**Submitted to *Management Science***
**manuscript**

**Authors are encouraged to submit new papers to INFORMS journals by means of a style file template, which includes the journal title. However, use of a template does not certify that the paper has been accepted for publication in the named journal. INFORMS journal templates are for the exclusive purpose of submitting to an INFORMS journal and should not be used to distribute the papers in print or online or to submit the papers to another publication.**

# INVARIANCE OF BUY-SELL SWITCHING POINTS

Kyoung-hun Bae
Hanyang University, South Korea, khbae@hanyang.ac.kr

Albert S. Kyle
Corresponding author. Robert H. Smith School of Business, University of Maryland, College Park, MD, USA, akyle@rhsmith.umd.edu

Eun Jung Lee
Hanyang University, South Korea, ejunglee@hanyang.ac.kr

Anna A. Obizhaeva
New Economic School, Moscow, Russia, aobizhaeva@nes.ru

Market microstructure invariance predicts business time to unfold at a rate proportional to the 2/3 power of the product of dollar volume and returns volatility. Define a "switching point" as an investor changing the direction of trading from buying to selling or selling to buying. For a specific market, the aggregate number of switching points is a good indicator of the pace of business time. Using data from the Korea Exchange (KRX) from 2008 to 2010, we calculate the number of switching points for each stock for each month. The estimated exponent is 0.675 (standard error 0.005, $R^2 = 0.93$) validates the business time clock predicted by invariance. Most variation reflects variation in the number of accounts trading a stock, not variation of switching points per account.

*Key words*: finance, market microstructure, asset pricing, invariance, trading volume, volatility, liquidity, price impact, market depth

*History*: First draft: June 26, 2014. This draft: June 19, 2019. Albert S. Kyle has worked as a consultant for various companies, exchanges, and government agencies. He is a non-executive director of a U.S.-based asset management company.

---

We use buy-sell switching points, measured from audit trail data of South Korean stocks, to validate the business time clock proposed by the hypothesis of market microstructure invariance.

Business time measures the rate at which traders incur costs from trading. Costs include both direct trading costs associated with the market impact of executing large bets and information costs associated with research and decision-making. It is intuitive that business time passes faster in stocks with higher trading activity and liquidity.

Within the paradigm of market microstructure invariance, trading stocks is the same game played at different speeds in different markets. It can be proved that business time passes at a rate proportional to the 2/3 power of the product of dollar volume and return volatility. Since both dollar volume and return volatility are easily measured, the invariance hypothesis implies that the business time clock can be easily calibrated from public information.

We measure business time as the frequency with which traders change the direction of their trading from buying to selling or from selling to buying. We call such changes “switching points.” Switching direction of trading is related to both trading costs and information costs. When new information flows more rapidly, new trading ideas that reverse old ideas arrive more quickly. When trading costs are lower, it is less costly to buy and then sell. Since the data set has unique masked trader IDs, counting switching points is straightforward. Market frictions such as lot size and tick size are expected to have a small effect on this measure. Switching points are thus a direct measure of the rate at which business time passes.

Other direct measures of business time are challenging to implement empirically because it is difficult to observe bets, trading costs, information costs directly. First, bets are often broken into pieces and executed gradually over time. Order shredding algorithms are adjusted to incorporate the effects of market frictions. It is difficult to infer when execution of one bet stops and execution of a new bet starts. Second, typical data bases usually do not have traders IDs; it is hard to disentangle multiple bets contemporaneously executed by different traders. Third, even when bets can be identified, measuring market impact costs is hard. The preferred method, implementation shortfall, compares the execution price with the price just before the order was placed. This requires a large number of observations to obtain accurate measures of trading costs. The quantities traded are subjects to selection biases, since cheap orders are executed fully and expensive bets are often cancelled. Fourth, it is also difficult to measure dollar costs of information because firms keep such information confidential.

As the intuition of microstructure invariance suggests, the number of switching points accumulates at a rate proportional to the 2/3 power of the product of dollar volume and return volatility. Following Kyle and Obizhaeva (2017), we provide in section 1 a simple derivation of the 2/3 exponent using dimensional analysis, leverage neutrality, and the invariance assumption that information and trading costs are the same across markets. Kyle and Obizhaeva (2016) derive the same hypothesis using ad hoc invariance conjectures. Kyle and Obizhaeva (2019) derive the same invariance hypothesis from a dynamic model of costly information acquisition, in which it is shown that trading costs and information costs are proportional. Kyle and Obizhaeva (2018) describe how invariance is related to adverse selection.

Markets are more liquid when business time pasts faster. When market impact costs are lower, traders take larger positions based on a given flow of costly information. Therefore, business time passes at a rate less than proportional to trading volume. In more liquid markets, traders also hold the positions for a shorter period of time, changing their direction of trading more frequently. This intuition suggests that switching points increase with dollar trading volume but less than proportionally.

Using account-level data from the Korea Exchange from 2008 to 2010, we examine this scaling law by presenting results of OLS and negative-binomial regressions of the log of the number of switching points on log-volume and log-volatility, testing whether the slope coefficient is close to predicted 2/3. We find the estimated log-exponent to be 0.675 with standard error of 0.005. The number of switching points exhibit significant variation in the data and varies by more than a factor of 150 across securities, but patterns of this variation are consistent with the predictions of invariance.

When only a log-intercept is estimated, invariance explains about 93% of the entire variation in the number of switching points across months and stocks. The fraction of explained variation increases to 98% by adding 36 month fixed effects, 686 stock fixed effects, and eight additional variables such as volume, volatility, price levels, fraction of retail investors, and dummies indicating inclusion into several indices. At a minimum, these results imply that invariance is a useful benchmark for investigating how microstructure details affect the business time clock.

Unlike trading in the U.S. market, which is dominated by institutional investors, trading in the South Korean market is dominated by retail investors. Even though the trading of retail investors is usually believed to be subject to various behavioral biases, the exponent of 2/3 is particularly prominent for the subset of Korean retail investors. Except for some government agencies, which are believed to follow arbitrage strategies, both domestic and foreign institutions also trade in a manner consistent with invariance. Most variation in the number of switching points is due to the variation in the number of unique accounts trading stocks, whereas the average number of switching points per account does not vary much.

The 2/3 exponent is not a regularity with an explanation based on a mechanical interdependence among variables. Although the results in this paper are so precise that they look like a law of physics and not a law of finance or economics, we do not expect a such near perfect fit to invariance predictions in all data sets. Instead, we conjecture that deviations from the predictions of invariance will occur. When they do occur, invariance provides a natural benchmark from which to interpret the deviations economically.

The empirical finance literature has long been interested in how to measure business time in financial markets. For example, Mandelbrot and Taylor (1967), Clark (1973), Tauchen and Pitts (1983), Harris (1987), Jones, Kaul and Lipson (1994), and Ané and Geman (2000) relate business time to the number of trades, trading volume, or returns variance. No consensus on a proper business clock emerged from this literature. Market microstructure invariance, with its 2/3 exponent, provides a simple, theoretically well-grounded, approach to measuring business time.

A number of power laws have been previously documented in the literature. For example, Gabaix et al. (2006) and Gabaix (2009) discuss relationships among exponents describing power laws in the tails of empirical probability distributions of financial variables like volume, trade size, and returns. The invariance relationships in this paper are quite different from these extensively studied empirical regularities. These are log-linear power-law relationships between variables applying to entire probability distributions generally, not just to their tails.

If substantial variation in financial data can indeed be better explained by understanding the appropriate market-specific business-time clock, this may have important implications for other areas of economics as well. For example, business cycles may be influenced by business-time clocks which operate on different scales in different countries and in different markets, such as the markets for consumers adjusting spending on homes and durable good; workers changing employment; manufacturing firms adjusting plant, equipment, and inventories; and banks raising new capital.

This paper is structured as follows. Section 1 develops the scaling law for cross-sectional and time-series variation in the number of switching points. Section 2 describes the data. Section 3 presents main results. Section 4 shows results for different groups of traders. Sections 5 and 6 provide more detailed analysis. Section 7 concludes.

## 1. Scaling Law for Switching Points

Let $S_{it}$ denote the aggregate number of switching points, summed across all accounts which traded stock $i$ during month $t$. Suppose the number of switching points $S_{it}$ is some unknown function of share price $P_{it}$, share volume $V_{it}$, returns volatility $\sigma_{it}$, and average dollar costs $C_{it}$ of executing bets:

$$S_{it} = f(P_{it}, V_{it}, \sigma_{it}, C_{it}). \tag{1}$$

Following the approach of Kyle and Obizhaeva (2017), the invariance-implied scaling exponent of $2/3$ can be derived from this equation in three steps by imposing restrictions of the function $f()$ implied by (1) dimensional analysis, (2) leverage neutrality, and (3) market microstructure invariance.

*First*, dimensional analysis reduces the dimensionality of the problem by enforcing consistency of units of measurement on the left- and right-hand sides of equation (1). Let brackets $[X]$ define an operator which gives the dimensions of a variable $X$. The quantities $P_{it}$, $V_{it}$, $\sigma^2_{it}$, $C_{it}$, and $S_{it}$ have the following dimensions:

$$[S_{it}] = \frac{1}{\text{time}}, \qquad [C_{it}] = \text{currency}, \qquad [P_{it}] = \frac{\text{currency}}{\text{shares}},$$
$$[\sigma^2_{it}] = \frac{1}{\text{time}}, \qquad [\sigma_{it}] = \frac{1}{\text{time}^{1/2}}, \qquad [V_{it}] = \frac{\text{shares}}{\text{time}}.$$

Return standard deviation $\sigma_{it}$ has dimension $1/\text{time}^{1/2}$, not $1/\text{time}$, because return variance $\sigma^2_{it}$ has dimension $1/\text{time}$.

Since $S_{it}$ and $\sigma^2_{it}$ have the same dimension ($1/\text{time}$) and there is exactly one more parameter (four) than the number of dimensions (three—currency, time, shares), the Buckhingham $\pi$ theorem says that $S_{it}$ may be expressed as the product of $\sigma^2_{it}$ and some function of exactly one dimensionless variable defined as the product of powers of the four arguments $P_{it}$, $V_{it}$, $\sigma^2_{it}$, and $C_{it}$. Leting $L_{it}$ denote this dimensionless variable, we have

$$S_{it} = f(P_{it}, V_{it}, \sigma_{it}, C_{it}) = \sigma^2_{it} \cdot g(L_{it}) = \text{const} \cdot \sigma^2_{it} \cdot L^{\beta}_{it}. \tag{2}$$

To simplify matters further, we assume (unnecessarily, see footnote 2) that $g$ is a power function defined by a scaling constant and an exponent $\beta$. Write the parameter $L_{it}$ as

$$L_{it} = \left( \frac{P_{it} \cdot V_{it}}{\sigma_{it}^2 \cdot C_{it}} \right)^{1/3}. \tag{3}$$

The parameter $L_{it}$ is dimensionless because the numerator $P_{it} \cdot V_{it}$ and the denominator $\sigma_{it}^2 \cdot C_{it}$ have the same dimensions, currency/time. Without loss of generality, we choose to add the exponent of 1/3; this allows $1/L_{it}$ to be interpreted as an asset-specific measure of illiquidity, which is proportional to percentage trading costs, as shown by Kyle and Obizhaeva (2017). In this way, dimensional analysis simplifies the problem from finding a function $f(\ldots)$ of four variables to finding a function $g(\ldots)$ of one variable.

*Second*, leverage neutrality—the market microstructure version of the Modigliani–Miller Theorem—imposes another restriction on equations (1) and (2). Leverage neutrality is based on the assumption that trading cash-equivalent assets is costless because it does not involve a transfer of risk. Thus, the economic cost of trading bundles made up of risky securities and positive or negative quantities of cash-equivalent assets must be the same as trading the unlevered risky securities themselves.

Suppose that shares of stock are levered up or down by combining each share of stock with cash worth $P_{it} \cdot (A-1)$ for some number $A$.[1] Consistent with the intuition of the Modigliani–Miller Theorem, changes in leverage lead to changes of the share price from $P_{it}$ to $P_{it} \cdot A$ and the return standard deviation from $\sigma_{it}$ to $\sigma_{it}/A$, thus leaving the dollar risks per share $P_{it} \cdot \sigma_{it}$ unchanged. Since trading a relevered share transfers the same dollar risk as trading an unlevered share, changing leverage does not lead to changes in the number of switching points $S_{it}$, trading volume $V_{it}$, or the dollar cost $C_{it}$, but changes liquidity from $L_{it}$ to $L_{it} \cdot A$.

$$\begin{array}{lll} P_{it} \to P_{it} \cdot A, & \sigma_{it} \to \sigma_{it} \cdot A^{-1}, & L_{it} \to L_{it} \cdot A, \\ S_{it} \to S_{it}, & V_{it} \to V_{it}, & C_{it} \to C_{it}. \end{array}$$

Since $S_{it}$ remains unchanged as a result of relevering by $A$, we must have $\beta = 2$ in equation (2):[2]

$$S_{it} = \text{const} \cdot \sigma_{it}^2 \cdot L_{it}^2. \tag{4}$$

The number of switching points is linearly proportional to the returns variance $\sigma_{it}^2$ and the square of asset-specific liquidity $L_{it}$.

[1] If $A > 1$, the asset is levered down by adding cash; if $A < 1$, the asset is levered up by adding debt (subtracting cash). For example, if $A = 2$, then each share is bundled with the same amount of cash, which doubles the price and makes a bundle less risky.

[2] The assumption that $g()$ is a power function is actually unnecessary. As a function of $A$, the function $g()$ satisfies $\sigma_{it}^2 \cdot g(L_{it}) = \sigma_{it}^2 \cdot A^{-2} \cdot g(L_{it} \cdot A^3)$; this implies $g()$ is homogeneous of degree 2/3 and therefore must be a power function with $\beta = 2$. To prove this, substitute $A = 1/L^3$ to obtain $\sigma_{it}^2 \cdot g(L_{it}) = \sigma_{it}^2 \cdot L^2 \cdot g(1)$.

Equation (4) may be written in an equivalent manner using the concept of trading activity instead of liquidity. Define trading activity $W_{it}$ as the product of dollar volume $P_{it} \cdot V_{it}$ and returns volatility $\sigma_{it}$:

$$W_{it} := P_{it} \cdot V_{it} \cdot \sigma_{it}. \tag{5}$$

The quantity $W_{it}$ measures aggregate risk transfer per time period. Then, equation (4) is equivalent to

$$S_{it} = \text{const} \cdot \left( \frac{W_{it}}{C_{it}} \right)^{2/3}. \tag{6}$$

Note that $W_{it}$ has dimension currency/time$^{3/2}$. Scaling by $C_{it}$ and raising to the power $2/3$ changes the dimension to $1/$time, exactly consistent with the dimension of switching points $S_{it}$.

Both invariance equations (4) and (6) have the parameter $C_{it}$, which measures the cost of executing a bet. In general, it is very difficult—perhaps impossible—to measure $C_{it}$ from public information.

*Third*, market microstructure invariance is an empirical hypothesis which simplifies empirical estimation of equation (6) by conjecturing that the average dollar costs $C_{it}$ of executing bets is approximately the same across assets and time. The invariance hypothesis is consistent with Occam's razor. Economically, it may be justified with an equilibrium model in which liquidity constrained traders trade on costly information in different markets and time periods, along the lines of Kyle and Obizhaeva (2019). The average dollar costs of trading are expected to be equalized among markets in the equilibrium. If $C_{it} = C$ is the same dollar quantity for all assets and time periods, then equation (6) has easily measurable variables and only one constant parameter to estimate.

Equation (6) says that the number of switching points $S_{it}$ is proportional to the $2/3$ power of trading activity $W_{it}$. For example, holding volatility constant, if trading volume increases by a factor of eight, the number of switching points is expected to increase by a factor of four $(= 8^{2/3})$. The number of bets and information flow, measured in similar units as $S_{it}$, satisfy similar scaling laws.

To provide an economic interpretation for the constant factor in equation (6), rescale the proportionality constant so that it is equal to the expected number of switching points per day for a hypothetical benchmark stock with a daily volume of one million shares, daily volatility of 2%, and price of 47,440 Korean won (KRW) per share.[3] With trading activity $W_* := P_* \cdot V_* \cdot \sigma_*$, this stock would hypothetically be at the bottom of the top 50 stocks in the Korean Composite Stock Price Index (KOSPI). With this rescaling, the invariance hypothesis implies

$$S_{it} = a \cdot \left( \frac{W_{it}}{W_*} \right)^{2/3}, \tag{7}$$

To summarize, assuming that the number of switching points is a function of the four variables price $P_{it}$, volume $V_{it}$, volatility $\sigma_{it}$, and the cost of a bet $C_{it}$, equation (7) follows directly from dimensional analysis, leverage neutrality, the invariance hypothesis. We test the prediction $S_{it} \sim W_{it}^{2/3}$ in the next section.

[3] This share price approximately equals to \$40 per share given the average exchange rate of 1,186 KRW per U.S. dollar (USD) between 2008 to 2010.

## 2. The South Korean Stock Market Data

This study is based on trade-level and account-level data provided by the Korea Exchange (KRX) for the period from February 2008 through November 2010. The Korea Exchange was created after the integration of the Korea Stock Exchange, the KOSDAQ Stock Exchange, and the Korea Derivatives Market in 2005. According to the World Federation of Exchanges, the South Korean stock market is ranked 17$^{\text{th}}$ in terms of market capitalization (about $1 trillion). The data set includes only the stocks listed in the KOSPI Market division at the Korea Exchange.

In contrast to the fragmented markets for U.S. stocks, the KRX operates a single central limit order book for each KOSPI stock. The data set contains records of all orders placed, canceled, or modified as well as all transactions executed. Records include trading codes for block trades, short-sale codes, trading system codes, and time stamps to the millisecond. Each message is linked to the specific accounts involved, and some additional information on account types is collected, such as whether accounts belong to domestic retail investors, domestic institutional investors (financial investment companies, insurance companies, private equity funds, government agencies, etc.), or foreign investors. The KRX database has about 2.69 billion messages and 1.29 billion distinct trade records during the sample period.

We analyze observations for each stock and each period of 20 trading days from February 2008 through November 2010. We refer informally to each period of 20 trading days as a "month" even though the 20-trading-day periods do not correspond precisely to calendar months. Using this definition, the data set covers 36 months. We begin with 24,441 observations, one observation for each KOSPI stock and each month from February 2008 through November 2010. We drop 2,506 stock-month observations because trading of some stocks was discontinued during particular months, thus introducing a downward bias into the number of switching points. Our final sample has 21,935 observations of stock-month pairs. There are on average 609 KOSPI stocks traded during each month.

The main variable of interest is the aggregate number of buy-sell switching point events per stock per month, which we use as a proxy for the pace of business time. For each month and each security, we count how many times each individual account changes its trading direction from buying to selling or from selling to buying and then sum these numbers across all accounts to find an aggregate number of switching points in a given stock in a given month. If an account trades a given stock in a given month but not in the previous month, then we count its number of switching point as at least one.[4] Each time an individual account changes the direction of its trading from buying to selling or from selling to buying, the number of switching points is increased by one.

[4] Our results do not change much if we count the number of switching points as zero in this case; for example, the estimated slope is 0.679 in regression (10) instead of 0.675.

For each stock $i$ and for each month $t$, let $N_{it}$ be the aggregate number of accounts which trade the stock and $S_{it}$ be the aggregate number of buy-sell switching points summed across all accounts. Dollar share price $P_{it}$ is the product of the closing KRW stock price and the exchange rate between the Korean won and the U.S. dollar (KRW–USD exchange rate). Share volume $V_{it}$ is obtained from the official daily public share volume report. Daily returns volatility $\sigma_{it}$ is the sample standard deviation of daily percentage returns during the same month. Market capitalization is based on the number of shares outstanding at the end of each year. The annualized turnover rate $\nu_{it}$ is based on share volume for stock $i$ in month $t$ and shares outstanding at the end of the previous year.

The data set identifies three broad categories of traders: domestic retail investors, domestic institutional investors, and foreign investors. Let $\alpha_{it}$ denote the fraction of share volume due to domestic retail investors.

In sharp contrast to the U.S. stock market, the South Korean stock market is dominated by retail investors. There are in total 425,440,260 switching points in the sample, on average 19,395 switching points per month per stock in the KOSPI universe: 94.2% from accounts of domestic retail investors, 4.7% from accounts of domestic institutions, and 1.1% from accounts of foreign investors. There are 5,886,557 distinct accounts in the sample: 94% domestic retail investors, 5.1% domestic institutions, and 0.8% foreign investors.

Table 1 shows summary statistics presented for the entire sample as well as the six volume subgroups defined by the 30$^{\text{th}}$, 60$^{\text{th}}$, 75$^{\text{th}}$, 85$^{\text{th}}$, 95$^{\text{th}}$, and 100$^{\text{th}}$ percentiles of average daily volume. The largest volume group is dominated by Samsung Electronics, the largest stock in the Korea Exchange, which accounts for about 5% of the total trading volume.

The average number of switching points per month increases by a factor of 147 from 930 for the lowest volume group to 136,710 for the highest volume group. Trading activity $W_{it} = P_{it} \cdot V_{it} \cdot \sigma_{it}$ increases by a factor 1,464 from the lowest to the highest group. These patterns are approximately consistent with invariance predictions, since a factor of 147 is not too different from $1464^{2/3} \approx 129$. Most of the variation in trading activity is due to variation in daily volume, which increases from 0.08 billion KRW to 94.88 billion KRW. Volatility varies much less across groups, and the changes are not monotonic. Daily volatility is 2.22 percent in the lowest group, 3.34 percent in the 75$^{\text{th}}$ percentile group, and 2.74 percent in the highest group.

The minimum lot size is equal to ten shares if the share price is below 50,000 KRW and to one share if share price is above 50,000 KRW; this corresponds to orders of about \$500 and \$50, respectively. The median size of trades is equal to 38 shares, implying that the minimum lot size constraint is often binding. Indeed, about 23.25% of trades are executed in the minimum size allowed; the fraction decreases from 28.78% for the low volume group to 17.51% for the high volume group. As in the U.S. market, extensive order shredding makes it difficult to test directly the invariance hypothesis by identifying bets in market data.

**Table 1 Summary Statistics.**

| | Volume Group | | | | | | |
|---|---|---|---|---|---|---|---|
| Variable | All | 30th | 60th | 75th | 85th | 95th | 100th |
| Price | 36,839 | 13,957 | 26,530 | 37,815 | 47,599 | 77,869 | 119,947 |
| Daily Volume (1B) | 8.50 | 0.08 | 0.50 | 2.08 | 6.68 | 23.55 | 94.88 |
| Volatility (%) | 2.79 | 2.22 | 2.88 | 3.34 | 3.21 | 2.97 | 2.74 |
| Capitalization (1T) | 1.32 | 0.07 | 0.15 | 0.33 | 0.99 | 3.39 | 14.62 |
| Annual Turnover (%) | 263.70 | 49.23 | 193.23 | 429.22 | 553.08 | 495.40 | 363.91 |
| Tick Size (BPS) | 22.10 | 21.53 | 22.25 | 22.44 | 22.80 | 22.30 | 21.69 |
| # Trades/Day | 5,659 | 255 | 1,170 | 3,574 | 7,893 | 17,033 | 41,400 |
| Avg Trade Size (1M) | 2.87 | 1.21 | 2.06 | 2.51 | 3.75 | 5.90 | 10.37 |
| Trades at Min Lot Size (%) | 23.25 | 28.78 | 23.42 | 20.27 | 19.21 | 18.46 | 17.51 |
| DR Volume (%) | 78.32 | 86.63 | 81.57 | 79.09 | 71.95 | 62.31 | 54.71 |
| DI Volume (%) | 13.93 | 10.11 | 12.08 | 13.78 | 17.87 | 21.45 | 23.91 |
| FI Volume (%) | 7.75 | 3.27 | 6.35 | 7.13 | 10.18 | 16.24 | 21.38 |
| Avg # Switches | 19,395 | 930 | 4,072 | 13,353 | 28,501 | 57,567 | 136,710 |
| Avg # Stock | 609 | 176 | 185 | 93 | 62 | 62 | 32 |
| # Observations | 21,935 | 6,330 | 6,669 | 3,341 | 2,220 | 2,235 | 1,140 |

The table shows the price (KRW), daily volume (1 billion KRW), daily volatility (%), market capitalization (1 trillion KRW), annual turnover (%), tick size (basis points or BPS), number of trades, average trade size (1 million KRW), percentage of trades of minimum lot size, the fraction of double-sided volume from domestic retail investors, the fraction of double-sided volume from domestic institutional investors, the fraction of double-sided volume from foreign investors, average number of switches per month, average number of stocks, and number of month-stock observations. The average exchange rate is 1,186 KRW per USD during the sample period.

The minimum tick size is determined according to a schedule.[5] The average tick size is about 22.10 basis points, approximately ten times larger than the typical tick size in the U.S. stock market.[6] The average tick size is relatively stable across volume groups, ranging from 21.53 basis points for low volume group to 22.83 basis points for high volume group. In principle, the large tick size may influence the trading behavior of market participants and have an effect on the aggregate number of switching points.

Another possibly important market friction is South Korea's transactions tax. The exchange collects a tax of about 30 basis points on the sale of securities, paid by the seller. Trading fees of about 1.50 basis points are paid to on-line brokers on executed orders.

In the Korean market, several stock indices are used as reference values for actively traded derivatives contracts. The Korea Composite Stock Price Index (KOSPI) includes all common stocks traded on the Korea Exchange, with weights proportional to market capitalization. The KOSPI includes about 688 stocks. The KOSPI 50 index includes the 50 largest companies listed on the Korea Exchange, approximately corresponding to the $95^{th}$ percentile and the $100^{th}$ percentile volume groups in table 1. The KOSPI

[5] The tick size is equal to 1 KRW if share price is below 1,000 KRW; 5 KRW if share price is between 1,000 KRW and 5,000 KRW; 10 KRW if share price is between 5,000 KRW and 10,000 KRW; 50 KRW if share price is between 10,000 KRW and 50,000 KRW; 100 KRW if share price is between 50,000 KRW and 100,000 KRW; 500 KRW if share price is between 100,000 KRW and 500,000 KRW; and 1,000 KRW if share price is above 500,000 KRW.

[6] For example, a tick size of one cent on a U.S. stock with a typical price of about $40 is only 2.5 basis points.

200 index includes the 200 largest companies listed on the Korea Exchange, approximately corresponding to the $75^{\text{th}}$ percentile to $100^{\text{th}}$ percentile volume groups in Table 1.

## 3. Results: Switching Points and Trading Activity

The main result of this paper concerns the empirical relationship between the logarithm of the aggregate number of buy-sell switching points $\ln(S_{it})$ and the logarithm of scaled trading activity $\ln(W_{it}/W^*)$ in the same month.

Figure 1 shows a scatter plot with all 21,935 observations lining up along a straight line whose fitted slope of 0.675 from an OLS regression is very close to the predicted slope of 2/3. Observations for stocks included in the KOSPI 50 universe (black points) and KOSPI 200 universe (blue points) are close to the fitted line as well. Even the observations for the largest South Korean stock, Samsung Electronics, at the far right corner of figure 1 do not deviate much from the fitted line. When Samsung Electronics is compared to the stock with the least amount of trading activity, the difference in trading activity between two stocks is a factor of about $\exp(10)$, or approximately 22,000; yet, observations for both of these stocks are consistent with invariance predictions.

It is apparent from visual observation that the data is relatively homoskedastic. For a given level of trading activity, the logarithm of the number of switching points for the less actively traded stocks deviates from the fitted line only slightly more than for the more actively traded stocks. This slightly higher deviation may indicate a larger estimation error in the estimates of expected trading activity for smaller stocks.

A similar conclusion can be drawn from an OLS regression analysis of the logarithm of the aggregate number of buy-sell switching points $\ln(S_{it})$ on the logarithm of scaled trading activity $\ln(W_{it}/W^*)$, clustering standard errors in the panel data regression at monthly levels:

$$\ln(S_{it}) = 11.156 + 0.675 \cdot \ln(W_{it}/W^*) + \epsilon_{it}. \tag{8}$$

The estimated coefficient of 0.675 has a clustered standard error of 0.005, implying that the hypothesis that the coefficient is equal to the predicted value of 2/3 is not rejected ($t = 1.67$). The non-clustered standard error is 0.0012. The constant term of 11.156 implies that the benchmark stock has on average about 53,000 buy-sell switching points per month. The $R^2$ of the regression is 0.935.

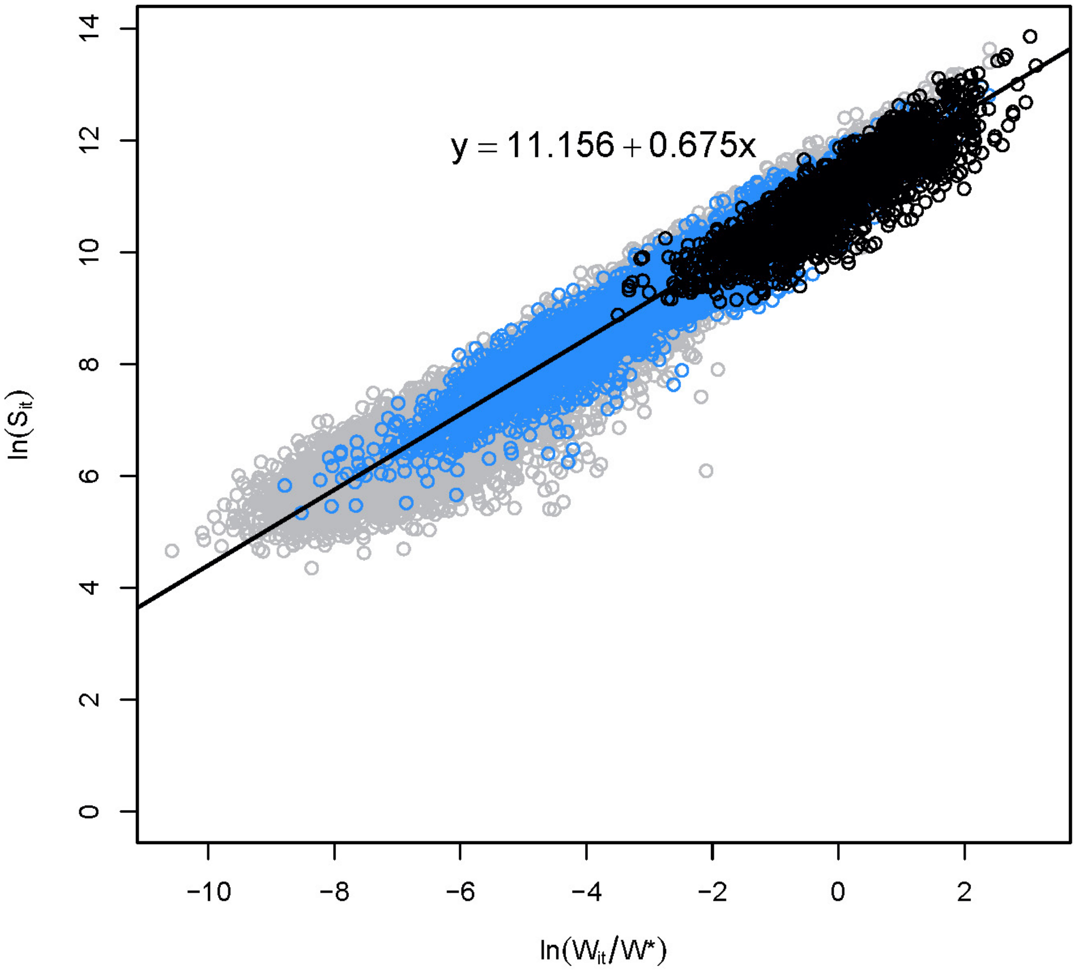
A negative binomial specification produces similar estimates of

$$\ln(S_{it}) = 11.234 + 0.673 \cdot \ln(W_{it}/W^*) + \epsilon_{it} \tag{9}$$

with clustered standard errors of 0.025 and 0.003 for both coefficients, respectively.

Figure 2 presents the time series of estimates from 36 monthly regressions of the logarithm of the aggregate number of switching points $\ln(S_{it})$ on the logarithm of scaled trading activity $\ln(W_{it}/W^*)$. To

**Figure 1** **Aggregate Number of Switching Points $\ln(S_{it})$ against Trading Activity $\ln(W_{it}/W^*)$.**

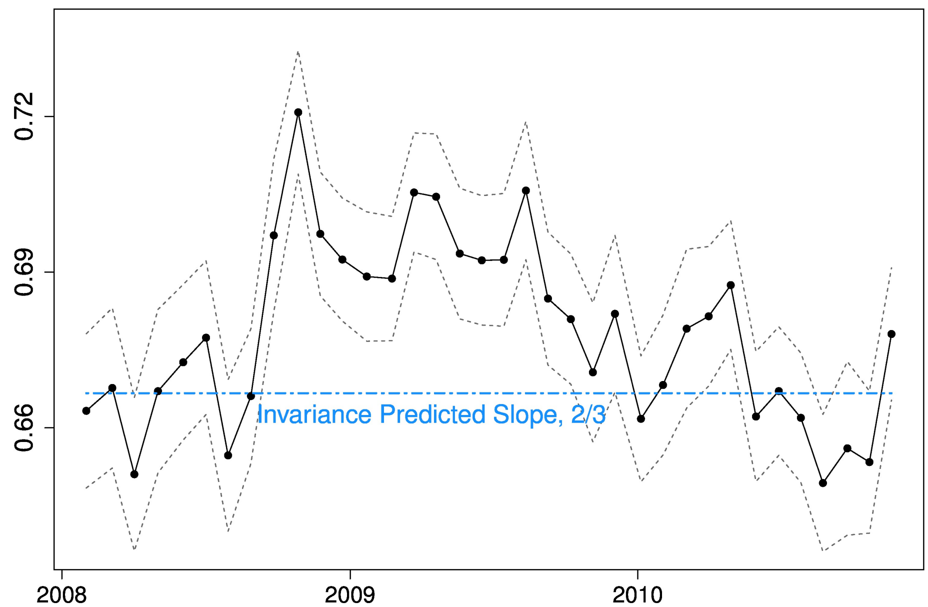



*Note.* The vertical axis is $\ln(S_{it})$. The horizontal axis is $\ln(W_{it}/W^*)$, where $W^* = 10^6 \cdot 40 \cdot 1186 \cdot 0.02$ and $W_{it} = V_{it} \cdot P_{it} \cdot \sigma_{it}$. The fitted line is $11.156 + 0.675 \cdot \ln(W_{it}/W^*)$. The invariance-implied slope is $2/3$.

make interpretation of results easier, the figure also contains a horizontal line indicating the regression coefficient of $2/3$ predicted by invariance.

All 36 point estimates of monthly regression coefficients are economically very close to $2/3$. The estimated coefficients exhibit persistence across months, fluctuating over time between 0.64 and 0.72. Only 15 out of 36 point estimates lie slightly outside of the 95%-confidence bounds. Most of these months occur between October 2008 and November 2009, when the South Korean market was most affected by the 2008 financial crisis.

**Figure 2** Time Series of Monthly Regression Coefficients




*Note.* The time series of estimates $\beta_s$ and their 95%-confidence intervals from 36 cross-sectional regressions $\ln(S_{it}) = \ln(a) + \beta_s \cdot \ln(W_{it}/W^*) + \epsilon_{it}$, where $S_{it}$ is the aggregate number of switching points and $W_{it}$ is expected trading activity for stock $i$ and month $t$. The time period is from February 2008 to November 2010. The invariance-implied slope is 2/3.

We conclude that even though there is enough variation in the time series of monthly regression coefficients to statistically reject the hypothesis that the coefficient is equal to 2/3 every month, the coefficient estimates are economically close to the predicted value.

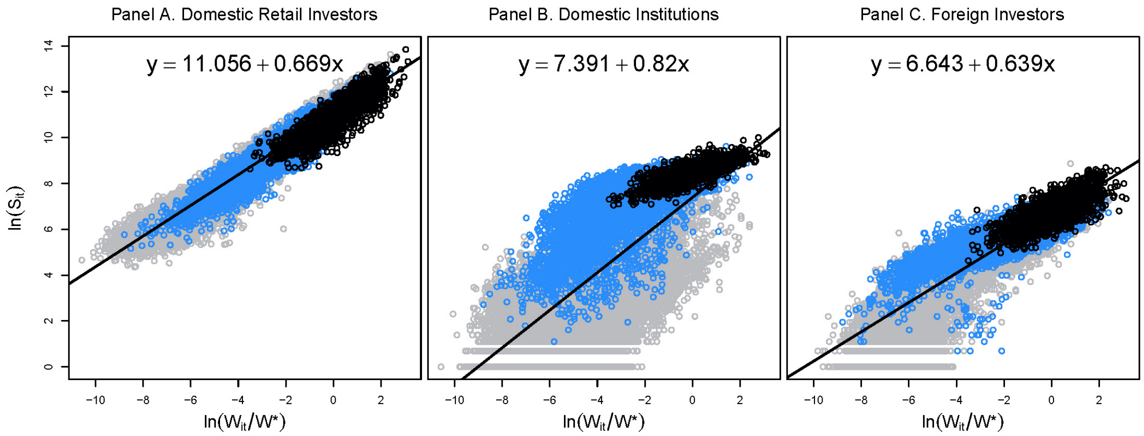
## 4. Results: Switching Points and Different Types of Traders

Figure 3 shows the relationship between the logarithm of buy-sell switching points and the logarithm of scaled trading activity for different types of traders: domestic retail investors, domestic institutional investors, and foreign investors.

Panel A of figure 3 shows results for the subset of domestic retail investors. These observations reveal a striking invariance relationship. The slope of the fitted line, 0.669 ($t = 0.4630$ using clustered standard error, t=1.7903 using non-clustered standard error), does not reject the hypothesis of equality to the predicted value of 2/3. Trades by retail investors dominate the results in figure 1 since domestic retail investors account for about 94.7% of switching points in the entire sample.

Panel B of figure 3 shows results for the subset of domestic institutional investors. These observations account only for about 4.7% of switching points of the entire sample. The small counts for less actively traded securities introduce discreteness as revealed by horizontal lines corresponding to one through

**Figure 3 Aggregate Number of Switching Points $\ln(S_{it})$ against Trading Activity $\ln(W_{it}/W^*)$ for Different Types of Investors.**




*Note.* The vertical axis is the log of the number of switching points for stock $i$ in month $t$, $\ln(S_{it})$. The horizontal axis is $\ln(W_{it}/W^*)$, where $W_{it}$ is trading activity in stock $i$ in month $t$, and $W^*$ is trading activity in the benchmark stock. Trading activity $W_{it}$ is defined as $W_{it} = V_{it} \cdot P_{it} \cdot \sigma_{it}$, where $V_{it}$ is average daily share volume during month $t$; $P_{it}$ is the dollar share price of stock $i$ at the end of month $t$, obtained by multiplying the KRW price by the exchange rate 1186 KRW/USD; and $\sigma_{it}$ is the daily volatility of stock $i$ in month $t$. The benchmark stock's trading activity $W^*$ is defined by $W^* = 10^6 \cdot 40 \cdot 1186 \cdot 0.02$, where $10^6$ represents share volume of one million shares per day; $40 \cdot 1186$ is the KRW price of a \$40 stock; and 0.02 represents volatility of 2% per day. Panel A presents results for domestic retail investors; the fitted line is $11.056 + 0.669 \cdot \ln(W_{it}/W^*)$. Panel B presents results for domestic institutional investors; the fitted line is $7.391 + 0.82 \cdot \ln(W_{it}/W^*)$. Panel C presents results for foreign investors; the fitted line is $6.643 + 0.639 \cdot \ln(W_{it}/W^*)$. The stocks are shaded black for the KOSPI 50, blue (medium) for the KOSPI 200, and light (gray) for other stocks.

ten switching points per month. The number of switching points deviates from the predictions of invariance in several respects. On the one hand, the slope of the fitted line, 0.82, for domestic institutions is steeper than the predicted coefficient of 2/3. The steeper-than-predicted coefficient may result from domestic institutions avoiding trading low-volume stocks. On the other hand, the number of switching points for stocks included in the KOSPI 50 universe (black points) is flatter than predicted by invariance; the estimated slope for these observations is only 0.332. The number of switching points for stocks in the KOSPI 200 universe but outside of the KOSPI 50 universe (blue points) is slightly steeper; the estimated slope for these observations is 0.532.

The flatness of the empirical distribution on the right side of the graph on Panel B is consistent with the interpretation that index trading plays an important role in generating trading patterns of some domestic institution, especially for stocks in the KOSPI 50 universe. For example, if index trader, including index arbitragers, tend to buy or sell all 50 stocks at the same time, this would lessen variation in the number of switching points across stocks and make the regression coefficient smaller. The identification of basket trades in the data set is complicated because the data set does not link accounts trading in the stock market to accounts trading in the derivatives market.

Figure 4 zooms further into the scaling relationship for domestic institutional investors. Some of these institutions are classified as government agencies, including the Korean Post Office. This large company is in charge of postal service, postal banking, and insurance services. It is known for implementing index arbitrage strategies.[7] Panel A shows the results with estimated slope of 0.82 for all domestic institutions as before. Panel B shows the results for domestic institutions but excluding government agencies; the estimated slope drops significantly from 0.82 to 0.685, the level consistent with invariance predictions. Panel C presents the results for solely government agencies; the estimated slope of 0.832 is close to the slope of 0.82 in panel A, suggesting that the sample for domestic institutions is dominated by trading by government agencies who tend to trade baskets of stocks and perhaps implement index arbitrage strategies. About a third of domestic institutional investors are government agencies, which generate about three quarters of all switching points in that subsample.

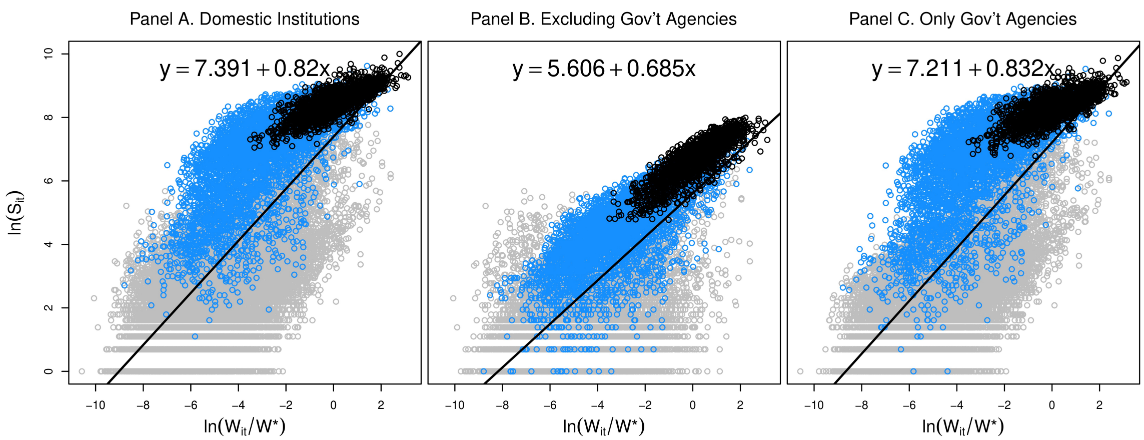
Panel C of figure 3 shows results for the subset of foreign investors. The slope of the fitted line, 0.639, is lower than the predicted slope of 2/3, but not by much. The points representing stocks included in KOSPI 50 and KOSPI 200 indices have much flatter slopes; the slopes of the fitted lines are 0.451 for the stocks in the KOSPI 50 universe and 0.35 for the stocks in the KOSPI 200 universe but outside of the KOSPI 50 universe. These slopes are similar in magnitude to the slopes for domestic institutions, suggesting that cross-market arbitrage affects trading patterns of both domestic institutions and foreign investors in a similar way. Since these observations account for about 0.6% of all switching points, these patterns are also influenced by small counts for less actively traded stocks, but this issue is less important for this subset than for the subset of domestic institutions.

Despite clearly visible data discreteness in figure 3, the results of the negative binomial regressions are very similar to the results of OLS regressions. The estimates of the slope coefficients are 0.669, 0.798, and 0.601 for samples in the three panels, respectively.

As developed by Kyle and Obizhaeva (2016), the invariance hypothesis is based on the idea that institutional investors choose their strategies for placing bets and professional intermediaries respond to these bets in a manner which leads to invariance relationships. The results in this paper suggest that the trading of retail investors leads to invariance relationships as well. Trading by retail investors, as measured by the rate at which switching points occur, reflects the passage of business time in a manner strikingly close to the predictions of market microstructure invariance. A conceptual issue raised by this result concerns whether invariance results from the trading behavior of institutional investors, retail traders, or both.

[7] See the article at `https://www.yna.co.kr/view/AKR20170328120800008`; Korean Post Office, `http://www.koreapost.go.kr/`.

**Figure 4** **Aggregate Number of Switching Points $\ln(S_{it})$ against Trading Activity $\ln(W_{it}/W^*)$ for Domestic Investors.**




*Note.* The vertical axis is the log of the number of switching points for stock $i$ in month $t$, $\ln(S_{it})$. The horizontal axis is $\ln(W_{it}/W^*)$, where $W_{it}$ is trading activity is stock $i$ in month $t$ and $W^*$ is trading activity in the benchmark stock. Trading activity $W_{it}$ is defined as $W_{it} = V_{it} \cdot P_{it} \cdot \sigma_{it}$, where $V_{it}$ is average daily share volume during month $t$; $P_{it}$ is the dollar share price of stock $i$ at the end of month $t$, obtained by multiplying the KRW price by the exchange rate 1186 KRW/USD; and $\sigma_{it}$ is the daily volatility of stock $i$ in month $t$. The benchmark stock's trading activity $W^*$ is defined by $W^* = 10^6 \cdot 40 \cdot 1186 \cdot 0.02$, where $10^6$ represents share volume of one million shares per day; $40 \cdot 1186$ is the KRW price of a \$40 stock; and 0.02 represents volatility of 2% per day. Panel A presents results for domestic institutional investors; the fitted line is $7.391 + 0.82 \cdot \ln(W_{it}/W^*)$. Panel B presents results for domestic institutional investors excluding government agencies; the fitted line is $5.606 + 0.685 \cdot \ln(W_{it}/W^*)$. Panel C presents results for government agencies; the fitted line is $7.211 + 0.832 \cdot \ln(W_{it}/W^*)$. The stocks are shaded black for the KOSPI 50, blue (medium) for the KOSPI 200, and light (gray) for other stocks.

The importance of retail investors may be an institutional characteristic specific to the South Korean stock market; retail investors account for a much larger share of trading than in most other countries, about 78.32% of double-counted trading volume, i.e., about 39.16% of buys and 39.15% of sells. Also, many large traders are classified as retail investors in the data, but they trade in a manner similar to institutional investors; South Koreans often refer to large retail investors as "super-ants".

## 5. Results: Adding Other Explanatory Variables

When the slope coefficient is fixed at the predicted value of 2/3 and only a constant term is estimated, we obtain

$$\ln(S_{it}) = 11.123 + \frac{2}{3} \cdot \ln(W_{it}/W^*) + \epsilon_{it}. \tag{10}$$

The mean squared error is 0.191, and the $R^2$ is 0.935, where $1 - R^2$ is defined as the variance of residuals divided by the variance of the demeaned data, i.e., 0.191/2.926. Since the data closely fit the invariance relationship to begin with, neither the mean squared error nor the $R^2$ are different from those of the regression equation (8) in an economically significant way. Thus, even though invariance has only one free parameter to be estimated, it explains about 93% of the variations in the logarithm of the number of buy-sell switching points.

We next study what can explain the remaining variation in the aggregate number of switching points.

We examine five model specifications for the logarithm of the number of switching points by month and stock:

1. A constant term only, with the coefficient on the logarithm of trading activity $\ln(W_{it}/W^*)$ fixed at 2/3.

2. A constant term and the logarithm of trading activity $\ln(W_{it}/W^*)$.

3. A constant term; the logarithm of trading activity $\ln(W_{it}/W^*)$; and the logarithm of effective relative tick size $\ln(e_{it}/e^*)$:

Let $\Delta_{it}$ denote the tick size in units of KRW for stock $i$ in month $t$; for example, $\Delta_{it}$ is 1 KRW if the share price is below 1,000 KRW. Following Kyle, Obizhaeva and Tuzun (2016), define effective relative tick size $e_{it}$ as the ratio of relative tick size $\Delta_{it}/P_{it}$ in basis points to the standard deviation of returns over one unit of business time and scaled by $e^* := \Delta_*/(P_* \cdot \sigma_* W_*^{-1/3})$ so that this ratio is equal to one for the benchmark stock. This yields[8]

$$\frac{e_{it}}{e^*} := \frac{\frac{\Delta_{it}}{P_{it} \cdot \sigma_{it} W_{it}^{-1/3}}}{\frac{\Delta_*}{P_* \cdot \sigma_* W_*^{-1/3}}}. \tag{11}$$

Similarly to the derivation of equation (6), it is possible to show that the arrival rate of bets (per units of calendar time), which determines the length of "business day," is proportional to $W_{it}^{2/3}$. Thus, the standard deviation of returns over one unit of business time is equal to $\sigma_{it} W_{it}^{-1/3}$; it is used to scale relative tick size $\Delta_{it}/P_{it}$ in the definition above.

4. A constant term; the logarithm of the three separate components of trading activity, share volume $\ln(V_{it}/V^*)$, share price $\ln(P_{it}/P^*)$, volatility $\ln(\sigma_{it}/\sigma^*)$); the logarithm of effective relative tick size $\ln(e_{it}/e^*)$; the logarithm of the stock's turnover rate $\ln(\nu_{it}/\nu^*)$; the logarithm of a fraction of volume executed by domestic retail investors $\ln(\alpha_{it}/\alpha^*)$; dummy variables for each stock in the KOSPI 50 and the KOSPI 200 universes; and 36 month fixed effects.

5. The logarithm of trading activity $\ln(W_{it}/W^*)$ and stock fixed effects.

6. The logarithm of effective relative tick size $\ln(e_{it}/e^*)$; the logarithm of the components of trading activity (share volume $\ln(V_{it}/V^*)$, share price $\ln(P_{it}/P^*)$, volatility $\ln(\sigma_{it}/\sigma^*)$); the logarithm of the turnover rate $\ln(\nu_{it}/\nu^*)$; the logarithm of the fraction of volume executed by domestic retail investors $\ln(\alpha_{it}/\alpha^*)$; dummy variables for the stocks in the KOSPI 50 and the KOSPI 200 universes; month and stock fixed effects.

[8] In principle, the benchmark quantities $e^*$ and $P^*$ might be adjusted for inflation or changes in growth or productivity and written with subscript $t$ as $e_t^*$ and $P_t^*$. These adjustments are not made here because such adjustments are not likely to matter over a brief span of three years. If the data spanned a longer period, say twenty years, or if inflation was extremely high over three years, then such adjustments would be appropriate.

All explanatory variables are scaled so that the estimated coefficients correspond to the benchmark stock with $V^* = 10^6$, $P^* = 40 \cdot 1186$, $\sigma^* = 0.02$, $\alpha^* = 1$, $\nu^* = 1/12$, and $W^* = V^* \cdot P^* \cdot \sigma^*$. The standard errors are clustered at the monthly level.

**Table 2 Explanatory Power of Other Variables.**

| Covariate | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Intercept | 11.123 | 11.156 | 11.358 | – | – | – |
| | (0.012) | (0.022) | (0.046) | – | – | – |
| $\ln(W/W^*)$ | 0.667 | 0.675 | 0.659 | – | 0.679 | – |
| | FIXED | (0.005) | (0.005) | – | (0.005) | – |
| $\ln(e/e^*)$ | – | – | 0.066 | – | – | −0.047 |
| | – | – | (0.012) | – | – | (0.007) |
| $\ln(P/P^*)$ | – | – | – | 0.539 | – | 0.617 |
| | – | – | – | (0.012) | – | (0.014) |
| $\ln(V/V^*)$ | – | – | – | 0.727 | – | 0.802 |
| | – | – | – | (0.016) | – | (0.018) |
| $\ln(\sigma/\sigma^*)$ | – | – | – | 0.245 | – | 0.228 |
| | – | – | – | (0.008) | – | (0.011) |
| $\ln(\nu/\nu^*)$ | – | – | – | 0.049 | – | −0.023 |
| | – | – | – | (0.018) | – | (0.020) |
| $\ln(\alpha/\alpha^*)$ | – | – | – | 0.590 | – | 0.562 |
| | – | – | – | (0.025) | – | (0.025) |
| KOSPI50 | – | – | – | −0.028 | – | −0.030 |
| | – | – | – | (0.020) | – | (0.017) |
| KOSPI200 | – | – | – | 0.120 | – | 0.127 |
| | – | – | – | (0.026) | – | (0.027) |
| F.E. Month | No | No | No | Yes | No | Yes |
| F.E. Stock | No | No | No | No | Yes | Yes |
| Nobs | 21,935 | 21,935 | 21,935 | 21,935 | 21,935 | 21,935 |
| Adj. $R^2$ | 0.935 | 0.935 | 0.936 | 0.973 | 0.969 | 0.984 |
| MSE | 0.191 | 0.190 | 0.188 | 0.078 | 0.091 | 0.047 |

The explanatory variables are trading activity $\ln(W_{it}/W^*)$, share volume $\ln(V_{it}/V^*)$, share price $\ln(P_{it}/P^*)$, volatility $\ln(\sigma_{it}/\sigma^*)$, effective relative tick size $\ln(e_{it}/e^*)$, turnover rate $\ln(\nu_{it}/\nu^*)$, the fraction of volume executed by domestic retail investors $\ln(\alpha_{it}/\alpha^*)$, and dummy variables for stocks in the KOSPI 50 and the KOSPI 200 universes. Some specifications have month and stock fixed effect. In the first column, $1 - R^2$ is defined as the variance of residuals divided by the variance of the demeaned data, which is 0.191/2.926 when the coefficient on $\ln(W_{it}/W^*)$ is fixed at the value $0.667 \approx 2/3$.

Table 2 presents results of OLS panel data regressions of the logarithm of the number of switching points by month and stock on five sets of explanatory variables. The most important results are the $R^2$ and the mean squared errors of each specification. The coefficients themselves are less important because they are heavily affected by multi-collinearity.

The initial variation of the logarithm of the number of switching points is equal to 2.926 (21,395 observations). The invariance model in the first column of table 2—where only a constant term is estimated and coefficient on $\ln(W_{it}/W^*)$ is fixed at a value of 2/3—has the $R^2$ of 0.935 and thus explains 93.5% of variation. The addition of other explanatory variables, including month and stock fixed effects, improves the $R^2$ in a statistically significant manner but not much in economic terms. The remaining five specifications have $R^2$ of only 0.935, 0.936, 0.973, 0.969, and 0.984, respectively. The highest value of 0.984 is achieved in the sixth specification which has 8 estimated parameters, 36 month fixed effects, and 686 stock fixed effects (20,665 degrees of freedom!). The mean squared errors of the regressions show similar variation across different specifications.

Of particular interest is the issue of tick size. In table 2, column 3 differs from column 2 by adding relative tick size $\ln(e_{it}/e^*)$ as an explanatory variable. Since the $R^2$ increases from 0.935 to only 0.936 and the estimated coefficient on the log of trading activity remains close to 2/3, changing from 0.675 to 0.659, the regression results indicate that relative tick size has limited effect on the number of switching points.

The estimated coefficients for negative binomial specifications are very similar and therefore not presented.

## 6. Results: Number of Accounts and Points per Account

By definition, the aggregate number of switching points $S_{it}$ is equal to the product of the number of unique accounts traded in a given month $N_{it}$ and the average number of switching points per account:

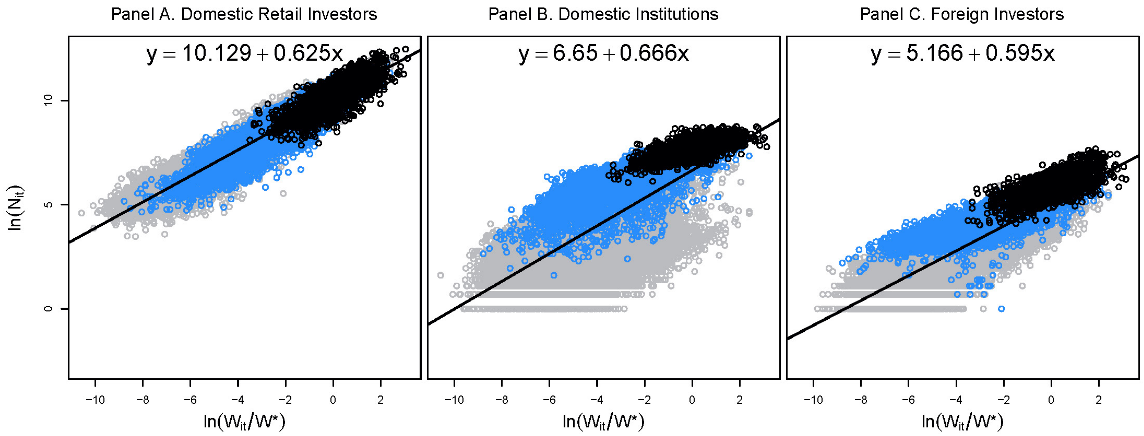
$$S_{it} = N_{it} \cdot \frac{S_{it}}{N_{it}}. \tag{12}$$

To the extent that the theory has been developed so far, market microstructure invariance does not predict whether changes in switching points will show up as changes in the number of accounts which trade stocks or the number of switching points per account.

We examine next the variation in these two factors empirically. Our results are presented here to provide benchmarks against which the future theoretical predictions may be compared.

Figure 5 shows the relationship between the logarithm of the number of unique accounts $\ln(N_{it})$ trading a given security $i$ during a given month $t$ and the logarithm of trading activity $\ln(W_{it})$. The OLS slopes of 0.625, 0.666, and 0.595 for domestic retail investors, domestic institutions, and foreign investors, respectively, are slightly lower than the value of 2/3 implied by invariance if the number of switching points per account is constant. The slopes of the corresponding negative binomial specifications of 0.624, 0.664, and 0.581 are very similar. The higher intercept for domestic retail investors reveals the exceptionally high level of retail participation in the South Korean stock market. Domestic institutions and foreign investors are much less active than retail investors. Many stocks were traded by only a few domestic institutions or foreign investors during a particular month, as reflected by clustering of data points around horizontal lines of $\ln(1)$, $\ln(2)$, $\ln(3)$, and $\ln(4)$.

**Figure 5** **The Number of Unique Accounts $\ln(N_{it})$ against Trading Activity $\ln(W_{it}/W^*)$ for different types of investors.**

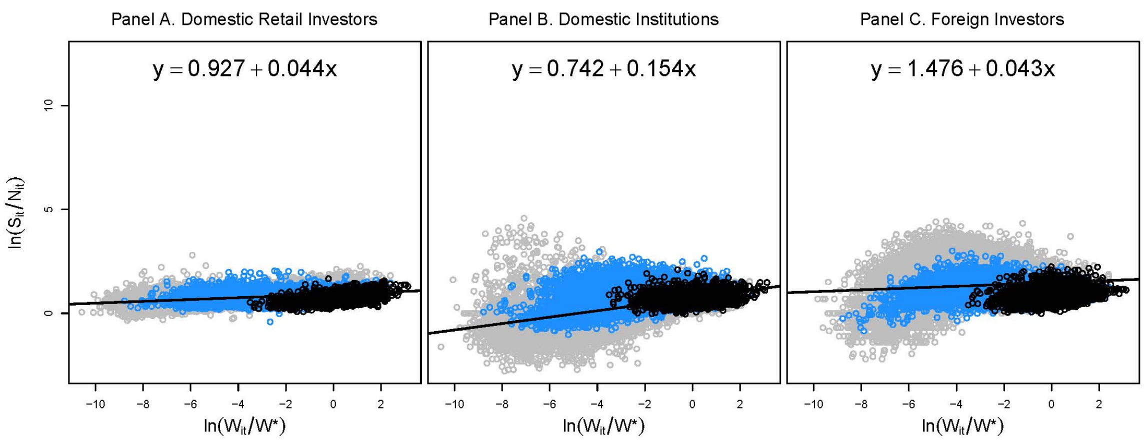



*Note.* The vertical axis is log of the number of accounts which trade stock $i$ in month $t$, $\ln(N_{it})$. The horizontal axis is $\ln(W_{it}/W^*)$, where $W_{it}$ is trading activity is stock $i$ in month $t$, and $W^*$ is trading activity in the benchmark stock. Trading activity $W_{it}$ is defined as $W_{it} = V_{it} \cdot P_{it} \cdot \sigma_{it}$, where $V_{it}$ is average daily share volume during month $t$; $P_{it}$ is the dollar share price of stock $i$ at the end of month $t$, obtained by multiplying the KRW price by the exchange rate 1186 KRW/USD; and $\sigma_{it}$ is the daily volatility of stock $i$ in month $t$. The benchmark stock's trading activity $W^*$ is defined by $W^* = 10^6 \cdot 40 \cdot 1186 \cdot 0.02$, where $10^6$ represents share volume of one million shares per day; $40 \cdot 1186$ is the KRW price of a \$40 stock; and 0.02 represents volatility of 2% per day. Panel A presents results for domestic retail investors; the fitted line is $10.129 + 0.625 \cdot \ln(W_{it}/W^*)$. Panel B presents results for domestic institutional investors; the fitted line is $6.65 + 0.666 \cdot \ln(W_{it}/W^*)$. Panel C presents results for foreign investors; the fitted line is $5.166 + 0.595 \cdot \ln(W_{it}/W^*)$. The stocks are shaded black for the KOSPI 50, blue (medium) for the KOSPI 200, and light (gray) for other stocks.

Figure 6 shows the analogous relationship for the average number of switching points per account, $\ln(S_{it}/N_{it})$. The clouds of data points for all three categories of traders—domestic retail, domestic institutions, foreign investors—are almost flat. The OLS slopes of 0.044, 0.154, and 0.043 for the three investor categories are close to zero. The corresponding estimated log-slopes in the negative binomial specifications of 0.045, 0.095, and 0.016 are very similar and close to zero. By construction, the sums of the slopes in figure 5 and figure 6 must be equal to the corresponding slopes in figure 3. There are more data points on the left side of the subplot for domestic retail investors than the other subplots, suggesting that domestic institutions and foreign investors may tend to avoid trading South Korean stocks with low trading activity.[9]

[9] The faintly clustering patterns along horizontal lines in figure 6 are less distinct than in figure 3 because the horizontal lines correspond to both integers (such as one switch for one account, two switches for one account, two switches for two accounts) and fractions (one switch for two accounts, one switch for three accounts, two switches for three accounts, etc.). Nevertheless, the horizontal clustering is still visible on panel B and panel C. Also, the data points in those two panels are somewhat symmetric relative to each other.

**Figure 6** **The Average Number of Switching Points per Account $\ln(S_{it}/N_{it})$ Against Trading Activity $\ln(W_{it}/W^*)$ for Different Types of Investors**




*Note.* The vertical axis is the log of the number of switching points per account for stock $i$ in month $t$, $\ln(S_{it}/N_{it})$. The horizontal axis is $\ln(W_{it}/W^*)$, where $W_{it}$ is trading activity is stock $i$ in month $t$ and $W^*$ is trading activity in the benchmark stock. Trading activity $W_{it}$ is defined as $W_{it} = V_{it} \cdot P_{it} \cdot \sigma_{it}$, where $V_{it}$ is average daily share volume during month $t$; $P_{it}$ is the dollar share price of stock $i$ at the end of month $t$, obtained by multiplying the KRW price by the exchange rate 1186 KRW/USD; and $\sigma_{it}$ is the daily volatility of stock $i$ in month $t$. The benchmark stock's trading activity $W^*$ is defined by $W^* = 10^6 \cdot 40 \cdot 1186 \cdot 0.02$, where $10^6$ represents share volume of one million shares per day; $40 \cdot 1186$ is the KRW price of a \$40 stock; and 0.02 represents volatility of 2% per day. Panel A presents results for domestic retail investors; the fitted line is $0.927 + 0.044 \cdot \ln(W_{it}/W^*)$. Panel B presents results for domestic institutional investors; the fitted line is $0.742 + 0.154 \cdot \ln(W_{it}/W^*)$. Panel C presents results for foreign investors; the fitted line is $1.476 + 0.043 \cdot \ln(W_{it}/W^*)$. The stocks are shaded black for the KOSPI 50, blue (medium) for the KOSPI 200, and light (gray) for other stocks.

We conclude that the invariance relationship arises mostly from cross-sectional variation in the number of unique accounts, not from variation in the number of switching points per account. This empirical fact is consistent with the spirit of the theoretical model in Kyle and Obizhaeva (2019), where the endogenously determined number of traders—each of whom makes decision to participate in the trading game, buy a signal of the same precision, and place exactly one bet—is shown to satisfy the invariance relationship.

Yet, this similarity should be taken with a word of caution. A slope slightly lower than 2/3 for the number of accounts may indicate that financial firms devote more resources, generate better signals, and place bigger bets when trading more active stocks. For example, domestic institutions and foreign investors may restrict their trading to stocks present in relevant benchmark indices such as the MSCI Emerging Markets Index, of which South Korea is one of the largest components. The empirical patterns may also be influenced by trades of cross-market arbitrageurs that flatten the average number of switching points across constituents of indices when trading their baskets.

## 7. Conclusion

The patterns documented in this paper strongly support the predictions of market microstructure invariance. This evidence complements the evidence on the invariance relationships in the U.S. market data documented by Kyle and Obizhaeva (2016), Kyle, Obizhaeva and Tuzun (2016), Andersen et al. (2016), and Kyle et al. (2014). It suggests that invariance relationships may hold outside of the U.S. markets as well. It also suggests that the trading of retail traders, not just institutions, exhibits an invariance relationship.

## References

**Andersen, Torben G., Oleg Bondarenko, Albert S. Kyle, and Anna A. Obizhaeva.** 2016. "Intraday Trading Invariance in the E-mini S&P 500 Futures Market." Working Paper, Available at SSRN 2693810.

**Ané, Thierry, and Hélyette Geman.** 2000. "Order Flow, Transaction Clock, and Normality of Asset Returns." *Journal of Finance*, 55(5): 2259–2284.

**Clark, Peter K.** 1973. "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices." *Econometrica*, 41(1): 135–155.

**Gabaix, Xavier.** 2009. "Power Laws in Economics and Finance." *Annual Review of Economics*, 121: 255–293.

**Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and Eugene H. Stanley.** 2006. "Institutional Investors and Stock Market Volatility." *Quarterly Journal of Economics*, 121: 461–504.

**Harris, Lawrence.** 1987. "Transaction Data Tests of the Mixture of Distributions Hypothesis." *Journal of Financial and Quantitative Analysis*, 22(02): 127–141.

**Jones, Charles M., Gautam Kaul, and Marc L. Lipson.** 1994. "Transactions, Volume, and Volatility." *Review of Financial Studies*, 7(4): 631–651.

**Kyle, Albert S., and Anna A. Obizhaeva.** 2016. "Market Microstructure Invariance: Empirical Hypothesis." *Econometrica*, 84(4): 1345–1404.

**Kyle, Albert S., and Anna A. Obizhaeva.** 2017. "Dimensional Analysis, Leverage Neutrality, and Market Microstructure Invariance." Working Paper. Available at SSRN 2785559.

**Kyle, Albert S., and Anna A. Obizhaeva.** 2018. "Adverse Selection and Liquidity: From Theory to Practice." Working Paper. Available at SSRN 2785559.

**Kyle, Albert S., and Anna A. Obizhaeva.** 2019. "Market Microstructure Invariance: A Dynamic Equilibrium Model." Working Paper. Available at SSRN 3326889.

**Kyle, Albert S., Anna A. Obizhaeva, and Tugkan Tuzun.** 2016. "Microstructure Invariance in U.S. Stock Market Trades." FEDS Working Paper No. 2016-034. Available at SSRN 2774039.

**Kyle, Albert S., Anna A. Obizhaeva, Nitish R. Sinha, and Tugkan Tuzun.** 2014. "News Articles and the Invariance Hypothesis." Working Paper. Available at SSRN 1786124.

**Mandelbrot, Benoît, and Howard M. Taylor.** 1967. "On the Distribution of Stock Price Differences." *Operations Research*, 15(6): 1057–1062.

**Tauchen, George E., and Mark Pitts.** 1983. "The Price Variability–Volume Relationship on Speculative Markets." *Econometrica*, 51(2): 485–505.