Retail Ebb and Flow and the Overnight–Intraday Return Gap*

Yongkil Ahn[†] Alfred Qi Fan[‡] Don Noh[§] Stella Yeayeun Park[¶]

This draft: December 19, 2025

Abstract

In most stock markets, average overnight returns substantially exceed average intraday returns. Using accurate and exhaustive investor-type-level flow data from Korea, we show that retail trading intensity drives this overnight–intraday return gap. We establish causality by instrumenting for retail trading proportion (RTP) with nominal share prices, which affect retail trading through accessibility constraints but carry no fundamental information. We attribute this relationship to the recurrent "retail ebb and flow": driven by sensation-seeking and illusion of control, retail investors seek daytime exposure by systematically net buying at the open and net selling at the close. Consistent with this mechanism, day trading concentrates among young, male investors using mobile and desktop platforms and intensifies following high volatility.

JEL classification: G11, G12, G40
Keywords: Overnight returns, Intraday returns, Retail investors, Day trading, Price pressure

*This research was supported by the Singapore Ministry of Education (MOE) Academic Research Fund (AcRF) Tier 1 Grant. We thank Jingi Ha, Michael Junho Lee, Rongchen Li, Simon Oh, Sean Shin, Wei Xiong, and Motohiro Yogo for feedback and suggestions. We also thank participants of the Princeton Finance Student Workshop, HKUST Finance Brown Bag, SMU Finance Brown Bag, 2023 Annual Conference on Asia-Pacific Financial Markets, 2023 Australasian Finance and Banking Conference, 2024 Hawaii Accounting Research Conference, 2024 FMA Annual Meeting, 2025 Asia-Pacific Association of Derivatives Conference, for helpful comments.

[†]Seoul National University of Science and Technology (yongkil.ahn@seoultech.ac.kr)
[‡]Hong Kong University of Science and Technology (alfredfan@ust.hk)
[§]Hong Kong University of Science and Technology (noh@ust.hk)
[¶]Sogang University (yeayeunpark@sogang.ac.kr)

1 Introduction

A robust feature of global equity markets is the sharp contrast between returns realized during trading hours and those accrued overnight. Stocks earn substantial premiums from market close to the next day's open, while intraday returns are frequently negligible or negative.[1] This overnight–intraday return gap persists across time and geographies, suggesting it stems from fundamental investor behavior interacting with market structure rather than being a spurious relationship. Recent literature proposes that investor heterogeneity drives this pattern: distinct clienteles with differing objectives dominate trading at the open versus the close, and their systematic flows generate predictable price pressure (Lou et al., 2019). However, identifying which investors drive this wedge and why requires granular, high-frequency data on investor identities.

In this paper, we show that retail trading plays a central role in generating the overnight–intraday return gap. We use trade-level data from the Korea Exchange (KRX), which offers two crucial advantages in making our claim: (i) investor types, including retail investors, are identified for each trade directly by the KRX and (ii) the KRX is the sole stock exchange in Korea throughout our sample period. In other words, we work with accurate and exhaustive flows that can be safely called *the* retail flow. This is an advantage over U.S. studies where retail flows are either imputed or partial (Han and Kumar, 2013, Lou et al., 2019, Boehmer et al., 2021, Barber et al., 2022a). Using these data, we first document that in the time series, periods of high market-level retail trading proportion (RTP) exhibit high overnight returns and low intraday returns. We then show that in the cross-section, stocks with high RTP subsequently earn high overnight returns and low intraday returns. Beyond establishing this predictive relationship, we demonstrate that it is causal. We use an instrumental variable (IV) approach that exploits retail investors' tendency to trade stocks with low nominal per-share prices more actively—a preference unrelated to firm fundamentals but strongly predictive of retail trading intensity (Du et al., 2022).

We document a key stylized fact that explains how retail trading drives the return gap: retail investors systematically net buy at the market open and net sell near the close, a pattern we refer to as "retail ebb and flow." High-RTP stocks and periods exhibit larger return gaps because they have a

[1] See Figure A1 for global evidence. A notable exception is China, where a $T + 1$ settlement rule prohibits same-day position reversals (Qiao and Dam, 2020). The daytime-exposure mechanism we propose is consistent with this pattern.

higher concentration of traders seeking daytime exposure. Consistent with this interpretation, we find that high RTP predicts both stronger retail net buying near the open and stronger net selling near the close in subsequent periods, with this predictive relationship especially pronounced for mobile retail flows. Notably, RTP predicts flows at the open and close with opposite signs—net buying at the open but *net selling* at the close. This pattern distinguishes our mechanism from attention-based explanations (Berkman et al., 2012): if retail investors simply concentrate demand at the earliest opportunity following a salient event, we would observe positive flows at both the next day's open and close, yet we document net selling at the close followed by renewed net buying at the subsequent open.

Why do retail investors prefer such daytime exposure? We argue this behavior reflects sensation-seeking (Grinblatt and Keloharju, 2009) and overconfidence (Barber and Odean, 2000, 2001a). Overconfidence in this context operates through the illusion of control (Barber and Odean, 2001b): retail investors derive utility from monitoring positions and managing them actively during market hours, believing such active engagement improves outcomes, but close them by day's end, as overnight positions offer neither monitoring opportunities nor the perception of control. Consistent with this narrative, we find that past volatility predicts retail net buying near the open in the aggregate KRX data, with this pattern especially pronounced among mobile retail flows.

We further test this interpretation using account-level data from a national Korean brokerage firm.[2] First, we decompose brokerage flows into day trades, which are positions opened and closed within the same session, and non-day-trades. We find that the ebb-and-flow pattern is entirely driven by day trades, while non-day-trade flows exhibit the opposite intraday pattern, net selling at the open and net buying at the close. Day trades account for 35–50% of volume in our brokerage sample, a surprisingly high share. Second, day trading propensity varies across trader characteristics in ways consistent with sensation-seeking and overconfidence: it concentrates among young, male investors using mobile or desktop applications.

The Korean stock market provides an ideal setting for studying retail trading and its effect on asset pricing dynamics. Beyond the data advantages discussed earlier, Korea serves as a harbinger for retail trading trends in developed markets. Korea's market-level RTP has hovered around 70% over our sample period, well above traditional U.S. estimates of 5–20%. However, recent data show U.S.

[2]The identity of the brokerage firm cannot be revealed, but we provide more details of the data in Section 2.2.

retail share rising rapidly, with some sources estimating 30–37% of daily trading volume.[3] Similarly, mobile trading has accounted for over 50% of retail volume in Korea since 2020, foreshadowing trends now emerging in the U.S. and other major stock markets. As retail intensity and mobile trading continue to rise globally, the patterns we document in Korea offer insight into the future of asset pricing dynamics.

Our empirical analysis proceeds in four steps. First, we document that both time-series and cross-sectional variation in RTP predict the overnight–intraday return gap. We split the sample into months with above- and below-median aggregate RTP and find that high-RTP months exhibit return gaps of 4.11% compared to 2.05% in low-RTP months. Time-series regressions confirm this pattern. We regress the forward aggregate market return gap on current market-level RTP and find that a 1pp increase in RTP is associated with a 17.87bps (0.62bps) increase in the monthly (daily) return gap. In the cross-section, stocks in the highest RTP decile on average earn a 14.72% monthly return gap (5.26% overnight returns and −5.60% intraday returns), compared to 1.77% (0.87% overnight returns and −0.28% intraday returns) in the lowest RTP decile.[4] These patterns survive double sorts on size and volatility and persist in Fama–MacBeth regressions after controlling for size, volatility, book-to-market, momentum, and other firm characteristics. In the most stringent specification with industry fixed effects and controls, a one-standard-deviation increase in RTP is associated with a 2.35pp higher return gap, a 43% effect relative to the unconditional mean of 5.52%.

Second, we establish causality using nominal share price as an instrument for RTP. The exclusion restriction requires that, conditional on market capitalization and firm characteristics, the nominal price affects returns only through its effect on RTP.[5] This restriction is plausible if nominal price influences RTP through accessibility constraints: in the absence of fractional share trading, stocks with lower per-share prices are more accessible to investors with small account balances, who are disproportionately retail. We validate this channel using brokerage data. Even after controlling for demographics and trading experience, a one-standard-deviation decrease in the log of average daily account balance is associated with a 6.03pp increase in the proportion of volume executed in stocks

[3] Boehmer et al. (2021) estimate U.S. retail share at 17.5% as of 2022, down from a COVID-era peak of 24%. More recent estimates from MEMX (2025) suggest retail has since rebounded to 30–37% of daily U.S. equity volume.

[4] The monthly average overnight return minus the average intraday return does not sum exactly to the return gap, due to compounding and the specific definition of the monthly gap outlined in Section 2 (see also Lou et al., 2019).

[5] While nominal price is mechanically linked to size, our inclusion of market capitalization as a control isolates the variation in price driven by share count.

with opening prices below KRW 10,000. A remaining concern is that nominal prices may be related to omitted characteristics such as lottery-like characteristics and volatility, or that it depends on endogenous corporate decisions. However, to the extent that these factors affect returns by attracting retail investors, our reasoning for instrumenting RTP with nominal prices remains valid. Empirically, the instrument is strong (Kleibergen-Paap F-statistic > 300), and the two-stage least squares (2SLS) coefficient exceeds the OLS estimate. This is consistent with the attenuation bias in the OLS from measurement error or feedback trading by sophisticated investors who trade against the predictable return gap.

Third, we test whether high RTP predicts stronger retail ebb and flow. We measure stock-level retail net buy as retail buy minus retail sell volume, normalized by shares outstanding. In daily Fama–MacBeth regressions, RTP predicts persistent flows near open and close over 5- and 20-day horizons. Over the next 5 days, a one-standard-deviation increase in daily RTP predicts, on average for each day, 0.017bps higher aggregate retail net buying at the open and 0.020bps net selling at the close. The pattern is more pronounced among mobile traders (0.029bps at open, −0.031bps at close) and also present in day-trade flows from our brokerage data (0.035bps at open, −0.025bps at close).[6] Importantly, RTP predicts flows at the open and close with opposite signs for 5- and 20-day horizons—net buying at the open but net selling at the close. To test whether these flows generate price pressure, we compare retail transaction prices to those of short-term institutions within the same stock, day, and 30-minute interval. Relative to institutional volume-weighted average prices (VWAPs), retail investors, particularly mobile traders, buy at prices significantly higher in the opening 30 minutes (5.44bps for mobile vs. 2.27bps for other retail) and sell at prices significantly lower in the final 30 minutes (4.50bps for mobile vs. 4.97bps for other retail). This behavior indicates aggressive, impatient trading consistent with price pressure from the retail ebb and flow.

Fourth, we examine the micro foundations using account-level brokerage data. Using trade-level timestamps and unique account identifiers, we define day trades as intraday round-trip positions—transactions where a position is opened and closed, completely or partially, within the same daily session. Decomposing aggregate retail activity into day trades and non-day-trades reveals that the ebb-and-flow pattern is driven entirely by day trades, while non-day-trade flows exhibit the

[6] These magnitudes are numerically small because flows are computed over narrow time windows—the first and last hour of trading—and normalized by total shares outstanding.

opposite intraday pattern. To test whether day-trading activity is motivated by sensation-seeking and overconfidence, we compute each investor's day-trade proportion as day-trade volume divided by total trading volume and regress this measure on demographics. After controlling for trading platform, we find that male traders engage in 2.94pp more day trading than female traders, younger traders engage in 3.68pp more, and those with high self-reported investment knowledge engage in 5.19pp more. Mobile and desktop platform users engage in 6.06pp and 13.05pp more day trading, respectively, consistent with technology-enabled platforms facilitating intraday position turnover. These demographic effects reinforce each other: young male traders with high self-reported knowledge exhibit the highest day-trading propensity. These effects are economically significant relative to the unconditional average day trading proportion of 17.01% (Table 3). Collectively, these results indicate that day trading concentrates among demographics associated with sensation-seeking and overconfidence (Barber and Odean, 2000, 2001a, Kumar, 2009, Grinblatt and Keloharju, 2009).[7]

We demonstrate that our mechanism complements, rather than substitutes for, alternative explanations. For instance, overnight risk may amplify the retail ebb-and-flow effect: when overnight risk is high, arbitrageurs face higher costs of holding positions overnight, making them less willing to trade against the predictable price pressure created by retail investors' preference for daytime exposure. In Fama–MacBeth regressions that progressively add controls for overnight systematic risk (Bogousslavsky, 2021), attention-driven trading (Berkman et al., 2012), fast arbitrageur liquidity provision (Lu et al., 2023), and other variables proposed in prior studies, the coefficient on RTP remains robust and stable (Table 11). We further test for heterogeneous effects by interacting RTP with indicators for high values of these other mechanism proxies. The results reveal that the influence of retail trading is significantly amplified in subsamples with high overnight risk, heightened attention, and tighter limits to arbitrage, confirming that our proposed mechanism complements these other established economic drivers.

Related literature The overnight–intraday return gap is well-documented across multiple markets (Kelly and Clark, 2011, Branch and Ma, 2015). The literature most closely related to our work attributes this pattern to price pressure from distinct investor groups. Specifically, Lou et al. (2019,

[7]The cross-sectional average of trader-level day trading share is lower than the overall share because traders with the highest turnover ratios tend to be the most active day traders.

2024) propose a "tug-of-war" between overnight and intraday clienteles; Berkman et al. (2012) link high overnight returns to retail attention, arguing that individual investors concentrate buy orders at the market open. While our framework aligns with this clientele-based perspective, we offer important refinements. In particular, our mechanism explains the *recurrent* retail ebb and flow driven by a systematic preference for daytime exposure. We further distinguish our study by employing accurate and exhaustive exchange-tagged flows—rather than the imputed proxies common in prior research—and by establishing a causal link between retail intensity and the return gap.

Beyond the price pressure narrative, recent studies explore various mechanisms driving overnight–intraday dynamics. Extending the tug-of-war framework, Akbas et al. (2022) link its intensity to future return predictability, while Wang (2025) documents an asymmetric cross-firm tug-of-war where peer overnight returns predict focal overnight returns due to retail persistence. Baltussen et al. (2024) document an end-of-day reversal where early-day returns negatively predict returns in the final 30 minutes, attributing this to retail buy-the-dip behavior and short-seller risk management. Other explanations emphasize risk and market frictions. For instance, Hendershott et al. (2020) argue for distinct day and night risk premia, while Bogousslavsky (2021) and Boyarchenko et al. (2023) attribute the gap to inventory risk and arbitrage constraints. Focusing on market microstructure, Lu et al. (2023) model heterogeneous liquidity providers, demonstrating that fast arbitrageurs can charge higher prices when absorbing retail order flow at the open, as cream-skimming risk prevents competition from slower, lower-cost liquidity providers. Other studies point to uncertainty resolution (Hu et al., 2022, Bondarenko and Muravyev, 2023), extrapolation (Jones et al., 2022), or sentiment (Aboody et al., 2018) as primary drivers. Barardehi et al. (2023) decompose past returns into day and night components, finding that momentum is driven exclusively by intraday returns, which reflect private information revealed through trading, while overnight returns generate no momentum. Similar patterns exist in options markets; Muravyev and Ni (2020) find that delta-hedged option returns are negative overnight but positive intraday, suggesting that option prices fail to fully account for the fact that volatility is significantly higher during trading hours.

Our work also contributes to the literature on retail investor behavior and its price impact (Barber et al., 2009). Seminal works establish that retail investors are overconfident (Barber and Odean, 2000, 2001a), attention-driven (Barber and Odean, 2008, Barber et al., 2022b), and prone to speculation (Kumar and Lee, 2006, Kumar, 2009, Han and Kumar, 2013). While recent studies utilize order flow data to link retail activity to return predictability (Boehmer et al., 2021, Barber et al., 2022a), we isolate *day trading* as a distinct and important channel. While prior studies typically view day trading through the lens of investor learning (Linnainmaa, 2011), skill (Barber et al., 2014), or noise (Dalvi et al., 2023), we position our findings alongside Grinblatt and Keloharju (2009) and Dorn and Sengmueller (2009), who link high-frequency trading to sensation-seeking and entertainment. We show that day trading is far from a niche activity; rather, it accounts for over half of retail volume at times, driving an aggregate ebb and flow that affects asset dynamics in a persistent and economically significant manner.

Finally, our results contribute to the literature on price pressure and limits to arbitrage (Coval and Stafford, 2007, Kaniel et al., 2008). We show that predictable, high-frequency flows generate persistent mispricing that sophisticated investors do not fully offset. These findings underscore the presence of limits to arbitrage even at intraday horizons (Heston et al., 2010).

2 Data and stylized facts

This section describes our primary data sources—trade-level data from KRX and account-level data from a large Korean discount brokerage—and documents three stylized facts that motivate our analysis. First, overnight returns substantially exceed intraday returns in the Korean market. Second, retail investors are net buyers in the morning and net sellers in the afternoon, while institutions exhibit the opposite pattern. Third, a large fraction of retail trading volume is driven by day trades, defined as positions opened and closed within the same day.

2.1 KRX data

We obtain trade-level data for all common stocks listed on the KRX from January 2009 to December 2023. The KRX operates two major markets: the KOSPI (the main board) and the KOSDAQ (analogous to NASDAQ). As of December 2023, these markets collectively hosted 2,541 listed common stocks.

Given that the average Korean listed firm is small relative to those listed on U.S. exchanges, we restrict our sample to the largest half of stocks ranked by beginning-of-year market capitalization to ensure that micro-cap stocks do not influence our results. This filter yields an average of 1,015 stocks per year and a total of 169,961 stock-month observations over the sample period, accounting for an average of 97% of total market capitalization.

We obtain book equity, earnings, dividends per share, shares outstanding, and foreign ownership from the KRX's public data portal.[8] We compute market capitalization, return volatility, the Amihud (2002) illiquidity measure, and market beta from the trade-level data. Appendix A provides detailed variable definitions.

Overnight and intraday returns Our main outcome variables are overnight and intraday returns. Before defining them precisely, we discuss two relevant institutional details. First, opening and closing prices are determined through opening and closing auctions in Korea, as in other stock exchanges. These auctions aggregate buy and sell orders during the 30 minutes before market open and 10 minutes before market close to determine uniform opening and closing prices. We use these opening and closing auction prices to compute overnight and intraday returns.[9] Second, the Korean stock market extended its closing time by 30 minutes on August 1, 2016, from 3:00 p.m. to 3:30 p.m.

[Figure 1 here]

We follow Lou et al. (2019) in decomposing the daily close-to-close return into overnight and intraday components. We define the intraday return as the open-to-close return:

$$r_{i,s}^{\text{intraday}} = \frac{P_{i,s}^{\text{close}}}{P_{i,s}^{\text{open}}} - 1,$$

where $P_{i,s}^{\text{open}}$ and $P_{i,s}^{\text{close}}$ denote the opening and closing auction prices of stock i on day s. We then define the overnight return as:

$$r_{i,s}^{\text{overnight}} = \frac{1 + r_{i,s}^{\text{close-to-close}}}{1 + r_{i,s}^{\text{intraday}}} - 1,$$

where $r_{i,s}^{\text{close-to-close}}$ is the full daily return of stock i on day s. Figure 1 illustrates this decomposition.

[Figure 2 here]

[8]Foreign ownership is reported for regulatory reasons, but ownership shares of other investor types are not reported.

[9]Our main results are unchanged if we instead use VWAPs during the first and last 30 minutes of trading. See Appendix Table A2.

Figure 2 plots market capitalization-weighted cumulative overnight, intraday, and close-to-close log returns for our sample. Overnight returns substantially exceed close-to-close returns, while intraday returns are negative on average, consistent with U.S. evidence in Berkman et al. (2012). Over our 15-year sample from the beginning of 2009 to the end of 2023, a $1 investment in the overnight component grows to approximately $21.84, while the same investment in the intraday component declines to roughly $0.12.

We define the overnight–intraday return gap at the daily level as the difference between the overnight and intraday returns:

$$\mathrm{RG}_{i,s} = r_{i,s}^{\text{overnight}} - r_{i,s}^{\text{intraday}}.$$

For our main analysis, we aggregate to the stock-month level. Following Lou et al. (2019), we compound daily returns within each month and define the monthly return gap as:

$$r_{i,t}^{\text{overnight}} = \prod_{s\in\text{month } t} \left(1 + r_{i,s}^{\text{overnight}}\right) - 1,$$

$$r_{i,t}^{\text{intraday}} = \prod_{s\in\text{month } t} \left(1 + r_{i,s}^{\text{intraday}}\right) - 1,$$

$$\mathrm{RG}_{i,t} = \frac{1 + r_{i,t}^{\text{overnight}}}{1 + r_{i,t}^{\text{intraday}}} - 1.$$

Investor type-level trade flows A key advantage of the KRX data is that each executed trade is tagged with the investor type of both the buyer and the seller. KRX classifies investors into nine categories: (i) financial investment companies, (ii) insurance companies, (iii) investment trust companies, (iv) private equity funds, (v) banks, (vi) other financial companies, (vii) government pension funds, (viii) other corporations, and (ix) individuals (retail investors). We label financial investment companies as short-term institutions and group all other non-retail institutions as long-term institutions.[10]

In addition to investor type, KRX tags each trade with the trading medium through which it was executed: (i) brokerage, (ii) landline, (iii) mobile trading system (MTS), (iv) home trading system (HTS), or (v) direct market access (DMA). We group brokerage, landline, and DMA into "other media"

[10]The KRX uses investor code 1000 for financial investment companies; we classify these as short-term institutions and group all other non-retail institutions as long-term institutions. We make this grouping because some institutional types, such as private equity and insurance companies, trade infrequently at intraday horizons. See the KRX data portal for additional details.

because, with the rise of mobile and desktop trading, they collectively account for less than 10% of retail volume since 2016. This classification yields five investor categories: short-term institutions, long-term institutions, MTS retail traders, HTS retail traders, and other retail traders.

We document intraday trading patterns using 30-minute intervals. For each stock i, investor type j, and day s, we aggregate trades into 30-minute intervals from 9:00 a.m. to 3:00 p.m. To address the August 2016 extension of trading hours from 3:00 p.m. to 3:30 p.m., we combine all trades from 2:30 p.m. to 3:30 p.m. into a single final interval for the entire sample period. This yields 12 intervals per day, indexed by $\tau \in \{1, \ldots, 12\}$. For each interval, we define:

$$\text{buy}_{i,j,s,\tau} = \text{shares of stock } i \text{ bought by investor type } j \text{ during interval } (s, \tau),$$
$$\text{sell}_{i,j,s,\tau} = \text{shares of stock } i \text{ sold by investor type } j \text{ during interval } (s, \tau).$$

We define net buy as buy minus sell, normalized by shares outstanding:

$$\text{net buy}_{i,j,s,\tau} = \frac{\text{buy}_{i,j,s,\tau} - \text{sell}_{i,j,s,\tau}}{\text{shares outstanding}_{i,s}}.$$

Normalizing by shares outstanding allows us to compare net buy across stocks and over time. We express net buy in basis points because trade flows are small relative to shares outstanding at the 30-minute frequency.

[Figure 3 here]

Figure 3 plots average net buy across stocks and days for each 30-minute interval, separately for retail investors, short-term institutions, and long-term institutions. Panel (a) shows that retail investors are net buyers in the morning (peaking in the first 30 minutes) and net sellers in the afternoon (especially in the final interval). Short-term and long-term institutions exhibit the opposite pattern, consistent with market clearing. Panel (b) decomposes retail net buy by trading medium. The morning-buying and afternoon-selling pattern is most pronounced among MTS users, suggesting that mobile traders drive the aggregate retail pattern.

An important independent variable in our analysis is RTP. Han and Kumar (2013) impute RTP from small trade sizes in U.S. data, but investor type identifiers allow us to compute it directly in this study. At the stock-month level, we define RTP as:

$$\mathrm{RTP}_{i,t} = \frac{\text{retail volume}_{i,t}}{\text{total volume}_{i,t}}.$$

Summary statistics Table 1 presents summary statistics for our main variables. As noted earlier, the average stock-month overnight return is 2.37%, while the average intraday return is −1.45%, yielding a mean return gap of 5.52%. These magnitudes are comparable to U.S. evidence: Berkman et al. (2012) report daily overnight and intraday returns of 0.098% and −0.066%, which compound to approximately 1.96% and −1.32% per month.

[Table 1 here]

The median market capitalization in our sample is KRW 300 billion (approximately USD 255 million), reflecting the relatively small size of Korean listed firms.[11] The average RTP is 72.98%, well above traditional U.S. estimates of 5–20%, though recent data show U.S. retail share has surged to over 35%.[12] Korea's high retail intensity reflects institutional features that constrain high-frequency institutional trading, most notably a 0.3% transaction tax per round-trip, and widespread adoption of mobile and desktop trading platforms.

We argue that such retail-dominance is in fact a strength, as Korea serves as a harbinger for equity market trends in developed markets. The surge in retail activity, mobile trading, and intraday position turnover occurred earlier and more intensely in Korea than in the U.S. and other markets. This makes the Korean equity market an ideal laboratory for studying retail trading behavior and its effects on asset prices—patterns that are likely to become more relevant globally as retail intensity continues to rise.

[Table 2 here]

Table 2 reports time-averaged cross-sectional correlations among our main variables. Monthly RTP has a positive correlation of 0.19 with the return gap, consistent with our hypothesis that retail trading drives the overnight–intraday return spread. However, RTP is strongly correlated with other

[11] To facilitate interpretation, we convert KRW values to USD using the sample period average exchange rate of 1,130 KRW per USD.

[12] Boehmer et al. (2021) estimate U.S. retail share at 17.5% as of 2022, down from a COVID-era peak of 24%. More recent estimates from MEMX (2025) suggest retail now accounts for 30–37% of daily U.S. equity volume.

firm characteristics: high-RTP stocks tend to be small ($\rho = -0.74$ with log market equity), illiquid ($\rho = 0.11$ with Amihud illiquidity), and volatile ($\rho = 0.31$), consistent with Laarits and Sammon (2025). This motivates our use of control variables and instrumental variables in Section 3.

[Figure 4 here]

Figure 4 shows that the positive relationship between RTP and the return gap persists within size terciles. We assign stocks to terciles based on average daily market capitalization each month, then create binscatter plots using 20 bins within each tercile. The monotonic relationship between RTP and the return gap appears in all three terciles, indicating that size alone does not explain this pattern.

2.2 Brokerage data

Our brokerage data are analogous to the U.S. discount-brokerage datasets in Barber and Odean (2000, 2001a) and the Chinese brokerage data in Feng and Seasholes (2005). We observe anonymized account-level trades, end-of-day positions, and investor demographics. Our sample covers approximately 235,000 active retail accounts at a large Korean discount broker from January 2015 to December 2018. We receive data on all traders who execute at least 10 trades during the sample period, which represents approximately 7.8% of total retail trading volume during the same period.

The trade file reports all executions (quantity, price, timestamp, and order medium), the balance file contains end-of-day holdings for all securities, and the demographic file includes sex, birth date, account opening date, and other self-reported information. Korean brokerages collect self-reported information on income, investment knowledge, investment aggressiveness, and derivative experience to determine investor suitability for different investment products, as required by Korean financial regulations. The specific questions asked are included in Appendix Table A1.

Using the trade file, we identify round-trip positions that qualify as day trades. For a given investor, stock, and day, we define a day trade as any round trip in which a position opened during the day is fully or partially closed during the same day. Each executed share is uniquely assigned to either a day-trade round trip or a non-day-trade position; hence, all trade volume is partitioned into day-trade and non-day-trade categories. For each day, we compute the fraction of total executed volume attributable to day trades.

[Figure 5 here]

Figure 5 plots the monthly time series of the day-trade volume share. Computing this quantity requires account-level data to track position-level round trips, which explains why day trade volume share is rarely discussed in the literature. The proportion is surprisingly high, averaging around 45% over the sample period. However, this figure is less surprising when we consider that longer-horizon holdings, by construction, generate less turnover: investors who hold positions for weeks or months contribute little to daily volume, while day traders who open and close positions multiple times per day account for disproportionate trading activity.

[Figure 6 here]

To examine whether day traders drive the intraday flow pattern documented in Figure 3, we decompose retail net buy into day-trade and non-day-trade components. Figure 6 shows the results. The morning-buying and afternoon-selling pattern is largely driven by day-trade flows, while non-day-trade flows are relatively flat throughout the day. This suggests that the retail ebb and flow we document reflects the behavior of investors who open and close positions within the same day.

The brokerage data also allow us to distinguish traders by interface. We classify traders as MTS users, HTS users, or others based on the order medium recorded at execution. For each investor, we define the main medium as the interface through which she executes the greatest number of trades. We use these categories to characterize differences in activity, day-trade intensity, and performance.

[Table 3 here]

Table 3 presents summary statistics for our brokerage sample. Panel A shows that average account balances and trade sizes are comparable to those in Seru et al. (2010) who use a Finnish dataset. The age and gender composition is also similar: the mean user age is 46.15, and 61% of users are male.

Panel B compares users across trading platforms. MTS users are the youngest (mean age 42) and hold the smallest account balances (mean KRW 23.26 million), while also reporting the lowest levels of income and investment experience. In contrast, HTS users are older (mean age 50) and manage the largest balances (mean KRW 68.68 million). Users of other traditional channels (labeled as "Other"), such as brokerage-assisted or web-based trading, are the oldest group (mean age 52) but are less active traders.

Both MTS and HTS users engage in substantially more day trading than traditional media users: 17.11% of MTS volume and 24.11% of HTS volume come from day trades, compared to only 10.31% for other media. This pattern suggests that technology-enabled trading platforms, whether mobile or desktop, facilitate intraday position turnover by providing real-time price information and low-friction execution. In addition, MTS users exhibit signs of lower sophistication: they self-report the lowest investment knowledge (mean 2.36 vs. 2.76 for HTS and 2.64 for traditional media) and realize the worst performance, with a mean return of −4.75% compared to −3.13% for HTS and −3.44% for traditional users.[13]

3 Empirical results

We begin by documenting time-series variation in the overnight–intraday return gap at the market level and show that aggregate RTP is positively associated with the gap. We then turn to cross-sectional analysis, showing that stocks with higher RTP earn higher overnight returns and lower intraday returns. Finally, we establish a causal link from RTP to the return gap using an instrumental variables approach that exploits variation in nominal stock prices.

3.1 Time-series evidence

Each day (month), we compute value-weighted average overnight returns, intraday returns, and return gaps, weighting by beginning-of-day (beginning-of-month) market capitalization. We compute aggregate RTP in period t as the ratio of retail trading volume to total trading volume (both measured in KRW) across all stocks in our sample. We then split the sample into high-RTP and low-RTP periods based on whether period t market-level RTP is above or below its time-series median, and examine returns in period $t+1$.

[Table 4 here]

Table 4 reports average overnight returns, intraday returns, and the return gap following high-RTP and low-RTP periods. Panel A splits the sample by whether daily (monthly) RTP is above or below its median. At the daily level, high-RTP periods exhibit an overnight return of 0.11% versus 0.06% in

[13] See Appendix Table A1 for how we compute realized returns.

low-RTP periods, a difference of 0.05pp (t-statistic = 2.37). The return gap is 0.19% during high-RTP days versus 0.11% during low-RTP days, a difference of 0.08pp (t-statistic = 2.37). Intraday returns are slightly more negative during high-RTP days (−0.06% vs. −0.04%), though the difference is not statistically significant (t-statistic = −0.78).

These patterns are more pronounced at the monthly level. High-RTP months exhibit overnight returns of 2.39% versus 1.20% in low-RTP months (difference = 1.19%, t-statistic = 1.87) and return gaps of 4.11% versus 2.05% (difference = 2.06%, t-statistic = 2.97). Intraday returns are −1.57% during high-RTP months versus −0.76% during low-RTP months (difference = −0.81%, t-statistic = −1.73).

Panel B presents results from predictive regressions where we regress next period returns on current-period aggregate RTP, denoted as RTP_t. Specifically, we estimate the following regression:

$$y_{t+1} = \alpha + \beta \text{RTP}_t + \gamma r_t + \delta \sigma_t + \varepsilon_{t+1} \tag{1}$$

where y_{t+1} may be overnight return, intraday return, and return gap of aggregate market in period $t+1$, r_t is the close-to-close market return, and σ_t is market volatility. In daily regressions, σ_t is the standard deviation of daily market returns over the preceding 20 days. The monthly analogue is estimated as the standard deviation of daily market returns over trading days in month t.

At the daily level, a 1pp increase in RTP is associated with a 0.33bps increase in next-day overnight returns (t-statistic = 1.80) and a 0.62bps increase in the return gap (t-statistic = 2.65). At the monthly level, the predictive power is stronger: a 1pp increase in RTP is associated with a 8.01bps increase in next-month overnight returns (t-statistic = 2.04) and a 17.87bps increase in the return gap (t-statistic = 4.33).

These results show that periods with more intense retail trading are associated with higher overnight returns, lower (more negative) intraday returns, and a larger overnight–intraday return gap.

3.2 Cross-sectional return predictability

We now turn to the cross-sectional relationship between retail trading and the overnight–intraday return gap. We show that stocks with higher RTP subsequently earn higher overnight returns and lower intraday returns.

Portfolio sorts Each month, we sort stocks into deciles based on RTP during month t and compute value-weighted portfolio returns over month $t + 1$. Table 5 reports average monthly overnight returns, intraday returns, and the return gap for each decile portfolio.

[Table 5 here]

Panel A shows that overnight returns increase monotonically from 0.87% in the lowest RTP decile to 5.26% in the highest RTP decile, a spread of 4.39pp (t-statistic = 8.64). The local Fama–French five-factor alpha (α_{FF5}) also increases monotonically, from 0.67% to 4.99%, with a high-minus-low alpha of 4.31pp (t-statistic = 10.18).[14] Panel B shows that intraday returns decline from −0.28% in the lowest RTP decile to −5.60% in the highest RTP decile, a spread of −5.32pp (t-statistic = −10.41). The α_{FF5} spread is −5.07pp (t-statistic = −11.03). Panel C reports the return gap—overnight minus intraday returns. The return gap increases from 1.77% in the lowest RTP decile to 14.72% in the highest RTP decile, a spread of 12.95pp (t-statistic = 14.14). The α_{FF5} spread is 12.63pp (t-statistic = 14.06). All t-statistics are computed using Newey–West standard errors with six lags to account for potential autocorrelation in monthly returns. These large, statistically significant spreads show that RTP reliably predicts the overnight–intraday return gap, with monotonic patterns across deciles for both overnight and intraday returns.

[Table 6 here]

One concern is that RTP is correlated with firm size and return volatility (Table 2). To address this, we perform dependent double sorts on RTP and size, and on RTP and volatility, sorting first by size (volatility) and then by RTP within each size (volatility) quintile. Table 6 reports the results. Panel A shows that the return gap monotonically increases with RTP within each size quintile. For example, among small stocks, the return gap increases from 2.37% in the low-RTP quintile to 9.35% in the high-RTP quintile, a spread of 6.98pp (t-statistic = 9.99). Among large stocks, the spread is 8.33pp (t-statistic = 10.86). Panel B shows similar results for double sorts on RTP and volatility. Among low-volatility stocks, the return gap spread is 2.64pp (t-statistic = 4.69); among high-volatility stocks,

[14] The local factors are obtained from Global Factor Data by Jensen et al. (2023).

the spread is 11.78pp (t-statistic = 12.07). These results confirm that the predictive power of RTP is not merely driven by its correlation with size or volatility. However, the RTP effect is notably stronger among high-volatility stocks, consistent with the view that overnight risk makes it more difficult to arbitrage away the price pressure created by the retail clientele.

Fama-MacBeth regressions To control for multiple firm characteristics simultaneously, we estimate Fama–MacBeth regressions. Specifically, we estimate the following cross-sectional regression each month:

$$y_{i,t+1} = \alpha_t + \beta_t \mathrm{RTP}_{i,t} + \Gamma_t' \mathbf{X}_{i,t} + \varepsilon_{i,t+1}, \tag{2}$$

where $y_{i,t+1}$ is the overnight return, intraday return, or return gap for stock i in month $t+1$, and RTPi, t is RTP in month t. The vector $\mathbf{X}i, t$ includes size, book-to-market, market beta, Amihud illiquidity, foreign ownership, momentum, idiosyncratic volatility, and maximum daily return. All independent variables are winsorized at the 1% and 99% levels and standardized cross-sectionally each month to have zero mean and unit standard deviation. We report time-series averages of the coefficients β_t and Γ_t with Newey–West standard errors using six lags. Appendix Table A1 provides detailed variable definitions.

[Table 7 here]

Table 7 reports the results. A one-standard-deviation increase in RTP is associated with a 1.24pp increase in next-month overnight returns (column 1, t-statistic = 9.52), a 1.22pp decrease in intraday returns (column 4, t-statistic = −11.55), and a 2.97pp increase in the return gap (column 7, t-statistic = 13.12). These magnitudes are economically significant: a one-standard-deviation increase in RTP accounts for approximately 54% of the unconditional mean return gap of 5.52% (Table 1).

Columns (2), (5), and (8) add current overnight and intraday returns, size, market beta, book-to-market, Amihud illiquidity, foreign ownership, momentum, idiosyncratic volatility, and maximum daily return as controls. The coefficient on RTP remains statistically significant across all three outcome variables, though the magnitudes are somewhat smaller than in the univariate specification, indicating that part of RTP's predictive power operates through correlated firm characteristics. Columns (3), (6), and (9) add industry fixed effects based on KRX classifications. Even with controls and industry fixed effects, the effect remains economically large: a one-standard-deviation increase in RTP is associated with a 2.35pp increase in the return gap (t-statistic = 12.99). The adjusted R^2 values indicate that RTP and controls explain substantial cross-sectional variation in returns: 4.37% to 15.69% for overnight returns, and 4.11% to 16.02% for the return gap. These magnitudes are comparable to those in Lou et al. (2019).

The controls reveal several additional patterns. Consistent with Berkman et al. (2012), idiosyncratic volatility (IVOL) is positively associated with the return gap, while past overnight returns ($r^{\text{overnight}}$) generate short-term continuation. Interestingly, the coefficient on size is positive, seemingly at odds with Berkman et al. (2012), who find smaller stocks have higher overnight returns. This reversal occurs because RTP is strongly negatively correlated with size ($\rho = -0.74$, Table 2): within each size tercile, the return gap increases with RTP, but smaller stocks have higher RTP on average (Figure 4). Once we control for RTP, the residual size effect is positive.

3.3 Instrumental variables approach

While the Fama–MacBeth regressions establish a strong association between RTP and the return gap, the relationship may still be subject to omitted variables bias or reverse causality. Sophisticated investors may trade against predictable return gaps, attenuating OLS estimates, or unobserved firm characteristics may simultaneously attract retail trading and generate return gaps through independent channels.

To establish a causal link, we use an IV approach that exploits retail investors' preference for low-priced stocks. We verify this pattern in our brokerage data and use log nominal share price on the first trading day of each month as our instrument. The per-share price of a stock should carry no economic meaning: conditional on market capitalization, a $100 stock with 1 million shares outstanding is economically identical to a $10 stock with 10 million shares. However, in the absence of fractional share trading, which was not readily available to Korean retail investors during our sample period, stocks with lower per-share prices are more accessible to investors with small account balances.

Accessibility constraints We first validate this mechanism using account-level data from our brokerage sample. For each investor, we compute the average daily account balance over the sample period and the proportion of trading volume (in KRW) executed in stocks with opening prices below three thresholds: KRW 1,000, KRW 5,000, and KRW 10,000. We also compute the volume-weighted average opening price across all stocks traded by each investor. We then residualize both variables with respect to investor demographics (sex, age, income, investment knowledge, derivative experience, and trading medium) to isolate variation in trading behavior that is not explained by observable characteristics.

[Figure 7 here]

Figure 7 plots binned scatter plots of these residualized variables against residualized log average daily balance. Panel (a) shows that investors with smaller account balances concentrate their trading in low-price stocks: a one-standard-deviation decrease in the log of average daily account balance is associated with a 6.0pp increase in the proportion of volume executed in stocks with opening prices below KRW 10,000 (t-statistic = 20.74). Panel (b) shows that investors with smaller balances trade stocks with lower prices: a 10% decrease in average daily balance is associated with a 1.8% decrease in the volume-weighted average opening price (t-statistic = 19.92). These patterns confirm that accessibility constraints drive retail investors' preference for low-priced stocks.

Two-stage least squares regressions We estimate the following two-stage least squares (2SLS) specification:

$$\mathrm{RTP}_{i,t} = \alpha_{\mathrm{FS}} + \gamma \log(\text{per-share price})_{i,t} + \Gamma_{\mathrm{FS}}' \mathbf{X}_{i,t} + \text{Month FE} + \text{Industry FE} + \varepsilon_{i,t}, \tag{3}$$

$$y_{i,t+1} = \alpha + \beta \widehat{\mathrm{RTP}}_{i,t} + \Gamma' \mathbf{X}_{i,t} + \text{Month FE} + \text{Industry FE} + \eta_{i,t+1}, \tag{4}$$

where $y_{i,t+1}$ is the overnight return, intraday return, or return gap for stock i in month $t+1$, per-share price$_{i,t}$ is the nominal closing price on the first trading day of month t, $\mathbf{X}_{i,t}$ includes the same controls as in Table 7, and $\widehat{\mathrm{RTP}}_{i,t}$ is the fitted value from the first stage. We include month and industry fixed effects, based on KRX classifications, and cluster standard errors by stock and month.

The exclusion restriction requires that, conditional on market capitalization and firm characteristics, nominal price affects returns only through its effect on RTP. This is plausible: nominal price carries no information about fundamentals or cash flows, and any effects on investor attention or trading behavior should operate through retail trading activity, which is captured by RTP. The first stage is strong: the Kleibergen-Paap (KP) rank Wald F-statistic is 318.43 (Table 8), well above conventional thresholds for weak instruments (Stock and Yogo, 2005, Kleibergen and Paap, 2006).

[Table 8 here]

Table 8 reports the results. The OLS coefficient on RTP is 2.36 (t-statistic = 14.34), indicating that a one-standard-deviation increase in RTP is associated with a 2.36pp increase in the return gap. The 2SLS coefficient is 4.51 (t-statistic = 10.08), larger in magnitude than the OLS estimate. This pattern is consistent with measurement error or feedback trading attenuating the OLS coefficient. Sophisticated investors—both institutional and retail—may respond to predictable return gaps by trading against them, partially closing the gap and biasing OLS estimates toward zero. The IV approach isolates exogenous variation in RTP driven by accessibility constraints, yielding a larger causal effect. These results establish that retail trading intensity causally drives the overnight–intraday return gap, rather than merely correlating with it due to omitted variables or reverse causality.

4 Mechanisms

We now explore why higher RTP causes a larger overnight–intraday return gap. First, we establish that high-RTP stocks exhibit stronger retail ebb and flow—larger net buying near the open and larger net selling near the close—in subsequent periods. Second, we examine why retail investors prefer daytime exposure by analyzing day trading behavior and trader characteristics. Third, we assess whether our mechanism is robust to controls for alternative explanations and test whether it operates as a substitute or complement to other proposed drivers of the return gap.

4.1 RTP predicts retail ebb and flow

We test whether RTP predicts retail net buying near the open and net selling near the close by estimating the following Fama–MacBeth regressions:

$$\text{Flow}_{i,t+\tau} = \alpha_t + \beta_t \text{RTP}_{i,t} + \Gamma_t' \mathbf{X}_{i,t} + \varepsilon_{i,t+\tau}, \tag{5}$$

where $\text{Flow}_{i,t+\tau}$ is retail net buy, mobile trader net buy, or day-trade net buy (from brokerage data) at the open or close for stock i, averaged over the subsequent $\tau \in \{1, 5, 20\}$ days, $\text{RTP}_{i,t}$ is RTP on day t, and $\mathbf{X}_{i,t}$ includes the same controls as in Table 7.[15]

[Table 9 here]

Table 9 reports the results. We first consider mobile traders (MTS) and day traders in Columns (2), (3), (5), and (6). RTP predicts net buying near the open and net selling near the close, with this pattern persisting over the subsequent 5 and 20 days. For mobile traders over the next 5 days, a one-standard-deviation increase in RTP predicts 0.029bps higher net buying at the open (t-statistic = 2.61) and 0.031bps net selling at the close (t-statistic = −4.04). Day traders exhibit similar patterns: 0.028bps higher net buying at the open (t-statistic = 11.24) and 0.020bps net selling at the close (t-statistic = −12.59).

For aggregate retail flows in Columns (1) and (4), the pattern is largely consistent except for the positive coefficient at the close in Column (4). This likely reflects two offsetting forces: aggregate retail includes both naive traders (predominantly mobile) who exhibit strong ebb and flow, and more sophisticated retail traders who may trade against this pattern. Additionally, the attention mechanism documented by Berkman et al. (2012) suggests that persistent retail attention to high-RTP stocks can generate net buying pressure at both the open and close in the immediate next day. However, this attention effect dissipates over longer horizons as Panels B and C show the expected negative sign emerging at the close.

We again note these opposite signs for flows near the open and close at longer horizons. If retail investors were simply attracted to high-attention stocks and sought to establish positions as quickly as possible, they would concentrate their net buying at the earliest opportunity—whether at the open or close—and this pattern would persist. Instead, we observe persistent net buying at the open but net *selling* at the close over 5- and 20-day horizons, indicating that retail investors systematically open positions at the start of the trading day and close them by the end. This intraday turnover is consistent with a preference for daytime exposure rather than merely attention-driven entry into positions.

Price pressure from retail traders We have established that retail investors exhibit net buying at the open and net selling at the close. However, because net buy across investor types sums to zero

[15] We obtain similar results in a panel regression set up (see Appendix Table A3).

within each stock and interval—for every buyer there is a seller—the net flow pattern alone does not establish price pressure. We provide supplementary evidence that retail flows are "impatient": if retail investors demand immediacy, they should buy at higher prices and sell at lower prices relative to institutional traders who provide liquidity.

[Figure 8 here]

Figure 8 plots the average deviation of retail transaction prices from short-term institutional transaction prices within the same stock, day, and 30-minute interval, measured in basis points. We use short-term institutions as the reference because most market makers are classified into this category in the KRX data. The solid lines show the deviation of retail buy prices from institutional buy prices, and the dashed lines show the deviation of retail sell prices from institutional sell prices. Red lines represent mobile traders, and gray lines represent other retail traders.

Mobile traders exhibit the largest price impact: they buy at prices approximately 5.44bps higher than short-term institutions in the opening 30 minutes and sell at prices approximately 4.50bps lower in the final 30 minutes. This pattern is consistent with mobile traders demanding immediacy and absorbing liquidity, thereby pushing prices up when they buy and down when they sell. Other retail traders exhibit a similar but less pronounced pattern (2.27bps at open, 4.97bps at close). These findings align with two patterns presented earlier. First, mobile traders appear less sophisticated in the brokerage data, with smaller account balances, lower self-reported investment knowledge, and worse performance (Table 3). Second, mobile traders exhibit the strongest retail ebb and flow, with the most pronounced net buying at the open and net selling at the close (Figure 3). Together, these results suggest that mobile traders are a key source of intraday price pressure.

4.2 Sensation-seeking and illusion of control

We next present suggestive evidence that the preference for daytime exposure stems from behavioral patterns like sensation-seeking and illusion of control.

Sensation-seeking investors are attracted to volatile stocks and frequent trading (Grinblatt and Keloharju, 2009). If day trading provides entertainment or excitement, investors with a taste for sensation may prefer to hold risky positions during market hours when they can monitor them, but close positions by day's end to avoid overnight exposure. Consistent with this hypothesis, in untabulated results from the same regressions used to produce Table 9, we find that retail flows chase past volatility: stocks with high idiosyncratic volatility attract stronger retail net buying at the open and stronger net selling at the close. This pattern is most pronounced among mobile traders, who are younger and less sophisticated (Table 3).

Illusion of control—the belief that active monitoring and frequent trading improve outcomes—reinforces this behavior: investors who overestimate their ability to time intraday price movements derive utility from the perception of active management (Barber and Odean, 2001a). Mobile and desktop trading platforms amplify both mechanisms by providing real-time price updates and low-friction execution, making it easier to open and close positions within a single day.

For each investor, we compute the proportion of day-trade volume out of total trading volume (DT prop). Recall from Section 2.2 that a position is opened when an investor first buys a stock and is closed when it is sold either partially or completely. We then estimate the following cross-sectional regression:

$$\text{DT prop}_i = \alpha + \beta \mathbf{1}(\text{male})_i + \gamma \mathbf{1}(\text{young})_i + \delta \mathbf{1}(\text{high inv. know.})_i + \varepsilon_i, \tag{6}$$

where DT prop$_i$ is investor i's proportion of day-trade positions (expressed in percentage points), $\mathbf{1}(\text{young})_i$ equals one if investor i is younger than 35 as of December 30, 2018, and $\mathbf{1}(\text{high inv. know.})_i$ equals one if investor i self-reported high investment knowledge.[16]

[Table 10 here]

Table 10 presents the results. Column (1) shows that male traders engage in 4.03pp more day trading than female traders (t-statistic = 44.78), consistent with Barber and Odean (2001a), who document that men trade more frequently due to overconfidence. Younger traders (defined as below-median age) engage in 3.94pp more day trading (t-statistic = 32.83), and those with high self-reported investment knowledge engage in 5.54pp more (t-statistic = 39.57). Column (2) adds controls for trading medium: MTS users engage in 6.80pp more day trading than other media (t-statistic = 61.82), and HTS users engage in 13.80pp more (t-statistic = 106.15), consistent with technology-enabled platforms facilitating intraday position turnover. Column (3) shows that all

[16] The available responses ranged from 1 (low) to 4 (high) on an ordinal scale. We define "high" as a response of 4; all other responses are coded as zero. See Appendix Table A1 for details.

demographic effects remain significant when controlling for trading medium. Column (4) includes interactions: the day-trading propensity is highest among young male traders with high self-reported knowledge, with the three-way interaction adding 2.21pp (t-statistic = 2.91).

Overall, we see that day trading is concentrated among young, male, technology-enabled traders who self-report high investment knowledge. Combined with the evidence that retail flows chase volatile stocks (Table 9) and mobile traders exhibit the strongest ebb-and-flow pattern (Figure 3), these findings support our interpretation that the overnight–intraday return gap reflects a clientele effect driven by a specific subset of retail investors who disproportionately demand daytime exposure, generating predictable intraday price pressure.

4.3 Alternative and complementary mechanisms

We do not claim that daytime exposure is the only driver of the return gap. We view many alternative explanations as potentially reinforcing rather than competing with our mechanism. We assess this in two ways. First, we test whether RTP's predictive power survives controls for variables proposed in prior studies. Second, we examine whether RTP's effect varies across subsamples where alternative mechanisms should be stronger or weaker.

Robustness tests We begin with Fama–MacBeth regressions that progressively add controls motivated by alternative explanations. We include all controls in Table 7 as well.

[Table 11 here]

Table 11 presents the results. Column (1) reproduces our baseline specification from Table 7, which already includes idiosyncratic volatility (IVOL) among the controls. IVOL captures retail attention as in Berkman et al. (2012) and reflects limits to arbitrage as documented in Stambaugh et al. (2015).

Column (2) adds abnormal turnover (Abn_Vol) from Barber and Odean (2008), computed as the difference between turnover in month t and average turnover over the prior 12 months. This variable provides an additional proxy for attention-driven trading activity beyond what IVOL captures.

Column (3) adds overnight systematic variance ($\sigma^2_{\text{sys, overnight}}$) from Bogousslavsky (2021). Following his approach, we estimate firm-specific overnight betas by regressing each stock's overnight return on the market overnight return using a rolling 12-month window with at least 100 daily observations. Overnight systematic variance is then computed as the product of squared beta and variance of market returns estimated from the regressions, capturing exposure to systematic overnight risk.

Column (4) adds retail order imbalance at the open (rOIB_open) from Lu et al. (2023). We compute this at the daily level as the buy-minus-sell volume of aggressive retail orders during the first hour of trading (9:00–10:00), scaled by total daily volume, then aggregate to the monthly level. In Lu et al. (2023), high retail opening order imbalance allows fast arbitrageurs to absorb inventory during periods of high information asymmetry when cream-skimming risk prevents competition from lower-cost slow arbitrageurs. Consequently, fast arbitrageurs become the marginal liquidity providers and set opening prices that deviate from fundamental values to compensate for their high inventory costs. Thus, high retail buying pressure at the open predicts positive overnight returns and the associated night-minus-day wedge.

Column (5) includes overnight market beta ($\beta_{\text{overnight}}$) from Hendershott et al. (2020), estimated using the same rolling regression as overnight systematic variance but requiring only 30 daily observations. This beta captures differential sensitivity to overnight versus intraday market movements.

Column (6) adds abnormal negative reversals (Abn_NRev) from Akbas et al. (2022). For each stock-month, we compute NRev as the proportion of trading days exhibiting negative reversals—days with positive overnight returns followed by negative intraday returns. Abn_NRev is the difference between current-month NRev and the average NRev over the prior 12 months, capturing abnormal intensification of the tug-of-war pattern.

Across all specifications, RTP remains statistically significant and economically large. In the fully saturated specification in Column (6), the coefficient on RTP is 0.94 for overnight returns (t-statistic = 10.87) and −1.07 for intraday returns (t-statistic = -11.69), which are very close to 1.05 and −1.26 in the baseline specification (Table 7). This indicates that RTP captures variation in the return gap beyond what these alternative mechanisms explain, though the modest attenuation suggests some overlap with complementary channels.

Several alternative mechanisms exhibit significant coefficients, consistent with their proposed roles in generating the overnight–intraday return gap. Column (3) shows that overnight systematic risk increases the return gap, as expected from Bogousslavsky (2021). Column (4) shows that the fast- and slow-arbitrageur mechanism from Lu et al. (2023) also plays a role in increasing the return gap. Some variables are absorbed by others as they capture similar concepts: for instance, abnormal turnover's effect (column 2) becomes small once other attention proxies are included, while overnight market beta (column 5) enters with expected signs but modest magnitudes. The fact that RTP retains substantial explanatory power even after controlling for these mechanisms suggests that retail ebb and flow operates as a distinct but potentially complementary driver of the return gap.

Heterogeneous effects We expect alternative mechanisms to amplify rather than substitute for the retail ebb and flow channel. For example, overnight risk makes it costlier for arbitrageurs to hold positions overnight, reducing their willingness to close predictable gaps created by retail price pressure (Bogousslavsky, 2021). If these complementarities hold, we should observe stronger RTP effects in subsamples where these mechanisms are most active.

Table 12 presents results from Fama–MacBeth regressions that interact RTP with indicator variables for high values of each mechanism variable:

$$y_{i,t+1} = \alpha_t + \beta_t \mathrm{RTP}_{i,t} + \gamma_t \mathrm{High}_{i,t} + \delta_t \mathrm{RTP}_{i,t} \times \mathrm{High}_{i,t} + \Gamma_t' \mathbf{X}_{i,t} + \varepsilon_{i,t+\tau} \tag{7}$$

where $\mathrm{High}_{i,t} = \mathbf{1}[Z_{i,t} > \mathrm{median}_t(Z_{i,t})]$ indicates whether the mechanism variable $Z_{i,t}$ exceeds its cross-sectional median in month t. The parameter of interest is δ_t, which captures whether RTP's effect on the return gap is amplified (if positive) or attenuated (if negative) when the alternative mechanism is strong.

[Table 12 here]

The interaction terms reveal how alternative mechanisms interact with RTP's effect on the return gap. Each variable does not exhaustively or exclusively capture a single concept. For instance, IVOL proxies for both attention and limits to arbitrage; both $\sigma^2_{\text{sys, overnight}}$ and $\beta_{\text{overnight}}$ capture overnight risk. Thus, we cannot interpret every single coefficient in isolation. However, the overall pattern of interaction terms suggests that our retail ebb and flow mechanism is complementary with overnight risk, attention, and limits to arbitrage, rather than operating as an independent channel. For example, Columns (2) and (5) show that RTP's effect is significantly amplified when idiosyncratic volatility and retail order imbalance at the open are high, with both interactions positive for overnight returns and negative for intraday returns.

5 Conclusion

In this paper, we provide comprehensive evidence that retail trading activity is a primary driver of the overnight–intraday return gap. Leveraging unique trade-level data from the KRX that explicitly identifies investor types, we document a robust positive relationship between RTP and the return gap. Using nominal share price as an instrument for retail accessibility, we establish that this relationship is causal: intense retail trading generates the gap.

We attribute this phenomenon to a recurrent behavioral pattern we term retail ebb and flow. Retail investors, particularly those engaging in day trading via mobile platforms, systematically exhibit net buying pressure at the market open and net selling pressure near the close. This behavior is consistent with a preference for daytime exposure driven by sensation-seeking and the illusion of control: investors seek the utility of monitoring and managing positions during active market hours but prefer to close positions before the overnight period when they cannot act. Our analysis of account-level brokerage data confirms that this intraday turnover is concentrated among young, male investors and those using technology-enabled trading interfaces.

Our findings demonstrate that retail trading exerts an economically significant and persistent influence on asset pricing dynamics, and that these effects interact with channels such as overnight risk and limits to arbitrage. This insight is particularly timely given the global surge in retail participation and the proliferation of mobile trading platforms. As zero-commission trading and gamified interfaces lower barriers to entry worldwide, the Korean market provides a valuable glimpse of what to expect in the future of global equity markets.

References

Aboody, D., Even-Tov, O., Lehavy, R., and Trueman, B. (2018). Overnight returns and firm-specific investor sentiment. *Journal of Financial and Quantitative Analysis*, 53(2):485–505.

Akbas, F., Boehmer, E., Jiang, C., and Koch, P. D. (2022). Overnight returns, daytime reversals, and future stock returns. *Journal of Financial Economics*, 145(3):850–875.

Baltussen, G., Da, Z., and Soebhag, A. (2024). End-of-Day Reversal.

Barardehi, Y. H., Bogousslavsky, V., and Muravyev, D. (2023). What Drives Momentum and Reversal? Evidence from Day and Night Signals.

Barber, B. M., Huang, X., Jorion, P., Odean, T., and Schwarz, C. (2022a). A (sub)penny for your thoughts: Tracking retail investor activity in TAQ.

Barber, B. M., Huang, X., Odean, T., and Schwarz, C. (2022b). Attention-induced trading and returns: Evidence from robinhood users. *The Journal of Finance*, 77(6):3141–3190.

Barber, B. M., Lee, Y.-T., Liu, Y.-J., and Odean, T. (2014). The cross-section of speculator skill: Evidence from day trading. *Journal of Financial Markets*, 18:1–24.

Barber, B. M. and Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance*, 55(2):773–806.

Barber, B. M. and Odean, T. (2001a). Boys will be boys: Gender, overconfidence, and common stock investment*. *The Quarterly Journal of Economics*, 116(1):261–292.

Barber, B. M. and Odean, T. (2001b). The Internet and the Investor. *Journal of Economic Perspectives*, 15(1):41–54.

Barber, B. M. and Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The Review of Financial Studies*, 21(2):785–818.

Barber, B. M., Odean, T., and Zhu, N. (2009). Do retail trades move markets? *The Review of Financial Studies*, 22(1):151–186.

Berkman, H., Koch, P., Tuttle, L., and Zhang, Y. (2012). Paying attention: Overnight returns and the hidden cost of buying at the open. *Journal of Financial and Quantitative Analysis*, 47(4):715–741.

Boehmer, E., Jones, C. M., Zhang, X., and Zhang, X. (2021). Tracking retail investor activity. *The Journal of Finance*, 76(5):2249–2305.

Bogousslavsky, V. (2021). The cross-section of intraday and overnight returns. *Journal of Financial Economics*, 141(1):172–194.

Bondarenko, O. and Muravyev, D. (2023). Market return around the clock: A puzzle. *Journal of Financial and Quantitative Analysis*, 58(3):939–967.

Boyarchenko, N., Larsen, L. C., and Whelan, P. (2023). The overnight drift. *The Review of Financial Studies*, page hhad020.

Branch, B. S. and Ma, A. J. (2015). Overnight return, the invisible hand behind intraday returns?

Coval, J. and Stafford, E. (2007). Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics*, 86(2):479–512.

Dalvi, M., Deuskar, P., Glosten, L. R., and Jagannathan, R. (2023). Day Traders, Noise, and Cost of Immediacy.

Dorn, D. and Sengmueller, P. (2009). Trading as Entertainment? *Management Science*, 55(4):591–603.

Du, J., Huang, D., Liu, Y.-J., Shi, Y., Subrahmanyam, A., and Zhang, H. (2022). Nominal Prices, Retail Investor Participation, and Return Momentum.

Feng, L. and Seasholes, M. S. (2005). Do investor sophistication and trading experience eliminate behavioral biases in financial markets? *Review of Finance*, 9(3):305–351.

Grinblatt, M. and Keloharju, M. (2009). Sensation seeking, overconfidence, and trading activity. *The Journal of Finance*, 64(2):549–578.

Han, B. and Kumar, A. (2013). Speculative retail trading and asset prices. *Journal of Financial and Quantitative Analysis*, 48(2):377–404.

Hendershott, T., Livdan, D., and Rösch, D. (2020). Asset pricing: A tale of night and day. *Journal of Financial Economics*, 138(3):635–662.

Heston, S. L., Korajczyk, R. A., and Sadka, R. (2010). Intraday patterns in the cross-section of stock returns. *The Journal of Finance*, 65(4):1369–1407.

Hu, G. X., Pan, J., Wang, J., and Zhu, H. (2022). Premium for heightened uncertainty: Explaining pre-announcement market returns. *Journal of Financial Economics*, 145(3):909–936.

Jensen, T. I., Kelly, B., and Pedersen, L. H. (2023). Is There a Replication Crisis in Finance? *The Journal of Finance*, 78(5):2465–2518.

Jones, C. S., Pyun, S., and Wang, T. (2022). Return extrapolation and day/night effects.

Kaniel, R., Saar, G., and Titman, S. (2008). Individual investor trading and stock returns. *The Journal of Finance*, 63(1):273–310.

Kelly, M. A. and Clark, S. P. (2011). Returns in trading versus non-trading hours: The difference is day and night. *Journal of Asset Management*, 12(2):132–145.

Kleibergen, F. and Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133(1):97–126.

Kumar, A. (2009). Who gambles in the stock market? *The Journal of Finance*, 64(4):1889–1933.

Kumar, A. and Lee, C. M. (2006). Retail investor sentiment and return comovements. *The Journal of Finance*, 61(5):2451–2486.

Laarits, T. and Sammon, M. (2025). The retail habitat. *Journal of Financial Economics*, 172:104144.

Linnainmaa, J. T. (2011). Why do (some) households trade so much? *The Review of Financial Studies*, 24(5):1630–1666.

Lou, D., Polk, C., and Skouras, S. (2019). A tug of war: Overnight versus intraday expected returns. *Journal of Financial Economics*, 134(1):192–213.

Lou, D., Polk, C., and Skouras, S. (2024). The day destroys the night, night extends the day: A clientele perspective on equity premium variation. *London School of Economics Working Paper*.

Lu, Z., Malliaris, S., and Qin, Z. (2023). Heterogeneous liquidity providers and night-minus-day return predictability. *Journal of Financial Economics*, 148(3):175–200.

Muravyev, D. and Ni, X. C. (2020). Why do option returns change sign from day to night? *Journal of Financial Economics*, 136(1):219–238.

Qiao, K. and Dam, L. (2020). The overnight return puzzle and the "T+1" trading rule in chinese stock markets. *Journal of Financial Markets*, 50:100534.

Seru, A., Shumway, T., and Stoffman, N. (2010). Learning by trading. *The Review of Financial Studies*, 23(2):705–739.

Stambaugh, R. F., Yu, J., and Yuan, Y. (2015). Arbitrage asymmetry and the idiosyncratic volatility puzzle. *The Journal of Finance*, 70(5):1903–1948.

Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear iv regression. In Andrews, D. W. K., editor, *Identification and Inference for Econometric Models*, pages 80–108. Cambridge University Press, New York.

Wang, H. (2025). Decoding Momentum Spillover Effects. *Journal of Financial and Quantitative Analysis*, pages 1–64.

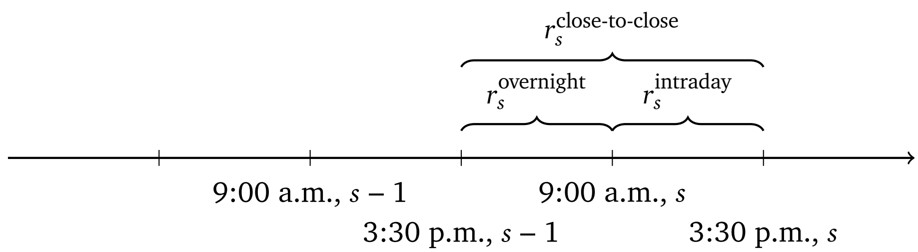


Figure 1. Overnight and intraday returns

This figure illustrates the decomposition of daily stock returns. We define the overnight return as the close-to-open return from 3:30 p.m. on day $s-1$ to 9:00 a.m. on day s, and the intraday return as the open-to-close return from 9:00 a.m. to 3:30 p.m. on day s. These two components combine to form the close-to-close return. The Korean stock market opens at 9:00 a.m. and closes at 3:30 p.m. Before August 1, 2016, the market closed at 3:00 p.m.

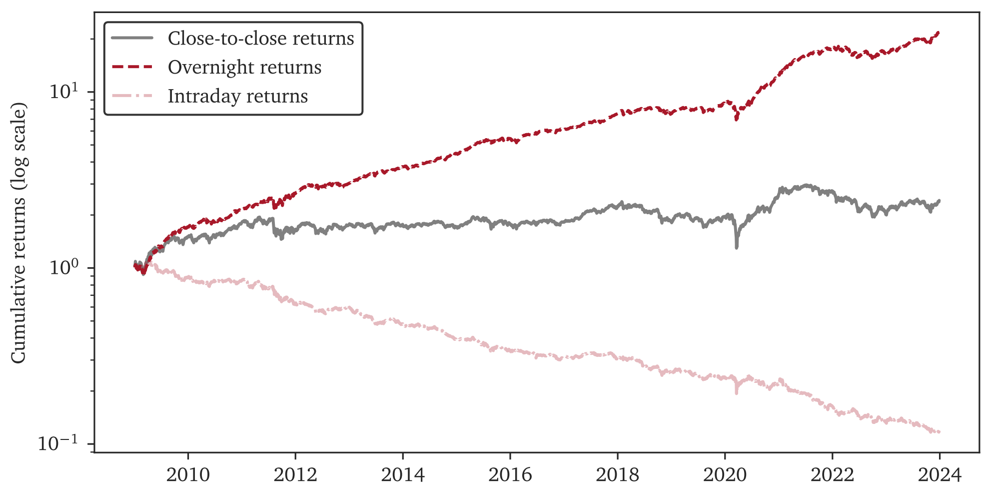
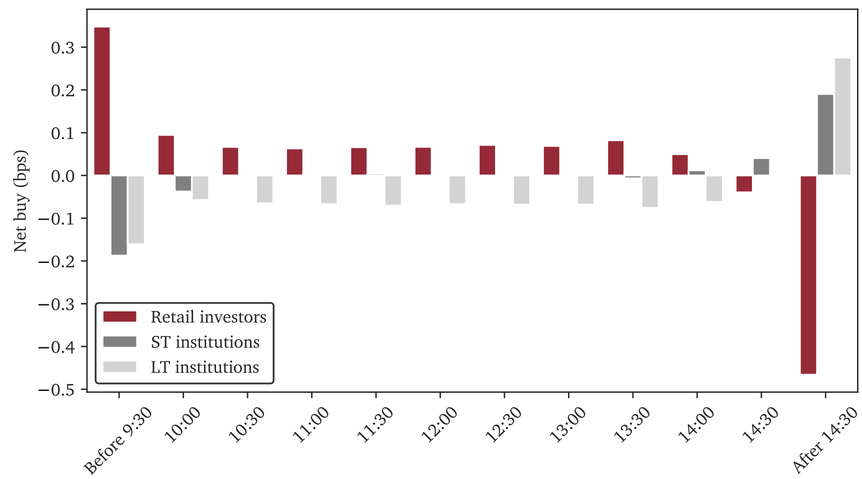
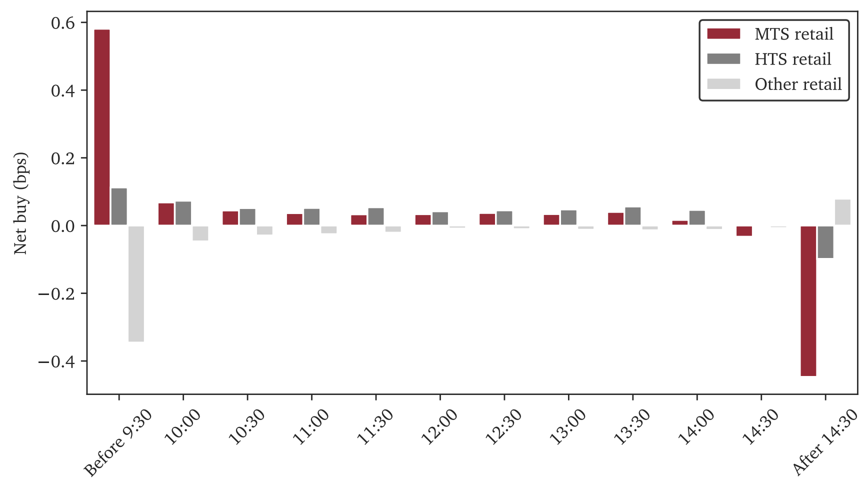


Figure 2. Cumulative overnight and intraday returns

This figure plots cumulative value-weighted returns for the Korean stock market from January 2009 to December 2023 on a log scale. We compute daily overnight returns (close-to-open), intraday returns (open-to-close), and close-to-close returns for each stock, then aggregate using the previous day's market capitalization as weights. A \$1 investment in the overnight component grows to approximately \$21.84, while the same investment in the intraday component declines to \$0.12.



(a) Retail and institutional flows



(b) Retail flows by trading medium

Figure 3. Intraday net buy pattern by investor type

This figure plots average net buy across all stocks and trading days for each 30-minute interval from market open to close, measured in basis points of shares outstanding. Panel (a) shows net buy separately for retail investors, short-term (ST) institutions, and long-term (LT) institutions. Panel (b) decomposes retail net buy by trading medium: mobile trading system (MTS), home trading system (HTS), and other media. "Before 9:30" aggregates all trades executed before 9:30 a.m. and "After 14:30" aggregates all trades executed after 2:30 p.m. The sample period is January 2009 to December 2023.

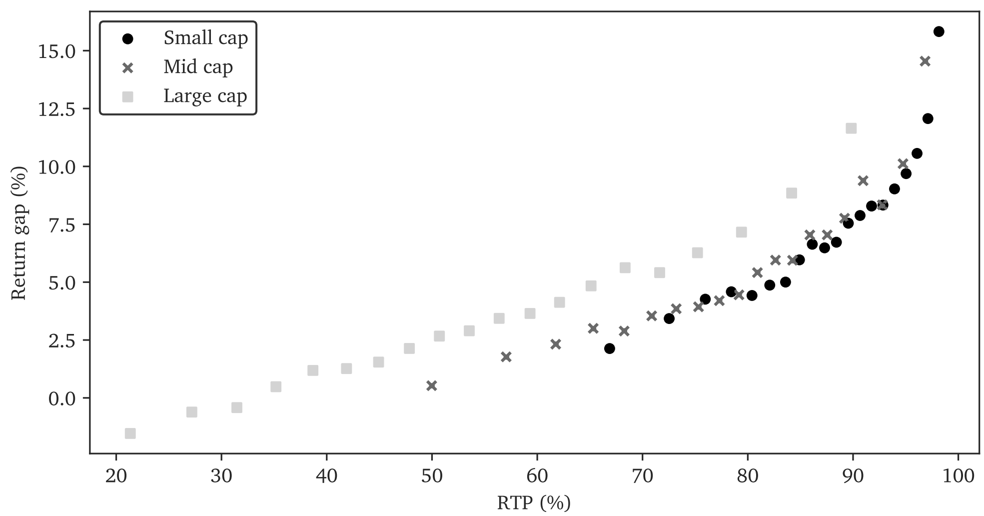


Figure 4. RTP and the overnight–intraday return gap

This figure plots binned scatter plots of the monthly return gap against retail trading proportion (RTP) within size terciles. Each month, we assign stocks to terciles based on average daily market capitalization, then create 20 bins of RTP within each tercile. Each point represents the average return gap and average RTP within a bin. The return gap is defined as the overnight return minus the intraday return, expressed in percentage points. RTP is the ratio of retail trading volume to total trading volume over the prior 20 trading days, expressed in percent. The sample period is January 2009 to December 2023.

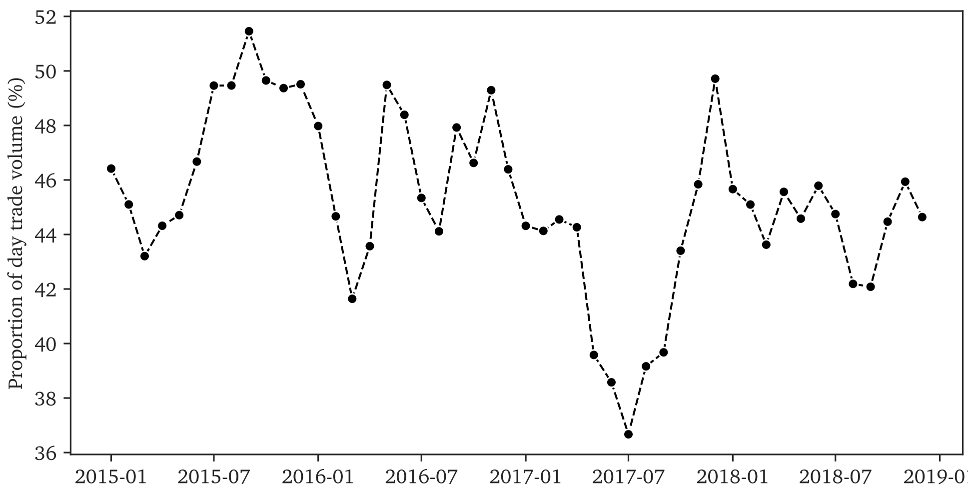


Figure 5. Proportion of day trade volume in the brokerage dataset

This figure plots the monthly proportion of trading volume attributable to day trades in the brokerage dataset. For each investor, stock, and day, we identify all positions that are opened and closed (fully or partially) within the same trading session as day trades. The proportion is computed as the total value of day trade volume divided by total trading volume in each month, expressed in percent. The sample covers approximately 235,000 active retail accounts at a large Korean discount brokerage from January 2015 to December 2018.

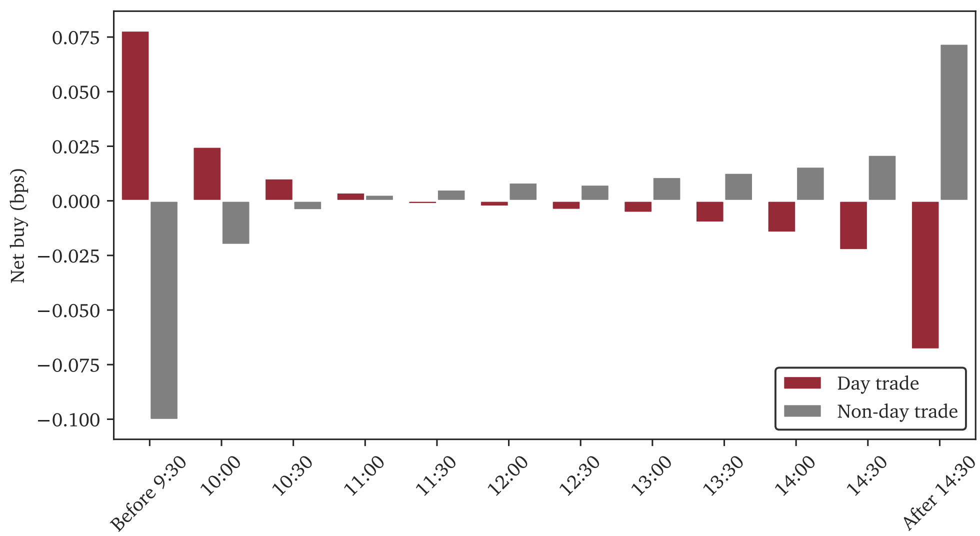
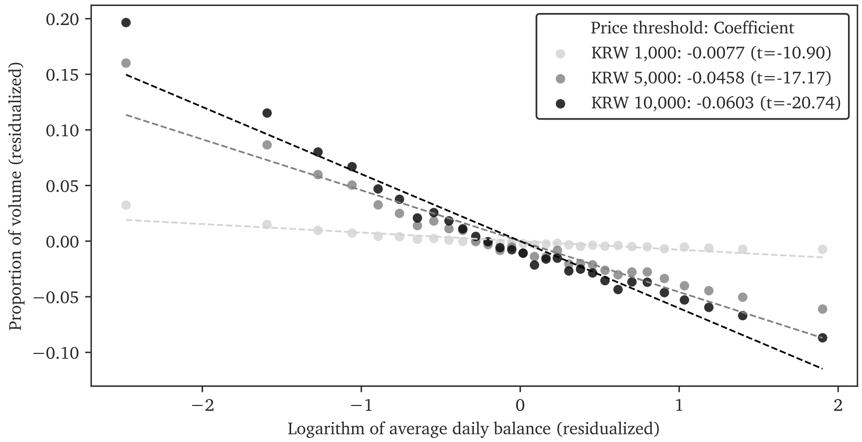
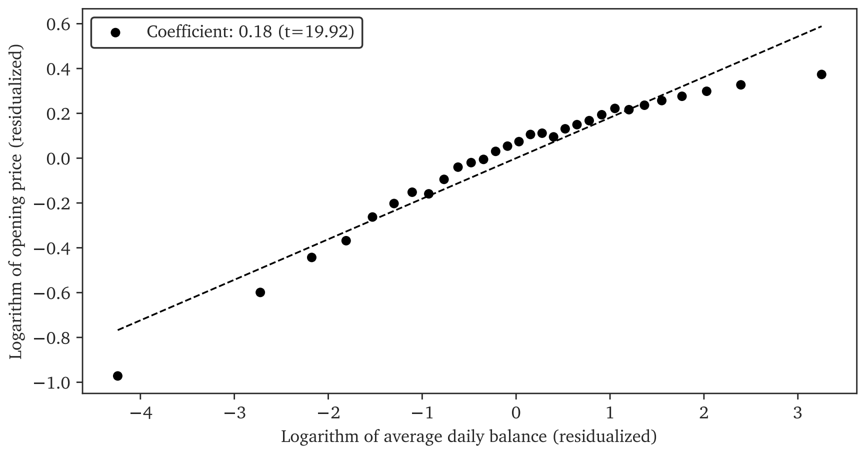


Figure 6. Intraday net buy pattern for day trades vs. non-day trades

This figure plots average net buy across all stocks and trading days for each 30-minute interval, separately for day trades and non-day trades, measured in basis points of shares outstanding. Day trades are positions opened and closed (fully or partially) within the same trading session. Non-day trades are all other trades. “Before 9:30” aggregates all trades executed before the 9:30 a.m., and “After 14:30” aggregates all trades executed after 2:30 p.m. The sample covers the brokerage dataset from January 2015 to December 2018.



(a) Proportion of volume in low-price stocks



(b) Volume-weighted average opening price

Figure 7. Tendency to trade low-price stocks and daily balance

This figure plots binned scatter plots of the relationship between average daily account balance and the tendency to trade low-price stocks, controlling for investor demographics. For each investor in the brokerage dataset, we compute the average daily account balance over the sample period and two measures of low-price stock trading: (a) the proportion of trading volume (in KRW) executed in stocks with opening prices below KRW 1,000, 5,000, and 10,000, and (b) the logarithm of the volume-weighted average opening price across all stocks traded. We residualize both variables with respect to sex, age, income, investment knowledge, derivative experience, and trading medium, then create 30 bins based on residualized log average daily balance. Each point represents the average residualized values within a bin. The dashed lines show fitted linear relationships, with coefficients and *t*-statistics reported in the legends. The sample covers approximately 235,000 active retail accounts from January 2015 to December 2018.

Figure 8. Intraday VWAP deviation relative to short-term institutions

This figure plots the average deviation of retail transaction prices from short-term institutional transaction prices within the same stock, day, and 30-minute interval, measured in basis points. For each 30-minute interval, we compute the volume-weighted average price (VWAP) separately for mobile retail traders (MTS), other retail traders, and short-term institutions. The solid lines show the percentage deviation of retail buy prices from institutional buy prices, and the dashed lines show the percentage deviation of retail sell prices from institutional sell prices. "Before 9:30" aggregates all trades executed before 9:30 a.m., and "After 14:30" aggregates all trades executed after 2:30 p.m. The sample period is January 2009 to December 2023.

Table 1. Summary statistics for the KRX data

This table presents summary statistics for the main variables used in the analysis. The sample consists of 169,961 stock-month observations covering common stocks listed on the Korea Exchange from January 2009 to December 2023, restricted to the largest half of stocks by beginning-of-year market capitalization. Overnight return is the close-to-open return, intraday return is the open-to-close return, and return gap is the difference between the two. Retail trading proportion (RTP) is the ratio of retail trading volume to total trading volume over all trading days in the month. Other variables include standard firm characteristics (size, book-to-market, beta, illiquidity, foreign ownership, and past return measures (120-day momentum, idiosyncratic volatility, maximum daily return). Detailed variable definitions are provided in Appendix Table A1.

	N	Mean	SD	25th	50th	75th
Overnight return (%)	169,961	2.37	8.41	-2.26	1.68	6.06
Intraday return (%)	169,961	-1.45	11.96	-8.44	-2.12	4.51
Return gap (%)	169,961	5.52	16.88	-4.68	3.80	13.46
RTP (%)	169,961	72.98	20.33	60.79	78.54	88.99
Log market equity	169,961	5.89	1.32	4.98	5.57	6.50
Book-to-market ratio	169,961	0.99	0.89	0.39	0.74	1.32
Market beta	169,961	0.89	0.50	0.54	0.87	1.22
Amihud illiquidity	169,961	0.07	0.91	0.00	0.01	0.03
Foreign ownership (%)	169,961	10.31	13.36	1.59	5.10	13.67
Past 120-day return (%)	169,961	4.05	38.20	-17.09	-2.72	15.73
Idiosyncratic volatility (%)	169,961	2.10	1.22	1.32	1.82	2.51
Maximum daily return (%)	169,961	5.96	4.47	3.16	4.75	7.17

Table 2. Cross-sectional correlation among firm characteristics

This table presents time-series averages of cross-sectional correlations among the main variables. For each month from January 2009 to December 2023, we compute pairwise correlations across all stocks in the sample, then average these correlations over time. The sample consists of common stocks listed on the Korea Exchange, restricted to the largest half of stocks by beginning-of-year market capitalization. Variable definitions are provided in Appendix Table A1.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) Overnight return (%)											
(2) Intraday return (%)	-0.22										
(3) Return gap (%)	0.67	-0.84									
(4) RTP (%)	0.20	-0.10	0.19								
(5) Log market equity	-0.03	0.10	-0.10	-0.74							
(6) Book-to-market ratio	-0.09	0.07	-0.11	-0.06	-0.10						
(7) Market beta	0.08	-0.08	0.10	0.12	0.07	-0.14					
(8) Amihud illiquidity	-0.02	0.03	-0.02	0.11	-0.19	0.18	-0.19				
(9) Foreign ownership (%)	-0.07	0.04	-0.08	-0.58	0.54	0.02	-0.05	-0.07			
(10) Past 120-day return (%)	0.20	0.27	-0.08	-0.02	0.15	-0.14	-0.07	-0.03	0.02		
(11) Idiosyncratic volatility (%)	0.31	0.13	0.11	0.31	-0.12	-0.22	0.15	-0.08	-0.16	0.25	
(12) Maximum daily return (%)	0.33	0.27	0.00	0.29	-0.08	-0.18	0.18	-0.07	-0.14	0.25	0.86

Table 3. Summary statistics for the brokerage dataset

This table presents summary statistics for approximately 235,000 active retail accounts at a large Korean discount brokerage from January 2015 to December 2018. Panel A reports statistics for all investors. Panel B reports statistics by main trading medium: mobile (MTS), desktop (HTS), and other traditional channels (Other) such as brokerage-assisted or web-based trading. Main trading medium is the interface through which the investor executes the most trades. Performance is realized return over the sample period, winsorized at 1% and 99%. Day trade volume proportion is the percentage of trading volume attributable to positions opened and closed within the same day. Approximately half of investors completed questionnaires on monthly income, investment knowledge, investment aggressiveness, and derivative experience. Detailed variable definitions are provided in Appendix Table A1.

Panel A: All investors

	N	Mean	SD	25th	50th	75th
Male	235,445	0.61	0.49	0.00	1.00	1.00
Age	235,445	46.15	11.57	37.00	44.00	54.00
Monthly income	143,119	2.86	1.06	2.00	3.00	3.00
Investment knowledge	143,119	2.53	0.93	2.00	2.00	3.00
Investment aggressiveness	143,119	3.99	0.94	4.00	4.00	5.00
Derivative experience	143,119	1.68	0.86	1.00	1.00	3.00
Performance (%)	235,445	-4.07	20.03	-9.51	-1.95	1.26
Avg. daily balance (KRW millions)	235,445	42.10	291.42	3.89	11.48	32.08
Avg. trade size (KRW millions)	235,445	4.24	8.40	0.73	2.00	4.84
Number of unique stocks traded	235,445	47.31	94.03	8.00	18.00	46.00
Number of trades	235,445	796.31	7577.56	32.00	98.00	352.00
Day trade volume proportion (%)	235,445	17.01	21.91	0.00	7.57	26.58

Panel B: Investors by trading medium

	MTS		HTS		Other	
	N	Mean	N	Mean	N	Mean
Male	125,498	0.64	52,485	0.65	57,455	0.48
Age	125,498	42.00	52,485	49.84	57,455	51.83
Monthly income	67,658	2.78	30,332	3.01	45,122	2.89
Investment knowledge	67,658	2.36	30,332	2.76	45,122	2.64
Investment aggressiveness	67,658	3.74	30,332	4.16	45,122	4.24
Derivative experience	67,658	1.54	30,332	1.81	45,122	1.79
Performance (%)	125,498	-4.75	52,485	-3.13	57,455	-3.44
Avg. daily balance (KRW millions)	125,498	23.26	52,485	68.68	57,455	59.00
Avg. trade size (KRW millions)	125,498	3.20	52,485	4.60	57,455	6.18
Number of unique stocks traded	125,498	40.32	52,485	78.10	57,455	34.46
Number of trades	125,498	473.99	52,485	2073.91	57,455	333.36
Day trade volume proportion (%)	125,498	17.11	52,485	24.11	57,455	10.31

Table 4. Time-series evidence

This table presents the time-series relation between aggregate retail trading proportion (RTP) and subsequent market returns at daily and monthly frequencies. Panel A reports average returns following high-RTP and low-RTP periods, where high-RTP periods are those with aggregate RTP above the sample median. Panel B reports regression coefficients from predictive regressions of next-period market returns on current-period aggregate RTP, controlling for contemporaneous market returns and market volatility. RTP is measured as a percentage, and coefficients are scaled by 100 for readability; they indicate the change in returns (in basis points) associated with a one-percentage-point increase in RTP. For daily (monthly) observations, Newey–West standard errors are reported in parentheses, computed with 20 lags (six lags). The sample period is January 2009 to December 2023. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Daily level			Monthly level		
	Overnight	Intraday	Return gap	Overnight	Intraday	Return gap
Panel A: Average returns in periods with high and low RTP						
High RTP	0.11***	-0.06***	0.19***	2.39***	-1.57***	4.11***
	(0.02)	(0.02)	(0.02)	(0.53)	(0.32)	(0.58)
Low RTP	0.06***	-0.04***	0.11***	1.20***	-0.76***	2.05***
	(0.01)	(0.02)	(0.02)	(0.28)	(0.26)	(0.34)
High – Low	0.05**	-0.02	0.08**	1.19*	-0.81*	2.06***
	(0.02)	(0.03)	(0.03)	(0.64)	(0.47)	(0.70)
Panel B: Predictive regressions with RTP						
RTP	0.33*	-0.29	0.62***	8.01**	-9.25**	17.87***
	(0.18)	(0.19)	(0.24)	(3.93)	(3.90)	(4.12)
Controls	✓	✓	✓	✓	✓	✓
Adjusted R^2 (%)	0.48	0.09	0.11	10.26	5.77	12.89

Table 5. Portfolios sorted by retail trading proportion (RTP)

This table presents the monthly average excess returns of portfolios sorted by retail trading proportion (RTP). Each month, stocks are sorted into deciles based on RTP. Panels A, B, and C report overnight returns, intraday returns, and return gaps, respectively, for each decile portfolio. H–L is a long-short portfolio that buys the highest RTP decile and sells the lowest RTP decile. α_{FF5} denotes alphas from the Fama–French five-factor model estimated for the Korean market. Newey–West t-statistics are reported in parentheses, computed with six lags. Portfolios are value-weighted and rebalanced monthly. The sample period is January 2009 to December 2023.

	Low	2	3	4	5	6	7	8	9	High	H-L
Panel A: Overnight returns											
Excess	0.87	1.92	2.60	2.96	3.29	3.50	3.49	3.62	3.94	5.26	4.39
	(2.15)	(4.09)	(5.37)	(5.43)	(6.31)	(6.08)	(6.14)	(6.39)	(6.53)	(7.95)	(8.64)
α_{FF5}	0.67	1.69	2.33	2.65	2.99	3.13	3.15	3.31	3.52	4.99	4.31
	(2.78)	(6.73)	(8.78)	(10.03)	(11.90)	(10.56)	(9.96)	(11.18)	(10.43)	(11.94)	(10.18)
Panel B: Intraday returns											
Excess	-0.28	-1.52	-1.67	-2.62	-2.26	-2.76	-2.78	-3.03	-3.35	-5.60	-5.32
	(-0.84)	(-4.62)	(-5.15)	(-7.53)	(-5.84)	(-8.07)	(-7.26)	(-8.90)	(-10.48)	(-12.91)	(-10.41)
α_{FF5}	-0.61	-1.84	-1.94	-2.93	-2.43	-2.98	-2.92	-3.16	-3.52	-5.68	-5.07
	(-3.00)	(-7.38)	(-7.25)	(-8.30)	(-7.17)	(-10.32)	(-7.27)	(-9.42)	(-10.36)	(-14.11)	(-11.03)
Panel C: Overnight-intraday return gap											
Excess	1.77	4.36	5.45	7.03	7.32	8.05	8.36	8.83	9.67	14.72	12.95
	(3.91)	(7.89)	(9.29)	(8.92)	(10.04)	(10.81)	(10.20)	(11.38)	(10.99)	(15.22)	(14.14)
α_{FF5}	1.97	4.53	5.52	7.12	7.29	7.95	8.21	8.71	9.46	14.60	12.63
	(4.37)	(9.04)	(9.90)	(10.70)	(11.78)	(12.80)	(11.59)	(12.59)	(13.05)	(16.64)	(14.06)

Table 6. Return gaps of RTP-sorted portfolios controlling for size and volatility

This table presents average monthly return gaps of portfolios sorted by retail trading proportion (RTP) conditional on market equity (Panel A) or past return volatility (Panel B). Each month, stocks are first sorted into quintiles based on the control variable, then sorted into quintiles based on RTP within each control quintile. H–L is a long-short portfolio that buys the highest RTP quintile and sells the lowest RTP quintile within each control group. Newey–West t-statistics are reported in parentheses for H–L portfolios. Portfolios are value-weighted and rebalanced monthly. The sample period is January 2009 to December 2023.

	Low RTP	2	3	4	High RTP	H-L	t-statistic
Panel A: Double sort on RTP and market cap							
Small	2.37	4.74	5.88	7.27	9.35	6.98	(9.99)
P2	2.06	4.70	5.91	7.59	12.34	10.28	(12.40)
P3	1.12	4.25	5.90	7.56	12.99	11.87	(12.48)
P4	-0.43	2.61	5.03	7.71	11.47	11.90	(17.44)
Big	1.36	2.85	4.10	6.46	9.69	8.33	(10.86)
Panel B: Double sort on RTP and return volatility							
Low Vol	0.73	2.39	2.00	2.83	3.37	2.64	(4.69)
P2	2.01	3.94	4.18	5.14	6.02	4.01	(7.53)
P3	2.65	5.54	5.99	6.64	8.18	5.53	(7.15)
P4	4.54	8.08	9.37	8.89	10.77	6.23	(8.99)
High Vol	7.70	11.17	12.89	14.33	19.48	11.78	(12.07)

Table 7. Fama–MacBeth regressions

This table presents results from Fama–MacBeth regressions of one-month-ahead returns on retail trading proportion (RTP). The dependent variables are one-month-ahead overnight returns in Columns (1)–(3), one-month-ahead intraday returns in Columns (4)–(6), and one-month-ahead return gaps in Columns (7)–(9). Returns are expressed in percentage points. The independent variable of interest is RTP. Detailed variable definitions are provided in Appendix Table A1. All independent variables are winsorized at the 1% and 99% levels and standardized cross-sectionally to have zero mean and unit standard deviation. Newey–West standard errors are reported in parentheses, computed with six lags. Industry fixed effects include 51 KRX industry classifications. The sample period is January 2009 to December 2023. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Overnight returns			Intraday returns			Return gap		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTP	1.24***	1.05***	1.00***	-1.22***	-1.26***	-1.22***	2.97***	2.43***	2.35***
	(0.13)	(0.10)	(0.09)	(0.11)	(0.11)	(0.11)	(0.23)	(0.18)	(0.18)
$r^{\text{overnight}}$		1.16***	1.14***		-1.15***	-1.21***		2.51***	2.55***
		(0.08)	(0.08)		(0.08)	(0.08)		(0.17)	(0.18)
r^{intraday}		-0.19***	-0.24***		0.17**	0.07		-0.26**	-0.20*
		(0.05)	(0.04)		(0.08)	(0.08)		(0.12)	(0.11)
Size		0.33***	0.42***		-0.71***	-0.76***		0.99***	1.15***
		(0.05)	(0.05)		(0.11)	(0.11)		(0.11)	(0.13)
beta		0.34***	0.30***		-0.23***	-0.22***		0.60***	0.54***
		(0.06)	(0.07)		(0.08)	(0.08)		(0.13)	(0.13)
BM		-0.25***	-0.16***		0.25***	0.30***		-0.64***	-0.60***
		(0.04)	(0.04)		(0.08)	(0.07)		(0.08)	(0.09)
ILLIQ		0.04	0.03		0.17***	0.14**		0.07	0.08
		(0.04)	(0.04)		(0.06)	(0.06)		(0.10)	(0.11)
Fown		0.11***	0.08**		-0.08**	-0.11***		0.14***	0.15***
		(0.03)	(0.03)		(0.03)	(0.03)		(0.05)	(0.05)
MOM		0.21***	0.17***		-0.02	-0.03		0.39***	0.38***
		(0.04)	(0.03)		(0.09)	(0.08)		(0.11)	(0.10)
IVOL		0.43***	0.47***		-1.09***	-1.05***		2.31***	2.29***
		(0.07)	(0.07)		(0.13)	(0.12)		(0.19)	(0.20)
MAX		0.04	0.02		0.29***	0.32***		-0.34**	-0.41**
		(0.05)	(0.06)		(0.09)	(0.10)		(0.14)	(0.16)
Observations	169,961	169,961	159,472	169,961	169,961	159,472	169,961	169,961	159,472
Industry FE			✓			✓			✓
Adjusted R^2 (%)	4.37	12.63	15.69	2.06	9.38	12.81	4.11	13.74	16.02

Table 8. IV regressions with nominal price as instrument

This table presents results from panel OLS and two-stage least squares (2SLS) regressions of one-month-ahead returns on retail trading proportion (RTP). Column (1) reports first-stage results using the log of the nominal closing price on the first trading day of month t as an instrument for RTP. KP F-stat is the Kleibergen–Paap rank Wald F-statistic testing for weak identification. Columns (2)–(7) report second-stage and OLS results with one-month-ahead overnight returns, intraday returns, and return gaps as dependent variables, expressed in percentage points. Control variables are the same as those in Table 7. All independent variables are winsorized at the 1% and 99% levels and standardized cross-sectionally to have zero mean and unit standard deviation. Standard errors are reported in parentheses, double-clustered by firm and month. Industry fixed effects include 51 KRX industry classifications. The sample period is January 2009 to December 2023. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	RTP	Overnight returns		Intraday returns		Return gap	
	1st stage	OLS	2SLS	OLS	2SLS	OLS	2SLS
RTP		0.99***	2.28***	-1.24***	-1.90***	2.36***	4.51***
		(0.08)	(0.24)	(0.11)	(0.30)	(0.16)	(0.45)
Nominal price	-0.19***						
	(0.01)						
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	159,472	159,472	159,472	159,472	159,472	159,472	159,472
Month FE	✓	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓	✓	✓	✓
KP F-stat	318.43***						
Adjusted R^2 (%)		24.54	4.95	12.00	3.11	15.79	7.54

Table 9. Fama–MacBeth regressions: RTP predicting net buy flows

This table presents results from Fama–MacBeth regressions of future net buy flows on retail trading proportion (RTP) using daily stock-level observations. The dependent variables are investor net buy flows (buy minus sell volume, scaled by shares outstanding) measured in basis points. Columns (1)–(3) measure net buy using trades executed between 9:00 and 10:00 (market open), and Columns (4)–(6) use trades executed after 14:00 (market close). Column headers indicate the investor group: "Retail trader" includes all retail investors, "Mobile trader" includes retail investors using mobile trading systems (MTS), and "Day trader" includes round-trip positions opened and closed (completely or partially) within the same day. Retail trader and Mobile trader flows are based on KRX data; Day trader flows are based on brokerage data. Panels A, B, and C report results when the dependent variable is the average net buy flow over the next 1, 5, and 20 days, respectively, following day t. RTP is measured on day t. All independent variables are winsorized at the 1% and 99% levels and standardized cross-sectionally to have zero mean and unit standard deviation. Coefficients are scaled by 100 for readability; a coefficient of 9.75 indicates that a one-standard-deviation increase in RTP is associated with a 0.0975 basis point increase in net buy flows. Newey–West standard errors are reported in parentheses, computed with 20 lags. The sample period is January 2009 to December 2023 for retail and mobile trades, and January 2015 to December 2018 for day trades. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Net buy at open			Net buy at close		
	(1) Retail trader	(2) Mobile trader	(3) Day trader	(4) Retail trader	(5) Mobile trader	(6) Day trader
Panel A: Predicting flows in next day						
RTP	9.75***	25.39***	3.49***	12.02***	-16.23***	-2.45***
	(1.51)	(1.74)	(0.32)	(1.18)	(0.20)	(0.19)
Controls	✓	✓	✓	✓	✓	✓
Observations	3,399,550	3,399,550	861,979	3,399,550	3,399,550	861,979
Adjusted R^2 (%)	4.20	4.61	3.59	2.61	3.65	3.58
Panel B: Predicting average flows in next 5 days						
RTP	1.70**	2.85***	3.53***	-2.02***	-3.11***	-2.48***
	(0.83)	(1.09)	(0.31)	(0.68)	(0.77)	(0.20)
Controls	✓	✓	✓	✓	✓	✓
Observations	3,399,550	3,399,550	861,979	3,399,550	3,399,550	861,979
Adjusted R^2 (%)	-0.00	0.02	6.11	0.04	0.04	6.94
Panel C: Predicting average flows in next 20 days						
RTP	2.08***	3.27***	3.41***	-2.21***	-3.38***	-2.40***
	(0.75)	(0.96)	(0.34)	(0.61)	(0.72)	(0.19)
Controls	✓	✓	✓	✓	✓	✓
Observations	3,399,550	3,399,550	861,979	3,399,550	3,399,550	861,979
Adjusted R^2 (%)	-0.01	0.02	9.14	0.13	0.14	11.01

Table 10. Day trading proportion and trader characteristics

This table presents results from cross-sectional OLS regressions of day trading proportion on trader characteristics. The dependent variable is the percentage of total trading volume (in KRW) attributable to positions opened and closed (completely or partially) within the same trading day. Male is an indicator for male investors. Young is an indicator for investors aged 35 or below. High invest. know. is an indicator for investors who self-report the highest level of investment knowledge (4 on a scale of 1 to 4). MTS and HTS are indicators for investors whose main trading medium is mobile or desktop, respectively. Standard errors are reported in parentheses. The sample covers approximately 235,000 active retail accounts from January 2015 to December 2018. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Day trade proportion (%)			
	(1)	(2)	(3)	(4)
Male	4.03***		2.94***	2.17***
	(0.09)		(0.09)	(0.11)
Young	3.94***		3.68***	1.23***
	(0.12)		(0.12)	(0.20)
High invest. know.	5.54***		5.19***	5.03***
	(0.14)		(0.14)	(0.25)
MTS		6.80***	6.06***	6.11***
		(0.11)	(0.11)	(0.11)
HTS		13.80***	13.05***	13.10***
		(0.13)	(0.13)	(0.13)
Male × Young				4.08***
				(0.25)
Male × High invest. know.				0.33
				(0.31)
Young × High invest. know.				-1.86***
				(0.62)
Male × Young × High invest. know.				2.21***
				(0.76)
Intercept	13.25***	10.31***	7.85***	8.27***
	(0.08)	(0.09)	(0.10)	(0.11)
Observations	235,445	235,445	235,445	235,445
Adjusted R^2 (%)	1.99	4.63	6.06	6.20

Table 11. Robustness to alternative mechanisms

This table presents results from Fama–MacBeth regressions of one-month-ahead returns on retail trading proportion (RTP) and other variables proposed in prior studies to explain overnight and intraday returns. Panel A reports one-month-ahead overnight returns and Panel B reports one-month-ahead intraday returns as dependent variables. Returns are expressed in percentage points. Control variables are the same as those in Table 7. Detailed definitions of all variables are provided in Appendix Table A1. All independent variables are winsorized at the 1% and 99% levels and standardized cross-sectionally to have zero mean and unit standard deviation. Newey–West standard errors are reported in parentheses, computed with six lags. The sample period is January 2009 to December 2023. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Overnight returns						
RTP	1.05***	1.03***	0.90***	0.97***	0.96***	0.94***
	(0.10)	(0.10)	(0.09)	(0.09)	(0.09)	(0.09)
IVOL	0.43***	0.46***	0.43***	0.40***	0.39***	0.37***
	(0.07)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)
Abn_Vol		-0.10*	-0.08	-0.09*	-0.09*	-0.09*
		(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
$\sigma^2_{\text{sys, overnight}}$			0.40***	0.39***	0.25	0.25
			(0.07)	(0.07)	(0.16)	(0.16)
rOIB_open				0.25***	0.25***	0.25***
				(0.02)	(0.02)	(0.02)
$\beta_{\text{overnight}}$					0.17	0.17
					(0.16)	(0.16)
Abn_NRev						-0.11***
						(0.02)
Controls	✓	✓	✓	✓	✓	✓
Observations	169,961	167,085	163,985	163,985	163,985	163,985
Adjusted R^2 (%)	12.63	12.94	13.40	13.53	13.74	13.78
Panel B: Intraday returns						
RTP	-1.26***	-1.25***	-1.08***	-1.09***	-1.08***	-1.07***
	(0.11)	(0.11)	(0.10)	(0.09)	(0.09)	(0.09)
IVOL	-1.09***	-1.08***	-1.00***	-1.00***	-1.00***	-0.97***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Abn_Vol		0.03	-0.02	-0.02	-0.02	-0.02
		(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
$\sigma^2_{\text{sys, overnight}}$			-0.55***	-0.54***	-0.76***	-0.76***
			(0.10)	(0.10)	(0.15)	(0.15)
rOIB_open				0.01	0.01	0.01
				(0.04)	(0.03)	(0.03)
$\beta_{\text{overnight}}$					0.24*	0.25*
					(0.14)	(0.14)
Abn_NRev						0.10**
						(0.04)
Controls	✓	✓	✓	✓	✓	✓
Observations	169,961	167,085	163,985	163,985	163,985	163,985
Adjusted R^2 (%)	9.38	9.62	9.95	10.00	10.07	10.10

Table 12. Heterogeneous effects of RTP across subsamples

This table presents results from Fama–MacBeth regressions testing whether RTP's predictive power varies with proxies for alternative mechanisms. Panel A reports one-month-ahead overnight returns and Panel B reports one-month-ahead intraday returns as dependent variables. Returns are expressed in percentage points. The parameter of interest is the interaction term between RTP and a high indicator for each mechanism variable, where the high indicator equals one when the mechanism variable exceeds its cross-sectional median and zero otherwise. Control variables are the same as those in Table 7. Detailed variable definitions are provided in Appendix Table A1. All independent variables are winsorized at the 1% and 99% levels and standardized cross-sectionally to have zero mean and unit standard deviation. Newey–West standard errors are reported in parentheses, computed with six lags. The sample period is January 2009 to December 2023. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Overnight returns							
RTP	1.05***	1.02***	0.98***	0.94***	0.88***	0.88***	0.90***
	(0.10)	(0.10)	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)
RTP × High IVOL		0.12***	0.13**	0.16***	0.15***	0.15***	0.15***
		(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
RTP × High Abn_Vol			0.02	-0.00	-0.04	-0.04	-0.05
			(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
RTP × High $\sigma^2_{\text{sys, overnight}}$				-0.10*	-0.10*	0.73	0.54
				(0.05)	(0.05)	(1.39)	(1.41)
RTP × High rOIB_open					0.21***	0.22***	0.22***
					(0.06)	(0.07)	(0.07)
RTP × High $\beta^{\text{overnight}}$						-0.84	-0.64
						(1.38)	(1.40)
RTP × High Abn_NRev							-0.05
							(0.04)
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	169,961	169,961	167,085	163,985	163,985	163,985	163,985
Adjusted R^2 (%)	12.63	12.74	12.86	13.04	13.19	13.19	13.22
Panel B: Intraday returns							
RTP	-1.26***	-1.20***	-1.07***	-0.98***	-0.81***	-0.82***	-0.83***
	(0.11)	(0.12)	(0.12)	(0.10)	(0.11)	(0.11)	(0.11)
RTP × High IVOL		-0.43***	-0.41***	-0.40***	-0.40***	-0.40***	-0.40***
		(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
RTP × High Abn_Vol			-0.07	-0.06	-0.02	-0.02	-0.01
			(0.07)	(0.06)	(0.07)	(0.07)	(0.07)
RTP × High $\sigma^2_{\text{sys, overnight}}$				-0.05	-0.06	-2.29	-1.90
				(0.09)	(0.09)	(1.44)	(1.37)
RTP × High rOIB_open					-0.27***	-0.27***	-0.28***
					(0.07)	(0.07)	(0.08)
RTP × High $\beta^{\text{overnight}}$						2.24	1.85
						(1.45)	(1.37)
RTP × High Abn_NRev							0.04
							(0.05)
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	169,961	169,961	167,085	163,985	163,985	163,985	163,985
Adjusted R^2 (%)	9.38	9.49	9.64	9.76	9.84	9.77	9.75

A Variable definitions

Table A1. Variable definitions

This table defines all variables used in the analysis. Variables are computed using KRX data (January 2009 to December 2023) unless otherwise noted. Brokerage variables are from a large Korean discount broker (January 2015 to December 2018).

Variable	Definition
Panel A: Return variables	
Overnight return	Close-to-open return. For day s, it is computed as $r_{i,s}^{\text{overnight}} = (1 + r_{i,s})/(1 + r_{i,s}^{\text{intraday}}) - 1$, where $r_{i,s}$ is the close-to-close return and $r_{i,s}^{\text{intraday}}$ is the intraday return. The monthly level overnight return is computed as $r_{i,t}^{\text{overnight}} = \prod_{s \in \text{month } t}(1 + r_{i,s}^{\text{overnight}}) - 1$
Intraday return	Open-to-close return. For day s, it is computed as $r_{i,s}^{\text{intraday}} = P_{i,s}^{\text{close}}/P_{i,s}^{\text{open}} - 1$, where $P_{i,s}^{\text{open}}$ and $P_{i,s}^{\text{close}}$ are the opening and closing auction prices. The monthly level intraday return is computed as $r_{i,t}^{\text{intraday}} = \prod_{s \in \text{month } t}(1 + r_{i,s}^{\text{intraday}}) - 1$
Return gap	Difference between overnight and intraday returns. For day s, it is computed as $\text{RG}_{i,s} = r_{i,s}^{\text{overnight}} - r_{i,s}^{\text{intraday}}$. The monthly level return gap is computed as $\text{RG}_{i,t} = (1 + r_{i,t}^{\text{overnight}})/(1 + r_{i,t}^{\text{intraday}}) - 1$
Close-to-close return	Daily return from close to close: $r_{i,t} = P_{i,t}^{\text{close}}/P_{i,t-1}^{\text{close}} - 1$
Panel B: Trading activity variables	
RTP	Retail trading proportion, computed as the ratio of retail trading volume to total trading volume over a specified period (typically a day or a month)
Net buy	Buy volume minus sell volume, normalized by shares outstanding (in basis points)
Panel C: Stock characteristics	
Size	Natural logarithm of market capitalization (stock price × shares outstanding), measured at the beginning of the month
BM	Book equity divided by market equity, measured at the beginning of the month
beta	Estimated using past 120 trading days of returns, with the KOSPI index as the market portfolio
ILLIQ	Average of $\lvert\text{ret}_{i,t}\rvert/\text{volume}_{i,t}$ over the past 120 trading days, where volume is measured in KRW billions
Fown	Percentage of shares outstanding held by foreign investors, measured at the beginning of the month
MOM	Cumulative return over the past 120 trading days
IVOL	Standard deviation of residuals in a regression of individual stock return on Fama-French three factors in Korean market, using daily data in current month
MAX	Maximum daily return over the trading days in current month

Table A1. Variable definitions (continued)

Variable	Definition
Panel D: Alternative mechanism variables	
Abn_Vol	Abnormal turnover: difference between month *t* turnover and average turnover over the prior 12 months
$\sigma^2_{\text{sys, overnight}}$	Overnight systematic variance: squared overnight beta times market overnight return variance, where overnight beta is estimated using a rolling 12-month window with at least 100 daily observations
rOIB_open	Retail order imbalance at open: retail buy-minus-sell volume during the first trading hour, scaled by daily volume, aggregated monthly
$\beta^{\text{overnight}}$	Overnight market beta: estimated from rolling regression of overnight stock returns on overnight market returns over 12 months with at least 30 observations
Abn_NRev	Abnormal negative reversals: current-month proportion of days with positive overnight and negative intraday returns minus the 12-month average
Panel E: Brokerage account variables	
Male	Binary indicator equal to 1 if the account holder is male
Age	Account holder's age in years as of December 31, 2018
Average daily balance	Mean daily account balance over the sample period (in KRW millions)
Average trade size	Mean trade size over the sample period (in KRW millions)
Number of unique stocks	Total number of unique stocks traded during the sample period
Number of trades	Total number of trades executed during the sample period
Performance	Realized return over the sample period, winsorized at 1% and 99%
Main trading medium	Trading interface (MTS, HTS, or Other) through which the trader executed the most trades
Panel F: Brokerage questionnaire variables	
Monthly income	Self-reported monthly income: 1 = less than KRW 1M, 2 = KRW 1–3M, 3 = KRW 3–6M, 4 = KRW 6–10M, 5 = more than KRW 10M
Investment knowledge	Self-reported investment knowledge: 1 = little understanding, 2 = partial understanding, 3 = deep understanding, 4 = understanding of most products including derivatives
Investment aggressiveness	Self-reported risky product experience: 1 = none, 2 = government bonds/MMFs, 3 = bond funds/high-credit bonds, 4 = mixed funds/medium-credit bonds, 5 = stocks/stock funds, 6 = derivatives/leveraged products
Derivative experience	Self-reported derivative trading experience: 1 = less than 1 year, 2 = 1–3 years, 3 = more than 3 years

B Additional figures and tables

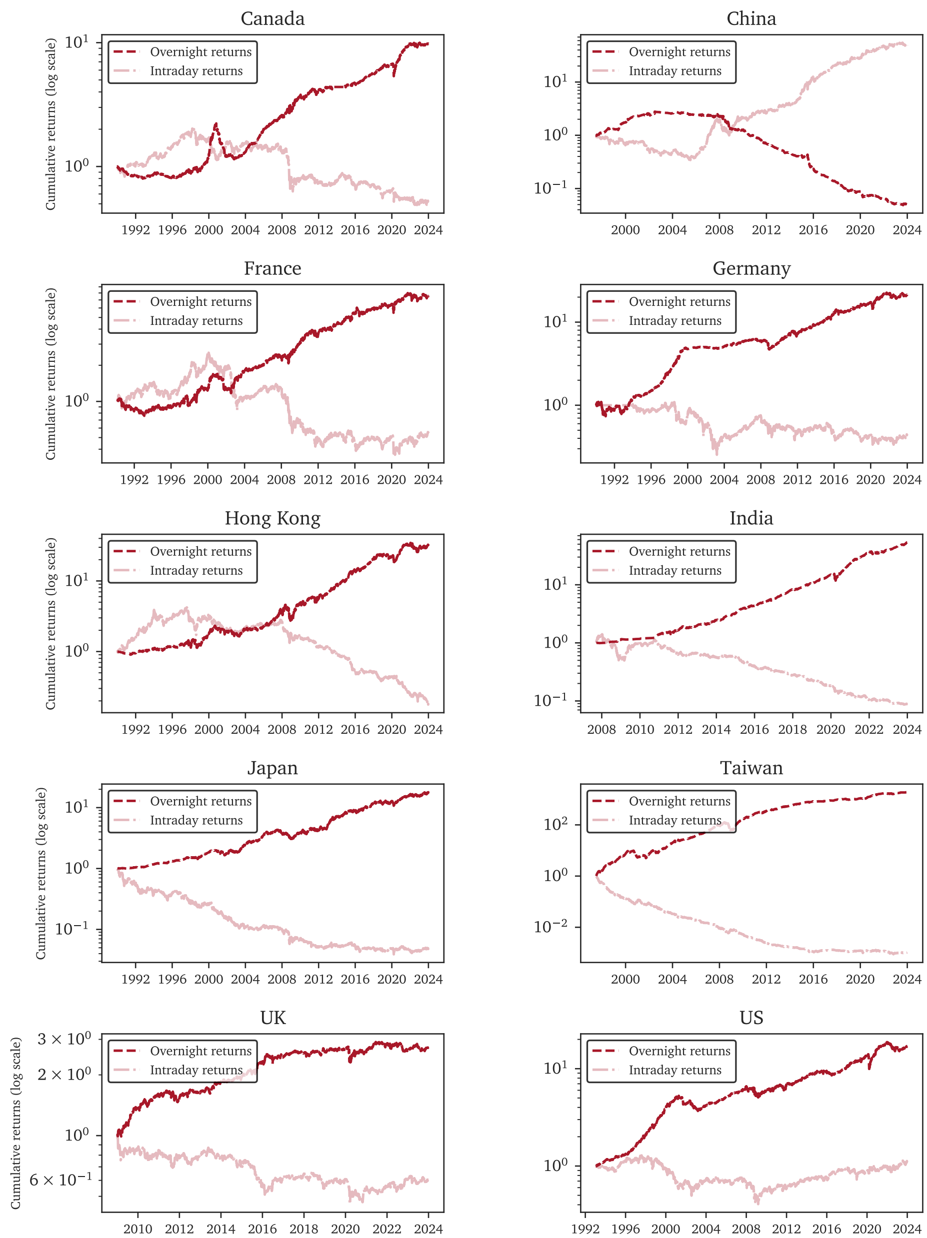


Figure A1. Cumulative overnight and intraday returns

This figure plots cumulative overnight and intraday returns for major equity market indices from January 2010 to December 2023 on a log scale. For the U.S. and U.K., we use liquid exchange-traded funds (ETFs) to compute intraday returns; for other markets, we use the underlying index tickers directly, as corresponding ETFs were either introduced later or have insufficient liquidity.

Table A2. Fama–MacBeth regressions using VWAP-based returns

This table presents results from Fama–MacBeth regressions using returns computed from volume-weighted average prices (VWAPs). The dependent variables are one-month-ahead overnight returns in Columns (1)–(3), one-month-ahead intraday returns in Columns (4)–(6), and one-month-ahead return gaps in Columns (7)–(9). Overnight and intraday returns are computed using VWAPs during the first and last 30 minutes of each trading day (market open and close), respectively. Returns are expressed in percentage points. The key independent variable is RTP. Detailed variable definitions are provided in Appendix Table A1. All independent variables are winsorized at the 1% and 99% levels and standardized cross-sectionally to have zero mean and unit standard deviation. Newey–West standard errors are reported in parentheses, computed with six lags. Industry fixed effects include 51 KRX industry classifications. The sample period is January 2009 to December 2023. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Overnight returns			Intraday returns			Return gap		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
RTP	2.22***	2.02***	2.00***	-2.14***	-2.20***	-2.21***	4.93***	4.36***	4.33***
	(0.15)	(0.11)	(0.10)	(0.12)	(0.11)	(0.11)	(0.28)	(0.19)	(0.19)
$r^{\text{overnight}}$		9.30***	8.97***		-8.90***	-9.26***		21.23***	21.47***
		(0.85)	(0.78)		(0.83)	(0.87)		(1.74)	(1.87)
r^{intraday}		-7.65***	-8.20***		7.23***	6.48***		-15.73***	-15.36***
		(0.67)	(0.60)		(0.68)	(0.62)		(1.21)	(1.19)
Size		0.52***	0.69***		-0.89***	-1.01***		1.32***	1.64***
		(0.09)	(0.08)		(0.08)	(0.09)		(0.14)	(0.13)
beta		0.54***	0.49***		-0.41***	-0.39***		1.01***	0.93***
		(0.07)	(0.07)		(0.06)	(0.06)		(0.11)	(0.12)
BM		-0.23***	-0.08		0.23***	0.21***		-0.56***	-0.39***
		(0.07)	(0.06)		(0.07)	(0.06)		(0.10)	(0.09)
ILLIQ		0.04	0.02		0.11**	0.08		-0.01	-0.01
		(0.04)	(0.04)		(0.05)	(0.05)		(0.09)	(0.09)
Fown		0.14***	0.10**		-0.10***	-0.13***		0.20***	0.19***
		(0.04)	(0.04)		(0.04)	(0.03)		(0.07)	(0.06)
MOM		0.21***	0.16**		-0.05	-0.05		0.34**	0.29**
		(0.07)	(0.06)		(0.09)	(0.09)		(0.15)	(0.15)
IVOL		-0.01	0.01		-0.60***	-0.58***		1.22***	1.21***
		(0.08)	(0.08)		(0.09)	(0.09)		(0.16)	(0.17)
MAX		0.37***	0.40***		-0.07	-0.05		0.45***	0.45***
		(0.06)	(0.06)		(0.08)	(0.09)		(0.14)	(0.15)
Observations	169,961	169,961	159,472	169,961	169,961	159,472	169,961	169,961	159,472
Industry FE	No	No	Yes	No	No	Yes	No	No	Yes
Adjusted R^2 (%)	6.43	13.30	17.29	5.84	13.27	16.38	9.08	17.86	20.50

Table A3. Panel regressions: RTP predicting net buy flows

This table presents results from panel regressions of future net buy flows on retail trading proportion (RTP) using daily stock-level observations. The dependent variables are investor net buy flows (buy minus sell volume, scaled by shares outstanding) measured in basis points. Columns (1)–(3) measure net buy using trades executed between 9:00 and 10:00 (market open), and Columns (4)–(6) use trades executed after 14:00 (market close). Column headers indicate the investor group: Retail trader" includes all retail investors, Mobile trader" includes retail investors using mobile devices, and "Day trader" includes round-trip positions opened and closed (completely or partially) within the same day, all based on KRX trade-level data except day trades, which come from a large Korean discount brokerage. Panels A, B, and C report results when the dependent variable is the average net buy flow over the next 1, 5, and 20 days, respectively, following day t. RTP is measured on day t. Control variables are the same as those in Table 7. All independent variables are winsorized at the 1% and 99% levels and standardized cross-sectionally to have zero mean and unit standard deviation. Coefficients are expressed in percentage points. Standard errors are reported in parentheses, double-clustered by firm and day. Industry fixed effects include 51 KRX industry classifications. The sample period is January 2009 to December 2023 for retail and mobile trades, and January 2015 to December 2018 for day trades. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Net buy at open			Net buy at close		
	(1) Retail trader	(2) Mobile trader	(3) Day trader	(4) Retail trader	(5) Mobile trader	(6) Day trader
Panel A: Predicting flows in next day						
RTP	14.39***	35.21***	3.78***	7.08***	-20.87***	-2.59***
	(1.33)	(1.90)	(0.35)	(1.44)	(1.57)	(0.29)
Controls	✓	✓	✓	✓	✓	✓
Observations	3,189,957	3,189,957	818,003	3,189,957	3,189,957	818,003
Industry FE	✓	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓	✓
Adjusted R^2 (%)	2.71	2.96	1.79	1.70	2.48	1.92
Panel B: Predicting average flows in next 5 days						
RTP	1.59	3.25**	3.67***	-3.00***	-4.10***	-2.57***
	(1.11)	(1.53)	(0.35)	(1.05)	(1.31)	(0.28)
Controls	✓	✓	✓	✓	✓	✓
Observations	3,189,957	3,189,957	818,003	3,189,957	3,189,957	818,003
Industry FE	✓	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓	✓
Adjusted R^2 (%)	0.10	0.11	4.77	0.06	0.12	5.75
Panel C: Predicting average flows in next 20 days						
RTP	2.16**	3.81***	3.54***	-3.04***	-4.28***	-2.47***
	(1.09)	(1.47)	(0.33)	(1.02)	(1.28)	(0.27)
Controls	✓	✓	✓	✓	✓	✓
Observations	3,189,957	3,189,957	818,003	3,189,957	3,189,957	818,003
Firm FE	✓	✓	✓	✓	✓	✓
Day FE	✓	✓	✓	✓	✓	✓
Adjusted R^2 (%)	0.30	0.31	8.95	0.17	0.30	10.45