Noise Traders Incarnate:
Describing a Realistic Noise Trading Process

Joel PERESS and Daniel SCHMIDT*

August 7, 2015

ABSTRACT

We estimate a realistic process for noise trading to help theorists calibrate their models. For this purpose we characterize the trades executed by individual investors, who are natural candidates for the role of noise traders because their trades are (on average) cross-correlated, loss making, and weakly correlated with stocks' future fundamentals. We use transactions data from a retail brokerage house and small TAQ trades, obtaining consistent results. We find that noise trading can be treated as approximately i.i.d. normal at the monthly frequency, thus conforming well to standard modeling assumptions. Weekly trades follow an AR(1) process, but their residuals are not normal. Daily trades require multiple lags and have nonnormal residuals. We provide a complete description of these processes, including estimates of their standard deviation. In line with theory, these estimates are higher for more liquid and volatile stocks.

---

* Joel Peress is at INSEAD, Boulevard de Constance, 77300 Fontainebleau, France. Email: joel.peress@insead.edu. Daniel Schmidt is at HEC Paris, 1 rue de la Libération, 78350 Jouy-en-Josas, France. Email: schmidt@hec.fr. For helpful comments, we thank Bernard Dumas, Thierry Foucault and Bart Yueshen as well as seminar participants at INSEAD. We are grateful to Terry Odean for sharing the large discount brokerage data. Joel Peress thanks the AXA Research Fund and the Institut Europlace de Finance for their financial support; he also thanks the Marshall School of Business at the University of Southern California for its hospitality while some of this research was developed.

Since its inception three decades ago, the noisy rational expectations equilibrium (NREE) paradigm has led to myriad of models of trading under asymmetric information.[1] "Noise" or "liquidity" trading is an essential ingredient of these models. Without it, asset prices would perfectly reveal traders' private information, thereby undermining the incentive to collect costly information in the first place (the Grossman-Stiglitz paradox). To avoid this paradox, NREE models commonly hypothesize an exogenous noise process for the residual stock supply available to speculators. Important properties of asset prices therefore depend crucially on features of this process. Yet, little is known about the *empirical properties of a realistic noise process*, so theorists are mostly in the dark regarding its broad features and how best to calibrate their models.

In this paper, we document the properties of a realistic noise trading process. While noise trading comes in many guises and is invoked in several literatures, our focus is specifically on NRRE models. Consistent with this literature, we define a noise trade as any trade unrelated (orthogonal) to fundamental information. We then estimate a process for noise trading from retail trading data under the identifying assumption that retail trades and noise trades are correlated . While we recognize that some retail trades may be informed (and hence do not qualify as noise trades according to our definition), we emphasize that our approach *only* requires the *average* retail trader to behave as a noise trader—an assumption consistent with extensive evidence in the literature as well as with our own analysis. Barring a better alternative, we believe that our study offers valuable guidance to theorists about an essential, yet little-known part of NRRE models.

To appreciate the importance our task, consider the *persistence* of noise trades. There are at least three reasons why this persistence plays a central part in NRRE models. First, it determines the degree to which arbitrageurs are willing to correct any mispricing and, in turn, the informativeness of asset prices—for example, whether they are better or worse predictors of fundamentals than the consensus opinion (e.g., Grundy and McNichols (1989), He and Wang (1995), Cespa and Vives (2012)). Second, noise persistence

[1] Grossman (1976), Grossman and Stiglitz (1980), and Hellwig (1980) laid the foundations for noisy rational expectations models in competitive markets. Kyle (1985) offered the seminal analysis of strategic markets. To date, Grossman and Stiglitz (1980) and Kyle (1985) have over 14,000 citations in Google Scholar.

controls the serial correlations of stock returns and of trading volume (Wang (1993)). As noted by Banerjee and Kremer (2010, pp. 1271–72), "one can generate serial correlation in volume by assuming serial correlation in the aggregate supply shocks [i.e., in noise trading], or [one] can generate trade without price changes by forcing aggregate supply shocks to perfectly offset aggregate information shocks. However, this is unappealing in terms of providing insight into what generates these patterns, since the noise process is assumed to be unexplained and exogenous."

Third, the persistence of noise trades is central to the debate on how the liquidity of financial markets should be measured. In a recent empirical study, Collin-Dufresne and Fos (2014a) document that standard measures of stock price liquidity—and, in particular, of the adverse selection component (e.g., estimates of Kyle's (1985) lambda)—fail to capture the presence of informed trading. These authors inspect trades executed by informed investors and uncover a strong *positive* relation between liquidity and the likelihood of informed trades; thus, contrary to what traditional models imply, informed trades are associated with high liquidity, not with low liquidity. The leading explanation, developed further in Collin-Dufresne and Fos (2014b), is that informed investors choose when to trade and participate only when they expect the market and/or the target stock to be liquid. Since liquidity is typically associated with the presence of noise traders, this explanation is based on the notion that noise trading is predictable, such as when noise trading is persistent.

Another fundamental aspect of noise trading is its *intensity* (i.e., standard deviation). It is a key input when calibrating or simulating models without which no quantitative predictions can be made. But because noise trading is not directly observable, theorists typically either pick an arbitrary value for its variance or choose it such that the model's predicted moments match sample moments estimated from market data.[2] Although matching moments is a sensible approach, it offers no way of gauging the plausibility of the chosen noise trading parameters. More importantly, once stock market moments are matched, the

[2] An example of the former strategy is offered by Watanabe (2008, p. 246), who argues: "Since no estimate is available for the variance of individual endowment noises, it is set somewhat arbitrarily at $\Sigma_{\zeta}^{1/2} \equiv 4\Sigma_{\eta}^{1/2}$ throughout the rest of the calibration." An example of the latter strategy is given by Campbell et al. (1993, p. 931): "The trickiest part of the calibration is to specify the dynamics of the Z_t process [Z_t is the marginal investor's risk aversion, which is subject to shocks and thus generates noise trading]. We would like to pick a process that generates realistic stock price behavior."

empirical validity of a model's predictions about those moments can no longer be evaluated. By pinning down a realistic noise process, we enable researchers to bring additional testable restrictions to the data.

Finally, consider the noise trading *distribution* and its *correlation with fundamentals*. Standard models assume that noise trades are both normally distributed and uncorrelated with the asset's fundamental value. Yet, recent theoretical work suggests that neither assumption is innocuous. In fact, both are required to rule out strategic complementarities in information acquisition and hence the possibility for multiple equilibria. In Breon-Drish (2010, 2014), complementarities arise because of departures from the normal distribution. The intuition is that the price signal's informativeness varies with the price level, which can lead to a backward-bending demand curve for uninformed traders (meaning that the demand for the asset can increase with its price). This, in turn, clouds the price signal and may render the value of information nonmonotonic in the number of informed traders. In Barlevy and Veronesi (2000, 2008), complementarities arise because there is a positive correlation between the asset's fundamental value and its supply. Indeed, a high price is associated not only with a high fundamental value (as in standard models with a zero correlation) but also with a low fundamental value through a low supply. Under these conditions, prices tend to be less informative when more traders become informed, thus spurring further information acquisition. Ravi and Zigrand (2014) reach similar conclusions in a model in which investors have interdependent private valuations. Thus, equilibrium uniqueness is fragile outside the independently-distributed and normal framework. We assess the plausibility of these assumptions.

In order to estimate a realistic noise trading process, we analyze trades executed by retail investors, who are natural candidates for the role of noise traders. Indeed, previous research has documented that retail investors perform poorly on average, even before transactions costs (Odean (1999), Barber and Odean (2000)),[3] and that they trade in concert. In other words, their trades contain a common systematic component that—far from washing out in the aggregate—can actually "blur" the price signal (Kumar and

[3] We do not argue that all retail investors lose from trading, only that they do so on average. Some may be skilled investors; see Coval et al. (2005), Kaniel et al. (2012), and Kelley and Tetlock (2013).

Lee (2006), Barber et al. (2009)).[4] We use two complementary data sources for our analyses. The first are retail trading data from a large discount brokerage house. The second are small trades from the New York Stock Exchange's Trade and Quote (TAQ) database, which until decimalization in 2001 were likely to have been initiated by retail investors (Barber et al. (2009), Hvidkjaer (2008)). Both datasets have pros and cons: the retail brokerage data allows us to track individual traders but represent only a fraction of all retail trading; TAQ small trades are more comprehensive but do not reveal traders' identities. Using these two datasets, we confirm that retail trades resemble noise trades: they lose on average and occur in concert as reported in previous studies. We document further that they are only weakly correlated with stocks' future fundamentals, which is another defining feature of noise trades. [5]

We iterate that our approach relies on the assumption that retail trades are correlated with noise trades. Of course, not all retail trades are noise trades–some are actually informed. Conversely, not all noise trades are retail trades, as some institutions trade on noise too. Still, the aforementioned evidence does suggest that retail investors—notwithstanding their heterogeneity—behave *on average* as noise traders (their trades tend to be unrelated to fundamentals and unprofitable). Institutions' trades, on the other hand, are more closely related to stocks' fundamentals (e.g., Cohen et al. (2002), Campbell et al. (2009)), and tend to be profitable (e.g., Wermers (2000), Chen, et al. (2000), Fama and French (2010)).[6] Ultimately, the validity of our identifying assumption cannot be directly assessed, short of asking investors why they traded (and

[4] Considerable evidence in the literature suggests that retail investors behave as noise traders. For example, Stambaugh (2014), in his Presidential address, analyses the influence of noise trading on investment management using the fraction of US equity owned directly by individuals as a proxy for noise trading. He finds that the decline in that fraction over the past three decades explains several concomitant trends, including the shift by active managers toward lower fees and the rise of more index-like investing. Foucault et al. (2011) study a reform of the French stock market that raised the cost of trading for retail investors. They document that the consequent decline in retail trading reduced stock return volatility while increasing the magnitude of return reversals and the price impact of trades. In most theoretical accounts, these stock characteristics are associated with noise trading.

[5] The higher the correlation between noise trades and future fundamentals, the more informative stock prices are. A perfect correlation makes prices fully revealing.

[6] Wermers (2000) reports that mutual funds, on average, hold stocks that outperform the market by 1.3% per year –but underperform after deducting all expenses (namely, trading costs, management fees, costs associated with non-stock holdings). Chen, et al. (2000) estimate that stocks managers buy outperform stocks that they sell by 2% per year. Fama and French (2010)) find that funds break even on average, net of trading costs but gross of management fees, implying that they earn positive returns gross of these expenses.

even then, they may be misled…). Nonetheless, we view our work as a first attempt to measure and calibrate a realistic noise trading process for NRRE models.

We serve the needs of theorists by providing an accurate description of the noise trading process within the canonical framework they employ. For tractability, models typically assume that investors are risk neutral or display constant absolute risk aversion (CARA), so that their demand—given in number of shares (or measured as turnover after dividing by the number of shares outstanding)—is linear in random variables, including prices. Hence these models assume that aggregate noise trader demand is also measured in number of shares (or turnover), perhaps because each noise trader trades a random number of shares or because noise traders randomly participate in the market. Accordingly, we analyze three variables: (1) a measure of the number of shares traded by households, or their share turnover (the aggregate value of their trades normalized by the total value of the market); (2) the number of trades executed by households; and (3) the number of households that trade. All variables are *net* in the sense that they measure the difference between buys and sells: respectively, (1) the buy turnover minus the sell turnover; (2) the number of buy trades minus the number of sell trades; and (3) the number of households buying minus the number of households selling.[7]

Models usually assume that the distribution of noise trades is normal and either i.i.d. or with an autoregressive component. We therefore seek to fit a parsimonious autoregressive process to households' aggregate trades. We find that noise trades are serially correlated at weekly and higher frequencies. The most parsimonious models for daily trades include at least three lags, but more than ten lags are needed in some specifications to render the residuals from these models indistinguishable from white noise. Weekly trades can be described with a single lag—that is, as a first-order autogressive or AR(1) process. In contrast, we can reasonably model monthly trades as being serially uncorrelated.[8] Focusing on AR(1) processes, we

[7] These data display seasonal patterns. In line with prior studies, we find that net buys are lower in December, consistent with households realizing losses for tax purposes, and over the summer (when households are on vacation); see Badrinath and Lewellen (1991) and Hong and Yu (2008). Our own analysis is performed *after* purging the data of such calendar effects.

[8] This result confirms the intuition in Banerjee (2011). Bringing his model to the data, Banerjee argues that "[f]rom an empirical perspective, while we may expect to find persistence in supply shocks at short

find that the first-order autocorrelation coefficient declines as the duration of time periods increases, as conjectured by He and Wang (1995) and Cespa and Vives (2012). More specifically, our results indicate that the coefficient drops by 0.7-1% for each additional trading day.

Turning to their parametric form, we find that residuals are roughly normal at the monthly frequency, whereas their distributions display fat tails at higher frequency (daily, weekly). We also document that noise trades are but weakly correlated with fundamentals. Together these results suggest that strategic complementarities in information acquisition and multiple equilibria are less likely to arise at monthly or lower frequencies.

In short, households' aggregate trades at the monthly frequency match standard model assumptions: they are serially uncorrelated and normally distributed. Weekly trades are governed by an AR(1) process, as also commonly assumed, but their residuals are not normal. At the daily frequency, both the AR(1) assumption and normality are rejected.

Next, we attempt to quantify the intensity of noise trading, by no means an easy task. Even assuming (as we do) that the trades in our samples are noise trades, we cannot say what fraction of total noise trading they account for. Do the traders in our brokerage sample represent 1/10,000 or 1% of the noise trading in a stock? We develop a procedure for answering this question. The idea is to regress total trading volume in the market (from CRSP) on retail investors' trading volume. We demonstrate that the regression coefficient provides bounds on the fraction of noise trading volume accounted for by our retail trades, which in turn enables us to derive bounds on the standard deviation of noise trading in the market. We find that the households in our brokerage sample account for at least 0.039%, 0.025%, and 0.024% of all noise trades at (respectively) the daily, weekly, and monthly frequency. The implication is that the standard deviation of noise trading represents *no less* than 38%, 44%, and 36% of (respectively) the standard deviation of total daily, weekly, and monthly trading volume in the market. The estimates using small TAQ trades are remarkably similar.

---

horizons (e.g., over days or weeks), the independence assumption is not likely to be restrictive over the monthly horizon at which the predictions are tested" (p. 3032).

We also measure the noise trading intensity over groups of stocks. This analysis serves a double purpose. First, it confirms that our approach to estimating the variance of noise trading is reasonable. Indeed, consistent with extant theory, we find that noise trader risk is greater among more liquid stocks (Kyle (1985)) and stocks exhibiting greater return volatility (Hellwig (1980), He and Wang (1995)). Second, the cross-sectional estimates reported here are of interest in their own right because they can help calibrate multi-stock NREE models.

Our paper speaks to the large stream of theoretical research that specifies an exogenous noise trading process. This stream comprises models building on the seminal works of Grossman and Stiglitz (1980) and Kyle (1985), which describe investors' trading behavior and price formation in the presence of asymmetric information. Our contribution is to suggest a plausible process for noise trading that will enable theorists (i) to make qualitatively realistic assumptions and (ii) to calibrate and simulate their models without having to choose parameters arbitrarily or match moments, thus freeing up testable restrictions.

We note that a theoretically appealing alternative is to endogenize noise trading. Indeed, a few papers (e.g., Dow and Gorton (1994), Wang (1994), Dow and Gorton (1997)) follow this approach. These models offer qualitatively interesting predictions, but they are too stylized to capture a realistic noise-trading process. For example, Wang (1994) assumes rational agents have access to a private investment opportunity, whose return is random but correlated to stock returns. Shocks to these private investment returns cause random shifts in investors' demand for stocks. These noise trades inherit all the time-series and cross-sectional properties assumed for private investment returns. But barring data on these returns, the noise trading process remains largely arbitrary.

The rest of our paper proceeds as follows. Section 1 describes the data. Section 2 provides evidence for our identifying assumption that the average retail trader behaves like a noise trader. In Section 3 we explore the time-series properties of noise trading, and in Section 4 we estimate its intensity. Section 5 concludes.

1. Data

We use two transactions datasets, one from a brokerage house and the other from TAQ.

a. Households' trading data

The first dataset consists of trades by retail investors or "households" at a large discount brokerage firm. These data are described in detail by Barber and Odean (2000) and amount to some 1.9 million common stock trades executed by 78,000 households between January 1991 and November 1996. Hirshleifer et al. (2008) argue that this dataset is representative of individual investors as a whole; with 1.25 million clients (from which the 78,000 households were randomly drawn), the broker accounts for 4% of the population of individual shareholders. Moreover, Ivković et al. (2005) document that the patterns of stock sales recorded in this dataset are similar to those reported by individuals on their income-tax returns. Because the number of households in this dataset displays structural breaks in January of each year, which are likely due to how the brokerage house recorded the data rather than to actual changes in its client base, we focus on the trades of 12,743 households with portfolio holdings throughout the 1991-1996 sample period (as in Barber and Odean (2002)). We obtain virtually identical results when we use instead all the households in the dataset (with one exception as documented below).

[INSERT Figure 1 about Here]

We consider three measures of households' net buys: (1) the net turnover (henceforth "turnover"), defined as the aggregate value of their buys minus the aggregate value of their sells divided by the market's total value; (2) the net number of trades, defined as the number of households' buy trades minus their number of sell trades; and (3) the net number of households buying shares, defined as the number of households buying minus the number of households selling. Each variable is constructed at the daily, weekly, and monthly frequency. Figure 1 displays the daily time series of households' aggregate trades, and Table 1 presents summary statistics for the different frequencies.

[[INSERT Table 1 about Here]]

b. TAQ data

Our second data source consists of transactions in NYSE/AMEX stocks recorded in the TAQ database. Research has revealed that, until decimalization was introduced in 2001 (and made order splitting cost-effective), small trades were likely to stem from individual investors whereas large trades were typically placed by institutions (Hvidkjaer (2008)). We therefore use small trades over the period 1991 to 2000 to identify retail trades.[9] Trades are classified as being buyer- or seller-initiated according to the Lee and Ready (1991) algorithm, and they are classified by size via a procedure described in Hvidkjaer (2006). This procedure sorts stocks into quintiles based on NYSE/AMEX firm-size cutoff points and uses the following small-trade (resp., large-trade) cutoff points within firm-size quintiles: \$3,400 (resp., \$6,800) for the smallest firms; \$4,800 (\$9,600), \$7,300 (\$14,600), and \$10,300 (\$20,600) for the three middle quintiles; and \$16,400 (resp., \$32,800) for the largest firms. We then aggregate dollar buys and dollar sells over the entire dataset separately for small and large trades and by day, week, and month. Next we calculate the difference between buys and sells and divide by the market's total value to obtain a measure of net turnover. Thus we produce three pairs of time series for net turnover—one pair of turnovers (representing small and large trades) for each frequency. Figure 1 displays the daily time series of net turnover estimated from small trades in TAQ, and Table 1 presents summary statistics at daily, weekly, and monthly frequencies.

c. Complementarity of data sources

The brokerage firm's data on households' trades and the TAQ small trades data complement each other. One advantage of the former is that it covers retail investors exclusively—that is, noise traders as we define them. Furthermore, investors are identified and followed over time, thus enabling the measurement of investor-level variables such as stock market participation and number of trades executed. A drawback of this dataset is that it covers only a subsample of the population of retail investors and stocks and so may not be representative of either.

[9] The data for 1991 and 1992 come from the ISSM database. In analyzing various transaction databases, including the one we use here, Lee and Radhakrishna (2000) and Barber et al. (2009) confirm that trade size is an effective proxy for identifying retail trades over the 1991–2000 period.

In contrast, the TAQ dataset covers all NYSE and AMEX stocks and offers a broad view of the market. It allows for the examination of small trades with less concern about the sample's representativeness. It also offers a natural benchmark—namely, large trades—against which to compare small trades. Indeed, finding (as we do) that small trades are less profitable, more cross-correlated, and less closely related to future fundamentals than are large trades suggests that the former are made by less sophisticated investors. One shortcoming of the TAQ data is that they do not contain traders' identities, which makes it impossible to confirm that small trades are executed by retail investors. Not only are some small trades likely made by informed investors breaking up their trades to "pass" as noise traders, but also some large trades may be driven by liquidity shocks (as when an institution is subject to large inflows or redemptions from clients) and therefore qualify as noise trades.

In conclusion, since the two datasets have different strengths, we report results for both of them.

d. Seasonality

The trading data display seasonal patterns. Regressing net buys on calendar month dummies yields results consistent with prior studies (we do not report these regressions for brevity). We find that net buys are lower in December, which is consistent with individual investors realizing losses for tax purposes (Badrinath and Lewellen (1991)), and in August and September, which coincides with summer vacation (Hong and Yu (2008)). We also find some evidence of day-of-the-week effects when we regress daily data on day-of-the-week dummies, but the coefficient estimates tend to be statistically insignificant. Throughout the analysis, we purge households' and TAQ trades of calendar effects and time trends using the residuals from regressions on indicator variables for day-of-the-week, months-in-year, and year.[10]

2. Are households' and small TAQ trades noise trades?

Before turning to the main analysis, we check whether households' trades and small TAQ trades resemble noise trades. Specifically, we examine whether net buys (a) are correlated, (b) perform poorly, and (c) correlate only weakly to stocks' future fundamentals.

[10] Results are qualitatively unchanged if we use the raw data instead.

a. Correlation among trades

We first check that households' and small TAQ net buys contain a common component that does not wash out in the aggregate and hence can blur the price signal (Kumar and Lee (2006), Barber et al. (2009)). We start by looking at the household data. Following Kumar and Lee (2006), we document two related findings. First, in a stock-month panel setting, a given stock is more likely to be bought by households at times when they are buying other stocks. Second, in a household-month panel setting, a given household tends to buy stocks at times when other households are buying stocks. To establish the first result, we regress a stock's net buys (measured as turnover, the number of trades, and the number of households trading) in a given month on the average net buys across all other stocks (where this average excludes the stock's own net buy to prevent inducing an automatic correlation). Following Kumar and Lee (2006), we include the market return as a control variable to remove the common component in investor net demand that is due to overall market movements. We proceed in a similar fashion for the second result. Namely, we run a household-month panel regression of a household's net buys of all stocks in a given month (in addition to the previous measure, we include now the number of distinct stocks bought by a household) on the average net buys across all other households (where this average again excludes the household's own net buy) and the market return. The estimation results displayed in Table 2 (Panels A and B) show positive and statistically significant coefficients for average net buys in both regressions and across all trading measures. These coefficients range from 0.5 to 1, which means that a one-unit increase in average net buys increases a given stock's or household's net buys by as much as one unit and by no less than half a unit.

[[INSERT Table 2 about Here]]

We conduct a similar stock-month panel analysis using the TAQ data. As before, we regress a stock's small-trade net turnover in a given month on the average of that turnover across all other stocks (here, too, the average excludes the stock's own net turnover). Panel C of Table 2 reports a positive and statistically significant coefficient estimate for the average small-trade net turnover. As a comparison, we run the same regression for large-trade net turnover. The estimated coefficient for the average large-trade net turnover is also positive and statistically significant, but its magnitude is only half of that for small trades.

These findings confirm the existence of a strong common directional component in the trades of households and in small TAQ trades.

b. Performance of trades

Here we investigate the performance of households' and small TAQ trades. As noted by Black (1986, p. 531), "most of the time, the noise traders as a group will lose money by trading, while the information traders as a group will make money." Following Odean (1999), we measure the post-trade return difference between buy and sell transactions. Specifically, we calculate the equal-weighted average return of all buy (sell) transactions over a horizon of 84 (252) trading days subsequent to the transaction date and then take the difference. Because noise traders lose by trading against informed investors, we expect this return difference to be negative; Table 3 confirms this expectation. The first column (based on raw returns) shows that households' average post-trade return difference is a marginally significant –0.5% (t-statistic of 1.7) after 84 trading days and a highly significant –2.6% (t = 3.9) after 252 trading days. Results are similar when the post-trade return difference is measured using market-adjusted returns. Overall, these figures are close to those reported in Odean (1999).[11]

[[INSERT Table 3 about Here]]

The next three columns in Table 3 report the findings of the corresponding analysis based on TAQ trades. Small trades underperform significantly at both horizons irrespective of the return adjustment, whereas the performance of large TAQ trades is not distinguishable from zero. Thus, small trades perform poorly but large trades do not.

[11] We reach a similar conclusion when looking at the average portfolio returns of our sample households (cf. Barber and Odean (2000)). Although households earn positive raw returns (thanks to the equity risk premium), they significantly underperform their own benchmark—that is, the return they would have earned had they simply held their beginning-of-the-year portfolio for the entire year.

Both households' trades and small TAQ trades yield losses: these investors would have earned superior returns if they had not sold the stocks they sold in order to buy the stocks they bought.[12] *Bringing transaction costs into the picture would only make that underperformance worse.*

c. Trades and firms' fundamentals

A final defining feature of noise trading is its low correlation with future fundamentals. Indeed, NREE models typically define noise trades as those that are orthogonal to fundamentals. If noise trades were perfectly correlated with fundamentals then prices would be fully revealing. We undertake an empirical assessment of how closely noise trades track firms' future fundamentals. To this end, we measure, for each stock and quarter, the *earnings surprise* as the difference between actual and expected earnings, where the latter are derived from a seasonal random walk with drift (as in Bernard and Thomas (1990)). To normalize earnings surprises, we divide them by their standard deviation and label the resulting variable *standardized unexpected earnings* (SUE):

$$\mathrm{SUE}_{i,q} = \frac{E_{i,q} - (E_{i,q-4} + \mathrm{drift}_{i,q})}{\sigma_{i,q}}, \quad \text{where } \mathrm{drift}_{i,q} = \frac{1}{8}\sum_{n=1}^{8}(E_{i,q-n} - E_{i,q-n-4}).$$

Here $E_{i,q}$ denotes the actual earnings of firm *i* in quarter *q* (Compustat's earnings per share, excluding extraordinary items) and $\sigma_{i,q}$ is the standard deviation of earnings surprises estimated over the preceding eight quarters. We sort SUE into deciles and use the decile number as the dependent variable to mitigate the effect of outliers. Then, for each firm and quarter, we aggregate households' and small TAQ net buys over windows of 40, 20, 10, and 5 days; these windows end on the day *before* the firm announces its earnings. We restrict the analysis of households' trades to stocks that were traded at least 100 times over the period 1991–1996; we restrict the analysis of TAQ trades to stocks with at least $100,000 worth of small trades over the period 1991–2000.[13] Finally, we estimate a panel regression model of (pre-announcement) net buys on (announcement) earnings surprise deciles. The regression includes firm, quarter, and month-in-year fixed effects; standard errors are clustered by firm.

[12] We are not arguing that all retail investors lose from trading, only that they do so on average. Some may be skilled investors (e.g., Coval et al. (2005), Kaniel et al. (2012), and Kelley and Tetlock (2013)).
[13] Our findings are not sensitive to the choice of these filters.

[[INSERT Table 4 about Here]]

The results for households are displayed in Panel A of Table 4. The estimated coefficients for net buys are negative and statistically significant across all measures of trading, consistent with previous findings that individuals tend to be contrarian in the short-term (Kaniel et al. (2008)). However, the coefficient estimates are small in terms of economic magnitude. For example, the coefficient of –0.553 in the first row and column of the table indicates that a decrease in earnings surprises from the top decile to the bottom decile is associated with a 5×10^{-6} (= 0.553 × [10–1]/1M) increase in net turnover over the 40-day pre-announcement window, or one fifth (= $[5 \times 10^{-6}]/[25.28 \times 10^{-6}]$) of a standard deviation. The weak economic significance is also reflected in the low R-squares of less than 2%. For small TAQ trades, the estimated regression coefficient for the earnings surprise decile is no longer negative but statistically insignificant (see Panel B of the table). Nevertheless, the coefficients are significantly lower than those obtained for large TAQ trades—which are, in contrast, positively associated with futures earnings surprises.

Thus, both households' trades and small TAQ trades are extremely poor predictors of future earnings news, as one would expect of noise trades.

This econometric setup can also be used to measure the *contemporaneous* correlation between noise trading and fundamentals. As argued in the introduction, this correlation is important for assessing whether information acquisition decisions display strategic substitutability, as in the standard Grossman-Stiglitz (1980) framework, or rather complementarity, as in Barlevy and Veronesi (2000, 2008) and Rahi and Zigrand (2014).

As dependent variables in our regressions, we now use households' and small TAQ net buys (in a stock and quarter) *on the day a firm announces its earnings*. We estimate, as before, a panel regression model of these announcement net buys on earnings surprises. The results, reported in Table 4, reveal coefficient estimates of different signs and statistical significances. Yet their economic significance is weak throughout, with R-squares of at most 0.3%. These regression results suggest that, contemporaneously, noise trades are only weakly correlated with fundamentals. Hence, the scope for complementarities in information acquisition through the correlation channel appears to be limited.

Overall, households' and small TAQ trades exhibit all the attributes that we associate with noise trading: they display a strong common component, lose money, and are only weakly related to fundamentals. In the next section we turn to a time-series analysis of aggregate trades.

3. Time-series properties of aggregate trades

a. In this section, we investigate the time-series properties of aggregate trades in the households and TAQ datasets.[14] Fitting an autoregressive process to the data

Models typically assume either that noise trading is i.i.d. or that it follows an autoregressive process. We evaluate these assumptions and determine the number of lags to include. We fit households' and TAQ small net buys to autoregressive models with up to 30 lags. In Figure 2 we plot the *p*-value of a white-noise *Q*-test for the residuals (left axis). High *p*-values indicate that we cannot reject the null hypothesis of residuals from the fitted process being serially uncorrelated. We also show the value of Akaike's information criterion (dashed line and right axis) as a function of the number of lags.[15] Lower values of this criterion correspond to better models.

[[INSERT Figure 2 about Here]]

A comparison of Panels A, B, and C reveals that fewer lags are required to fit the data at lower frequencies. At the daily frequency, multiple lags are needed to eliminate serial dependence in the residuals. The number of lags ranges from 3 for small TAQ net buys to 15 for the number of households trading (using a 10% significance level). At the weekly and monthly frequencies, in contrast, one lag or less is sufficient to produce uncorrelated residuals. An AR(0) model offers a reasonable approximation for monthly data in particular, as we cannot reject the hypothesis that monthly net buys are serially uncorrelated.[16] This is good

[14] Unreported Dickey-Fuller tests confirm that these time series are stationary.

[15] Akaike's information criterion is used to discriminate among nested econometric models. It trades off goodness of fit against model complexity (in our case, the number of lags).

[16] Strictly speaking, this statement is valid only at the 10% significance level. For the net number of trades, the *p*-value for the white-noise *Q*-test on the raw monthly data (i.e., with no lag) is 6.4%. For all other variables, this value is at least 20%. When we use all households in the brokerage dataset as opposed to only those with portfolio holdings throughout the sample period, the *p*-values for net turnover and net number of households also fall just below 10%.

news for theorists because fewer lags entail less complexity in their models. The information criterion usually selects at least one lag, so an AR(1) model may prove to fit the data best.[17]

[[INSERT Figure 3 about Here]]

We now examine the performance of AR(1) processes in more detail. Indeed, several theoretical papers model noise trading as an AR(1) process and argue that the magnitude of the first-order autocorrelation coefficient decreases with the duration of a period (see e.g. He and Wang (1995), Cespa and Vives (2012)). This conjecture is consistent with our previous analysis of the lag order. It is also consistent with Figure 3, which displays the first-order autocorrelation coefficient as a function of the time period's duration (in days). A downward trend is visible in all four panels, as hypothesized by theorists. For households' turnover (upper left panel), the fitted line has a slope of −0.0066, which means that extending the period by one day reduces the coefficient by 0.0066. The slopes for the number of households' trades (lower left panel), the number of households trading (lower right panel), and small trades turnover in TAQ (lower right panel) are in that neighborhood: −0.0067, −0.0102, and −0.0086 (respectively). The solid circles in Figure 3 mark coefficients that are statistically significant at the 5% level. The plot becomes noisier as duration increases (rightward movement in the graph) because the number of periods decreases, magnifying variations in the coefficient and reducing the number of statistically significant coefficients.

In summary: Daily trades require multiple lags, but weekly and monthly trades can be accurately described with either no lag or a single lag.

b. Parametric form

Here we examine the parametric shape of aggregate trades in the households and TAQ datasets. Figure 4 plots their histograms. The curves are hump-shaped like a normal distribution, yet fat tails are also visible. Figure 5 displays quantile-to-quantile (Q-Q) plots. That is, this figure plots quantiles of trades against quantiles of a normal distribution. Points along the 45-degree line conform to a normal distribution. The

[17] When we use households' trades, the first-order autocorrelation coefficients for net turnover equal 0.156, 0.199, and 0.039 at (respectively) the daily, weekly, and monthly frequency. The corresponding values for using small TAQ trades are 0.489, 0.456, and 0.292.

daily and weekly data deviate from the 45-degree line in the tails across all trading measures; behavior that is symptomatic of the presence of extreme values (first two columns of graphs). In contrast, the monthly data are better aligned with the 45-degree line for all households' variables, suggesting that households' aggregate trades are approximately normally distributed at that frequency (last column, top three rows). However, small TAQ trades continue to display deviations from the 45-degree line even at the monthly frequency (last column, bottom row).

[[INSERT Figure 4 about Here]]

[[INSERT Figure 5 about Here]]

We formally test the hypothesis that the *residuals* from the fitted AR(1) process are normally distributed using the Shapiro-Wilk test. Table 5 presents the results. Consistent with the visual inspection of Figure 4 (though the figure is plotted for the raw variables, not for the AR(1) residuals), we find that the null hypothesis of normal residuals is rejected across all measures of noise trading at the daily and weekly frequencies; however, it is typically not rejected at the monthly frequency (except for some measures for small TAQ trades).

[[INSERT Table 5 about Here]]

We have shown that households' aggregate trades conform well to model assumptions at the monthly frequency: these trades can be considered i.i.d. normal or governed by an AR(1) process with serially uncorrelated and normally distributed residuals. The AR(1) assumption can be maintained at the weekly frequency but then normality fails. At the daily frequency, both the AR(1) assumption and normality are rejected. Small TAQ trades yield similar conclusions except that residuals do not appear normal (even at the monthly frequency).

4. Noise trading intensity

An essential aspect of noise trading is its intensity, parameterized as the *variance* of the stock's net supply in NRRE models. If noise trading follows an autoregressive process, then its variance is determined by the variance of the residual (and by the autocorrelation coefficients). Measuring this variance is a challenge and a long-standing question in finance, as reflected by the vast literature on stock market efficiency. Even when one assumes (as we do here) that households' or small TAQ trades are noise trades, we do not know what fraction of total noise trading they account for. Do they represent a small percentage, or the majority of noise trading in a stock?[18, 19] To answer this question, we relate the trading volume that originates from our households and from small TAQ trades to total trading volume in the market. After providing an overview of our strategy, we formalize it in terms of He and Wang's (1995) canonical framework. For concreteness, we show how our procedure works in the case of households; it applies equally well to small TAQ trades.

a. Overview

Trading volume consists of both noise trades and rational trades, where the latter consist in turn of two components, noninformational trades and informational trades. *Noninformational* trades are made by rational investors to accommodate supply shocks—that is, they sell (buy) when noise traders want to buy (sell), just as a market maker would do. *Informational* trades are instead motivated by speculation about future price changes and are driven by rational agents' private information. Thus:

$$\text{Total trading volume}_t = \frac{1}{2}\{\text{Noise trading volume}_t + (\text{Noninformational trading volume}_t + \text{Informational trading volume}_t)\},$$

where the factor $\frac{1}{2}$ avoids the double-counting of trades. Given that noise and noninformational trades are mirror images, the total trading volume can be expressed as follows:[20]

$$\text{Total trading volume}_t \approx \text{Noise trading volume}_t + \frac{1}{2}(\text{Informational trading volume}_t).$$

[18] Estimating predictive regressions of future fundamentals on current stock prices cannot answer this question. The reason is that the predictive power of stock prices for future fundamentals depends not only on the intensity of noise trading but also on unobservable features of rational traders such as their capital, risk aversion, and accuracy of private information. For example, low predictive power could reflect intense noise trading, binding capital constraints, high risk aversion and/or imprecise private signals.

[19] The aspects of noise trading discussed previously (e.g., lag order, autocorrelation coefficients, shape of the distribution) are independent of scale, so this question does not arise for them.

[20] We use the " $\approx$ " sign because, as discussed in what follows, noninformational and informational trades cannot be easily separated. Our formal analysis below accounts for this complication.

Our key identifying assumption is that households' trades are equal to noise trades up to a scaling factor *m*, which is our unknown. Hence

$$\text{Total trading volume}_t \approx \frac{1}{m}\,(\text{Households' trading volume}_t) + \frac{1}{2}\,(\text{Informational trading volume}_t).$$

Finally, if informational and noise trading are uncorrelated (as NRRE models typically assume), then a regression of total trading volume on households' trading volume will yield a slope coefficient equal to the inverse of the scaling factor m. This estimate, together with the standard deviation of retail trading, allows to back out the intensity of noise trading.

Our formal analysis, which we develop below in He and Wang (1995)'s model of disperse information, is made slightly more complicated by the fact that noninformational trades cannot be easily separated from informational trades. It is important to note, however, that our approach is fairly general and can be applied in most NRRE frameworks, including those with hierarchical information sets (e.g., Grossman and Stiglitz (1980)) and those with strategic traders (e.g., Kyle (1985)).[21]

b. Formalization using the He-Wang framework

He and Wang (1995) develop a dynamic rational expectations model of competitive trading volume. In their model, trading is performed by two groups of investors: noise traders and rational traders.

Noise traders have an inelastic (exogenous) demand for stocks that induces supply shocks. The residual supply of shares available to rational agents, θ_t, follows an AR(1) process:

$$\theta_t = a_\theta \theta_{t-1} + \varepsilon_{\theta,t}, \quad \text{where } -1 < a_\theta < 1 \text{ and } \varepsilon_{\theta,t} \sim N(0, \sigma_\theta^2).$$

We relate this process to our empirical analysis by normalizing the supply of shares to 1; then θ_t can be interpreted as the fraction of shares held by rational agents. Market clearing requires the change in rational

[21] Our approach only requires noise and fundamental shocks to be uncorrelated and to enter separately in investors' demand functions (which is always the case under CARA or risk neutral utility and normally distributed shocks). In Kyle (1985) for instance, total trades are made up of informed trades (a function of fundamentals only) and noise trades, which are assumed uncorrelated. Assuming again that households' trades account for a fraction m of all noise trades, we can back out *m,* and hence the intensity of noise trading, by regressing total trading volume on households' trading volume. See footnote 23 for more details on how our procedure applies in the Grossman and Stiglitz (1980) model.

agents' stock holdings to equal the noise traders' aggregate net buys. Let $\Delta\theta_t^h$ denote the net number of shares *purchased* by households in our dataset, and assume that these trades account for a fraction *m* of all noise trading in the economy. That is, *m* represents the ratio of our households' trades to total noise trades in the market. It follows that the change in the residual stock supply (which needs to be absorbed by rational agents) can be written as $\Delta\theta_t = -\frac{1}{m}\Delta\theta_t^h$.[22] Our purpose is to quantify *m.*

Under CARA utility, rational investors maximize the expected utility from consuming their wealth at the terminal date. There is a continuum of such agents, who are indexed by *i* and have unit mass. The agents receive both private and public information about a stock's fundamental value. Private signal errors are i.i.d. across investors, and public information includes the market price. Rational investors trade either to accommodate supply shocks (noninformational trading) or to speculate on future price changes based on their information (informational trading). Formally, the change in the holdings of agent *i* can be expressed as the sum of two *uncorrelated* components, $\Delta\theta_t + \Delta x_t^i$, which represent (respectively) noninformational and informational trading; see He and Wang (1995, p. 942).[23]

Total trading volume in the market comprises both noise trades and rational trades, as displayed in the following equation, where the factor $\frac{1}{2}$ prevents trades from being double-counted:

$$\text{Total volume}_t = \frac{1}{2}\left(|\Delta\theta_t| + \int_i \left|\Delta\theta_t + \Delta x_t^i\right|\right) = \frac{1}{2}\left(\left|\frac{1}{m}\Delta\theta_t^h\right| + \int_i \left|-\frac{1}{m}\Delta\theta_t^h + \Delta x_t^i\right|\right).$$

In our empirical analysis, we regress the total trading volume in the market (from CRSP), Total volume_t, on households' trading volume, $\left|\Delta\theta_t^h\right|$. In computing the regression coefficient we note that, for two jointly normal random variables z and ε and a scalar *a*,

[22] The minus signs accounts for the fact that $\Delta\theta_t$ denotes the number of shares bought by rational traders (which equals the number of shares *sold* by noise traders), while $\Delta\theta_t^h$ denotes the number of shares bought by households (which equals the number of shares bought by noise traders multiplied by *m*).

[23] In fact, our procedure extends to other models as long as we can make such a decomposition, which, again, only requires that noise and fundamentals are uncorrelated, and that investors' demand functions are separable. For example, if some investors are informed while others are not (as in Grossman and Stiglitz (1980)), then the demand of informed and uninformed investors can be expressed as $\Delta\theta_t + \Delta x_t^I$ and $\Delta\theta_t + \Delta x_t^U$, respectively, where $\Delta\theta_t$ is uncorrelated with Δx_t^I and Δx_t^U. (Of course, in equilibrium $\Delta\theta_t$, Δx_t^I and Δx_t^U are tied together by market clearing.) Nothing changes in our derivations below except that Δx_t^i is now replaced by either Δx_t^I or Δx_t^U.

$$\text{cov}(|az|, |z+\varepsilon|) = \left(1-\tfrac{2}{\pi}\right)\left(1-\sqrt{1-\text{corr}^2(az, z+\varepsilon)}\right)\sqrt{\text{var}(az)}\sqrt{\text{var}(z+\varepsilon)};$$

see Wang (1994, Apx. B). If z and ε are uncorrelated then $\text{var}(z+\varepsilon) = \text{var}(z) + \text{var}(\varepsilon)$ and $\text{corr}(az, z+\varepsilon) = \text{corr}(z, z+\varepsilon) = \sqrt{\frac{\text{var}(z)}{\text{var}(\varepsilon)+\text{var}(z)}}$, from which we can infer that

$$\text{cov}(|az|, |z+\varepsilon|) = \left(1-\tfrac{2}{\pi}\right)|a|\ \text{var}(z)\left(\sqrt{1+\frac{\text{var}(\varepsilon)}{\text{var}(z)}} - \sqrt{\frac{\text{var}(\varepsilon)}{\text{var}(z)}}\right).$$

Since $\text{var}(|az|) = \left(1-\frac{2}{\pi}\right)a^2\ \text{var}(z)$, it follows that regressing $|z+\varepsilon|$ on $|az|$ yields the regression coefficient

$$b = \frac{1}{|a|}\left(\sqrt{1+\frac{\text{var}(\varepsilon)}{\text{var}(z)}} - \sqrt{\frac{\text{var}(\varepsilon)}{\text{var}(z)}}\right).$$

If we now substitute $z = -\frac{1}{m}\Delta\theta_t^h$, $\varepsilon = \Delta x_t^i$, and $a = -m$ and then sum over all agents *i*, we obtain the coefficient from regressing rational trading volume, $\frac{1}{2}\int_i \left|-\frac{1}{m}\Delta\theta_t^h + \Delta x_t^i\right|$, on households' trading volume, $\left|\Delta\theta_t^h\right|$:

$$b_{\text{rational}} = \frac{1}{2m}\left(\sqrt{1+\frac{\text{var}(\Delta x_t^i)}{\text{var}\left(-\frac{1}{m}\Delta\theta_t^h\right)}} - \sqrt{\frac{\text{var}(\Delta x_t^i)}{\text{var}\left(-\frac{1}{m}\Delta\theta_t^h\right)}}\right) = \frac{1}{2m}\left(\sqrt{1+u} - \sqrt{u}\right),$$

where $u \equiv \frac{m^2\ \text{var}(\Delta x_t^i)}{\text{var}(\Delta\theta_t^{hh})}$ is always positive. Thus the estimated coefficient for the regression of total trading volume on households' trading volume is

$$b = \frac{1}{2m} + b_{\text{rational}} = \frac{1}{2m}\left(1+\sqrt{1+u}-\sqrt{u}\right).$$

Therefore, $m = \frac{1}{2b}\left(1+\sqrt{1+u}-\sqrt{u}\right)$. In principle, *u* can take any positive value as it depends on unobservable investor parameters such as their risk aversion or their signal precision. Nonetheless, observing that $\sqrt{1+u}-\sqrt{u}$ lies between 0 and 1 for any positive *u,* allows to bound *m* as follows:

$$\frac{1}{2b} \leq m \leq \frac{1}{b}.$$

Hence the standard deviation of noise trading is $\frac{1}{m}\sqrt{\text{var}(|\Delta\theta_t^h|)}$, where $\text{var}(|\Delta\theta_t^h|)$ is the time-series variance of households' aggregate trades.

In sum: The standard deviation of noise trading is bounded from below by the standard deviation of our households' aggregate trades multiplied by the regression coefficient of CRSP trading volume on households' trading volume, and from above by twice that product.

c. Noise trading intensity in the overall market

Table 6 displays the results of our estimation procedure for the market at large.[24] The share turnover for the overall market is defined, analogously to that for households, as the value of shares traded in the market (obtained from CRSP) divided by the value of the market. The 12,743 households in our sample (i.e., those with with 71 consecutive months of common stock positions) account for 0.039–0.078%, 0.025–0.049%, and 0.024–0.049% of all noise trades at (respectively) the daily, weekly, and monthly frequency (as indicated by the estimates for the scaling parameter *m*). Because these traders represent about 1% of the broker's clients, our figures are consistent with Hirshleifer et al.'s (2008) back-of-the-envelope estimates that the broker's clients account for approximately 4% of all US retail traders.

[[INSERT Table 6 about Here]]

The standard deviation of noise trading is in the range 0.029–0.057%, 0.150–0.301%, and 0.456–0.912% at the daily, weekly, and monthly frequency (respectively) when we use households' trades, which constitute anywhere from one third to three quarters of the standard deviation of total trades in the market. These estimates are in the same ballpark as those obtained using small TAQ trades (last three columns of Table 6). At the daily frequency, for example, the bounds on the standard deviation of noise trades are 0.021% and 0.042%, or 27% and 55% of the standard deviation of total trades in the market. This table also reports bounds on the residual's standard deviation when we assume that noise trading follows an AR(1) process

[24] We perform this analysis only for turnover, as we do not have data on the number of trades and traders in the stock market as a whole.

(bottom four rows).[25] The lower bounds for daily data are 29% and 24% of the standard deviation of total trades for, respectively, households' trades and small TAQ trades. The consistency of results across different datasets provides comfort about our procedure.

It is not clear whether our estimates of noise trading are biased upward or downward. To the extent that some noise trades are uncorrelated to households' or small TAQ trades, our approach underestimates the variance of noise trading since these trades are ignored in our calculations. On the other hand, if some households' and small TAQ trades reflect information rather than noise, then our approach overestimates the variance of noise trading by treating these as noise trades.

c. Noise trading intensity by groups of stocks

Having estimated the noise trading intensity for the market as a whole, we now extend our analysis to groups of stocks. We continue to assume that households' trades and small TAQ trades are a scaled-down version of noise trades, but we allow the scaling factor to vary over groups of stocks. Formally, for each stock *k* (or group of stocks) and day *t*, we have

$$\Delta\theta_t^k = -\frac{1}{m^k}\Delta\theta_t^{h,k};$$

here $\Delta\theta_t^{h,k}$ and $\Delta\theta_t^k$ denote (respectively) households' trades and noise trades in stock *k* on day *t,* and m^k is a stock-specific constant. A larger scaling factor m^k indicates that a stock is more underrepresented in our sample of households' trades relative to its noise-trading intensity.[26]

---

[25] If noise trading follows a stationary AR(1) process, $\theta_t = a_\theta\theta_{t-1} + \varepsilon_{\theta,t}$ where $-1 < a_\theta < 1$ and $\varepsilon_{\theta,t} \sim N(0, \sigma_\theta^2)$, then $\operatorname{var}(\Delta\theta_t) = 2\sigma_\theta^2/(1 + a_\theta)$.

[26] That scaling factors can vary across stocks has no bearing on our previous analysis of the noise-trading process's scale-independent aspects (e.g., lag order, autocorrelation coefficients, shape of the distribution). Consider, for example, the persistence of noise trading. Suppose that noise trades follow an AR(1) process whose coefficient of autocorrelation *r* is identical across stocks:

$$\Delta\theta_{t+1}^k = r\Delta\theta_t^k + \epsilon_{t+1}^k, \quad \text{where } E[\epsilon_{t+1}^i | \Delta\theta_t^k, \Delta\theta_t^j] = 0.$$

It follows that $-\frac{1}{m^k}\Delta\theta_{t+1}^{h,k} = -r\frac{1}{m^k}\Delta\theta_t^{h,k} + \epsilon_{t+1}^k$ or, equivalently, that $\Delta\theta_{t+1}^{h,k} = r\Delta\theta_t^{h,k} - m^k\epsilon_{t+1}^k$. Summing over all stocks (*k* = 1, …, *N*), we obtain

$$\textstyle\sum_k^N \Delta\theta_{t+1}^{h,k} \equiv \Delta\theta_{t+1}^h = r\Delta\theta_t^h - \sum_k^N m^k\epsilon_{t+1}^k, \quad \text{where } E[\sum_k^N m^k\epsilon_{t+1}^k \,|\Delta\theta_t^h] = 0.$$

Thus, the autocorrelation coefficient can be estimated equivalently at the market level or the stock level.

Our procedure for measuring the market-wide scaling factor can be readily applied on a stock-by-stock basis. For each stock k, we regress total trading volume on households' trading volume and use b^k to denote the resulting regression coefficient. The standard deviation of noise trading in stock k equals $\frac{1}{m^k}\sqrt{\mathrm{var}(|\Delta\theta_t^{h,k}|)}$, where $\mathrm{var}\left(\left|\Delta\theta_t^{h,k}\right|\right)$ is the time-series variance of households' aggregate trades in stock k and m^k is bounded by $\frac{1}{2b^k} \leq m^k \leq \frac{1}{b^k}$.

Our estimation of the noise trading intensity for groups of stocks proceeds in four steps. First, for each month we sort stocks into deciles on their capitalization, share price, turnover, Amihud illiquidity ratio (a measure of price impact of trades), (closing) bid-ask spread, return volatility, and return autocovariance. All 7 variables are estimated every month from daily observations. A stock's capitalization, price, turnover, and bid-ask spread are its monthly average. The Amihud illiquidity ratio is the monthly average of the daily ratio of a stock's absolute return to its dollar trading volume. The return volatility is the standard deviation of the stock's daily raw return, and the return autocovariance is the autocovariance of the stock's daily returns.

Second, within each decile, we aggregate trading volume in our sample of households and in CRSP over daily, weekly, and monthly frequencies to produce a $7 \times 10 \times 3 \times 2$ time series (one for each sorting variable, decile, frequency and trading measure). Third, we obtain the coefficients b^k (k = 1, …, 10) by regressing, decile by decile, CRSP trading volume on households' trading volume. Finally, to derive bounds on the standard deviation of noise trading in each decile, we multiply the regression coefficient b^k by the standard deviation of households' trades for that decile.

[[INSERT Figure 6 about Here]]

The results of this procedure are graphed in Figure 6 and detailed in Table 7.[27] For almost all sorting variables, the variance of households' trades and of small TAQ trades as well as the scaling factor b vary across deciles. This finding suggests that stocks differ not only in the intensity with which they are traded in our datasets but also in the fraction of noise trading for which they account. Panel E of the table, where

[27] Negative estimates of the noise trading intensity correspond to estimates of the slope coefficient that are statistically insignificant; hence they can safely be ignored.

sorting is in terms of the Amihud illiquidity ratio, illustrates the importance of scaling the variance of households' and small TAQ trades decile by decile. The standard deviation of households' and small TAQ trades is greater for stocks that are less liquid, which reflects the prevalence of retail investors among small stocks. However, this does not imply that noise trading is higher among more illiquid stocks. Indeed, the regression coefficient b^k is considerably lower for these stocks, too, which suggests that the high standard deviation of households' and small TAQ trades is scaled by a smaller factor. The bounds on noise trading depend on the product of the two, so the final effect of illiquidity on the standard deviation of noise trading is unclear a priori. Table 7 actually shows that the scaling factor's effect dominates: noise trading is less volatile for stocks that are more illiquid. Sorts by other measures of liquidity—CRSP turnover in Panel B and bid-ask spreads in Panel C—confirm the positive association between noise trading and liquidity. This is precisely what adverse selection models in the spirit of Kyle (1985) predict.

[[INSERT Table 7 about Here]]

Greater volatility in returns is associated with greater volatility in noise trading, as shown in Panel F of Table 7 and implied by most NREE models (e.g., Hellwig (1980), He and Wang (1995)). In contrast, the standard deviation of noise trading does not vary much with stock size (Panel A), most likely because size is related to many stock characteristics and sometimes in opposite ways (as with, e.g., liquidity and volatility). The standard deviation of noise trading appears to be somewhat higher for high-priced stocks (Panel C), but this tendency is likewise weak. Indeed, even though the standard deviation of households' and small TAQ trades is an order of magnitude higher in the bottom price decile than in the top price decile, our scale adjustment mitigates the stock price's effect on the standard deviation of noise trading. With respect to daily households' trades, for example, the lower bounds are $271 = 2.39 \times 113.32$ in the bottom decile versus $277 = 0.23 \times 1219.39$ in the top decile. Once again, these results underscore that our estimates do not simply mirror households' preferences for some stocks, but capture noise trading more broadly—that is, by agents not in our sample (e.g., other retail investors and institutions).

The autocovariance results in Panel G of the table are less clear-cut. Whereas noise trading intensifies with increasing autocovariance of daily stock returns when the TAQ data are used, the pattern is U-shaped when household data are used. Measured as a fraction of total trades, noise trading is more volatile in the top deciles than in the bottom deciles in both datasets. This tendency is consistent with most theoretical models (e.g., Grossman and Stiglitz (1980), Kyle (1985)), where noise trades induce temporary price shifts that encourage investors or market makers to accommodate these trades (e.g., a price increase after noise buying to encourage the sale of those shares). Such price shifts subsequently revert (here, resulting in a price reduction) because they are unrelated to fundamentals; hence they generate a negative autocorrelation in returns.

To summarize: The bounds on noise trading vary in ways that are consistent with extant theory. In particular, greater liquidity and return volatility are associated with greater noise-trading volatility—as predicted by virtually all NREE models.

5. Conclusion

In this paper we breathe life into noise trading, an essential but mostly impalpable component of trading models. We characterize the trades executed by investors who are natural candidates for the role of noise traders: individual (retail) investors. Using two different data sources (a brokerage house and the TAQ database), we estimate a realistic process for noise trading that can help theorists make qualitatively plausible assumptions about noise trading, perform comparative statics that account for any potential effect on noise trading, and—just as importantly—calibrate their models.

Our data sources yield remarkably consistent findings in spite of their dissimilarity. First, they confirm that retail trades behave just as expected from noise trades. That is, on average these trades are cross-correlated, loss making, and weakly correlated with stocks' future fundamentals. Second, we find that noise trading can be treated as i.i.d. normal at the monthly frequency, in conformance with theorists' assumptions. Weekly trades follow an AR(1) process, as is also commonly assumed, but their residuals are not normal. Daily trades require multiple lags and are not normal.

We also take up the challenge of measuring the intensity of noise trading. We estimate that the standard deviation of noise trading represents from a fifth to a third of the standard deviation of total trading volume in the market, where the exact proportion depends on the trading measure and data frequency. We also quantify the noise trading intensity over groups of stocks in order to validate our estimation strategy and to help calibrate multi-stock NREE models. We find that our estimates vary across stocks in ways that are largely consistent with the predictions of NREE models. In particular, our results confirm that noise trader risk is higher for stocks that are more liquid and/or exhibit greater return volatility.

References

Badrinath, S. G., and W. G. Lewellen, 1991, Evidence on Tax-Motivated Securities Trading Behavior, *Journal of Finance* 46, 369–382.

Banerjee, S., 2011, Learning from Prices and the Dispersion in Beliefs, *Review of Financial Studies* 24, 3025-3068.

Banerjee, S., and I. Kremer., 2010, Disagreement and Learning: Dynamic Patterns of Trade, *Journal of Finance* 65, 1269–302.

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

Barber, B. M., and T. Odean, 2002, Online investors: Do the slow die first? *Review of Financial Studies* 15, 455–89.

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

Barlevy, G., and P. Veronesi, 2000, Information Acquisition in Financial Markets, *Review of Economic Studies* 67, 79–90.

Barlevy, G. and P. Veronesi, 2008, Information Acquisition in Financial Markets: a Correction, Working Paper.

Bernard, V. L., and J. Thomas, 1990, Evidence that stock prices do not fully reflect the implications of current earnings for future earnings, *Journal of Accounting and Economics* 13, 305–340.

Black, F., 1986, Noise. *Journal of Finance* 41, 529–43.

Breon-Drish, B., 2010, Asymmetric Information in Financial Markets: Anything Goes, Working Paper.

Breon-Drish, B., 2014, On Existence and Uniqueness of Equilibrium in a Class of Noisy Rational Expectations Models, Working Paper.

Campbell, J. Y., S. J. Grossman, and J. Wang, 1993, Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics* 108, 905–939.

Campbell, J. Y., T. Ramadorai, and A. Schwartz, 2009, Caught on tape: Institutional trading, stock returns, and earnings announcements, *Journal of Financial Economics* 92, 66-91.

Cespa, G., and X. Vives, 2012, Dynamic Trading and Asset Prices: Keynes vs. Hayek, *Review of Economic Studies* 79, 539–580.

Chen, Hsiu-Lang, Narasimhan Jegadeesh, and Russ Wermers, 2000, The value of active mutual fund management: An examination of the stockholdings and trades of fund managers. *Journal of Financial and quantitative Analysis* 35.03, 343-368.

Collin-Dufresne, P., and V. Fos, 2014a, Do prices reveal the presence of informed trading? *Journal of Finance*, forthcoming.

Collin-Dufresne, P., and V. Fos, 2014b, Insider Trading, Stochastic Liquidity, and Equilibrium Prices, NBER Working Paper No.18451.

Cohen, R. B., P. A. Gompers, and T. Vuolteenaho, 2002, Who underreacts to cash-flow news? Evidence from trading between individuals and institutions, *Journal of Financial Economics* 66, 409-462.

Coval, J. D., D. A. Hirshleifer, and T. Shumway, 2005, Can Individual Investors Beat the Market? HBS Finance Working Paper No. 04-025. Available at SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=364000 or http://dx.doi.org/10.2139/ssrn.364000.

Dow, J., and G. Gorton, 1994, Arbitrage Chains, *Journal of Finance* 49, 819–849.

Dow, J., and G. Gorton, 1997, Noise Trading, Delegated Portfolio Management, and Economic Welfare, *Journal of Political Economy* 105, 1024–1050.

Fama, E. F., and K. R. French, 2010, Luck versus skill in the cross-section of mutual fund returns, *Journal of Finance* 65, 1915-1947.

Foucault, T., D. Sraer, and D. Thesmar, 2011, Individual Investors and Volatility*. Journal of Finance* 66, 1369–1406.

Grossman, S. J., 1976, On the Efficiency of Competitive Stock Markets Where Trades Have Diverse Information*, Journal of Finance* 31, 573–585.

Grossman, S. J. and J. E. Stiglitz, 1980, On the Impossibility of Informationally Efficient Markets, *American Economic Review* 70, 393–408.

Grundy, B., and M. McNichols, 1989, Trade and Revelation of Information through Prices and Direct Disclosure, *Review of Financial Studies* 2, 495–526.

He, H. and J. Wang, 1995, Differential Information and Dynamic Behavior of Stock Trading Volume, *Review of Financial Studies* 8, 919–972.

Hellwig, M., 1980, On the Aggregation of Information in Competitive Markets, *Journal of Economic Theory* 22, 477–98.

Hirshleifer, D. A., J. N. Myers, L. A. Myers, and S. H. Teoh, 2008*,* Do Individual Investors Cause Post-Earnings Announcement Drift? Direct Evidence from Personal Trades, *Accounting Review* 83, 1521–1550.

Hong, H.G. and J. Yu, 2008, Gone Fishin': Seasonality in Trading Activity and Asset Prices, *Journal of Financial Markets* 12, 672–702.

Hvidkjaer, S., 2006, A Trade-based Analysis of Momentum, *Review of Financial Studies* 19, 457–491.

Hvidkjaer, S., 2008, Small Trades and the Cross-section of Stock Returns, *Review of Financial Studies* 21, 1123–1151.

Ivković, Z., J. Poterba, and S. Weisbenner, 2005, Tax-Motivated Trading by Individual Investors, *American Economic Review* 95, 1605–1630.

Kaniel, R., S. Liu, G. Saar, and S. Titman, 2012, Individual Investor Trading and Return Patterns around Earnings Announcements, *Journal of Finance* 67, 639–680.

Kaniel, R., Saar, G. and Titman, S. 2008, Individual Investor Trading and Stock Returns, *Journal of Finance* 63, 273–310.

Kelley, E., and P. Tetlock, 2013, How Wise are Crowds? Insights from Retail Orders and Stock Returns, *Journal of Finance* 68, 1229–1265.

Kumar, A., and C. Lee, 2006, Retail Investor Sentiment and Return Comovements, *Journal of Finance* 61, 2451–2486.

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

Lee, Charles M. C., and Balkrishna Radhakrishna, 2000, Inferring investor behavior: Evidence from TORQ data, *Journal of Financial Markets* 3, 83–111.

Lee, Charles and Mark J. Ready, 1991, Inferring Trade Direction from Intraday Data, *Journal of Finance* 46, 733-46.

Odean, T., 1999, Do Investors Trade Too Much? *American Economic Review* 89, 1279–98.

Rahi, R., and J. Zigrand, 2014, Information Aggregation in a Competitive Economy, Working Paper. Available at SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2516500.

Stambaugh, R. F., 2014, Investment Noise and Trends, *Journal of Finance* 69, 1415–1453.

Wang, J., 1993, A Model of Intertemporal Asset Prices Under Asymmetric Information, *Review of Economic Studies* 60, 249–282.

Wang, J., 1994, A model of competitive stock trading volume, *Journal of Political Economy* 102, 127–168.

Watanabe, M., 2008, Price Volatility and Investor Behavior in an Overlapping Generations Model with Information Asymmetry, *Journal of Finance* 63, 229–272.

Wermers, R., 2000, Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses, *Journal of Finance*, 55(4),1655-1703.

Figure 1: Time series of households' daily aggregate trades

The graphs in this figure display time series of households' daily aggregate trades at a large discount broker (between January 1991 and November 1996) as well as of small TAQ trades (between January 1991 and December 2000). The household data are for those holding common stock positions for 71 consecutive months. The TAQ data include all small trades in NYSE/AMEX stocks, where trades are classified as small or large based on the procedure described in Hvidkjaer (2006); see Section 1.b. The top panel considers households' net turnover, defined as the aggregate value of their buys minus the aggregate value of their sells divided by the market's total value. The second panel considers households' net number of buy trades, defined as the number of households' buy trades minus the number of sell trades. The third panel considers the net number of households buying shares, defined as the number of households buying minus the number of households selling. The bottom panel considers the net turnover for small trades in the TAQ dataset, which is defined as the aggregate value of small buys minus the aggregate value of small sells divided by the market's total value. We adjust all variables for seasonality and time trends by regressing them on day-of-the week, month-in-year, and year dummy variables and then taking the residuals.

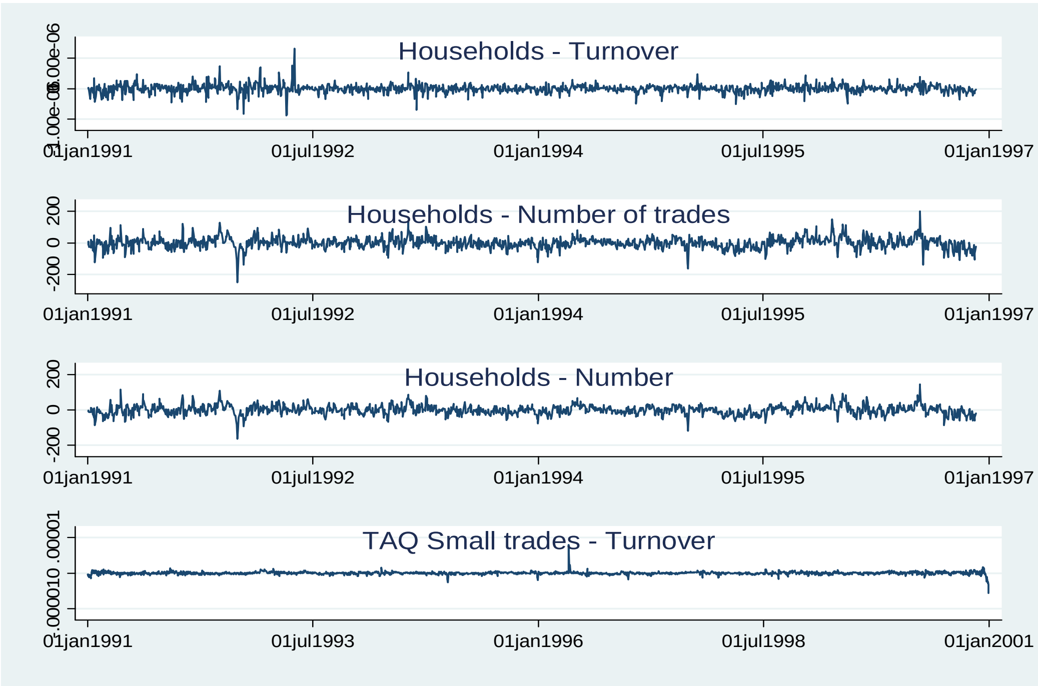
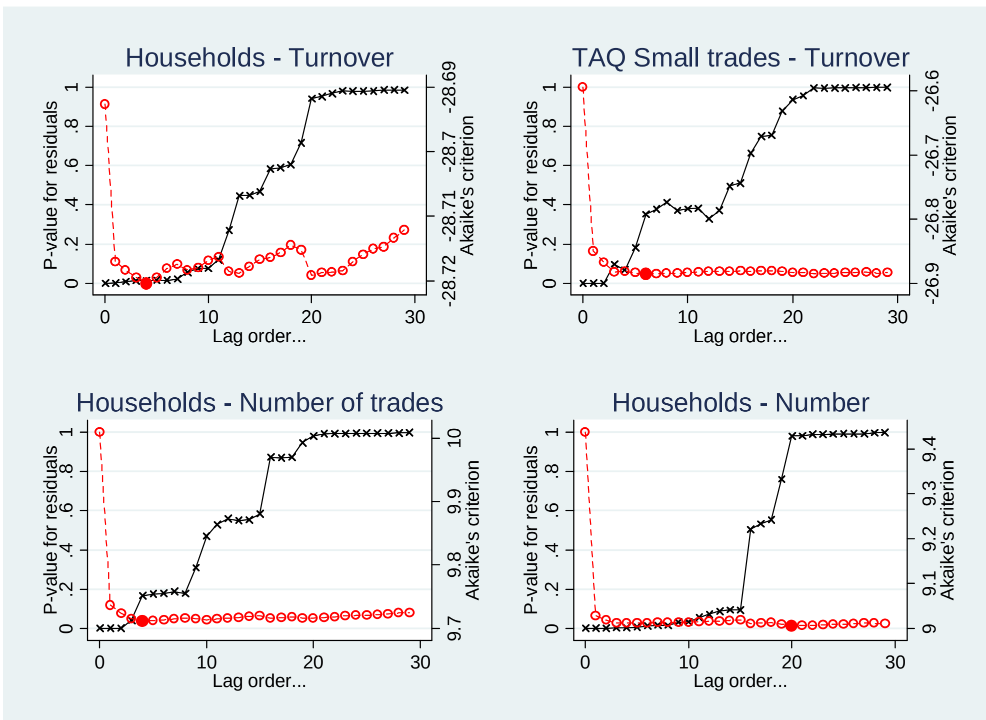


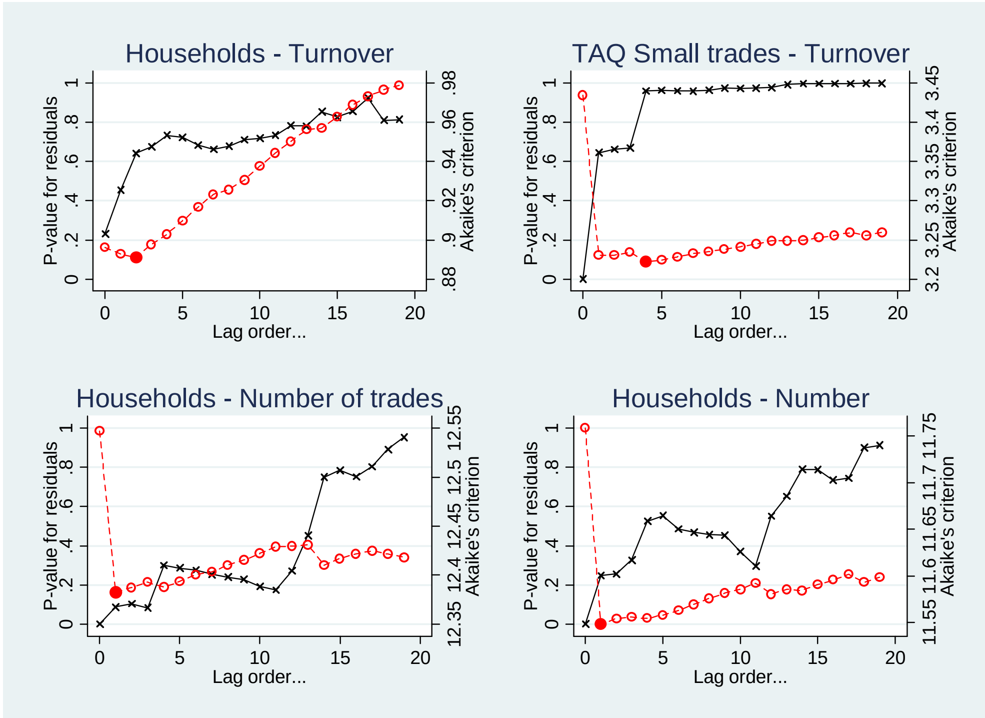
Figure 2: Lag-order selection

This figure displays the performance of autoregressive models fitted to households' aggregate trades (top left and bottom left and right panels) and to TAQ small trades (top right panel) as a function of the number of lags. The number of lags ranges from 0 to 30, 0 to 20, and 0 to 10 at (respectively) daily (Panel A), weekly (Panel B), and monthly (Panel C) frequencies. The graphs' crosses and left axes mark *p*-values of a white-noise *Q*-test for residuals of the fitted data. High *p*-values indicate that we cannot reject the null hypothesis of the residuals being serially uncorrelated. The circles and right axes mark the value of Akaike's information criterion, where lower values correspond to better models; a solid circle marks the lag order that is optimal by this criterion. We adjust all variables for seasonality and time trends by regressing them on day-of-the-week (as applies), month-in-year, and year dummy variables and then taking the residuals.

Panel A: Daily frequency



Panel B: Weekly frequency



Panel C: Monthly frequency

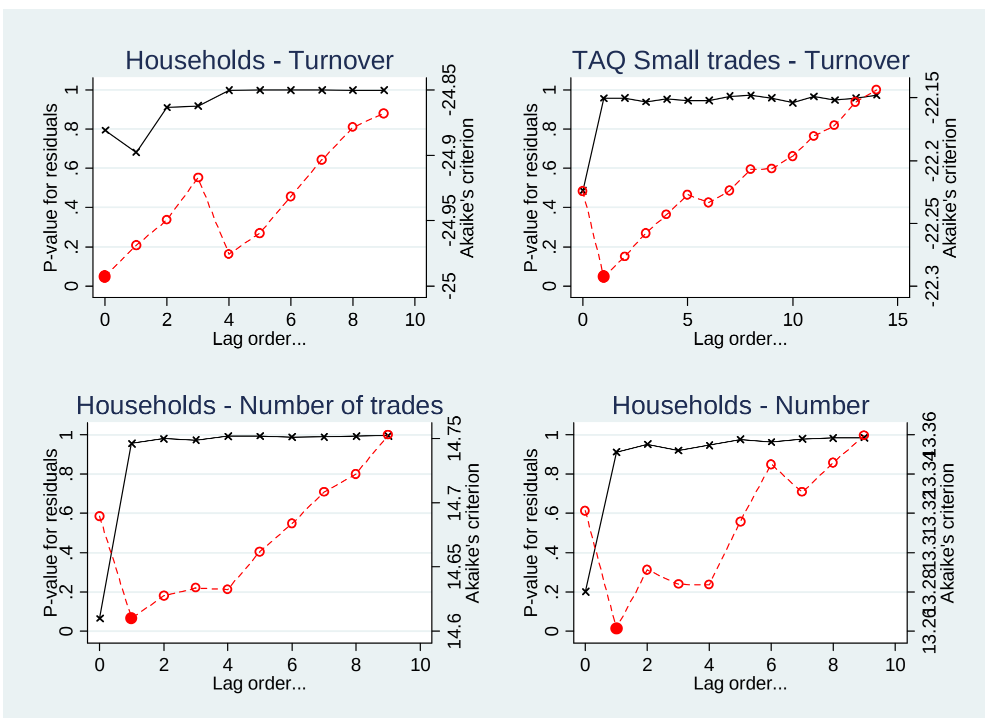


Figure 3: Fitting an AR(1) process to households' aggregate trades

The graphs in this figure plot the first-order autocorrelation coefficient of aggregate trades as a function of the duration of a time period in days. Solid circles mark coefficients that are statistically significant at the 5% level. The top left panel considers households' net turnover, defined as the aggregate value of their buys minus the aggregate value of their sells divided by the market's total value. The bottom left panel considers households' net number of buy trades, defined as the number of households' buy trades minus the number of sell trades. The bottom right panel considers the net number of households buying shares, defined as the number of households buying minus the number of households selling. The top right panel considers the net turnover for small trades in the TAQ dataset, which is defined as the aggregate value of small buys minus the aggregate value of small sells divided by the market's total value. The TAQ data include all small trades in NYSE/AMEX stocks, where trades are classified as small or large based on the procedure described in Hvidkjaer (2006); see Section 1.b. We adjust all variables for seasonality and time trends by regressing them on month-in-year and year dummy variables and then taking the residuals.

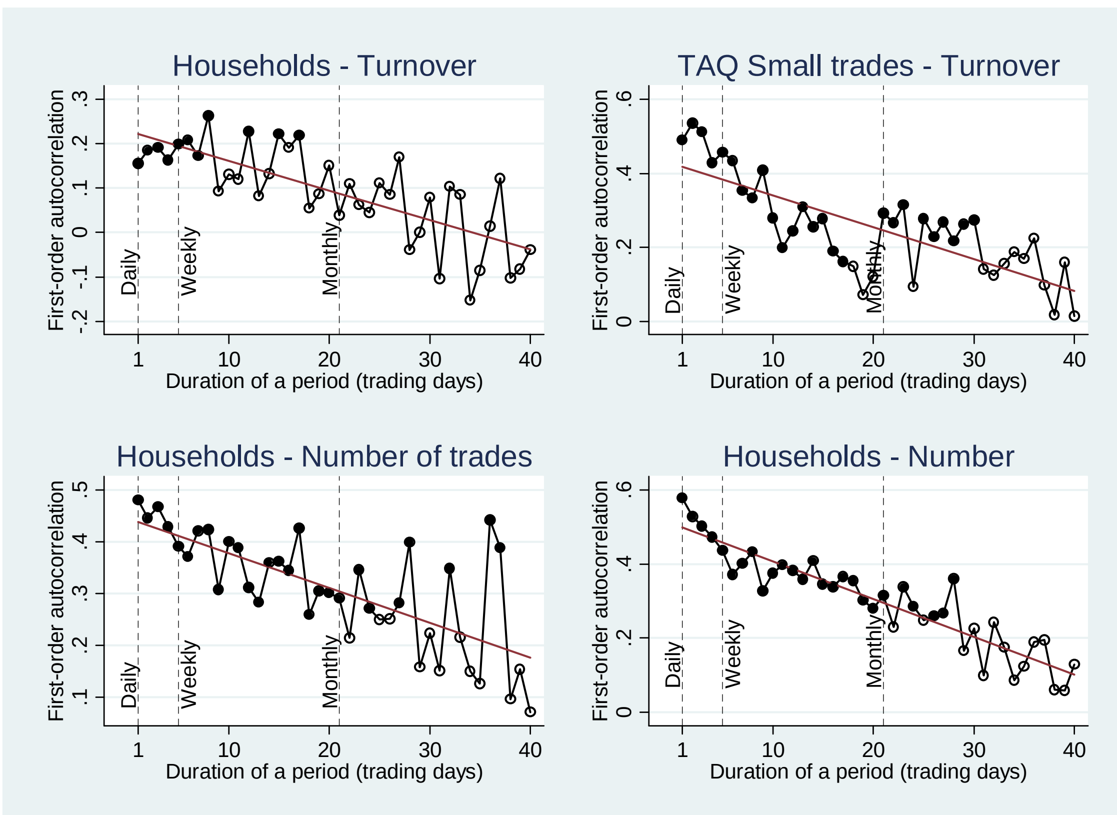


Figure 4: Histograms of households' daily aggregate trades

The graphs in this figure are histograms of aggregate trades. The left column considers daily data, the middle column considers weekly data, and the right column considers monthly data. The top row considers households' net turnover, defined as the aggregate value of their buys minus the aggregate value of their sells divided by the total value of the market. The second row considers households' net number of buy trades, defined as the number of households' buy trades minus the number of sell trades. The third row considers the net number of households buying shares, defined as the number of households buying minus the number of households selling. The bottom row considers the net turnover for small trades in the TAQ dataset, which is defined as the aggregate value of small buys minus the aggregate value of small sells divided by the market's total value. The TAQ data include all small trades in NYSE/AMEX stocks, where trades are classified as small or large based on the procedure described in Hvidkjaer (2006); see Section 1.b. We adjust all variables for seasonality and time trends by regressing them on day-of-the-week, month-in-year, and year dummy variables and then taking the residuals.

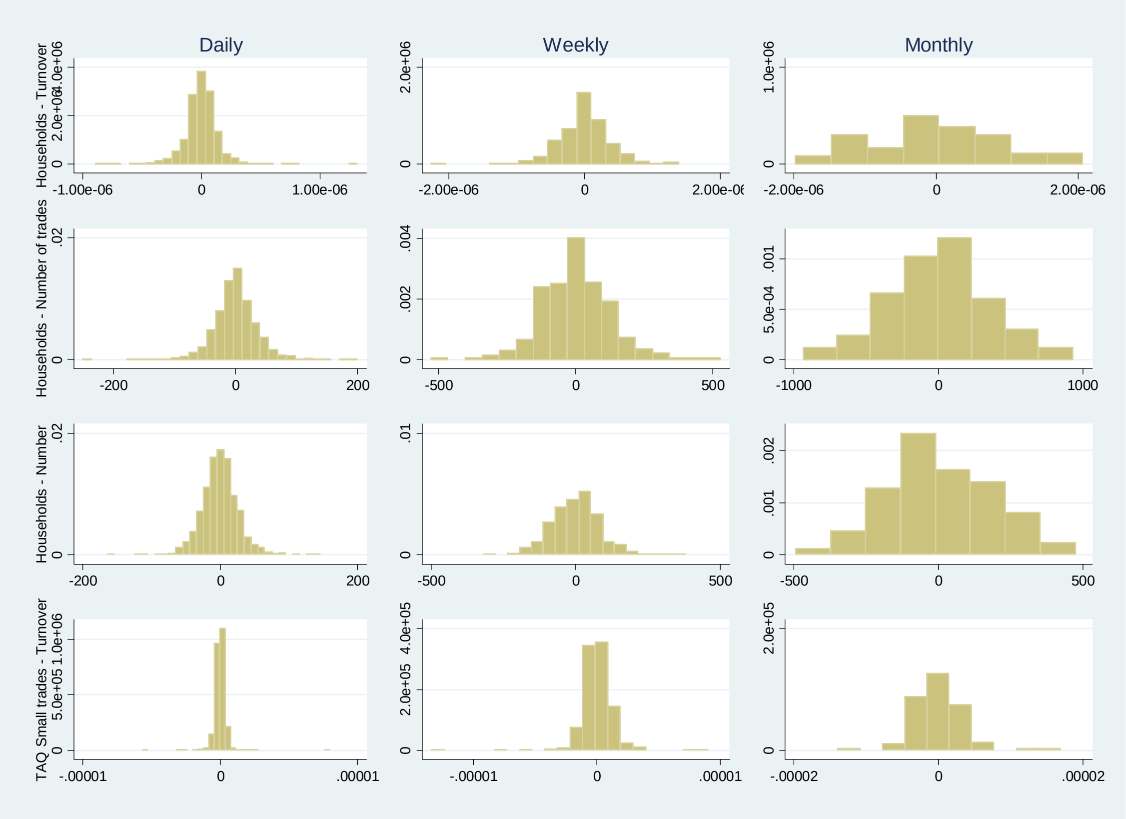


Figure 5: Probability (Q-Q) plots of households' aggregate trades

The graphs in this figure plot quantiles of households' aggregate trades against quantiles of a normal distribution at various frequencies. The left column considers daily data, the middle column considers weekly data, and the right column considers monthly data. The top row considers households' net turnover, defined as the aggregate value of their buys minus the aggregate value of their sells divided by the market's total value. The second row considers households' net number of buy trades, defined as the number of households' buy trades minus the number of sell trades. The third row considers the net number of households buying shares, defined as the number of households buying minus the number of households selling. The bottom row considers the net turnover for small trades in the TAQ dataset, which is defined as the aggregate value of small buys minus the aggregate value of small sells divided by the total value of the market. The TAQ data include all small trades in NYSE/AMEX stocks, where trades are classified as small or large based on the procedure described in Hvidkjaer (2006); see Section 1.b. We adjust all variables for seasonality and time trends by regressing them on day-of-the-week, month-in-year, and year dummy variables and then taking the residuals.

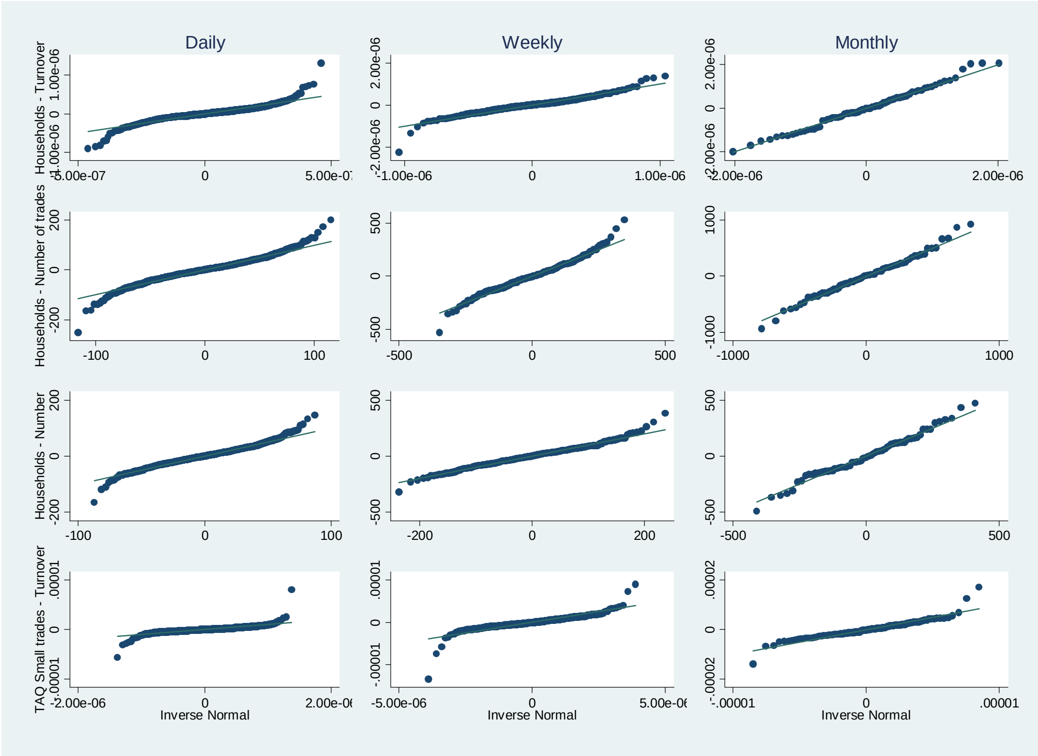


Figure 6: Estimating the intensity of noise trading for stock groups

The graphs in this figure plot the lower bound on the standard deviation of noise trades, across stock characteristic deciles, as measured using households' trades (solid lines, square markers) and small TAQ trades (dashed lines, circles). For each month, we sort stocks into deciles according to their capitalization, share price, turnover, closing bid-ask spread, Amihud illiquidity ratio, return volatility, and return autocovariance. All variables are estimated every month from daily observations. For a stock's capitalization, price, bid-ask spread, and turnover we use its respective monthly average. The Amihud illiquidity ratio is the monthly average of the daily ratio of the stock's absolute return to its dollar trading volume. The return volatility is the standard deviation of the stock's daily raw return over a month, and the return autocovariance is the autocovariance of the stock's daily returns over a month. Then, decile by decile, we regress daily total turnover (CRSP trading volume divided by the market's total value) on the daily retail turnover (the sum of buys and sells divided by the market's total value) as measured using households' trades and small TAQ trades. The lower bound on the standard deviation of noise trading in any decile is given by the time-series standard deviation of turnover in that decile multiplied by the regression coefficient. The upper bound (not displayed) is equal to twice the lower bound. All variables are adjusted for seasonality and time trends before running the regressions.

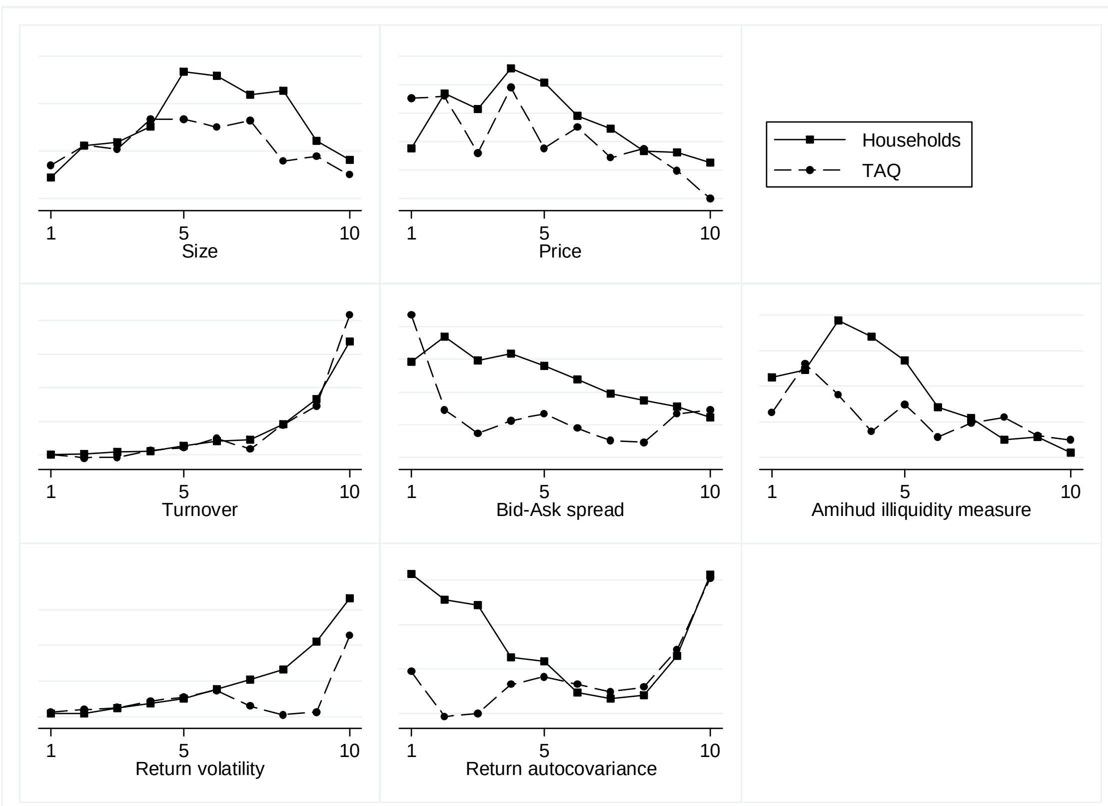


Table 1: Descriptive statistics for aggregate trades

This table presents summary statistics for the time series of households' daily aggregate trades at a large discount broker (between January 1991 and November 1996) and for the time series of small TAQ trades (between January 1991 and December 2000). The household data are for those holding common stock positions for 71 consecutive months. The TAQ data include all small trades in NYSE/AMEX stocks, where trades are classified as small or large based on the procedure described in Hvidkjaer (2006); see Section 1.b. Panel A considers households' net turnover, defined as the aggregate value of their buys minus the aggregate value of their sells divided by the market's total value (in millions). Panel B considers households' net number of buy trades, defined as the number of households' buy trades minus the number of sell trades. Panel C considers the net number of households buying shares, defined as the number of households buying minus the number of households selling. Panel D considers the net turnover for small trades in the TAQ dataset, which is defined as the aggregate value of small buys minus the aggregate value of small sells, divided by the market's total value (in millions); Panel E does likewise for large TAQ trades. We adjust all variables for seasonality and time trends by regressing them on day-of-the-week, month-in-year, and year dummy variables and then taking the residuals.

Frequency	Obs.	min.	mean	median	max.	std. dev.	skewness	kurtosis
	Panel A: Households - Turnover x 1M							
Daily	1497	-0.890	0.000	0.003	1.319	0.143	0.111	12.850
Weekly	309	-2.260	0.000	0.007	1.382	0.382	-0.421	7.750
Monthly	71	-1.979	0.000	0.005	2.069	0.919	0.183	2.693
Number of firms : 9,158								
	Panel B: Households - Number of trades							
Daily	1497	-250.899	0.000	0.471	200.707	36.136	-0.121	6.832
Weekly	309	-528.151	0.000	-4.841	528.259	127.143	0.195	4.971
Monthly	71	-936.576	0.000	10.050	925.758	356.475	0.070	3.393
Number of firms : 9,158								
	Panel C: Households - Number							
Daily	1497	-164.540	0.000	-0.247	145.291	27.130	0.135	5.700
Weekly	309	-318.232	0.000	-0.216	380.303	87.002	0.281	4.677
Monthly	71	-493.814	0.000	-14.064	473.152	186.214	0.100	3.223
Number of firms : 9,158								
	Panel D: TAQ Small trades - Turnover x 1M							
Daily	2526	-5.636	0.000	0.007	8.006	0.409	1.361	80.450
Weekly	522	-13.445	0.000	-0.032	9.052	1.346	-1.402	28.643
Monthly	120	-13.988	0.000	-0.093	16.983	3.539	0.762	8.769
Number of firms : 11,828								
	Panel E: TAQ Large trades - Turnover x 1M							
Daily	2526	-22.136	0.000	0.155	31.857	4.176	-0.078	6.229
Weekly	522	-43.300	0.000	0.448	38.708	12.232	-0.226	3.241
Monthly	120	-76.881	0.000	-0.068	70.290	30.073	-0.145	2.774
Number of firms : 11,790								

Table 2: Correlation among trades

This table examines whether households' trades (Panels A and B) and small TAQ trades (Panel C) have a systematic component. In Panel A we regress, in a stock-month panel setting, a stock's aggregate trades in a given month on the average aggregate trades across all other stocks (where this average excludes the stock's own trades), labeled "Mean Dep. Var." in the table, and on the market return. The results indicate that a given stock is more likely to be bought by households at times when they are buying other stocks. In Panel B we estimate a household-month panel regression; here the dependent variable is a household's trades in all stocks in a given month, and the independent variables are the average trades across all other households (where this average excludes the household's own trades) and the market return. Panels A and B consider households' net turnover (defined as the aggregate value of their buys minus the aggregate value of their sells divided by the total value of the market), households' net number of buy trades (defined as the number of households' buy trades minus the number of sell trades), and the net number of households buying shares (defined as the number of households buying minus the number of households selling). In addition, Panel B includes as a dependent variable the number of distinct stocks bought by a given household. The results indicate that a given household tends to buy stocks at times when other households are buying stocks. Panel C is similar to Panel A and reports results from a stock-month panel regression of a stock's aggregate trades—measured in the TAQ dataset in a given month—on the average aggregate trades across all other stocks (where this average excludes the stock's own trades) and on the market return. The table displays estimates for small and large trades as well as for their difference. The results indicate that a given stock is more likely to be bought at times when other stocks are bought, and this effect is stronger (by a factor of 2) for small trades than for large trades. Standard errors are double-clustered by month and by either firm (Panels A and C) or household (Panel B). ** and *** indicate statistical significance at (respectively) the 5% and 1% level.

Panel A: Correlation among stocks traded by households

	Number of trades	Number of households	Turnover
Mean Dep. Var.	0.988***	1.017***	0.498***
	(13.227)	(13.099)	(3.919)
Mkt return	-0.203	-0.074	-0.000***
	(-0.587)	(-0.233)	(-2.740)
Firms	9 158	9 158	9 158
Obs.	123 133	123 133	123 133
R-square	9.8%	9.8%	2.4%

Panel B: Correlation among households

	Number of trades	Number of households	Turnover	Number of stocks
Mean Dep. Var.	1.025***	1.033***	0.589***	0.941***
	(25.114)	(31.595)	(5.705)	(13.843)
Mkt return	0.201	0.197	-0.000**	0.075
	(0.664)	(1.536)	(-2.064)	(0.277)
Households	11 268	11 268	11 268	11 268
Obs.	159 305	159 305	159 305	159 305
R-square	10.9%	16.5%	3.3%	32.7%

Panel C: Correlation among stocks traded in TAQ

	Small trades - Turnover	Large trades - Turnover	Small minus Large trades - Turnover
Mean Dep. Var.	0.922***	0.514***	0.500***
	(14.229)	(9.170)	(8.133)
Mkt return	-0.000	0.000***	-0.000***
	(-0.111)	(4.727)	(-4.537)
Firms	11 850	11 850	11 850
Obs.	454 467	454 467	454 467
R-square	0.2%	0.4%	0.3%

Table 3: Performance of trades

We estimate the post-trade buy-sell return difference as in Odean (1999). That is, for each day, we first calculate the average return across all buy (sell) transactions executed on that day over the 84 (252) subsequent trading days and then take the difference. Average post-trade return differences are estimated using both raw returns and market-adjusted returns. Households' trade returns are equal-weighted; TAQ trade returns are weighted according to the value of the trade. The *t*-statistics (in parentheses) test whether the mean return differs from zero. To account for overlap in the return window, standard errors are adjusted for autocorrelation of up to 252 lags via the Newey-West correction. * and *** indicate statistical significance at (respectively) the 10% and 1% level.

	Households	TAQ Small trades	TAQ Large trades	Small minus Large trades
	Post-trade buy-sell return difference over the next 84 trading days			
Raw Returns	-0.0054*	-0.0242***	-0.0028	-0.0213***
	(-1.71)	(-5.75)	(-1.37)	(-4.16)
Market-adjusted Return	-0.0052*	-0.0226***	-0.0020	-0.0206***
	(-1.75)	(-5.51)	(-0.94)	(-4.11)
	Post-trade buy-sell return difference over the next252 trading days			
Raw Returns	-0.0260***	-0.0965***	-0.0012	-0.0953***
	(-3.92)	(-9.67)	(-0.14)	(-8.27)
Market-adjusted Return	-0.0221***	-0.0856***	-0.0017	-0.0839***
	(-3.74)	(-8.11)	(-0.21)	(-8.27)

Table 4: Trades and firms' fundamentals

We estimate a stock-quarter panel regression model of trades on earnings surprises. The independent variable is a firm's quarterly standardized unexpected earnings (SUE) decile . Such earnings "surprises" are defined, for each stock and quarter, as the difference between actual and expected earnings (where expected earnings are derived from a seasonal random walk with drift and are divided by their standard deviation):

$$\mathrm{SUE}_{i,q} = \frac{E_{i,q}-(E_{i,q-4}+drift_{i,q})}{\sigma_{i,q}}, \quad \text{where } \mathrm{drift}_{i,q} = \frac{1}{8}\sum_{n=1}^{8}(E_{i,q-n} - E_{i,q-n-4}).$$

$E_{i,q}$ denotes the actual earnings for firm *i* in quarter *q* (Compustat's earnings per share, excluding extraordinary items) and $\sigma_{i,q}$ the standard deviation of earnings surprises estimated over the last eight quarters. The dependent variables in Panel A are households' aggregate trades (net turnover among households and number of households trading), where the analysis is restricted to stocks that have at least 100 trades over the 1991–1996 sample period. The dependent variables in Panel B are TAQ trades (net turnover among small and large trades), where the analysis is restricted to stocks that have at least $100,000 worth of small trades over the 1991–2000 sample period. In all regressions, turnover is scaled by one million. In both panels, the dependent variables are trades—in a stock and quarter, aggregated over windows of 40, 20, 10, and 5 days ending on the day *before* the firm announces its earnings and on the announcement day (day 0). The regressions include firm, quarter, and month-in-year fixed effects; standard errors (in parentheses) are clustered by firm. *, **, and *** indicate statistical significance at (respectively) the 10%, 5%, and 1% level.

Panel A: Households' trades and firms' fundamentals

	Households - Turnover x 1M	Households - Number
	40-day window	
SUE	-0.553***	-0.260***
	(-5.063)	(-6.940)
R-square	0.9%	1.9%
	20-day window	
SUE	-0.332***	-0.128***
	(-3.280)	(-6.840)
R-square	0.7%	1.5%
	10-day window	
SUE	-0.184**	-0.067***
	(-2.515)	(-6.059)
R-square	0.5%	1.2%
	5-day window	
SUE	-0.073**	-0.026***
	(-2.078)	(-5.565)
R-square	0.4%	0.9%
	Day 0	
SUE	-0.012	-0.012***
	(-0.805)	(-3.565)
R-square	0.1%	0.3%
Obs.	12 841	12 841
Firms	670	670

Panel B: TAQ small trades and firms' fundamentals

	TAQ Small trades - Turnover x 1M	TAQ Large trades - Turnover x 1M	Small minus Large trades - Turnover
		40-day window	
SUE	0.084	0.482**	-0.398*
	(1.047)	(2.216)	(-1.763)
R-square	0.3%	0.3%	0.2%
		20-day window	
SUE	0.049	0.275**	-0.226*
	(0.794)	(1.966)	(-1.710)
R-square	0.3%	0.2%	0.2%
		10-day window	
SUE	0.048	0.240**	-0.191**
	(1.112)	(2.355)	(-2.012)
R-square	0.2%	0.2%	0.2%
		5-day window	
SUE	-0.022	0.06	-0.083
	(-0.506)	(0.909)	(-1.059)
R-square	0.2%	0.1%	0.2%
		Day 0	
SUE	0.076***	0.104***	-0.028
	(4.326)	(4.376)	(-0.939)
R-square	0.2%	0.1%	0.1%
Obs.	54 495	54 495	54 495
Firms	2 750	2 750	2 750

Table 5: Shapiro-Wilk test for normality

This table reports results of a Shapiro-Wilk test that households' aggregate trades, small TAQ trades, and their residuals from a fitted AR(1) process are normally distributed at daily, weekly, and monthly frequencies. The null hypothesis is that these series are normal, and the alternative is that they are not normal. The top panel considers households' net turnover, defined as the aggregate value of their buys minus the aggregate value of their sells divided by the market's total value. The second panel considers households' net number of buy trades, defined as the number of households' buy trades minus the number of sell trades. The third panel considers the net number of households buying shares, defined as the number of households buying minus the number of households selling. The bottom panel considers the net turnover for small trades in the TAQ dataset, which is defined as the aggregate value of small buys minus the aggregate value of small sells *divided by* the market's total value. The TAQ data include all small trades in NYSE/AMEX stocks, where trades are classified as small or large based on the procedure described in Hvidkjaer (2006); see Section 1.b. We adjust all variables for seasonality and time trends by regressing them on day-of-the-week, month-in-year, and year dummy variables and then taking the residuals.

	Variables		Residuals from fitted AR(1)	
	Test Statistic	*p*-value	Test Statistic	*p*-value
	Households - Turnover		Households - Turnover	
Daily	11.213	0.000	11.077	0.000
Weekly	5.706	0.000	5.914	0.000
Monthly	-0.363	0.642	0.044	0.482
	Households - Number of trades		Households - Number of trades	
Daily	8.787	0.000	8.013	0.000
Weekly	3.651	0.000	3.096	0.001
Monthly	-1.445	0.926	-1.579	0.943
	Households - Number		Households - Number	
Daily	7.609	0.000	7.287	0.000
Weekly	2.811	0.002	1.910	0.028
Monthly	0.926	0.177	0.353	0.362
	TAQ Small trades - Turnover		TAQ Small trades - Turnover	
Daily	14.899	0.000	14.799	0.000
Weekly	10.148	0.000	9.781	0.000
Monthly	4.762	0.000	4.937	0.000

Table 6: Estimating the intensity of noise trading

We regress total turnover (CRSP trading volume divided by the market's total value) on the turnover (sum of buys and sells divided by the market's total value) as measured using households or small TAQ trades and over different frequencies. All variables are first adjusted for seasonality and time trends by regressing them on day-of-the-week, month-in-year, and year dummy variables and taking residuals. The regression coefficient b determines bounds on the fraction m of noise turnover accounted for by the households in our sample: $\frac{1}{2b} \leq m \leq \frac{1}{b}$. The table also reports bounds on the standard deviation of noise trades and on the standard deviation of the residual—under the assumption that noise trading follows an AR(1) process with first-order autocorrelation coefficients drawn from Figure 3. These bounds are displayed in terms of levels and also as a fraction of the standard deviation of total turnover in the market.

			Households Turnover			TAQ Small trades Turnover		
	Frequency		Day	Week	Month	Day	Week	Month
	Std. dev. of turnover (x1 million)		0.222	0.741	2.232	1.239	5.473	17.463
	Regression coefficient b		1289	2032	2044	170	233	141
Scaling factor m (x1,000)		Lower bound	0.388	0.246	0.245	2.948	2.148	3.547
		Upper bound	0.776	0.492	0.489	5.897	4.296	7.093
Std. dev. of noise trading	Lower bound	(x1,000)	0.286	1.505	4.562	0.210	1.274	2.462
		% of std. dev. of total turnover	38%	44%	36%	27%	36%	20%
	Upper bound	(x1,000)	0.572	3.011	9.123	0.420	2.548	4.924
		% of std. dev. of total turnover	75%	87%	73%	55%	73%	39%
	First-order auto-correlation (Fig. 3)		0.156	0.199	0.039	0.489	0.456	0.292
Std. dev. of residual under AR(1)	Lower bound	(x1,000 for number) (in $b for value)	0.218	1.166	3.287	0.181	1.087	1.979
		% of std. dev. of total trades	29%	34%	26%	24%	31%	16%
	Upper bound	(x1,000 for number) (in $b for value)	0.435	2.331	6.575	0.363	2.174	3.957
		% of std. dev. of total trades	57%	68%	53%	47%	62%	32%

Table 7: Estimating the intensity of noise trading for stock groups

The panels in this table report bounds on the standard deviation of noise trades across groups of stocks sorted on various characteristics. For each month, we sort stocks into deciles according to their capitalization, share price, turnover, closing bid-ask spread, Amihud illiquidity ratio, return volatility, and return autocovariance. All variables are estimated every month from daily observations. For a stock's capitalization, price, bid-ask spread, and turnover we use its respective monthly average. The Amihud illiquidity ratio is the monthly average of the daily ratio of the stock's absolute return to its dollar trading volume. The return volatility is the standard deviation of the stock's daily raw return over a month, and the return autocovariance is the autocovariance of the stock's daily returns over a month. Then, decile by decile, we regress total turnover (CRSP trading volume divided by the market's total value) on the turnover (sum of buys and sells divided by the market's total value) as measured using households' trades (columns 2–6) or small TAQ trades (columns 7–11). We use b^k to denote the regression coefficient. All variables are first adjusted for seasonality and time trends by regressing them on day-of-the-week, month-in-year, and year dummy variables and taking residuals. The standard deviation of noise trading in decile *k* is given by the time-series standard deviation of trades in decile *k* divided by a scaling factor m^k, where $\frac{1}{2b^k} \leq m^k \leq \frac{1}{b^k}$. The bounds on the standard deviation of noise trades are reported in terms of levels and also as a fraction of the standard deviation of total trades in a decile. Negative estimates of the noise trading intensity correspond to statistically insignificant estimates of the slope coefficient.

Panel A: By stock size

	Households - Turnover x 1M					TAQ Small trades - Turnover x 1M				
				Lower bound on SD of noise trading					Lower bound on SD of noise trading	
Decile	Median stock size ($M)	SD of trades (x1M)	Regression coef. *b*	(x1M)	(% of SD of total trades)	Median stock size ($M)	SD of trades (x1M)	Regression coef. b	(x1M)	(% of SD of total trades)
	Daily frequency									
1	5.2	4.91	20.01	98.26	8%	5.6	6.74	120.51	812.79	36%
2	12.0	1.41	130.34	184.15	15%	14.3	2.38	345.06	820.74	45%
3	20.9	1.63	77.83	127.12	11%	25.9	1.28	433.04	554.06	38%
4	34.5	1.28	208.52	266.81	22%	43.2	0.94	732.30	686.24	38%
5	54.6	1.56	241.86	377.01	29%	70.0	0.59	1095.63	649.12	39%
6	87.9	1.04	398.08	414.89	31%	114.0	0.28	2144.96	593.98	37%
7	148.0	0.85	485.63	413.96	31%	189.0	0.29	1283.10	366.68	25%
8	281.0	0.73	644.09	473.02	38%	351.0	0.08	5032.34	408.81	32%
9	710.0	0.43	802.86	341.44	33%	823.0	0.04	12232.32	439.25	42%
10	6310.0	0.22	1180.89	259.00	34%	3380.0	0.01	43194.94	354.21	49%
	Weekly frequency									
1	5.1	9.70	70.28	681.79	12%	5.5	27.56	105.20	2898.89	34%
2	12.1	3.74	359.65	1343.88	24%	14.2	9.40	292.90	2753.29	40%
3	21.0	3.97	259.75	1032.03	18%	26.0	5.24	424.23	2224.24	37%
4	34.3	3.56	547.73	1951.30	35%	43.2	4.17	632.55	2638.24	36%
5	54.9	4.36	706.40	3080.50	51%	69.9	2.64	1046.57	2760.24	39%
6	88.2	3.14	974.82	3059.59	49%	114.0	1.20	2166.85	2602.18	37%
7	148.0	2.57	1157.73	2972.00	50%	189.0	1.10	1779.45	1955.32	30%
8	281.0	2.38	1299.69	3086.94	54%	351.0	0.38	4691.16	1768.56	31%
9	707.0	1.24	1636.46	2035.93	44%	823.0	0.17	12206.56	2092.62	44%
10	6260.0	0.68	1985.80	1344.30	40%	3380.0	0.01	43194.94	354.21	49%
	Monthly frequency									
1	5.2	19.28	114.41	2206.13	9%	5.5	89.87	39.01	3505.54	15%
2	12.4	10.52	529.40	5571.40	24%	14.2	31.33	178.36	5587.48	26%
3	21.7	9.50	626.91	5953.41	27%	26.0	18.74	279.60	5240.15	25%
4	34.6	10.10	749.45	7570.97	35%	43.3	15.52	537.38	8340.77	31%
5	54.9	12.15	1099.31	13357.78	57%	70.0	9.45	889.60	8405.68	33%
6	89.2	9.52	1355.68	12908.08	56%	114.0	4.32	1740.61	7521.31	29%
7	147.0	7.42	1472.93	10923.42	50%	189.0	3.27	2522.10	8254.33	36%
8	281.0	7.42	1532.77	11367.09	56%	351.0	1.38	2874.97	3963.37	20%
9	712.0	3.62	1680.30	6077.67	39%	825.0	0.64	7048.16	4480.06	33%
10	6310.0	1.86	2197.11	4092.48	34%	3390.0	0.14	17798.83	2495.94	30%

Panel B: By stock turnover

Decile	Households - Turnover x 1M					TAQ Small trades - Turnover x 1M				
	Median turnover	SD of trades (x1M)	Regression coef. b	Lower bound on SD of noise trading		Median turnover	SD of trades (x1M)	Regression coef. b	Lower bound on SD of noise trading	
				(x1M)	(% of SD of total trades)				(x1M)	(% of SD of total trades)
Daily frequency										
1	0.02%	0.15	1.52	0.23	0%	0.03%	0.01	738.27	7.17	12%
2	0.06%	0.20	34.73	7.10	4%	0.07%	0.02	-583.56	-10.72	-6%
3	0.10%	0.16	185.19	29.70	9%	0.11%	0.02	-474.32	-7.39	-3%
4	0.14%	0.16	379.84	61.96	15%	0.16%	0.01	15148.20	107.13	31%
5	0.19%	0.17	562.36	96.32	18%	0.22%	0.01	33810.98	199.68	46%
6	0.26%	0.25	651.11	164.38	24%	0.30%	0.01	32678.92	303.31	51%
7	0.34%	0.39	499.83	194.43	21%	0.40%	0.02	11161.75	205.93	26%
8	0.48%	0.47	809.68	377.89	28%	0.55%	0.03	17906.00	523.89	41%
9	0.74%	1.00	649.57	647.47	30%	0.83%	0.06	14232.04	811.25	39%
10	1.71%	2.09	896.96	1871.49	39%	1.65%	0.27	6405.78	1708.07	38%
Weekly frequency										
1	0.02%	0.29	33.30	9.80	4%	0.03%	0.04	952.70	36.44	16%
2	0.06%	0.42	112.98	47.59	5%	0.07%	0.08	-1246.19	-102.40	-15%
3	0.10%	0.40	435.19	173.32	13%	0.11%	0.07	-724.25	-52.79	-5%
4	0.14%	0.47	649.46	305.81	17%	0.16%	0.03	16947.44	578.47	39%
5	0.19%	0.45	1327.54	592.10	25%	0.22%	0.03	33634.29	1014.04	53%
6	0.26%	0.71	1439.66	1025.10	33%	0.30%	0.05	35054.59	1627.33	60%
7	0.34%	1.11	1111.16	1235.76	30%	0.40%	0.09	13453.05	1196.35	33%
8	0.47%	1.37	1675.52	2289.12	38%	0.55%	0.14	19040.10	2731.05	49%
9	0.74%	3.08	1210.04	3723.79	39%	0.83%	0.27	15710.60	4257.12	46%
10	1.71%	6.96	1455.47	10124.86	47%	1.65%	1.15	8539.49	9813.21	50%
Monthly frequency										
1	0.02%	0.63	174.57	110.20	10%	0.03%	0.13	85.30	11.32	2%
2	0.06%	0.85	376.58	318.73	9%	0.07%	0.33	-2505.35	-814.36	-36%
3	0.10%	1.02	956.36	977.45	19%	0.11%	0.32	-2109.73	-665.53	-22%
4	0.14%	1.32	849.21	1120.99	16%	0.16%	0.13	9624.83	1291.86	31%
5	0.19%	1.20	2216.93	2656.40	29%	0.22%	0.10	23079.16	2195.54	40%
6	0.26%	1.93	2104.06	4061.78	34%	0.30%	0.16	31698.87	4980.00	61%
7	0.34%	3.53	1307.67	4614.55	29%	0.40%	0.31	5367.30	1689.34	15%
8	0.47%	3.80	2419.67	9186.96	40%	0.55%	0.53	17238.86	9071.81	48%
9	0.74%	8.45	1959.13	16554.54	46%	0.83%	1.03	14210.11	14616.22	44%
10	1.71%	21.50	1575.82	33886.86	44%	1.64%	4.14	10063.06	41658.80	67%

Panel C: By stock price

	Households - Turnover x 1M					TAQ Small trades - Turnover x 1M				
Decile	Median stock price	SD of trades (x1M)	Regression coef. b	Lower bound on SD of noise trading (x1M)	Lower bound on SD of noise trading (% of SD of total trades)	Median stock price	SD of trades (x1M)	Regression coef. b	Lower bound on SD of noise trading (x1M)	Lower bound on SD of noise trading (% of SD of total trades)
	Daily frequency									
1	0.99	2.39	113.32	271.20	18%	1.09	9.38	92.68	869.45	52%
2	2.67	1.67	198.78	332.88	27%	2.76	2.20	364.95	801.58	47%
3	4.39	1.46	197.99	289.22	25%	4.61	0.78	777.28	603.48	45%
4	6.56	1.29	358.21	462.41	37%	6.82	0.40	2205.93	878.38	58%
5	9.23	1.05	344.93	361.74	31%	9.52	0.22	2411.93	523.04	44%
6	12.65	0.81	476.02	385.79	34%	12.90	0.12	5045.47	581.30	53%
7	16.44	0.73	429.41	312.81	29%	17.04	0.06	8689.63	489.55	52%
8	21.54	0.45	635.05	285.96	34%	22.52	0.03	16091.56	464.39	55%
9	28.76	0.35	686.17	240.03	29%	30.19	0.02	24653.95	386.27	54%
10	49.95	0.23	1219.39	277.70	34%	47.29	0.00	90540.63	410.03	50%
	Weekly frequency									
1	1.02	5.82	320.41	1863.47	27%	1.09	39.64	89.41	3544.24	51%
2	2.70	4.89	457.59	2237.92	38%	2.76	9.67	327.22	3163.91	45%
3	4.40	4.09	512.06	2092.22	38%	4.63	3.45	664.74	2293.40	39%
4	6.56	3.99	720.02	2874.15	50%	6.83	1.74	1810.60	3159.06	53%
5	9.21	3.14	829.48	2607.96	49%	9.57	1.00	2230.27	2233.64	43%
6	12.61	2.50	926.63	2319.78	46%	12.90	0.54	4708.05	2549.86	54%
7	16.51	1.96	1068.49	2099.08	45%	17.04	0.27	8282.97	2196.37	53%
8	21.50	1.32	1257.63	1659.85	46%	22.52	0.14	15376.38	2114.38	57%
9	28.76	1.03	1386.36	1423.19	39%	30.19	0.08	24220.19	1856.08	58%
10	50.11	0.72	2022.22	1446.22	40%	47.27	0.02	78724.11	1715.22	48%
	Monthly frequency									
1	1.02	12.79	432.73	5533.63	21%	1.09	128.52	70.36	9042.24	40%
2	2.71	14.29	655.38	9365.37	40%	2.76	32.11	286.15	9187.45	38%
3	4.37	11.53	718.99	8292.79	39%	4.61	12.27	422.23	5180.96	26%
4	6.58	12.10	921.93	11151.29	52%	6.83	6.46	1520.16	9818.10	49%
5	9.26	9.90	1025.24	10145.18	52%	9.57	3.75	1476.29	5543.26	34%
6	12.60	7.95	983.65	7815.47	43%	12.90	2.04	3445.81	7031.46	49%
7	16.46	5.04	1375.48	6931.24	43%	17.04	0.97	5028.39	4884.62	44%
8	21.47	3.84	1387.82	5325.67	46%	22.52	0.49	11218.97	5504.92	51%
9	28.80	2.90	1802.38	5230.28	39%	30.27	0.28	14349.41	3980.70	48%
10	49.29	2.00	2263.47	4518.69	34%	47.27	0.07	27582.39	2011.41	20%

Panel D: By stock bid–ask spread

Decile	Households - Turnover x 1M					TAQ Small trades - Turnover x 1M				
	Median bid-ask spread (basis points)	SD of trades (x1M)	Regression coef. b	Lower bound on SD of noise trading (x1M)	Lower bound on SD of noise trading (% of SD of total trades)	Median bid-ask spread (basis points)	SD of trades (x1M)	Regression coef. b	Lower bound on SD of noise trading (x1M)	Lower bound on SD of noise trading (% of SD of total trades)
	Daily frequency									
1	74	0.98	448.19	440.13	17%	84	0.03	30781.63	998.65	57%
2	143	0.63	807.38	506.84	32%	134	0.04	3499.10	153.78	20%
3	201	0.72	481.32	346.68	29%	184	0.14	-180.99	-25.07	-3%
4	268	0.80	324.87	260.16	26%	239	0.13	497.05	63.40	8%
5	348	0.99	223.64	222.10	22%	299	0.09	1858.06	166.19	21%
6	438	0.83	177.77	146.82	16%	374	0.16	330.07	52.03	6%
7	556	1.15	155.35	178.80	20%	468	0.24	-1.38	-0.32	0%
8	723	1.19	86.09	102.24	14%	608	0.39	-71.90	-28.03	-3%
9	1001	0.97	86.97	84.40	12%	859	2.19	116.13	254.44	31%
10	1502	0.84	88.33	74.23	15%	1442	1.72	117.40	202.32	29%
	Weekly frequency									
1	73	2.79	855.04	2383.96	20%	84	0.15	30779.70	4604.64	58%
2	143	1.99	1718.76	3423.73	48%	134	0.19	3072.38	590.86	17%
3	201	2.04	1152.88	2352.44	45%	184	0.56	-498.79	-280.83	-8%
4	268	1.97	1201.56	2361.46	52%	239	0.61	575.41	351.97	10%
5	349	2.57	772.42	1982.99	45%	299	0.42	1958.50	812.93	24%
6	439	2.23	616.31	1375.56	32%	376	0.70	383.21	269.67	6%
7	561	3.10	406.65	1260.89	31%	468	1.06	-43.14	-45.82	-1%
8	734	3.08	277.86	855.58	25%	608	1.76	-90.81	-160.08	-4%
9	1036	2.42	210.63	509.94	17%	859	8.64	112.83	974.29	29%
10	1538	2.05	195.38	400.66	20%	1442	6.96	126.95	883.55	33%
	Monthly frequency									
1	72	7.71	1241.52	9573.28	19%	84	0.59	28154.31	16733.05	57%
2	143	5.95	2261.80	13449.41	48%	135	0.57	4048.68	2296.18	20%
3	200	6.20	1585.81	9828.27	53%	185	1.73	-752.91	-1300.15	-13%
4	268	5.08	2144.60	10895.33	70%	241	1.73	344.99	595.91	6%
5	347	7.31	1237.84	9043.59	57%	300	1.65	1036.54	1707.76	16%
6	440	6.56	1052.48	6903.36	43%	377	2.64	-212.40	-560.12	-4%
7	558	8.48	558.79	4738.15	29%	468	4.48	-524.20	-2350.27	-17%
8	738	8.42	439.59	3701.84	27%	610	7.34	-367.16	-2693.57	-18%
9	1016	5.00	562.43	2812.96	23%	861	28.01	61.08	1710.72	14%
10	1533	4.36	268.71	1170.93	14%	1442	23.63	95.15	2248.28	24%

Panel E: By stock Amihud illiquidity measure

Decile	Households - Turnover x 1M: Median Amihud illiquidity measure (x1M)	SD of trades (x1M)	Regression coef. b	Lower bound on SD of noise trading (x1M)	Lower bound on SD of noise trading (% of SD of total trades)	TAQ Small trades - Turnover x 1M: Median Amihud illiquidity measure (x1M)	SD of trades (x1M)	Regression coef. b	Lower bound on SD of noise trading (x1M)	Lower bound on SD of noise trading (% of SD of total trades)
	Daily frequency									
1	0.00	0.25	1157.69	285.52	33%	0.00	0.01	44742.69	376.25	49%
2	0.01	0.33	741.40	248.08	29%	0.01	0.04	11445.15	420.95	47%
3	0.03	0.49	626.51	307.07	35%	0.02	0.08	4122.48	315.09	33%
4	0.08	0.60	366.15	220.81	30%	0.06	0.14	1803.64	248.68	29%
5	0.21	0.54	297.79	159.90	23%	0.15	0.27	1187.27	314.65	42%
6	0.53	0.52	192.68	100.11	18%	0.34	0.32	589.04	187.22	34%
7	1.26	0.51	125.14	63.59	15%	0.83	0.81	195.17	159.03	34%
8	3.02	0.48	103.92	49.94	16%	1.98	0.66	392.16	259.76	44%
9	8.04	0.54	65.21	35.52	13%	5.34	0.91	197.72	179.23	41%
10	38.20	3.35	9.53	31.89	11%	28.90	1.82	74.11	135.09	24%
	Weekly frequency									
1	0.00	0.76	1962.03	1496.38	39%	0.00	0.04	42576.58	1763.75	52%
2	0.01	1.03	1578.29	1625.02	41%	0.01	0.17	11829.20	2065.03	51%
3	0.03	1.52	1366.56	2081.80	52%	0.02	0.35	4240.02	1462.88	34%
4	0.08	1.66	1049.83	1744.27	52%	0.06	0.61	1609.18	978.19	26%
5	0.21	1.55	783.73	1215.79	37%	0.15	1.11	1092.54	1210.60	39%
6	0.54	1.36	588.26	798.51	31%	0.34	1.36	549.01	748.57	31%
7	1.28	1.28	439.58	564.12	31%	0.83	3.15	220.47	694.91	37%
8	3.03	1.28	243.41	312.13	25%	1.98	2.52	369.67	931.13	40%
9	8.33	1.34	191.18	255.38	24%	5.35	3.79	159.17	603.28	36%
10	39.20	6.44	27.10	174.55	16%	28.90	7.33	68.99	505.34	28%
	Monthly frequency									
1	0.00	2.09	2151.54	4500.26	33%	0.00	0.14	17673.60	2533.73	29%
2	0.01	3.10	1589.54	4931.83	34%	0.01	0.66	7998.34	5253.44	40%
3	0.03	4.82	1597.33	7705.66	53%	0.02	1.33	2645.36	3515.69	22%
4	0.08	4.78	1422.63	6793.22	56%	0.06	2.26	644.82	1455.19	11%
5	0.22	4.31	1264.52	5454.56	44%	0.15	3.82	784.24	2999.11	28%
6	0.53	3.75	756.65	2841.20	30%	0.34	4.79	243.85	1168.83	14%
7	1.31	3.59	624.04	2237.62	33%	0.82	9.54	204.27	1948.46	32%
8	3.03	3.20	314.91	1008.87	23%	1.97	8.54	264.89	2262.73	30%
9	8.27	2.85	408.55	1163.12	30%	5.21	13.48	91.11	1228.50	23%
10	36.70	12.21	23.73	289.64	7%	28.90	25.94	38.26	992.46	23%

Panel F: By stock return volatility

Decile	Households - Turnover x 1M					TAQ Small trades - Turnover x 1M				
	Median SD of stock returns	SD of trades (x1M)	Regression coef. b	Lower bound on SD of noise trading		Median SD of stock returns	SD of trades (x1M)	Regression coef. b	Lower bound on SD of noise trading	
				(x1M)	(% of SD of total trades)				(x1M)	(% of SD of total trades)
	Daily frequency									
1	1.02%	0.19	288.21	54.64	14%	1.03%	0.01	24117.12	177.07	45%
2	1.47%	0.20	552.30	113.12	21%	1.51%	0.01	52911.33	284.09	58%
3	1.89%	0.34	697.49	234.39	31%	1.94%	0.01	36409.21	355.44	54%
4	2.34%	0.72	428.84	308.16	22%	2.37%	0.02	32577.72	577.65	56%
5	2.81%	0.98	472.50	462.58	26%	2.84%	0.03	24322.01	736.30	46%
6	3.34%	1.15	717.97	826.36	33%	3.41%	0.06	15889.72	952.97	43%
7	3.99%	1.63	691.67	1128.69	35%	4.05%	0.12	5123.88	590.36	20%
8	4.89%	2.05	628.85	1291.12	30%	4.95%	0.22	2358.78	517.42	12%
9	6.32%	3.50	572.00	2004.77	39%	6.41%	0.59	984.25	581.74	11%
10	10.13%	7.03	557.45	3916.48	43%	10.21%	3.26	802.21	2614.83	22%
	Weekly frequency									
1	1.04%	0.47	874.42	413.11	25%	1.03%	0.04	24053.00	843.62	52%
2	1.47%	0.57	1088.34	619.96	27%	1.51%	0.03	52850.57	1366.88	64%
3	1.89%	1.00	1197.61	1199.17	37%	1.94%	0.05	34558.29	1627.42	57%
4	2.34%	1.78	1084.88	1927.47	32%	2.37%	0.09	32571.40	2820.15	62%
5	2.81%	2.95	822.45	2426.81	32%	2.85%	0.14	24325.60	3485.57	50%
6	3.34%	3.54	1211.60	4285.95	40%	3.41%	0.28	16199.82	4568.54	48%
7	3.99%	4.86	1362.73	6621.94	48%	4.07%	0.55	5013.72	2768.04	23%
8	4.88%	6.27	1258.21	7886.46	42%	4.95%	1.00	2770.38	2774.01	16%
9	6.29%	10.60	1072.47	11371.11	52%	6.41%	2.72	981.51	2664.83	12%
10	10.03%	24.33	803.15	19538.46	52%	10.22%	14.18	844.97	11978.77	25%
	Monthly frequency									
1	1.03%	1.10	1589.85	1748.82	31%	1.03%	0.14	17097.52	2451.10	55%
2	1.47%	1.46	1266.59	1847.99	23%	1.51%	0.10	41449.22	3995.07	61%
3	1.89%	2.80	1685.55	4711.51	41%	1.94%	0.18	27543.91	5014.07	54%
4	2.34%	4.58	1649.30	7552.35	34%	2.37%	0.33	26146.08	8667.63	59%
5	2.80%	7.40	1343.66	9941.16	35%	2.84%	0.55	20041.36	10933.99	45%
6	3.34%	10.87	1402.67	15246.15	38%	3.41%	1.05	14138.62	14776.11	44%
7	3.98%	14.64	1409.66	20633.93	41%	4.06%	2.22	2673.18	5932.66	15%
8	4.89%	18.27	1443.71	26373.04	37%	4.95%	3.85	272.52	1049.01	2%
9	6.30%	31.33	1339.36	41958.78	52%	6.41%	10.61	248.57	2637.65	4%
10	10.13%	75.39	881.85	66482.60	48%	10.21%	49.18	920.84	45282.39	26%

Panel G: By stock return autocovariance

Decile	Households - Turnover x 1M: Median return autocovariance (basis points)	Households: SD of trades (x1M)	Households: Regression coef. b	Households: Lower bound on SD of noise trading (x1M)	Households: Lower bound on SD of noise trading (% of SD of total trades)	TAQ Small trades - Turnover x 1M: Median return autocovariance (basis points)	TAQ: SD of trades (x1M)	TAQ: Regression coef. b	TAQ: Lower bound on SD of noise trading (x1M)	TAQ: Lower bound on SD of noise trading (% of SD of total trades)
Daily frequency										
1	-3.06	3.18	463.37	1474.34	32%	-3.16	2.47	338.47	834.39	11%
2	-1.07	1.95	521.11	1018.51	29%	-0.99	0.54	409.12	219.82	6%
3	-0.51	1.40	815.03	1139.30	36%	-0.48	0.29	626.14	184.06	7%
4	-0.28	0.93	580.66	539.78	25%	-0.26	0.10	4049.12	411.36	25%
5	-0.15	0.64	773.62	497.36	33%	-0.14	0.03	16811.02	505.90	45%
6	-0.07	0.42	431.85	181.49	20%	-0.06	0.02	25556.35	394.85	53%
7	-0.03	0.22	655.54	147.01	25%	-0.02	0.01	43515.93	339.39	51%
8	0.00	0.24	710.39	172.27	28%	0.01	0.01	41983.20	381.33	55%
9	0.05	0.39	1454.13	568.20	54%	0.07	0.03	32352.17	811.50	60%
10	0.29	1.66	873.22	1448.38	46%	0.32	0.36	3936.36	1400.60	38%
Weekly frequency										
1	-3.07	9.19	905.47	8325.60	45%	-3.16	10.70	273.65	2929.25	10%
2	-1.06	5.45	1168.22	6364.75	41%	-1.01	2.36	423.44	998.16	6%
3	-0.52	4.40	1510.74	6651.84	48%	-0.49	1.35	740.87	1001.36	9%
4	-0.28	2.52	1275.34	3215.67	34%	-0.26	0.47	4509.41	2125.58	30%
5	-0.15	1.95	1394.01	2714.63	40%	-0.14	0.14	16482.44	2365.20	49%
6	-0.07	1.06	1007.68	1066.54	27%	-0.06	0.07	25765.57	1905.79	58%
7	-0.03	0.62	1510.54	939.05	36%	-0.02	0.04	42198.73	1602.91	54%
8	0.00	0.75	1366.47	1031.41	38%	0.01	0.04	43372.15	1866.26	61%
9	0.05	1.37	2303.63	3166.10	67%	0.07	0.12	32644.76	3818.87	63%
10	0.28	6.22	1266.15	7880.95	57%	0.33	1.48	4838.60	7167.73	45%
Monthly frequency										
1	-2.93	30.07	1043.06	31369.16	43%	-3.16	40.54	230.60	9348.16	9%
2	-1.07	15.19	1686.68	25617.28	42%	-1.01	8.93	-85.61	-764.83	-1%
3	-0.52	12.49	1950.53	24357.56	47%	-0.48	5.90	-12.35	-72.90	0%
4	-0.28	7.98	1577.95	12598.54	34%	-0.26	2.01	3204.98	6436.37	26%
5	-0.15	5.82	2006.83	11682.28	44%	-0.14	0.59	13913.84	8219.63	49%
6	-0.07	2.63	1804.05	4740.02	31%	-0.06	0.31	21289.28	6508.01	59%
7	-0.03	1.72	1940.95	3330.87	37%	-0.02	0.14	34323.48	4855.30	48%
8	0.00	2.21	1823.95	4025.66	42%	0.01	0.17	34892.50	5926.02	56%
9	0.05	4.97	2609.27	12964.26	72%	0.07	0.47	30223.67	14288.46	64%
10	0.28	21.76	1432.46	31164.14	57%	0.33	5.60	5430.81	30399.20	53%