard errors are corrected for clustering at the stock level and are presented in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	Dependent variable: $Ret[1, 16]$					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
VIX $\times$ $Imb[0]$	0.0020**	0.0017**				
	(0.0008)	(0.0008)				
VIX $\times$ $Ret]-1, 0]$	-0.0282*	-0.0209				
	(0.0149)	(0.0144)				
VIX $\times$ $Ret[-5, -1]$	-0.0385***	-0.0308***				
	(0.0106)	(0.0103)				
VIX $\times$ $Ret[-26, -6]$	-0.0317***	-0.0230**				
	(0.0113)	(0.0112)				
VIX $\times$ $LogME$	0.0003	0.0015**				
	(0.0006)	(0.0006)				
Idiosyncratic volatility $\times$ $Imb[0]$			0.0024**	0.0023**		
			(0.0010)	(0.0010)		
Idiosyncratic volatility $\times$ $Ret]-1, 0]$			-0.0625***	-0.0568***		
			(0.0196)	(0.0188)		
Idiosyncratic volatility $\times$ $Ret[-5, -1]$			-0.0636***	-0.0548***		
			(0.0134)	(0.0129)		
Idiosyncratic volatility $\times$ $Ret[-26, -6]$			-0.0526***	-0.0455***		
			(0.0122)	(0.0119)		
Idiosyncratic volatility $\times$ $LogME$			-0.0013**	-0.0006		
			(0.0006)	(0.0006)		
$Crisis$ $\times$ $Imb[0]$					0.0047**	0.0041*
					(0.0023)	(0.0022)
$Crisis$ $\times$ $Ret]-1, 0]$					-0.0340	-0.0132
					(0.0360)	(0.0350)
$Crisis$ $\times$ $Ret[-5, -1]$					-0.0307	-0.0124
					(0.0288)	(0.0283)
$Crisis$ $\times$ $Ret[-26, -6]$					-0.0145	0.0064
					(0.0278)	(0.0281)
$Crisis$ $\times$ $LogME$					0.0000	0.0023
					(0.0015)	(0.0015)
$Imb[0]$	-0.0007	-0.0003	-0.0040	-0.0038	0.0016***	0.0016***
	(0.0010)	(0.0010)	(0.0025)	(0.0023)	(0.0005)	(0.0005)
$Ret]-1, 0]$	-0.0179	-0.0271	0.1117**	0.0979*	-0.0559***	-0.0573***
	(0.0258)	(0.0257)	(0.0554)	(0.0532)	(0.0134)	(0.0137)
$Ret[-5, -1]$	0.0187	0.0028	0.1366***	0.1089***	-0.0353***	-0.0416***
	(0.0172)	(0.0166)	(0.0363)	(0.0349)	(0.0085)	(0.0082)
$Ret[-26, -6]$	0.0448***	0.0259	0.1441***	0.1188***	0.0010	-0.0072
	(0.0168)	(0.0159)	(0.0338)	(0.0323)	(0.0071)	(0.0065)
$LogME$	0.0008	-0.0274***	0.0045***	-0.0235***	0.0011***	-0.0256***
	(0.0008)	(0.0026)	(0.0015)	(0.0030)	(0.0004)	(0.0025)
Day fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Stock fixed effects	No	Yes	No	Yes	No	Yes
Number of observations	217,511	217,511	217,511	217,511	217,511	217,511
R^2	0.363	0.389	0.364	0.390	0.362	0.389

Table 7

Weekly rebalanced portfolio returns.

 This table presents the excess returns (Alpha) on a weekly rebalanced portfolio that is long in the (value-weighted) stocks purchased and short in the (value-weighted) stocks sold by retail investors. Each week over the sample period running from 2002 to 2010, we aggregate individual trades at the stock level. We sort stocks based on their net retail aggregate position into two subsets of stocks sold and stocks purchased. We form a long and a short portfolio by value-weighting the stocks in each of these two subsets. In each of these two portfolios, returns are value-weighted based on the dollar values of weekly aggregate retail traders' volumes in each stock. We rebalance each of the two portfolios at the end of each week. The weekly returns on the long-short portfolio are regressed on the weekly returns on the market, small-minus-big (SMB), high-minus-low (HML), and momentum factors. Panel A presents the results for the whole sample; Panel B restricts the sample to weeks when the VIX (Chicago Board Options Exchange Volatility Index) is above 20, its 2002-2010 median; Panel C restricts the sample to stock-weeks with extreme retail imbalances. We define an imbalance as extreme if the stock-week lies either in the top or the bottom tercile of the distribution of the stock's retail imbalances between 2002 and 2010. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	(1)	2)	(3)	(4)	(5)
Panel A: All stock-weeks, all orders					
Alpha	0.0026**	0.0023**	0.0029***	0.0033***	0.0033***
	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)
Market		0.3711***	0.2673***	0.2400***	0.2276***
		(0.0656)	(0.0691)	(0.0740)	(0.0784)
SMB			-0.2008***	-0.1925***	-0.1893***
			(0.0480)	(0.0487)	(0.0492)
HML				-0.0754	-0.0618
				(0.0731)	(0.0784)
Momentum					-0.0280
					(0.0579)
Number of observations	467	467	467	467	467
R^2	0.000	0.064	0.098	0.100	0.101
Panel B: High VIX weeks, all orders					
Alpha	0.0052***	0.0058***	0.0059***	0.0064***	0.0064***
	(0.0019)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
Market		0.4197***	0.2670***	0.2130*	0.1962
		(0.0909)	(0.0984)	(0.1102)	(0.1223)
SMB			-0.2378***	-0.2291***	-0.2263***
			(0.0665)	(0.0669)	(0.0676)
HML				-0.1102	-0.0978
				(0.1013)	(0.1086)
Momentum					-0.0265
					(0.0830)
Number of observations	234	234	234	234	234
R^2	0.000	0.084	0.132	0.137	0.137
Panel C: All stock-weeks, extreme imbalances					
Alpha	0.0037***	0.0034***	0.0042***	0.0047***	0.0047***
	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
Market		0.3819***	0.2488***	0.2073**	0.1925**
		(0.0731)	(0.0764)	(0.0818)	(0.0867)
SMB			-0.2576***	-0.2449***	-0.2412***
			(0.0531)	(0.0538)	(0.0543)
HML				-0.1145	-0.0984
				(0.0808)	(0.0867)
Momentum					-0.0333
					(0.0639)
Number of observations	467	467	467	467	467
R^2	0.000	0.056	0.101	0.105	0.105

38

Table 8

Correlations.

    This table presents the correlations between variables in the order-level sample. *Ret* is the holding period return of any given order. *Ret*[0] is the difference between the closing price on the day of the order and the execution price. *AdjRet* and *AdjRet*[0] are the risk-adjusted four factor-model counterparts. Days to reversal is the number of days from the day the order was placed until the earliest date at which the order is at least partially reversed. Panel A presents the correlations based on raw holding period returns, and Panel B presents those correlations based on risk-adjusted holding period returns.

Panel A: Raw holding period returns

	Ret	*Ret*[0]	Days to reversal
Ret	1		
Ret[0]	0.103***	1	
Days to reversal	-0.197***	-0.117***	1

Panel B: Risk-adjusted holding period returns

	AdjHPR	*HPR*[0]	Days to reversal
AdjRet	1		
AdjRet[0]	0.101***	1	
Days to reversal	-0.0322***	-0.118***	1

Table 9

Decomposition of returns.

This table presents the decomposition of retail traders' returns between the return on day 0, the returns between days 1 and 16, and the returns between day 17 and the date R of reversal. Ret is the order holding period return. $Ret[0]$ is the percentage change from the execution price to the closing price on the day the order is placed. $Ret[x, y]$ is the raw cumulative returns from day x to day y. $IRR[x, y]$ is the raw internal rate of return from day x to day y. Variables with an *Adj* prefix are the corresponding risk-adjusted (four-factor model) quantities. IRR variables are the corresponding internal rate of return quantities. Days to reversal is the number of days from the day the order was placed until the earliest date at which the order is at least partially reversed.

Variable	Mean	25th percentile	50th percentile	75th percentile
Days to reversal	310	7	40	346
Raw holding period returns				
Ret	-0.0269	-0.0951	0.0045	0.0775
$Ret[0]$	-0.0090	-0.0166	-0.0066	0.0011
$Ret[0, 16]$	-0.0094	-0.0651	-0.0082	0.0472
$Ret[1, 16]$	-0.0004	-0.0546	-0.0000	0.0539
$Ret[1, R]$	-0.0171	-0.0838	0.0093	0.0841
$Ret[17, R]$	-0.0268	-0.1663	0.0067	0.1359
Risk-adjusted holding period returns (four-factor model)				
$AdjRet$	-0.0093	-0.0666	0.0034	0.0660
$AdjRet[0]$	-0.0066	-0.0169	-0.0051	0.0059
$AdjRet[0, 16]$	-0.0062	-0.0502	-0.0052	0.0388
$AdjRet[1, 16]$	0.0004	-0.0415	0.0005	0.0424
$AdjRet[1, R]$	-0.0026	-0.0565	0.0062	0.0680
$AdjRet[17, R]$	-0.0053	-0.1067	0.0083	0.1162
Raw internal rates of returns				
IRR	0.0004	-0.0017	0.0001	0.0021
$IRR[0]$	-0.0090	-0.0166	-0.0066	0.0011
$IRR[0, 16]$	-0.0009	-0.0039	-0.0005	0.0027
$IRR[1, 16]$	-0.0003	-0.0035	-0.0000	0.0033
$IRR[1, R]$	0.0014	-0.0008	-0.0000	0.0013
$IRR[17, R]$	0.0003	-0.0011	0.0001	0.0011
Risk-adjusted internal rates of returns (four-factor model)				
$AdjIRR$	0.0003	-0.0014	0.0001	0.0015
$AdjIRR[0]$	-0.0066	-0.0169	-0.0051	0.0059
$AdjIRR[0, 16]$	-0.0006	-0.0030	-0.0003	0.0022
$AdjIRR[1, 16]$	-0.0001	-0.0026	0.0000	0.0026
$AdjIRR[1, R]$	0.0008	-0.0007	0.0000	0.0009
$AdjIRR[17, R]$	0.0003	-0.0008	0.0001	0.0009

Table 10

Heterogeneity in returns.

This table presents the heterogeneity in returns across investor types. To construct Panel A, we sort orders in our sample based on the experience of the trader, measured with the total number of prior orders placed. We compute the average of all our return variables for all deciles of experience. In Panel B, we sort orders in the sample based on the propensity of each trader to quickly flip their positions. We compute the average holding period of the trader. We then compute the average of all our return variables for all deciles of average days to reversal. $Ret[x,y]$ is the raw cumulative return from day x to day y. $IRR[x,y]$ is the raw internal rate of return from day x to day y. $Ret]-1,0]$ is the percentage change from the execution price to the closing price on the day the order is placed. Ret is the order holding period return. Variables with an Adj prefix are the corresponding risk-adjusted (four-factor model) quantities. IRR variables are the corresponding internal rate of return quantities. Days to reversal is the number of days from the day the order was placed until the earliest date at which the order is at least partially reversed. ***,**, and * indicate that differences are significant at the 1%, 5%, and 10% level, respectively.

Decile	Number of traders	Number of orders	Cumulative number of orders	Days to reversal	Raw					Risk-adjusted				
					$Ret]-1,0]$	$Ret[0]$	$Ret[1,16]$	Ret	IRR	$AdjRet]-1,0]$	$AdjRet[0]$	$AdjRet[1,16]$	$AdjRet$	$AdjIRR$
Panel A: Sorted by experience														
1		462,019	12	652	-0.007	-0.017	-0.001	-0.044	0.000	-0.004	-0.015	0.001	-0.009	0.000
2		468,814	41	515	-0.008	-0.011	-0.001	-0.033	0.000	-0.005	-0.009	0.000	-0.010	0.000
3		452,279	79	429	-0.008	-0.010	-0.001	-0.031	0.000	-0.006	-0.007	0.000	-0.011	0.000
4		460,676	129	364	-0.008	-0.009	0.000	-0.030	0.000	-0.006	-0.007	0.000	-0.010	0.000
5		458,042	198	308	-0.008	-0.008	0.000	-0.028	0.000	-0.006	-0.006	0.000	-0.009	0.000
6		461,210	294	259	-0.008	-0.008	-0.001	-0.026	0.000	-0.006	-0.006	0.000	-0.010	0.000
7		460,452	436	213	-0.009	-0.007	0.000	-0.024	0.001	-0.007	-0.005	0.000	-0.009	0.000
8		459,786	665	167	-0.009	-0.007	0.000	-0.022	0.001	-0.007	-0.004	0.000	-0.009	0.001
9		460,308	1,102	126	-0.010	-0.006	0.000	-0.018	0.001	-0.007	-0.004	0.000	-0.008	0.001
10		460,021	3,323	79	-0.011	-0.007	0.001	-0.015	0.001	-0.008	-0.004	0.001	-0.007	0.001
All		4,639,850	632	310	-0.009	-0.009	0.001	-0.027	0.000	-0.006	-0.007	0.001	-0.009	0.000
9+10 versus 1+2		2186***	-481***		-0.003***	0.008***	0.001***	0.022***	0.000***	-0.002***	0.008***	0.000*	0.002***	0.001***
10 versus 1		3311***	-573***		-0.004***	0.010***	0.002***	0.028***	0.000***	-0.003***	0.011***	0.000	0.002***	0.001***
Panel B: Sorted by days to reversal														
10	9,129	18,573	33	1986	-0.005	-0.042	-0.005	-0.195	-0.001	-0.003	-0.040	-0.002	-0.052	0.000
9	9,189	33,847	36	1539	-0.005	-0.031	-0.005	-0.225	-0.001	-0.003	-0.029	0.002	-0.008	0.000
8	9,165	53,599	54	1241	-0.006	-0.023	-0.006	-0.107	0.000	-0.004	-0.021	0.002	-0.011	0.000
7	9,175	81,995	67	1028	-0.007	-0.018	-0.005	-0.040	0.000	-0.005	-0.016	0.001	-0.013	0.000
6	9,164	134,651	94	858	-0.008	-0.015	-0.003	-0.029	0.000	-0.005	-0.013	0.000	-0.010	0.000
5	9,165	213,725	96	706	-0.008	-0.013	-0.001	-0.032	0.000	-0.005	-0.011	0.000	-0.012	0.000
4	9,165	326,482	135	565	-0.007	-0.011	0.000	-0.030	0.000	-0.005	-0.009	0.000	-0.009	0.000
3	9,165	523,789	189	427	-0.008	-0.010	0.000	-0.031	0.000	-0.005	-0.008	0.000	-0.013	0.000
2	9,165	960,769	315	280	-0.008	-0.009	0.000	-0.028	0.001	-0.006	-0.007	0.000	-0.011	0.000
1	9,165	2,292,420	1065	111	-0.009	-0.007	0.000	-0.018	0.001	-0.007	-0.004	0.000	-0.007	0.001
All	9,1647	4,639,850	632	310	-0.009	-0.009	0.000	-0.027	0.000	-0.006	-0.007	0.000	-0.009	0.000
9+10 versus 1+2		808***	-1537***		-0.004***	0.028***	0.005***	0.193***	0.001***	-0.003***	0.028***	0.000	0.015***	0.001***
10 versus 1		1032***	-1875***		-0.004***	0.035***	0.005***	0.177***	0.001***	-0.003***	0.036***	0.003***	0.045***	0.001***

Table 11
Picking-off effect and time to reversal with individual characteristics.

This table presents the results of order-level ordinary least squares regressions of the return on day 0, $Ret[0]$, and the number of days to reversal on vectors of time-varying order characteristics, trader characteristics, and trader, day, and stock×day fixed effects. The dependent variable, $Ret[0]$, is computed as the percentage change from the execution price to the closing price on the day the order is placed. Days to reversal is the number of days from the day the order was placed until the earliest date at which the order is at least partially reversed. *Purchase* is a dummy equal to one if the order is a purchase and zero if it is a sale. *Gender* is a dummy equal to one if trader is a man. Standard errors are corrected for clustering at the stock level and are presented in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Log number of days to reversal						
Log cumulative number of orders	-0.1100***	-0.0858***	-0.1040***	-0.0828***	-0.0404**	-0.0418**
	(0.021)	(0.022)	(0.021)	(0.022)	(0.017)	(0.016)
Squared log cumulative number of orders	-0.0183***	-0.0207***	-0.0200***	-0.0219***	0.0232***	0.0235***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Log monthly volume traded	-0.2520***	-0.2457***	-0.2494***	-0.2442***	-0.1006***	-0.0991***
	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
Log account size	0.2174***	0.1846***	0.2069***	0.1781***	0.0253***	0.0246***
	(0.007)	(0.005)	(0.007)	(0.005)	(0.002)	(0.002)
Purchase	-1.1849***	-1.1809***	-1.1851***	-1.1813***	-1.2076***	-1.2108***
	(0.041)	(0.043)	(0.041)	(0.043)	(0.041)	(0.044)
Log order size	-0.2464***	-0.2174***	-0.2453***	-0.2164***	-0.2921***	-0.2803***
	(0.012)	(0.010)	(0.012)	(0.010)	(0.008)	(0.005)
Log age in 2002			0.2740***	0.1944***		
			(0.014)	(0.011)		
Gender			-0.0052	0.0065		
			(0.007)	(0.007)		
Day fixed effects	Yes	No	Yes	No	Yes	No
Stock × day fixed effects	No	Yes	No	Yes	No	Yes
Trader fixed effects	No	No	No	No	Yes	Yes
Number of observations	4,639,091	4,639,091	4,639,091	4,639,091	4,639,091	4,639,091
R^2	0.331	0.403	0.332	0.404	0.461	0.516
Panel B: Ret[0]						
Log cumulative number of orders	0.0035***	0.0037***	0.0035***	0.0037***	0.0010***	0.0009***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Squared log cumulative number of orders	-0.0002***	-0.0002***	-0.0002***	-0.0002***	-0.0001***	-0.0001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log monthly volume traded	-0.0006***	-0.0007***	-0.0006***	-0.0007***	-0.0002***	-0.0002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log account size	0.0004***	0.0003***	0.0004***	0.0003***	0.0001***	0.0001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Purchase	0.0024***	0.0024***	0.0024***	0.0024***	0.0015***	0.0016***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log order size	0.0050***	0.0052***	0.0050***	0.0052***	0.0050***	0.0054***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log age in 2002			-0.0004***	-0.0010***		
			(0.000)	(0.000)		
Gender			-0.0002***	-0.0002***		
			(0.000)	(0.000)		
Day fixed effects	Yes	No	Yes	No	Yes	No
Stock × day fixed effects	No	Yes	No	Yes	No	Yes
Trader fixed effects	No	No	No	No	Yes	Yes
Number of observations	4,639,840	4,639,840	4,639,840	4,639,840	4,639,840	4,639,840
R^2	0.096	0.206	0.096	0.206	0.182	0.285

Table 12

Picking-off effect and time to reversal with individual characteristics, alternative proxy for experience.

This table presents the results of order-level OLS regressions of the return on day 0, $Ret[0]$, and the number of days to reversal on vectors of time-varying order characteristics, trader characteristics, and trader, day, and stock $\times$ day fixed effects. The dependent variable, $Ret[0]$, is computed as the percentage change from the execution price to the closing price on the day the order is placed. Days to reversal is the number of days from the day the order was placed until the earliest date at which the order is at least partially reversed. *Purchase* is a dummy equal to one if the order is a purchase and zero if it is a sale. *Gender* is a dummy equal to one if trader is a man. Standard errors are corrected for clustering at the stock level and are presented in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Log number of days to reversal						
Log cumulative volume traded	-0.3921***	-0.3689***	-0.3947***	-0.3710***	-0.0501***	-0.0441***
	(0.023)	(0.024)	(0.023)	(0.024)	(0.012)	(0.012)
Squared log cumulative volume traded	0.0140***	0.0123***	0.0137***	0.0122***	0.0144***	0.0139***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Log monthly volume traded	-0.2745***	-0.2711***	-0.2729***	-0.2703***	-0.1037***	-0.1020***
	(0.005)	(0.004)	(0.005)	(0.004)	(0.002)	(0.002)
Log account size	0.2139***	0.1821***	0.2067***	0.1785***	0.0280***	0.0270***
	(0.007)	(0.005)	(0.007)	(0.005)	(0.002)	(0.002)
Purchase	-1.1752***	-1.1715***	-1.1752***	-1.1717***	-1.2082***	-1.2112***
	(0.041)	(0.044)	(0.041)	(0.044)	(0.041)	(0.044)
Log order size	-0.1191***	-0.0834***	-0.1149***	-0.0806***	-0.3000***	-0.2891***
	(0.013)	(0.011)	(0.013)	(0.011)	(0.008)	(0.005)
Log age in 2002			0.1802***	0.1031***		
			(0.014)	(0.012)		
Gender			-0.0044	0.0049		
			(0.008)	(0.008)		
Day fixed effects	Yes	No	Yes	No	Yes	No
Stock $\times$ day fixed effects	No	Yes	No	Yes	No	Yes
Trader fixed effects	No	No	No	No	Yes	Yes
Number of observations	4,639,091	4,639,091	4,639,091	4,639,091	4,639,091	4,639,091
R^2	0.324	0.396	0.324	0.396	0.461	0.516
Panel B: Ret[0]						
Log cumulative volume traded	0.0055***	0.0058***	0.0055***	0.0058***	0.0019***	0.0018***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Squared log cumulative volume traded	-0.0004***	-0.0004***	-0.0004***	-0.0004***	-0.0002***	-0.0002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log monthly volume traded	-0.0000	-0.0001**	-0.0000	-0.0001***	-0.0001***	-0.0002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log account size	0.0003***	0.0002***	0.0003***	0.0002***	0.0001***	0.0001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Purchase	0.0022***	0.0023***	0.0022***	0.0023***	0.0015***	0.0016***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log order size	0.0047***	0.0048***	0.0047***	0.0048***	0.0050***	0.0054***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log age in 2002			0.0002*	-0.0003***		
			(0.000)	(0.000)		
Gender			-0.0001	-0.0000		
			(0.000)	(0.000)		
Day fixed effects	Yes	No	Yes	No	Yes	No
Stock $\times$ day fixed effects	No	Yes	No	Yes	No	Yes
Trader fixed effects	No	No	No	No	Yes	Yes
Number of observations	4,639,840	4,639,840	4,639,840	4,639,840	4,639,840	4,639,840
R^2	0.102	0.212	0.102	0.212	0.182	0.285

Appendix A. A simple model of liquidity provision

We illustrate how liquidity provision is affected by uncertainty in the context of a simple model in the spirit of Grossman and Miller (1988).

There are T periods of trading and overlapping generations of retail investors. Retail investors in the cohort t are born in period t with initial wealth W_t. They invest by purchasing assets from cohort $t-1$ and sell their portfolios to cohort $t+1$ at date $t+1$. Cohort t has a mass μ and has CARA (constant absolute risk aversion) preferences. There are two assets: (1) a risky asset in fixed supply (normalized to zero) that delivers a risky dividend in period t of $\tilde{d}_t \sim \mathcal{N}(d, \sigma_d^2)$ and (2) a risk-free asset in infinite supply delivering a rate of return r. We assume that there is a stochastic demand shock $z_t \sim \mathcal{N}(0, \sigma_z^2)$, which can be interpreted as a demand for immediacy by institutions.

Retail investors in cohort t maximize date $t+1$ expected wealth:

$$\max_y y(E_t[\tilde{d}_{t+1} + \tilde{P}_{t+1}] - (1+r)P_t) - \frac{\gamma}{2}y^2\left(\sigma_d^2 + \sigma_P^2\right). \tag{A.1}$$

We guess that the equilibrium price is $P_t = P^* + \lambda z_t$. Then, the previous program becomes

$$\max_y y(d + P^* - (1+r)P_t) - \frac{\gamma}{2}y^2\left(\sigma_d^2 + \lambda^2\sigma_z^2\right). \tag{A.2}$$

Thus, retail investors demand is simply

$$y_t = \frac{d + P^* - (1+r)P_t}{\gamma\left(\sigma_d^2 + \lambda^2\sigma_z^2\right)}. \tag{A.3}$$

The market clearing condition in period t is $\mu y_t = -z_t$, which gives the date-t price

$$P_t = \frac{d}{r} + \gamma\frac{\sigma_d^2 + \lambda^2\sigma_z^2}{(1+r)\mu}z_t. \tag{A.4}$$

Let σ_r^2 be the variance of the asset return, then

$$P_t = \frac{d}{r} + \gamma\frac{\sigma_r^2}{(1+r)\mu}z_t. \tag{A.5}$$

This last expression highlights the sensitivity of period t price P_t to the liquidity shock z_t. When the volatility of returns is low (because, e.g., the volatility of dividend is low), retail traders are willing to accommodate the liquidity shock, so that the sensitivity of the price to the liquidity shock goes down.

The imbalances created by retail traders are by definition $\mu y_t = -z_t = -\frac{1}{\lambda}\left(\frac{d}{r} - P_t\right)$.

Define $\tilde{\epsilon}_t^d = \tilde{d}_t - d$, then

$$\mu y_t = -\frac{(1+r)\mu}{\gamma}\frac{R_t}{\sigma_r^2} + \frac{(1+r)\mu}{\gamma}\frac{\tilde{\epsilon}_t^d}{\sigma_r^2} + (1+r)\mu_{t-1}y_{t-1} \tag{A.6}$$

Retail order imbalances are negatively correlated with contemporaneous returns. Importantly, this negative correlation is smaller when σ_r^2 is larger. Absorbing the demand shock is then riskier for retail traders so that they require a larger swing in prices for a given demand shock z. Therefore, to evaluate retail investors' risk-bearing ability we need to estimate the elasticity of y to $\frac{R_t}{\sigma_r^2}$, not to R_t.

Eq. (A.6) also shows that risk-bearing capacity of retail investors as a group ($\frac{\mu}{\gamma}$) can be high for two reasons: (1) A low γ, each retail investor's risk-bearing capacity is high or (2) a high μ, a larger fraction of investors are in a position to absorb the liquidity shock.

Appendix B. Supplementary tables

Table B1

Returns to liquidity provision, alternative clustering.

This table presents the results of stock×day-level ordinary least squares regressions of future returns on retail order imbalances, past returns, controls, and day and stock fixed effects. There are 730 distinct stocks traded between 2002 and 2010. $Imb[0]$ is our measure of retail imbalances, defined for a stock-day as the number of shares purchased minus sold over the number of shares purchased plus sold. $LogME$ is the log of the market capitalization of the stock. $Ret[x, y]$ is the cumulative holding period return from day x to day y. In Columns 1 and 2, standard errors are clustered by day. In Columns 3 and 4, they are clustered two ways, by stock and day. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Variable	$Ret[1, 16]$	$Ret[17, 100]$	$Ret[1, 16]$	$Ret[17, 100]$
$Imb[0]$	0.0018***	-0.0021**	0.0018***	-0.0021*
	(0.0004)	(0.0009)	(0.0005)	(0.0013)
$Ret] - 1, 0]$	-0.0611***	0.0499*	-0.0611***	0.0499*
	(0.0126)	(0.0268)	(0.0150)	(0.0297)
$Ret[-5, -1]$	-0.0390***	0.0416***	-0.0390***	0.0416**
	(0.0070)	(0.0144)	(0.0096)	(0.0210)
$Ret[-26, -6]$	-0.0004	0.0482***	-0.0004	0.0482**
	(0.0044)	(0.0085)	(0.0078)	(0.0193)
$LogME$	0.0011***	0.0041***	0.0011***	0.0041**
	(0.0002)	(0.0004)	(0.0004)	(0.0016)
Cluster	Day	Day	Day and stock	Day and stock
Day fixed effects	Yes	Yes	Yes	Yes
Number of observations	217,511	217,511	217,511	217,511
R^2	0.362	0.406	0.002	0.003

Table B2
Returns to liquidity provision, robustness.

This table presents the results of stock×day-level ordinary least squares regressions of future returns on retail order imbalances, past returns, controls, and day and stock fixed effects. There are 730 distinct stocks traded between 2002 and 2010. $Imb[0]$ is our measure of retail imbalances, defined for a stock-day as the number of shares purchased minus sold over the number of shares purchased plus sold. $LogME$ is the log of the market capitalization of the stock. $Ret[x, y]$ is the cumulative holding period return from day x to day y. Standard errors are corrected for clustering at the stock level and are presented in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

Variable	$Ret[1, 16]$	$Ret[17, 100]$	$Ret[1, 16]$	$Ret[17, 100]$
$Imb[0]$	0.0019***	-0.0018	0.0019***	-0.0010
	(0.0005)	(0.0012)	(0.0005)	(0.0011)
Max($Ret]-1,0],0$)	-0.1102***	-0.1060**	-0.0832***	-0.0175
	(0.0194)	(0.0506)	(0.0183)	(0.0370)
Min($Ret]-1,0],0$)	0.0117	0.2805***	-0.0265	0.1821***
	(0.0222)	(0.0546)	(0.0225)	(0.0505)
Max($Ret[-5,-1],0$)	-0.0400***	0.0484*	-0.0417***	0.0455*
	(0.0110)	(0.0278)	(0.0106)	(0.0245)
Min($Ret[-5,-1],0$)	-0.0358***	0.0362	-0.0451***	0.0315
	(0.0137)	(0.0368)	(0.0138)	(0.0343)
Max($Ret[-26,-6],0$)	0.0069	0.0509*	0.0010	0.0287
	(0.0093)	(0.0269)	(0.0092)	(0.0222)
Min($Ret[-26,-6],0$)	-0.0140	0.0431	-0.0198	0.0426
	(0.0144)	(0.0342)	(0.0136)	(0.0340)
$LogME$	0.0010***	0.0036**	-0.0251***	-0.1209***
	(0.0004)	(0.0015)	(0.0025)	(0.0127)
Day fixed effects	Yes	Yes	Yes	Yes
Stock fixed effects	No	No	Yes	Yes
Number of observations	217,511	217,511	217,511	217,511
R^2	0.362	0.406	0.389	0.496