Centre for Economic and Financial Research at New Economic School

October 2016

Invariance of buy-sell switching points

Kyoung-hun Bae
Albert S. Kyle
Eun Jung Lee
Anna A. Obizhaeva

Working Paper No 232

CEFIR / NES Working Paper series

INVARIANCE OF BUY-SELL SWITCHING POINTS

Kyoung-hun Bae* Albert S. Kyle† Eun Jung Lee‡ Anna A. Obizhaeva§

First draft: June 26, 2014
This draft: October 26, 2016

Abstract

Define the number of buy-sell "switching points" as the number of times that individual traders change the direction of their trading. Based on the hypothesis that switching points take place in business time, market microstructure invariance predicts that the aggregate number of switching points is proportional to the $2/3$ power of the product of dollar volume and volatility. Using trading data from the Korea Exchange (KRX) from 2008 to 2010, we estimate the exponent to be 0.675 with standard error of 0.005. Invariance explains about 93% of the variation in the logarithm of the number of switching points each month across stocks. Most of the variation represents changes in the number of accounts trading the stock and not the number of switching points per account.

JEL codes: G00, G12, G14.

Keywords: Finance, market microstructure, invariance, trading volume, volatility, liquidity, price impact, market depth.

*Ulsan National Institute of Science and Technology, South Korea, kbae@unist.ac.kr.

†Corresponding author. Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, USA, akyle@rhsmith.umd.edu. 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.

‡Hanyang University, South Korea, ejunglee@hanyang.ac.kr.

§New Economic School, Moscow, Russia, aobizhaeva@nes.ru.

The voluminous data generated by financial markets leave little doubt that trade size, trade frequency, bid-ask spreads and other microstructure quantities vary significantly across securities. The market microstructure invariance hypothesis developed by Kyle and Obizhaeva (2016*c*) claims that such differences across markets look similar when viewed from the perspective of an appropriate business-time clock. Market microstructure invariance predicts similarities across markets in the dollar amounts expected to be at stake, the scale of risk transferred, the magnitude of transaction costs, and the size of profits.

In this paper, we test the market microstructure invariance hypothesis by examining variation in the aggregate number of buy-sell switching points across stocks. Define the number of buy-sell "switching points" as the number of times that individual traders change the direction of their trading. We hypothesize that the number of switching points is proportional to the rate at which business time passes. Under this hypothesis, market microstructure invariance predicts that the aggregate number of switching points is proportional to the $2/3$ power of the product of dollar volume and volatility.

Using account-level data from the Korea Exchange from 2008 to 2010, we estimate the exponent to be 0.675 with standard error of 0.005. Invariance explains about 93% of the variation in the number of switching points each month across stocks. Invariance patterns are especially pronounced for the subset of domestic retail investors who dominate trading in the South Korean market. A decomposition into the number of unique accounts and the average number of switching points per account shows that it is the cross-sectional variation in the number of accounts that exhibits the invariance patterns, while the number of switching points per account is relatively stable.

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. In contrast, microstructure invariance relates business time to the number of bets, or investment ideas. Since invariance also postulates that the dollar size and dollar costs of bets are constant when measured in business time, invariance links trading volume and returns variance to the size and number of bets. This further leads to something like a structural model which makes the sharp prediction that business time passes at a rate proportional to the $2/3$ power of the product of dollar volume and returns variance. Since bets are difficult to observe and trade sizes are influenced by tick size, minimum lot size, and numerous other market microstructure details, switching points provide an excellent way to identify how fast business time passes.

If variation in financial data can indeed be better explained by understanding the appropriate 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.

1 Market Microstructure Invariance, Business Time, and Switching Points

According to the market microstructure invariance hypothesis of Kyle and Obizhaeva (2016*c*), the business-time clock is governed by the frequency at which independent ideas—referred to as "bets"—are expected to arrive into the marketplace. In more active markets, bets arrive more frequently and the business-time clock runs faster. As bets are placed at a faster rate, trading costs decrease and the average distance between the market price and unobserved fundamental value decreases by an amount proportional to the square root of the arrival rate of bets. For informed traders to make the same expected dollar profits per bet when trading costs fall and price efficiency increases, they scale up the dollar size of their bets proportionally. Holding volatility constant, invariance implies that the dollar size of bets increases at a rate proportional to the square root of the number of bets per day; this implies that, as trading volume varies across securities, the number of bets increases twice as fast as the size of bets. Thus, if trading volume increases by a factor of 8, the number of bets increases by a factor of 4 and the dollar size of bets increases by a factor of 2. Since the business-time clock—which ticks at the rate bets arrive—effectively speeds up by a factor of 4, the speed with which business time passes is proportional to the $2/3$ power of trading volume.

To adjust for differences in percentage returns volatility, define "trading activity" W as the product of daily dollar volume $P \cdot V$ (dollar share price times share volume per day) and daily percentage returns volatility σ,

$$W := P \cdot V \cdot \sigma. \tag{1}$$

The quantity W measures aggregate risk transfer per calendar day. Invariance predicts that the expected number of bets per calendar day—and thus the rate at which trading unfolds—is proportional to $W^{2/3}$. Invariance implies that specific exponents of $1/3$ and $2/3$ govern relationships between trading activity W and various market characteristics such as bet size, bid-ask spreads, market impact costs, speed of mean reversion, and the accuracy of prices. Kyle and Obizhaeva (2016*a*) explain how to derive invariance relationships based on dimensional analysis, leverage neutrality, and a microstructure invariance hypothesis. Kyle and Obizhaeva (2016*b*) explain how to derive invariance relationships from a theoretical model of informed trading with different beliefs.

The variable of interest in this paper is the aggregate number of buy-sell "switching" points. For each month and each security, we count how many times individual traders change their trading direction from buying to selling or from selling to buying and then aggregate those numbers across all accounts to find an aggregate number of switching points for all traders 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. 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. We denote the aggregate number of switching points, summed across all accounts which traded stock i during month t, as S_{it}.

We expect to find invariance relationships in the cross-sectional patterns of switching points. More precisely, consistent with the invariance hypothesis, we hypothesize that S_{it} is proportional to $W_{it}^{2/3}$,

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

where a is the same "invariant" constant for all stocks i and all months t. The constant a is scaled by W^* so that it quantifies the expected number of switching points per calendar day for a hypothetical benchmark stock with trading activity W^*. To match the benchmark stock of Kyle and Obizhaeva (2016*c*), we define the benchmark stock to have a daily volume of one million shares, daily volatility of 2%, and price of 47,440 Korean won (KRW) per share (approximately equal to $40 per share given the average exchange rate of 1,186 KRW per U.S. dollar (USD) between 2008 to 2010). This hypothetical stock would be at the bottom of the top 50 stocks in the Korean Composite Stock Price Index (KOSPI).

In addition to trades implementing buy-bets and sell-bets, trading volume also consists of trading by intermediaries who attempt to profit by taking the other side of bets. These intermediaries have shorter holding horizons than bets. While bets are likely to generate positions held for months or years, intermediation trades are likely to generate positions liquidated with a day, an hour, or even a minute. An intermediary turning over inventories at relatively high frequency is likely to have many more switching points than a long-term investor placing a bet. Therefore, our hypothesis that switching points take place in business time also assumes that intermediation trades take place in business time as well.

The invariance hypothesis is motivated by the trading behavior of institutional investors who dominate the U.S. market. The South Korean market has much greater participation by retail investors. It is not clear a priori whether the invariance hypothesis should apply to institutional and retail investors separately or to both institutional and retail investors combined. When the invariance hypothesis is applied to both domestic and foreign institutions separately, we find an exponent different from $2/3$ for each separate group. When the invariance hypothesis is applied to retail and institutional investors combined—or to retail investors separately—the exponent is close to $2/3$. This suggests that invariance applies to market-wide trading behavior but not to retail and institutional investors separately because institutions avoid trading smaller, less liquid stocks and tend to trade baskets of stocks in proportion to weights in indices.

Our tests have a number of advantages over other tests for invariance relationships in trading data because the hypothesis that switching points take place in business time does not require additional assumptions to make the hypothesis realistic.

Kyle and Obizhaeva (2016*c*) document invariance relationships for the size distributions of portfolio transition orders. These tests require the identifying assumption that portfolio transition orders of institutional investors be proportional to bets.

Andersen et al. (2016) find invariance patterns in trading data in the E-mini S&P 50 futures market. Kyle, Obizhaeva and Tuzun (2016) document invariance relationships for the size distributions of "prints" of quantities traded in the Trade and Quote data set (TAQ). These tests rely on the even stronger assumption that print sizes are proportional to bets. This assumption broke down after tick size was reduced to one cent in 2001 and electronic order handling algorithms motivated traders in the earlier 2000s to shred their larger "meta-orders" into trades equal in size to the minimum lot size of 100 shares or even smaller odd lots. Evidently, the hypothesis that invariance relationships in tick data are stable over time also requires that tick size, minimum lot size, and perhaps also trading technology remain stable as well.

Kyle and Obizhaeva (2016*c*) develop invariance hypotheses using the concept of "bets." In theory, a portfolio manager places a bet when a statistically independent decision is made to accumulate a position of a particular size. In practice, the concept of a bet is difficult to map into data. Bets do not map easily into orders or trades, since one bet might be broken into many orders and each order might be executed as multiple trades. Thus, bets do not map easily into public data on trades, such as the TAQ data available from public data feeds. Even in consolidated audit trail data with trades matched to accounts, bets may not be readily identifiable because it is difficult to infer from trades when one bet stops and another bet starts.

The aggregate number of switching points is an interesting economic variable because the number of switching points may not be heavily influenced by tick size, minimum lot size, and other institutional details. Switching points may be a relatively clean indicator of the rate at which business time passes.

In contrast to the concept of a bet, the concept of a switching point can be given an relatively unambiguous definition which maps into data in a straightforward manner, provided trading data are available by individual account. There is some ambiguity concerning the possibility that bets are spread across multiple accounts or multiple bets are merged together. While these possibilities may affect the number of switching points, the effect is likely to be small and proportional across stocks. Empirical tests of cross-sectional variation based on the number of switching points only require the structure of trading to be approximately preserved across securities, regardless of the specifics of how the flow of bets in the marketplace is expressed as a flow of trades.

Examining switching points is potentially a powerful way to test market microstructure invariance because the number of switching points is likely to be less affected by institutional details than other quantitative summaries of market activity. In principle, switching point results may be affected by various market frictions and institutional features such as minimum tick size, minimum lot size, the level of cross-market arbitrage, and the industrial organization of entities participating in trading financial securities. We show below that tick size has only a minor effect on the number of switching points.

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.

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, etc.), or foreign investors. The KRX database has about 2.69 billion messages and 1.29 billion distinct trade records during our sample period.

For our analysis, one observation is associated with each stock for 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 biasing downwards the number of switching points calculated for those observations. Our final sample has 21,935 observations of stock-month pairs. There are on average 609 KOSPI stocks traded during each month.

For each stock i and for each month t, N_{it} is the aggregate number of accounts which trade the stock, and S_{it} is the aggregate number of buy-sell switching points (summed across accounts). For each observation, the dollar share price P_{it} is the product of the exchange rate between the Korean won and the U.S. dollar (KRW–USD exchange rate) and the closing KRW stock price. Share volume V_{it} is obtained from the official daily public share volume report. Daily returns volatility σ_{it} is the sample standard deviation of daily percentage returns during the same month. Trading activity W_{it} is the product of daily dollar volume and volatility, $W_{it} := P_{it} \cdot V_{it} \cdot \sigma_{it}$. Market capitalization is based on the number of shares outstanding at the end of each year. The annualized turnover rate ν_{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. The number of accounts N_{it} and number of switching points S_{it} represent sums across these three investor types. Let α_{it} denote the fraction of share volume due to domestic retail investors.

The South Korean stock market has large retail participation. 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 I shows summary statistics 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 in KRW.

<<PLACE TABLE I HERE.>>

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 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. Monthly volatility is 2.22 percent in the lowest group, 3.34 percent in the 75^{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. In our sample, 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.

The minimum tick size is determined according to a schedule.[1] 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.[2] 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.

Let Δ_{it} denote the tick size in units of KRW for stock i in month t (e.g., Δ_{it} is 1 KRW if the share price is below 1,000 KRW). Following Kyle, Obizhaeva and Tuzun (2016), define effective

[1] 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.

[2] 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.

relative tick size e_{it}/e^* as the ratio of tick size in basis points Δ_{it}/P_{it} to the standard deviation of returns over one unit of business time (which is proportional to $\sigma_{it}/W_{it}^{1/3}$), scaled so that this ratio is equal to one for the benchmark stock. This yields[3]

$$\frac{e_{it}}{e^*} := \frac{\Delta_{it}}{P_{it}} \cdot \frac{P^*}{\Delta^*} \cdot \frac{W_{it}^{1/3}}{\sigma_{it}} \cdot \frac{\sigma^*}{W^{*1/3}}. \tag{3}$$

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.

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$^{\text{th}}$ percentile and the 100$^{\text{th}}$ percentile volume groups in table I. The KOSPI 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 I.

The largest 200 stocks are often traded by investors engaging in cross-market and index arbitrage strategies. The resulting basket trades may tend to affect the number of switching points across stocks in the KOSPI 50 and KOSPI 200 universes. 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.

3 Trading Activity and Switching Points

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.

[3] 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.

Figure I shows that all 21,935 observations line 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. At the far right corner of figure I, the observations for the largest South Korean stock, Samsung Electronics, 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 is a factor of about exp(10), or approximately 22,000. It is apparent from visual observation that the data is relatively homoskedastic. For a given level of the logarithm 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.

<<**PLACE FIGURE I HERE.**>>

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:

$$\ln(S_{it}) = 11.156 + 0.675 \cdot \ln(W_{it}/W^*) + \epsilon_{it} \tag{4}$$

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 leads to similar estimates of $\ln(S_{it}) = 11.234 + 0.673 \cdot \ln(W_{it}/W^*) + \epsilon_{it}$ with clustered standard errors of 0.025 and 0.003 for both coefficients, respectively.

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

<<PLACE FIGURE II HERE.>>

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

4 Switching Points for Different Types of Traders

Figure III 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 III 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 I since domestic retail investors account for about 94.7% of switching points in the entire sample.

Panel B of figure III 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 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, is steeper than the predicted coefficient of $2/3$. The steeper-than-predicted coefficient appears to 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 is flatter than predicted by invariance; the estimated slope for these observations is 0.332.

The number of switching points for stocks in the KOSPI 200 universe but outside of the KOSPI 50 universe 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 is consistent with the interpretation that cross-market arbitrage plays an important role in trading patterns of domestic institution, especially for stocks in the KOSPI 50 universe. For example, if index arbitragers tend to buy or sell all 50 stocks at the same time, this would lessen variation in the number of switching points and make the regression coefficient smaller.

The small counts for less actively traded securities (as revealed by horizontal lines corresponding to one through ten switching points per month) introduces further distortions.

Panel C of figure III 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.

<<**PLACE FIGURE III HERE.**>>

Despite clearly visible data discreteness in figure III, the results of the negative binomial regressions are 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.

The main lesson from these results is that 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 or retail traders. As developed by Kyle and Obizhaeva (2016*c*), 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.

The importance of retail investors may be an institutional characteristic specific to the South Korean stock market. In 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. 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 Effective Relative Tick Size, Index Inclusion, and Other Explanatory Variables

When the slope is fixed at the predicted value of 2/3 and only a constant term is estimated, we obtain $\ln(S_{it}) = 11.123 + 2/3 \cdot \ln(W_{it}/W^*) + \epsilon_{it}$; 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). Neither the mean squared error nor the R^2 are different from the regression equation (4) in an economically significant way, since the data closely fit the invariance relationship to begin with. Thus, invariance explains about 93% of the variations in the logarithm of the number of buy-sell switching points. We next study what explains the remaining variation in the aggregate number of switching points.

Table II 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:

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 $\ln(e_{it}/e^*)$;

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 stocks in the KOSPI 50 and the KOSPI 200 universes; and 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.

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

<<**PLACE TABLE II HERE.**>>

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 main lesson of table II is that the addition of other explanatory variables, including month and stock fixed effects, improves the R^2 in a statistically significant manner but nevertheless leaves some economically significant variation unexplained. The initial variation of the dependent variable is equal to 2.926 (21,395 observations). In comparison with the R^2 of 0.935 in the first column (where only a constant term is estimated and coefficient on $\ln(W_{it}/W^*)$ is fixed at a value of 2/3), the remaining five specifications have R^2 of 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 II, 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 a small effect on the number of switching points.

6 Decomposition into the Number of Accounts and the Number of Switching Points per Account

By definition, the aggregate number of switching points is equal to the product of the number of unique accounts traded in a given month and the average number of switching points per account. The cross-sectional variation in those two factors is the question we examine next.

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. Empirical results are presented here to provide benchmarks against which future theoretical predictions can be compared.

Figure IV 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 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 $\ln(4)$.

<<PLACE FIGURE IV HERE.>>

Figure V 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 slopes in the negative binomial specifications of 0.045, 0.095, and 0.016 are very similar. The sums of the slopes in figure IV and figure V are by construction equal to the corresponding slopes in figure III. There are more data points on the left side of the subplot for domestic retail investors rather than the other subplots since domestic institutions and foreign investors avoid trading South Korean stocks with low trading activity.[4]

<<**PLACE FIGURE V HERE.**>>

We conclude that the invariance relationship arises mostly from cross-sectional variation in the number of unique accounts, not from number of switching points per account. This empirical fact is consistent with the spirit of the theoretical model in Kyle and Obizhaeva (2016*b*), 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 tend to flatten the average number of switching points across stocks in indices.

[4] The faintly clustering patterns along horizontal lines in figure V are less distinct than in figure III 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.).

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*c*), Kyle, Obizhaeva and Tuzun (2016), Andersen et al. (2016), and Kyle et al. (2014). It suggests that invariance relationships hold in all markets, not just the U.S. markets. It also suggests that the trading of retail traders, not just institutions, exhibits an invariance relationship.

Many scaling laws have been documented in the finance literature. For example, Gabaix et al. (2006) and Gabaix (2009) discuss relationships among exponents describing power laws in the tails of empirical distributions of financial variables like volume, trade size, and returns. While the findings in this paper can be used to validate documented power laws, our findings are also different from empirical regularities that describe power laws in the tails of probability distributions of financial variables. The invariance relationships in this paper are log-linear power-law relationships applying to entire probability distributions generally, not just to their tails. They are derived from empirical invariance hypotheses motivated by theory.

The results in this paper are so precise that they look like laws of physics and not laws of finance or economics. Yet, the empirical patterns reported in this paper are not regularities which have an explanation based on a mechanical interdependence among variables. 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. If there is an alternative to the market microstructure invariance hypothesis which provides a more natural benchmark for explaining market microstructure data, we leave it to other researchers to discover it.

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." *The Annual Review of Economics*, 121: 255–293.

Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and Eugene H. Stanley. 2006. "Institutional Investors and Stock Market Volatility." *The 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*a*. "Dimensional Analysis and Market Microstructure Invariance." Working Paper. Available at SSRN 2785559.

Kyle, Albert S., and Anna A. Obizhaeva. 2016*b*. "Market Microstructure Invariance: A Dynamic Equilibrium Model." Working Paper. Available at SSRN 2749531.

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

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.

Table I

Summary Statistics

	Volume Group						
Variable	All	30^{th}	60^{th}	75^{th}	85^{th}	95^{th}	100^{th}
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), 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 of domestic retail investors, the fraction of double-sided volume for domestic institutional investors, the fraction of double-sided volume of 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.

Table II

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, i.e., 0.191/2.926, when the coefficient on $\ln(W_{it}/W^*)$ is fixed at the value $0.667 \approx 2/3$.

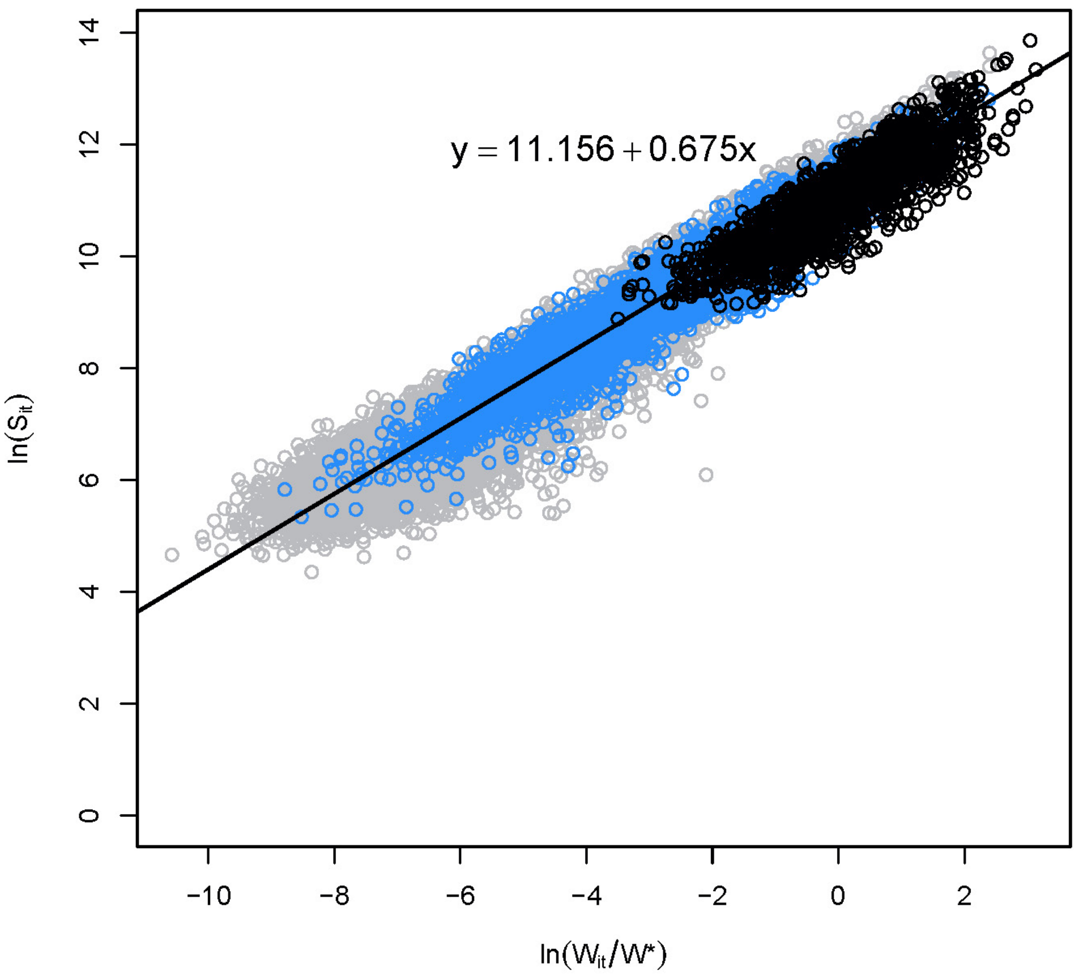

Figure I

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

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 $2/3$.

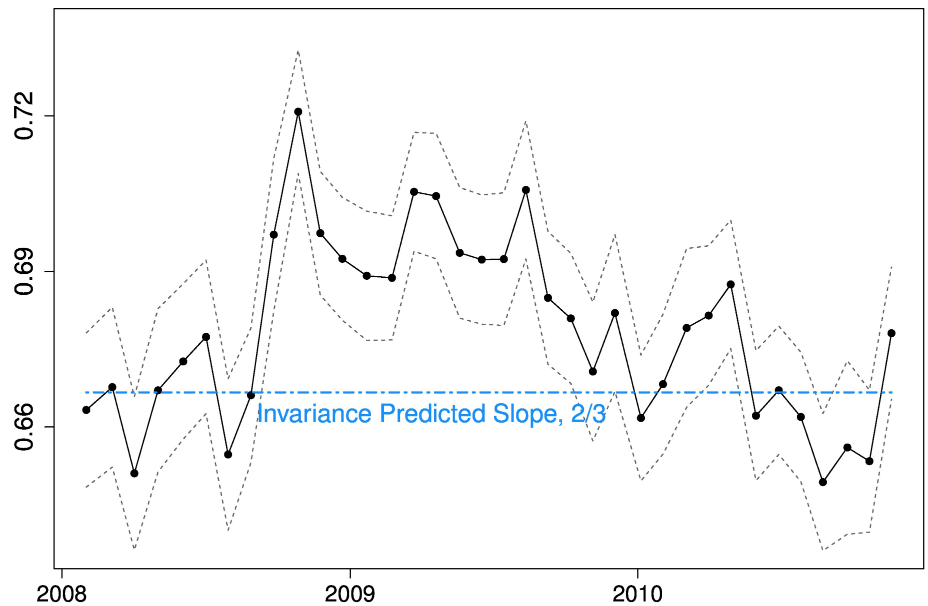

Figure II

Time Series of Monthly Regression Coefficients

The time series of estimates β_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 $2/3$.

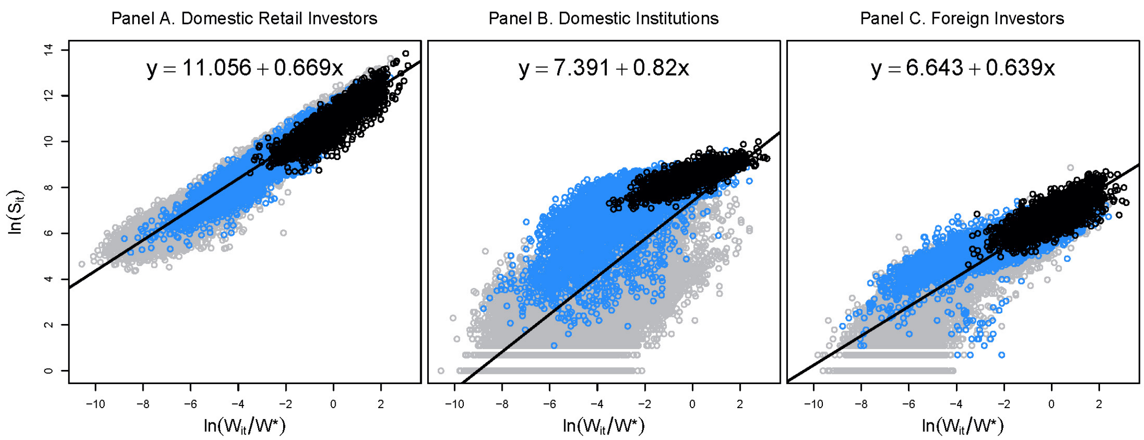

Figure III

Aggregate Number of Switching Points $\ln(S_{it})$ against Trading Activity $\ln(W_{it}/W^*)$ for different types of investors.

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 σ_{it} is the daily volatility of stock i in month t. The benchmark stocks' 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.

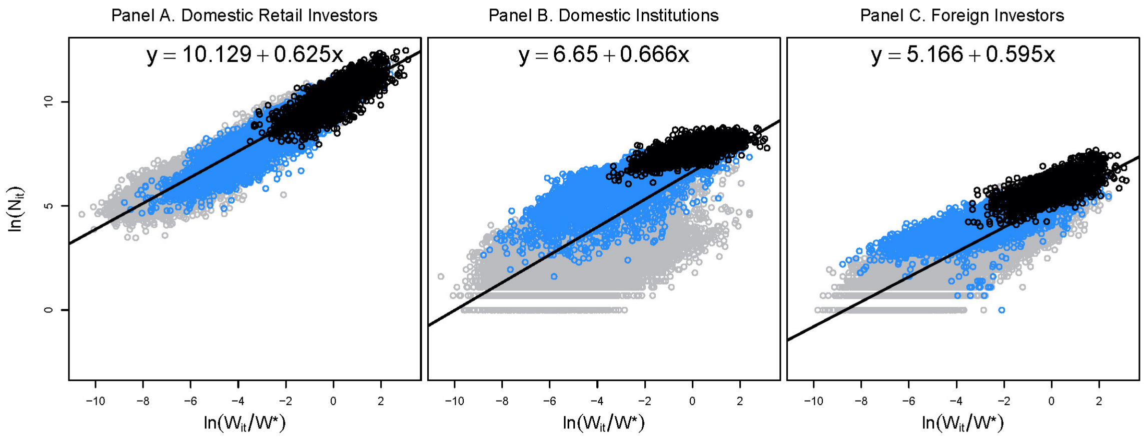

Figure IV

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

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 σ_{it} is the daily volatility of stock i in month t. The benchmark stocks' 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.

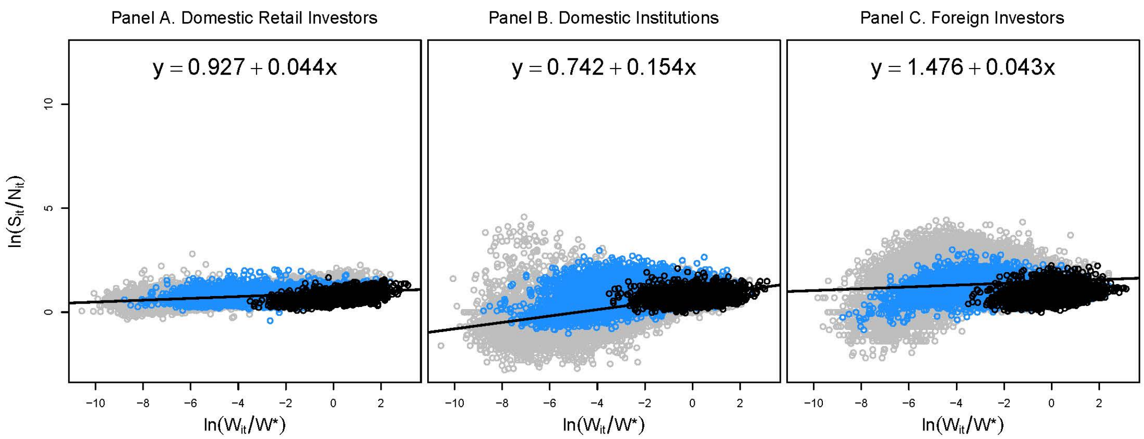

Figure V

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

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 σ_{it} is the daily volatility of stock i in month t. The benchmark stocks' 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.