Return seasonalities and investor clientele: Evidence from the China "twin-market"[1]

This draft: June 23, 2022

DONGXU CHEN, YUE DAI, ZHIGANG QIU

School of finance, Renmin University of China

Chen: cdx2792@ruc.edu.cn; Dai: yuedai@ruc.edu.cn; Qiu: zhigang.qiu@ruc.edu.cn

[1] We would like to thank seminar participants at Southwestern University of Finance and Economics and Zhongnan University of Economics and Law for helpful comments. This work is supported by Public Computing Cloud, Renmin University of China. All errors are our own.

Return seasonalities and investor clientele: Evidence from the China “twin-market”

Dongxu Chen, Yue Dai, Zhigang Qiu

ABSTRACT

This paper thoroughly investigates investor clientele's effects on time-series and cross-sectional return seasonalities at monthly, daily, and intraday level. A- and B-share stocks issued by the same firms in China form into a “twin-market” that serves as a natural experiment to explore investor clienteles’ effect. By comparing pricing performance of the twins, we find the magnitude of both time-series and cross-sectional return seasonalities are greater in A-shares due to the dominance of retail investors. The magnification effect brought by the dominance of noise trader is in accordance with our model prediction. Further analysis provide evidence in support of this argument, during times when clientele difference between the twins is large, the A- and B-shares’ gap in return seasonalies is wider than the average. Return volatility and trading volume exhibit similar seasonalities in both markets, and these seasonalities are also stronger in A-shares. From the perspective of return seasonalities, our study provides an insight into how the dominance of retail investors can affect liquidity-shock-driven mispricing.

Keywords: Anomalies, market efficiency, market microstructure
JEL classification: G11, G14

1. Introduction

Investor structure is one of the essential topics in market microstructure studies. Past research has mixed understanding about retail investors' role in market efficiency. Some studies discover that the presence of retail investors can create noise traders' risk, magnifying mispricing and harming market quality (Andradea, Chang, and Seasholes, 2008; Barber, Odean, and Zhu, 2009; Liao, Peng, and Zhu, 2021). Other research document that retail investors can provide liquidity for informed traders, increasing informed trading and making prices more informative (Barrot, Kaniel, Sraer, 2016; Ozik, Sadka, and Shen, 2021). In this study, we probe into the role of retail investors from the perspective of return seasonalities. We believe this study can reveal the role of retail investors in market efficiency and factors that can affect the magnitude of return seasonalities.

Return seasonality is an important market anomaly that exists widely across various assets[2] at monthly, daily and intraday level. Both time-series and cross-sectional return seasonalities are considered in this study. Time-series return seasonalities are mainly discussed before 2000s. For example, the January effect (Rozeff and Kinney, 1976; Thaler, 1987), the Monday effect (Cross, 1973; Miller, 1988), and the overnight return premium (Lockwood and Mcinish, 1990; Masulis and Ng, 1995). Since the pioneering studies of Heston and Sadka (2008, 2010) and Heston, Korajczyk, and Sadka (2010), seasonal patterns of cross-sectional returns are found to be existed at monthly, daily, and intraday levels. According to Keloharju, Linnainmaa, and Nyberg (2021), cross-sectional return seasonalities contain two parts: the "seasonalities" and "seasonal reversals". Seasonalities denote that future cross-sectional returns can be positively predicted by the past same-calendar-period returns in the last 20 years.[3] Seasonal reversals denote that future cross-sectional returns can be negatively predicted by the past other-calendar-period returns. Both time-series and cross-sectional return seasonalities are statistically and economically significant[4].

In this study, we try to explore how the dominance of retail investors can affect return seasonalities. Past studies have mentioned the effect of clientele changes on January

---

[2] Past studies documented seasonalities in stocks, bonds, commodities, and country indexes.

[3] For example, returns of the following January / Monday / first half-hour of the trading day can be positively predicted by the average of historical returns in January / Monday / first half-hour of trading over the past 20 years.

[4] In Rozeff and Kinney (1976), the average monthly return in January was about 3.5%, while other months averaged about 0.5%. In Heston and Sadka (2008), a trading strategy that chooses stocks based on their historical same-month returns earns excess returns as high as 12% per year.

effect and Monday effect (Lakonishok and Maberly, 1990; Kamara, 1997; Wang, Li, and Ericksone, 2012). However, none of them has thoroughly examined the role of investor clientele in the time-series and cross-sectional return seasonality at multiple frequencies. Another limitation is, these studies cannot control the potential effect of firm risks and news releases, although firm risks and news release could affect return seasonalities (Tinic and West, 1984; Rogalski and Tinic, 1986; Kim, 2006; Kelly and Tetlock, 2013; Gillain, Ittooa, and Lambert, 2022). In this study, we use a "twin-market" setting to cleanly isolate the effect of investor clientele. The "twin market" in China refers to the A- and B-share stocks issued by the same dual-share firms. These stocks have identical firm fundamentals, voting rights, and news releases. According to Chui, Subrahmanyam, and Titman (2022), China A- and B-share market have different investor clientele, A-share market is more dominated by retail investors, and B-share market has more institution investors. Combining these facts above, we claim that this twin market in China serves as a natural experiment to explore the effect of investor clientele.

We thoroughly investigate the existence and performance of return seasonalities in A- and B-share market. First, we build a model based on the framework of Bogousslavsky (2016) to theoretically generate both time-series and cross-sectional return seasonalities. Whether real data performance can fit in our model in multiple ways is carefully checked in this study. Specifically, we assume a risk-averse seasonal trading investor (or to say an infrequently rebalancing investor) exist in the economy, this is analogy to institutions who trade seasonally in the market. The presence of this special trader would generate time-series and cross-sectional return seasonality in the equilibrium. In time-series, when this investor is active, expected returns would be negative. On the contrary, when this investor is absent, the expected returns would be positive. Similar to seasonalities and seasonal reversals documented by the previous studies, in the equilibrium, same-calendar returns have positive autocovariances and other-calendar returns have negative autocovariances. To prove that the real performance of stock returns is in accordance with this model, we need to see negative returns during calendar periods when institutions are active, and positive returns when institutions are absent. Also, both seasonalities and seasonal reversals should exist in the cross-section of stock returns.

Empirically, we examine the existence of return seasonalities in A- and B-share markets, respectively. Similar patterns of return seasonalties are found in both markets. In time-series, at monthly level, returns in February are significantly positive and are higher than returns of other months, while returns in June are significantly negative. At daily level, returns in early-of-the-week are highly positive, while returns in later-of-the-week are significantly negative. At intraday level, overnight returns are significantly negative, while intraday returns are highly positive. The calendar periods when returns are negative (positive) is exactly the time when institutional are relatively active (absent) according to previous studies (Ali and Ülkü, 2018; Ülkü and Rogers, 2018; Gao, Han, Li, and Xiong, 2021). For cross-sectional return seasonalities, we first follow methodology of Heston and Sadka (2008), ploting the coefficients of regressing future returns on a series of lagged returns, clear periodical patterns are found. Further, following the methodology in Keloharju et al. (2021), we find same-calendar-period (other-calendar-period) returns in the past 120 months can positively (negatively) predict the future cross-sectional returns, which suggests the existence of seasonalities and seasonal reversals.

To investigate the role of retail investors in return seasonalities, we provide both theoretical and empirical evidences. Testable predictions are made by our model about the role of noise traders: when the market is more dominated by noise traders, both time-series and cross-sectional return seasonalities should become more prevalent. This prediction is empirically tested by comparing seasonalities of the A- and B-share twins. The result suggests that return seasonalities of all types are stronger in A-shares, which is consistent with our hypothesis.

To strengthen our findings, we present a series of additional results. To support the existence of the seasonal trading investor in China stock market, which is the key assumptions in our model, we examine the seasonal patterns in institutional trading. We find institutional trading is more pronounced in June, December, late-of-the-week and overnight, which is exactly when returns are on average negative. Further, we examine the seasonal performance of market demand. Stocks that are concentratedly bought in June, later-of-the-week and overnight are usually of high quality, these stocks have bigger market value and higher returns on equity. On the contrary, stocks that are concentratedly bought in February, early-of-the-week and intraday are usually speculative. These stocks have higher trading volume, higher idiosyncratic volatility, and lower share prices. The above results confirm the reasonability of our model assumptions and the link between seasonal trading behavior and return seasonalities.

To prove that the return seasonality difference between A- and B-shares are truly caused by difference in investor clientele but not other systematic differences, we focus on the time-varying clientele difference between the two markets. Empirical result suggests that during times when retail investors in A-shares are more active than those in B-shares, seasonality gap between the two markets becomes wider.

Seasonalities patterns in return volatility and trading volume further support our analysis. By derivation and simulation, our model suggests that time-series and cross-sectional seasonality should exist in return volatility and trading volume as well. In time-series, when the seasonally trading investor is active, return volatility and trading volume should be higher. Cross-sectionally, same-calendar return volatility and trading volume should have positive autocovariances, other-calendar return volatility and trading volume should have negative autocovariances. Time-series and cross-sectional seasonalities in volatility and volume should be stronger in A-shares due to the dominance of retail investors. Empirical results support these hypotheses above.

Our study contributes to several strands of research. We contribute to the literature on return seasonalities. First, we find return seasonalities exist in China stock market and the seasonality performance fits in our model in multiple pricing implications, this confirms that the presence of seasonal trading investor should be the reason why return seasonalities exist. Our model suggests return seasonalities are liquidity-shock-driven mispricing, seasonal clientele change is the source of liquidity shock. The position that should have been rebalanced by the seasonal trading investor when he is actually absent becomes an excess demand that act as a liquidity shock when he is active, this seasonally emerging excess demand drives return seasonalities. Moreover, our study proves that seasonal demand variation is the common cause of time-series and cross-sectional return seasonalities at monthly, daily and intraday level. Due to the different time preference of institutions in China stock market[5], our results also serve as an out-of-sample test for the relation between seasonal trading and return seasonalities.

This study also contributes to the literature on how price anomalies are affected by the dominance of retail investors. The closest existing study is Chui, Subrahmanyam, and Titman (2022), who finds that when retail investors dominate the market, momentum would be mitigated and short-term reversal would be magnified. We complement this type of study by first showing both theoretically and empirically that the dominance of retail investors can also magnify return seasonalities. Our theoretical framework suggest that the magnification effect brought by retail investors may also be applicable for other liquidity-shock-driven anomalies, for example the broad families of anomalies related to price trends and inventory risks. Further, this study enriches the literature on how to predict anomaly returns. As practitioners usually construct portfolios based on anomalies, our study reminds that they should be aware of the time-varying investor clientele and the related market policies.

[5] For example, past studies show that institutions in U.S. would like to trade near the close (Lou, Polk, and Sadka, 2021), but institutions in China would like to trade in the 10-min long open call auction (Gao, Han, Li, and Xiong, 2019). This is corresponding to positive (negative) overnight premium in the U.S (China).

To regulators, our study reveals that the dominance of retail investors can harm market efficiency by magnifying mispriced return seasonalities. Previous studies show that price discovering ability of A-share market is worse than that of B-shares due to market segmentation. Stocks in A-shares have higher premiums, A-share prices are often led by B-shares, and herding behaviors are more pronounced in A-share market (Tan, Chiang, Mason, and Nelling, 2008; Chang, Luo and Ren, 2013; Chui, Subrahmanyam, and Titman, 2022). We complement these studies from the perspective of mispriced return seasonalies. All these evidences above suggest that the artificial entry limits in A-share market set to protect domestic investors can actually harm price efficiency. To improve the quality of A-share market, gradually lifting the entry limits of foreign institutions and easing market segmentation between A- and B-share market is needed.

The remainder of the paper is structured as follows. Section 2 describes the institutional background for the twin-market in China. Based on the clientele difference of A- and B-share market, we build a model to make testable predictions. Section 3 introduces the sample selection and variable constructions for empirical tests. Section 4 empirically examines the existence of time-series and cross-sectional return seasonalities in both A-share and B-share market. The twin-market setting is used to identify the role of retail investors. Section 5 further discusses the seasonal trading patterns of institutional investors, the effect of time-varying clientele differences and seasonal patterns in return volatility and trading volume. Section 6 is the concluding remark.

2. Institutional background and empirical predictions

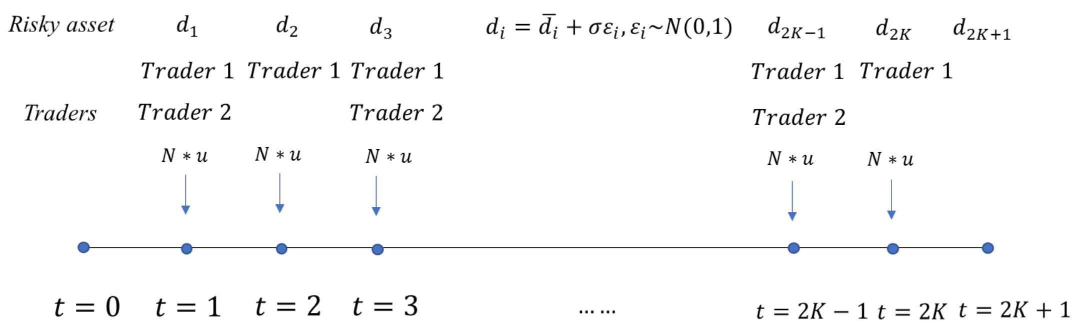
2.1. “Twin-market” of dual-share firms

Before the establishment of the Beijing Stock Exchange in 2021, Shanghai stock Exchange (SHSE) and the Shenzhen Stock Exchange (SZSE) were the only two stock exchanges in mainland China. Both A- and B-shares are allowed to trade on SHSE and SZSE. Only companies registered in China are allowed to issue A- and B-share stocks on SHSE and SZSE. A-shares stocks are mainly issued for raising domestic capital, while B-shares are mainly used for raising foreign capital. A- and B-share stocks have the same firm fundamentals, voting rights, and news releases.

Before 2001, A- and B-share market is completely separated, only domestic investors were allowed to trade A-shares and only foreign investors were allowed to trade B-shares. The market separation was lifted after 2001, domestic traders were allowed to trade B-shares since 19/2/2001. Foreign investors were allowed to trade A-shares since 1/12/2002. However, as Chui et al. (2022) noted, only a temporary change in investor structure occurred since this liberalization. Due to multiple reasons, such as foreign exchange restrictions, and quota limitations on foreign institutions' entrance into A-share market, severe market segmentation remains. Retail investors are more dominated in A-share market, while B-share market have relatively more institutions. According to Chui et al. (2022), during 2008 to 2018, the average ownership of institutions in A-share market was no more than 0.25%. Blair and Co (2018) documents that in 2015, the trading volume of retail investors accounted for more than 85% of the total volume. In contrast, in B-share market, the average ownership of institutional investors is around 5% with a peak of 8% in 2007 (Chui et al., 2022). Despite the difference in investor structure, A- and B-share market have similar trading schedule. In terms of market close, every year in China, there is a 7-day holiday for the Lunar Spring Festival (in February), Labor Day (in May) and National Day (in October). Besides, both A- and B-shares allow a 10-minute (9:15-9:25) call auction before the market open at 9:30. In addition, The ±10% price-hitting limit is set for both A- and B-share markets. (Kedar-Levy, Yu, Kamesaka and Ben-Zion, 2010).

2.2 Model predictions of investor clientele and return seasonalities

Referencing the framework of Bogousslavsky (2016), we build a model to generate return seasonalities. The model assumes a seasonal trading investor exist in the economy. Seasonal trading investors is analog to the institutions who would like to trade seasonally in China stock market. Institutions in China are more likely to trade in the year-end and half-year-end for window dressing or tax-considerations (He, Ng and Wang, 2004; Hu, McLean, Pontiff, and Wang, 2014). Besides, institutions would like to trade on a fixed day of the week and in a fixed time of the day (Ülkü and Rogers, 2018; Lou et al., 2019; Gao et al., 2021). With this model, we try to explain how seasonal trading investors can generate return seasonalities and make predictions about the effect of investor clientele on the magnitude of return seasonalities.



Consider an economy with discrete time periods of $T = 0,1,2,\ldots,2K+1$, $t = 0$ is the initial time period when no dividend is paid. $T = 2K+1$ is the terminal time period that all dividends are revealed. A risky asset exists in this economy, the asset pays off dividend $d_i = \bar{d_i} + \sigma\varepsilon_i$, $\varepsilon_i \sim N(0,1)$ and ε_i is i.i.d in each period i $(i \in \{1,2,\ldots,2K+1\})$. $P_{2K+1} = d_1 + \cdots + d_{2K} + d_{2K+1}$. For simplicity, we assume risk-free return is 0. A frequent trader (trader 1) exists and would trade at each time period $T = 1,\ldots,2K$. The frequent trader is analog to the market maker presents to clear the market demand. Trader 1 has CARA utility with risk-aversion of α_1, the asset demand of trader 1 is denoted by $x_{1,t}$. A seasonal trading investor (trader 2) exists and would trade only at odd time period $t = 2k-1$ $(1 \leq k \leq K+1)$. In this setting, "same-calendar" periods refer to odd vs. odd periods and even vs. even periods. "Other-calendar" periods refer to odd vs. even periods. The infrequent trader also has a CARA utility with risk-aversion of α_2, his demand of risky-asset is $x_{2,t}$. There are N noise traders exist in $T = 1,\ldots,2K$, the net demand of each noise trader is u. Assume $E(u) = \bar{u} > 0$ due to short-selling constraint in China. N represent the dominance of noise traders.

At $t = 2(K-i)$, only trader 1 is active, he would face the problem of maximizing his utility as follows:

$$\max_{x_{1,2(K-i)}} E\left(W_{2(K-i)+1}\right) - \frac{\alpha_1}{2} Var(W_{2(K-i)+1})$$

$$s.t. W_{2(K-i)+1} = W_{2(K-i)} + x_{1,2(K-i)}\left(P_{2(K-i)+1} - P_{2(K-i)}\right)$$

At $t = 2(K-i)-1$, both trader 1 and trader 2 are active. Similarly, trader 1's problem is:

$$\max_{x_{1,2(K-i)-1}} E\left(W_{2(K-i)}\right) - \frac{\alpha_1}{2} Var(W_{2(K-i)})$$

$$s.t. W_{2(K-i)} = W_{2(K-i)-1} + x_{1,2(K-i)-1}\left(P_{2(K-i)} - P_{2(K-i)-1}\right)$$

At the same time period, trader 2 joins the market, as he only cares about his utility in $2(K-i)+1$, he faces the problem as follows:

$$\max_{x_{2,2(K-i)-1}} E\left(W_{2(K-i)+1}\right) - \frac{\alpha_2}{2} Var(W_{2(K-i)+1})$$

$$s.t. W_{2(K-i)+1} = W_{2(K-i)-1} + x_{2,2(K-i)-1}\left(P_{2(K-i)+1} - P_{2(K-i)-1}\right)$$

Proposition 1: The equilibrium prices in this economy is unique:

$$P_{2(K-i)-1} = \Sigma_{j=1}^{2(K-i)-1} d_j + \Sigma_{j=2(K-i)}^{2K+1} \bar{d_j} + (i+1)(2i+3)\alpha_2\left[\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2 u\right]$$

$$P_{2(K-i)} = \Sigma_{j=1}^{2(K-i)} d_j + \Sigma_{j=2(K-i)+1}^{2K+1} \bar{d_j} + (2i+1)[\alpha_1 + (i+1)\alpha_2]\left[\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2 u\right]$$

(Please see Appendix A for proof.)

In this model, **time-series seasonality** behaves as follows:

1. Expected returns should be higher w/o. the presence of seasonal trading investor (trader 2). For $i \in \{0,1,2,\dots,K-1\}$, $E[r_{2(K-i)}] > 0$.

2. Expected returns should be lower w/i. the presence of seasonal trading investor. For $i \in \{0,1,2,\dots,K-1\}$, $E[r_{2(K-i)+1}] < 0$.

Seasonalities and seasonal reversals behaves as: (As only one risky asset exist in this economy, we consider autocovariances here).

1. The autocovariances of same-calendar-period returns are positive. In the equilibrium, autocovariances between any two odd periods and any two even periods are positive.
(For $\forall i,j \in \{0,1,2,\dots,K-1\}$ and $i<j$, $Cov(r_{2(K-i)+1}, r_{2(K-j)+1}) > 0$, $Cov(r_{2(K-i)}, r_{2(K-j)}) > 0$).

2. The autocovariances of other-calendar-period returns are negative. In the equilibrium, autocovariances between any odd and even period is negative.
(For $\forall i,j \in \{0,1,2,\dots,K-1\}$ and $i<j$, $Cov(r_{2(K-i)+1}, r_{2(K-j)}) > 0$, $Cov(r_{2(K-i)}, r_{2(K-j)+1}) > 0$).
(Please see Appendix A for proof.)

Our model suggests return seasonalities are liquidity-shock-driven mispricing, seasonal clientele change is the source of liquidity shock. The position that should have been rebalanced by the seasonal trading investor when he is actually absent becomes an excess demand that act as a liquidity shock when he is active, this seasonally emerging excess demand drives return seasonalities.

The testable predictions about investor clientele in this model are:
Hypothesis 1: The magnitude of time-series return seasonality should be greater in A-share market than B-share market, as A-share market has more noise traders.

$$\frac{\partial E[r_{2(K-i)}]}{\partial N} > 0, \qquad \frac{\partial E[r_{2(K-i)+1}]}{\partial N} < 0$$

Hypothesis 2: The magnitude of cross-sectional seasonalities and seasonal reversals should be greater in A-share market. For $\forall i,j \in \{1,2,\dots,K\}$ and $i<j$:

$$\frac{\partial Cov(r_{2(K-i)+1}, r_{2(K-j)+1})}{\partial N} > 0, \frac{\partial Cov(r_{2(K-i)}, r_{2(K-j)})}{\partial N} > 0$$

$$\frac{\partial Cov(r_{2(K-i)+1}, r_{2(K-j)})}{\partial N} < 0, \frac{\partial Cov(r_{2(K-i)}, r_{2(K-j)+1})}{\partial N} < 0$$

(Please see Appendix A for proof.)

3. Data description and variable construction

3.1 Sample selection

Our sample includes all stocks issued on Shanghai Stock Exchange (SHSE) and Shenzhen Stock Exchange (SZSE), sample period spans the time July 1999 to December 2019 due to limitation of intraday trading data in China. Intraday trading data comes from RESSET high frequency database. The sample distinguishes between collective bidding and continuous bidding[6]. Transactional data such as daily stock price comes from WIND database, accounting data such as stock market value comes from CSMAR database. We excluded: (1) stocks belonging to the financial industry according to the 2012 securities regulatory commission guidelines; (2) stocks that have been listed for less than 6 months; (3) stocks with less than 15 trading days in the previous month; (4) stocks with trading days less than 120 days in the past 12 months.

3.2 The construction of returns

The trading hours in China stock market is 9:30-11:30 and 13:00-15:00, we split them into 8 half-hour periods and calculate returns referencing Gao, Han, Li and Zhou (2018). The methodology is as follows: for the 8 intraday period returns ($r_{1,t}$~$r_{4,t}$ in the morning, $r_{6,t}$~$r_{9,t}$ in the afternoon) and return at 11:30-13:00 noon break ($r_{5,t}$) on trading day t, $r_{j,t} = \frac{p_{j,t}}{p_{j-1,t}} - 1, j = 1, \dots, 9$, where $p_{0,t}$ denotes opening price, $p_{j,t}$ denotes transaction price at the end of period j, $p_{9,t}$ is the closing price. The overnight return is $r_{0,t} = \frac{p_{0,t}}{p_{9,t-1}} - 1$. To unify the calculation of returns, we use closing prices to calculate daily and monthly returns.

3.3 Stock characteristics

Table 1 presents descriptive information of characteristic variables mainly used in this paper.

(Insert Table 1 Here)

A-share market has 3156 stocks, B-share market has 109 stocks. 81 of B-shares are

[6] Continuous bidding period is 9：30 to 11：30, 13：00 to 15：00 for SHSE and 9：30 to 11：30, 13：00 to 14：57 for SZSE. The sample does not include the after-hours trading, which is not allowed in China.

issued by dual-share firms. Companies with B-share stocks usually have larger market capitalization, higher book-to-market ratio and higher profitability in comparison to A-share firms. Trading in A-share market is more active, A-shares has greater trading volume and return volatility. In terms of share prices, B-shares is relatively cheap.

4. Investor clientele and return seasonalities

4.1 Time-series return seasonality

We conduct the following Fama-Macbeth regression to analyze time-series return seasonalities in China stock market.

$$r_{i,t} = \Sigma_{k \in \mathcal{K}} \mu_k Dummy_{t,k} + \varepsilon_{i,t} \tag{1}$$

$\mathcal{K}_{intraday} = \{Overnight, 9{:}30 - 10{:}00, \ldots, 14{:}30 - 15{:}00\}$, $\mathcal{K}_{daily} = \{Mon, \ldots, Fri\}$, $\mathcal{K}_{monthly} = \{Jan, Feb, \ldots, Dec\}$, $Dummy_{t,k}$ is the dummy that equals 1 if interval t belongs to period k. The t-statistics is adjusted by Newey and West (1987) correction with 12 lags for monthly regression (1 year), 5 lags for daily regression (1 week) and 10 lags for intraday regression (1 day).

(Insert Table 2 here)

Stock returns at monthly, daily and intraday frequency show obvious time-series seasonality in both A-shares and B-shares. At monthly level, we find that February returns are significantly positive and are higher than returns of the other month, returns in June and December returns are low, which is consistent with previous findings (Wen, Lin, Cao, Zhang and Yin, 2021). At daily level, early-in-the-week returns are relatively high, while the returns on Thursday and Friday is low, which is obviously different from the U.S. market. (Ali and Ulku, 2020; Luo and Tian, 2020; Zhang, Lai and Lin, 2017). The returns in the first half-hour of trading is extremely high, in addition, overnight return were negative, reflecting the short selling pressure caused by the unique "T+1" policy at the market open (Qiao and Dam, 2020). A strong and similar time-series seasonality effect exists in China A- and B-share stock market.

4.2 Time series return seasonalities under a “twin-market” setting

To empirically test our first hypothesis that time-series return seasonlities should be stronger when the market is more dominated by noise traders, we compare return seasonalities between A- and B-shares twins in our sample. We conduct the following

Fama-Macbeth regression on twin stocks to compare the difference in return seasonalities.

$$r_{i,t} = \alpha + \mu_k Dummy_{t,k} + \gamma_k Dummy_{t,k} * B_i + X_t + \varepsilon_{i,t} \tag{2}$$

$Dummy_{t,k}$ equals to 1 for the calendar period we are interested. At monthly level, we include dummies for February (**Feb**), June and December (**Jun &Dec**). At daily level, we include dummy for early-of-the-week (**Early**) that equals to 1 if the weekday is between Monday to Tuesday and dummy later-of-the-week (**Later**) that equals 1 if the weekday is between Thursday to Friday. At intraday level, we include dummies for overnight (**OV**), first half hour of trading at the morning open (**P1**). B is the dummy that equals to 1 if this is a B-share stock. Control variables $X_{i,t}$ include all other calendar dummies. The t-statistics is adjusted by Newey and West (1987) correction with 12 lags for monthly regression, 5 lags for daily regression and 10 lags for intraday regression.

(Insert Table 3 here)

Table 3 suggests that the time-series return seasonalitie are similar among A-share and B-shares of dual-share firms. The dummies of February, early-of-the-week and intraday have positive coefficients in regressions of both A-shares and B-shares. In contrast, the dummies of June and December, late-of-the-week and overnight have negative coefficients in regressions of both A-shares and B-shares. The time-series return seasonalities are stronger in A-share twins. Interaction terms of $Dummy_{t,k}$ dummy and B dummy always have coefficients of reversed sign with respective to the coefficients of $Dummy_{t,k}$. This suggest that time-series return seasonalities are stronger in A-share among twin stocks.

4.3 The existence of cross-sectional return seasonalities

To investigate the existence of seasonality and seasonal reversals in the cross-sectional stock returns, we conduct a standard methodology as in Heston and Sadka (2008, 2010) and Heston et al. (2010). The methodology is to regress current return $r_{i,t}$ on l period lag returns $r_{i,t-1} \sim r_{i,t-k}$ as the following Fama-Macbeth regression.

$$r_{i,t} = \alpha_{k,t} + \Sigma_{k=1}^{l} b_{k,t}\, r_{i,t-k} + \varepsilon_{k,t} \tag{3}$$

Where $r_{i,t}$ denotes current monthly, daily and intraday returns, coefficient $b_{k,t}$ is called return responses. They also consider single regressions as follows:

$$r_{i,t} = \alpha_{k,t} + b_{k,t}\, r_{i,t-k} + \varepsilon_{k,t}$$

In their studies, the single regression and multifactor regression has similar results, so we only report multiple regression result here, for multiple regressions can partly diminish biases caused by missing fixed effects as in Kamstra (2022). We separately run regressions in all A-shares and all B-shares at all frequency, we take $k = 60$ simply to present the trend and results of longer horizons will be reported in our next section.

(Insert Figure 1 here)

As left panel of Figure 1a presents, the monthly level regression has a periodical pattern in return responses in A-shares, the length of the cycle is exactly 1 year. Both seasonality and seasonal reversal exist. Return responses at annual lags are not significantly negative as those in non-annual lags. The pattern can persist for at least 60 months. But the periodical pattern is not that clear in B-shares as right panel of Figure 1a shows. Left panel of Figure 1b shows that seasonality is not obviously existed at daily level in A-shares, but we can see a strong reversal in the first 25 lags. This is quite different from the results in other stock markets. For example, in U.S. stock market, daily seasonality and seasonal reversals exist and can persist for a long time (Keloharju, Linnainmaa and Nyberg, 2016). This might due to overlap of a strong short-term reversal in China stock market. (Chui et al, 2010, 2022) Also, we can't see a clear periodical patterns in B-shares as right Panel of Figure 1b presents. Figure 1c represents a strong periodical pattern at the intraday level in A-shares, but again, this periodical pattern is not clear in B-shares.

To empirically test whether seasonality and seasonal reversal exist in a longer horizon, we use averages of past same-calendar and other-calendar returns to predict future cross-sectional return with the following Fama-Macbeth regression:

$$r_{i,t} = \alpha_t + \ b_t\, \bar{u}_{i,t} + X_{i,t} + \varepsilon_{k,t} \tag{4}$$

Where $\bar{u}_{i,t}$ denotes averages of past returns, it can be: (a) average of all past returns during formation period (**All**), (b) average of past same-calendar return (**Same-calendar**), (c) average of past other-calendar return (**Other-calendar**), (d) the difference of same-calendar average and other-calendar average (**Dif**). $X_{i,t}$ denotes firm characteristics including past 1-month return, past 2-to-12-month return, and past 13-to-60-month return, size and book-to-market ratio To test the persistency of the predictability, multiple formation periods are aligned here, we take year (-1), year (-3,-2), year (-5,-3), year (-10,-5) and year(-10,-1) as time window, we separately run regression (4) in all A-shares and all B-shares. The regression result is as follows:

(Insert Table 4 here)

We can see clear evidence of seasonality and seasonal reversal exist in both A-shares and B-shares from Table 4. From Panel A, reversal effect exists in A-shares when considering all returns in each formation period as in column **All**. A significant seasonality exists especially in loner horizon, coefficients in column **Same-calendar** are all positive. Also, we can see seasonal reversals as negative coefficients in column **Other-calendar** suggests, and as column **Dif** shows the difference between same-calendar and other-calendar can positively predict future returns. In comparison, the seasonalities in B-shares stocks are not that significant especially in the near term.

In daily level regression in Panel B, both seasonality and seasonal reversal exist in A-shares, but seasonality is stronger in the near term, seasonal reversal is stronger in past 5 to 10 years. And the difference between the same and other-calendar average returns can also positively predict future daily returns. We can see similar patterns exist in B-shares, but the seasonalities are not that significant. Again, the seasonalities in year (-1) is almost insignificant. Similar patterns exist in intraday level regression in Panel C.

In conclusion, both seasonality and seasonal reversal exist in A-shares and B-shares and it can persist for more than 120 months. But the cross-sectional seasonalities are not that significant in B-shares. Seasonality and seasonal reversal are not driven by certain individual stocks, in 25 portfolios sorted by Size and Book-to-Market ratio, the seasonality and seasonal reversal still exists (See Table B1). And surprisingly, the phenomenon goes stronger recently, in Table B2, we compare result during pre-2010 period and post-2010 period, we find stronger pattern in the later period.

4.4 Cross-sectional return seasonalities

We further examine the difference of cross-sectional return seasonalities between A- and B-share twin stocks. We focus on the predictability of **Same-calendar** and **Other-calendar** returns during the past 120 months. We separately run regressions in A-shares and B-shares of the dual-share firms (represented by **A|AB** and **B|AB**) and calculate the difference between the coefficient (represented by **A|AB – B|AB**), also, we report the difference between coefficients of all A-shares and all B-shares (represented by **A-B**).

(Insert Table 5 here)

Table 5 shows a markedly reduced seasonality and seasonal reversals in type B stocks relative to type A stocks of dual-share firms. As column (a) and column (b) shows, seasonality is very significant in A-share stocks, but not that significant in B-share stocks. Both mean and t value of coefficient in column (a) is bigger than that in column(b). As Panel A shows, and the seasonality at monthly and daily level is almost vanished. The differences between in A-share and B-share are significantly positive in column (c) and column (d). The seasonal reversal is also very significant in A-share stocks, but the significance falls in B-share market, the seasonal reversal is completely reversed at intraday level. And the coefficient differences between A-share and B-share are significantly negative, which suggest that seasonal reversal is stronger in A-share stocks. The above results confirm the hypothesis 2 that cross-sectional seasonalities should be stronger in A-shares.

5. Further discussion of cross-sectional return seasonalities

5.1 Evidence on seasonal institutional trading

As in this study, we assume seasonalities are generated by seasonal trading of institutions, in this section, we provide some direct evidence on seasonal institutional trading. As in Lou et al. (2019), the relation between the yearly change of institutional ownership (△ IO) and the contemporaneous returns can reflect the timing when institutions would like to trade. If institutions trade mostly in June and December, then a positive relation should be observed between △ IO and June and December returns. For returns in daily and intraday, we cumulate them into yearly data and figure out their relation to yearly △ IO.

(Insert Table 6 here)

The results above suggest that institutions do like to trade seasonally. Institutions are more likely to trade June and December and are less likely to trade in February. They are more likely to trade in the late-of-the-week and overnight. And are less likely to trade in intraday, especially in first half hour of trading.

5.2 The seasonal pattern of market demand

We also try to explore whether seasonal trading is accompanied with fixed but different trading preferences as well. Referencing Gao, Han, Li, Xiong (2019), we use the stock returns during a certain time interval as a proxy for investors' concentrated buying and selling. We sort stocks into 10 groups by their returns in each calendar month, day and period, the decile with highest return is called winners and decile with lowest returns are called losers. We graphically show the characteristic mean and differences between winners and losers in Figure B1, B2, B3 in both A-share market and B-share market.

The results suggest that investors are trading seasonally, and this seasonality are similar in both A-share and B-share markets. In June and December, the trading seems to be more rational, they tend to trade stocks with high quality (big size, high profitability). In February, speculative investors are more active, February winners usually have speculative characteristics (high volume, high idiosyncratic volatility, low share prices). At daily frequency, investors prefer to trade speculative stocks in early-of-the-week, and trade stock with high quality on late-of-the-week. At intraday frequency, most overnight winners are high-quality stocks, which may be related to China's 10-minute opening call auction, which attracts mostly institutions to trade during this period. Stocks during intraday, especially those around the morning and noon opening, are highly speculative.

The above results suggest that investors with different preferences are active at heterogeneous calendar periods, and this clientele is similar in both A-share and B-share markets. The active trading of speculative stocks is corresponded to the highly positive expected returns in February, early-of-the-week and the first half hour of trading. In contrast, the active trading of high-quality stocks is corresponded to the highly negative expected returns in June, late-of-the-week and the overnight. This confirms our model that expected returns should be positive (negative) when the infrequently rebalanced investor is absent (active).

5.3 Time-varying clientele difference

The clientele difference of A-share and B-share market is time-varying for the new arrival of retail investors. To make sure the time-varying change in investor clientele is also affective to return seasonalities, we use the difference between number of newly created retail accounts (NRA) in A-shares and B-shares as a proxy for the clientele differences. We split the sample by the median of NRA difference between A-shares and B-shares (Δ NRA = NRA of A-shares - NRA of B-shares), and we expect that the differences of return seasonalities should be stronger when Δ NRA is large. We conduct regression analysis on pooled A-share and B-share stocks of dual-share firms as follows:

$$r_{i,t} = \alpha + \gamma_{i,t} * Dummy_{k,t} * A_i * High_t + X_{i,t} + \varepsilon_{i,t} \tag{5}$$

Where $High_t$ is the dummy that equals 1 when △NRA is above the time-series median. The control variable $X_{i,t}$ includes the first order and second order intersections of $Dummy_{k,t}$, A_i and $High_t$.

(Insert Table 7 here)

The time-varying clientele difference △NRA can cause a difference between return seasonalities in A-shares and B-shares. When retail investors in A-share market are more active than those in B-share market, the returns in Februrary, early-of-the-week and the intraday are more positive than that of B-shares, also, the returns in June, December and overnight is more negative in A-shares than that in B-shares. This further confirms that time-series return seasonalities are affected by the activeness of retail investors.

To test whether △NRA can affect cross-sectional return seasonalities in A-share and B-shares, We split the dual-share sample by the median of △NRA, and run the following regression respectively:

$$r_{i,t} = \alpha + \gamma_t * \bar{u}_{i,t} * A_i + X_{i,t} + \varepsilon_{i,t} \tag{6}$$

Where $\bar{u}_{i,t}$ represents R_{same} and R_{other}. The control variable $X_{i,t}$ includes $\bar{u}_{i,t}$, A_i.

(Insert Table 8 here)

Table 8 suggests that when △NRA is high, the cross-sectional return seasonality difference between A-share and B-shares is getting bigger. For monthly return, the coefficient of R_{Same}*A is 0.163 (t =2.04), the coefficient of R_{other}*A is -0.284 (t = -1.19) when △NRA is above the median, these two coefficients represent the magnitude of difference between cross-sectional return seasonalities between A-share and B-share. When △NRA is below the median, the coefficient of R_{Same}*A is -0.027 (t =-0.42), the coefficient of R_{other}*A is -0.181 (t = -0.64) , the magnitude of seasonality differences shrinks when clientele differences is not prominent. The same pattern exist in daily and intraday level data, the magnitude of coefficient γ_t is greater when ΔNRA is above the median.

5.4 Seasonalities in trading volume and volatility

In this section, we focus on the seasonalities in trading volume and volatility. This is intuitive as return seasonalities is a trading-driven pricing anomaly. First, we focus on time-series properties of return volatility and trading volume. By model derivation, we find time-series seasonalities exist in return volatility and trading volume.

Hypothesis 3: Return volatility and trading volume should be higher when infrequent trader is active. And when the market has more noise traders, the presence of infrequent can cause greater return volatility and trading volume.
(Please see Appendix for proof).

And we do a model simulation according to our model to generate autocovariances of return volatility and trading volume.

Hypothesis 4: Same-calendar return volatility (trading volume) should have positive autocovariances; other-calendar return volatility (trading volume) should have negative autocovariances. And this phenomenon is more obvious when the market has more noise traders.

(Insert Figure 2 here)

Finally, we focus on the real data to check whether these hypothesis from our model is true. We test these results in real data.

(Insert table 9 here)

The time-series regression shows that in June and December, the return volatility and trading volume is higher than the other months, and this phenomenon is stronger in A-share of dual-share firms. Similarly, return volatility and trading volume is higher in Later-of-the-week and overnight and the effect is also more prominent in A-shares.

(Insert table 10 here)

The cross-sectional regression shows that same-calendar return volatility can positively predicts future, and the predictability is stronger in A-shares of dual-share firms. Other-calendar return volatility can negatively predicts the future, and the predictability is stronger in A-share as well. The same is true for trading volume.

6. Conclusion

In this paper, we thoroughly examine the effect of investor clientele on time-series and cross-sectional return seasonalities at monthly, daily and intraday level. First, we build a model with seasonal active investor to generate return seasonalities. In this model, we find the expected returns should be positive (negative) when the seasonal trading investor is absent (active). We define this as time-series return seasonality. Also, with the persistency of trading, same-calendar returns covariates positively, and other-calendar returns covariates negatively. We define this as cross-sectional return seasonalities. We make empirical prediction about the dominance of retail investor with this framework, when the market is more dominated by noise traders like in A-share market, both time-series and cross-sectional return seasonalities should be more pronounced.

We check empirically whether performance of return seasonalities is in accordance with our model. We find similar time-series return seasonalities exist in A-and B-share market, returns in February, early-of-the-week and first half hour of trading is highly positive, this is exactly when institutions in China are relatively absent and market demand is concentrated on buying speculative stocks. In contrast, returns in June, late-of-the-week and overnight is negative, this is exactly when institutions in China are relatively active and market demand is concentrated on buying high quality stocks. Next, we turn to the cross-sectional return seasonalities. We find seasonality and seasonal reversal exist in both A-share and B-share markets. Besides, the phenomenon can persist for over 120 months. In terms of the role of retail investors, we use a "twin-market" setting of A-share and B-share stocks issued by dual-share firms. Both time-series and cross-sectional return seasonalities at multiple frequencies in A-share have greater magnitude.

As additional supportive evidence, we focus on the time-varying clientele difference between A- and B-shares. Empirical result suggests that when retail investors in A-shares are more active than those in B-shares, seasonality gap between the two markets becomes wider. Further, return volatility and trading volume has the same seasonal patterns and are stronger in A-shares, which is in accordance with our model prediction.

This study provides evidence on the role of retail investors to return seasonalities. It's worth to note that dominance of retail investors is not the culprit to all types of mispricing, for example, the momentum and PEAD. In this way, revealing the effect of retail investors on anomalies can help understand their formation and can also provide useful insight into return predictability.

7. References

Ali, F., & Ülkü, N. (2020). Weekday seasonality of stock returns: The contrary case of China. *Journal of Asian Economics*, 68, 101201.

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, 21(2), 129–152.

Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, 159–178.

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

Bernard, V. L., & Thomas, J. K. (1990). Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting And Economics*, 13(4), 305–340.

Blair, W., & Co. (2018). Why china a, why now? *Perspectives on Global Equity.*

Bogousslavsky, V. (2016). Infrequent rebalancing, return autocorrelation, and seasonality. *The Journal of Finance*, 71(6), 2967–3006.

Chan, K. C., Chen, N.-f., & Hsieh, D. A. (1985). An exploratory investigation of the firm size effect. *Journal of Financial Economics,* 14(3), 451– 471.

Chen, D., Wu, K., & Zhu, Y. (2022). Stock return asymmetry in china. *Pacific-Basin Finance Journal*, 73, 101757.

Chui, A. C., Subrahmanyam, A., & Titman, S. (2022). Momentum, reversals, and investor clientele. *Review of Finance*, 26(2), 217–255.

Cross, F. (1973). The behavior of stock prices on fridays and mondays. *Financial Analysts Journal*, 29(6), 67–69.

Gao, Y., Han, X., Li, Y., & Xiong, X. (2019). Overnight momentum, informational shocks, and late informed trading in china. *International Review of Financial Analysis*, 66, 101394.

Gao, Y., Han, X., Li, Y., & Xiong, X. (2021). Investor heterogeneity and momentum-based trading strategies in china. *International Review of Financial Analysis*, 74, 101654.

Gillain, C., Ittoo, A., & Lambert, M. (2019). News-induced style seasonality.

Goetzmann, W. N., Kim, D., Kumar, A., & Wang, Q. (2015). Weather induced mood, institutional investors, and stock returns. *The Review of Financial Studies*, 28(1), 73–111.

Golder, S. A., & Macy, M. W. (2011). Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, 333(6051), 1878–1881.

Gu, M., Kang, W., & Xu, B. (2018). Limits of arbitrage and idiosyncratic volatility: Evidence from China stock market. *Journal of Banking & Finance*, 86, 240–258.

Han, C., & Shi, Y. (2022). Chinese stock anomalies and investor sentiment. *Pacific-Basin Finance Journal*, 73, 101739.

He, J., Ng, L., & Wang, Q. (2004). Quarterly trading patterns of financial institutions. *The Journal of Business*, 77(3), 493–509.

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

Heston, S. L., & Sadka, R. (2008). Seasonality in the cross-section of stock returns. *Journal of Financial Economics*, 87(2), 418–445.

Heston, S. L., & Sadka, R. (2010). Seasonality in the cross section of stock returns: The international evidence. *Journal of Financial and Quantitative Analysis*, 45(5), 1133–1160.

Hirshleifer, D., Lim, S. S., & Teoh, S. H. (2011). Limited investor attention and stock market misreactions to accounting information. *The Review of Asset Pricing Studies*, 1(1), 35–73.

Hirshleifer, D. A., Peng, L., & Wang, Q. (2021). Social networks and market reactions to earnings news. Available at SSRN 3824022.

Hu, G., McLean, R. D., Pontiff, J., & Wang, Q. (2014). The year-end trading activities of institutional investors: Evidence from daily trades. *The Review of Financial Studies*, 27(5), 1593–1614.

Kamstra, M. J., Kramer, L. A., & Levi, M. D. (2000). Losing sleep at the market: The daylight saving anomaly. *American Economic Review*, 90(4), 1005–1011.

Kamstra, M. J., Kramer, L. A., & Levi, M. D. (2003). Winter blues: A sad stock market cycle. *American Economic Review*, 93(1), 324–343.

Kedar-Levy, H., Yu, X., Kamesaka, A., & Ben-Zion, U. (2010). The impact of daily return limit and segmented clientele on stock returns in China. *International Review of Financial Analysis*, 19(4), 223–236.

Kelley, E. K., & Tetlock, P. C. (2013). How wise are crowds? insights from retail orders and stock returns. *The Journal of Finance*, 68(3), 1229– 1265.

Keloharju, M., Linnainmaa, J. T., & Nyberg, P. (2016). Return seasonalities. *The Journal of Finance*, *71*(4), 1557-1590.

Keloharju, M., Linnainmaa, J. T., & Nyberg, P. (2021). Are return seasonalities due to risk or mispricing? *Journal of Financial Economics*, 139(1), 138–161.

Kim, D. (2006). On the information uncertainty risk and the January effect. *The Journal of Business*, 79(4), 2127–2162.

Liu, J., Stambaugh, R. F., & Yuan, Y. (2019). Size and value in China. *Journal of Financial Economics*, 134(1), 48–69.

Lockwood, L. J., & McInish, T. H. (1990). Tests of stability for variances and means of overnight/intraday returns during bull and bear markets. *Journal of Banking & Finance*, 14(6), 1243–1253.

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

Luo, K., & Tian, S. (2020). The "black Thursday" effect in Chinese stock market. *Journal of Behavioral and Experimental Finance*, 27, 100367.

Masulis, R. W., & Ng, V. K. (1995). Overnight and daytime stock-return dynamites on the london stock exchange: The impacts of "big bang" and the 1987 stock-market crash. *Journal of Business & Economic Statistics*, 13(4), 365–378.

Miller, E. M. (1988). Why a weekend effect? *Journal of Portfolio Management*, 14(4), 43.

Murphy, D. P., & Thirumalai, R. S. (2017). Short-term return predictability and repetitive institutional net order activity. *Journal of Financial Research*, 40(4), 455–477.

Ning, Y. (2009). Research on the measurement of investor sentiment in Chinese stock market: The CICSI's construction. *Journal of Financial Research*, 11, 014.

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

Rogalski, R. J., & Tinic, S. M. (1986). The January size effect: Anomaly or risk mismeasurement? *Financial Analysts Journal*, 42(6), 63–70.

Rozeff, M. S., & Kinney Jr, W. R. (1976). Capital market seasonality: The case of stock returns. *Journal of Financial Economics*, 3(4), 379–402.

Thaler, R. H. (1987). Anomalies: The January effect. *Journal of Economic Perspectives*, 1(1), 197–201.

Tinic, S. M., & West, R. R. (1984). Risk and return: January vs. the rest of the year. *Journal of Financial Economics*, 13(4), 561–574.

Ülkü, N., & Rogers, M. (2018). Who drives the Monday effect? *Journal of Economic Behavior & Organization*, 148, 46–65.

Xiong, X., Meng, Y., Li, X., & Shen, D. (2019). An empirical analysis of the adaptive market hypothesis with calendar effects: Evidence from China. *Finance Research Letters*, 31.

Zhang, J., Lai, Y., & Lin, J. (2017). The day-of-the-week effects of stock markets in different countries. *Finance Research Letters*, 20, 47–62.

Table 1 Descriptive analysis of stock characteristics

Table 1 presents the variable construction and descriptive statistics of our sample. Panel A resents the construction methodology, Panel B is the descriptive statistics. Full sample includes all stocks issued on SSE and SZSE. AB includes all dual-share stocks.

Panel A Variable construction

Variable	Construction
LnME	Logarithm of market size. For dual-share firms, the market size is the sum of A-share market size and B-share market size.
BtM	Book-to-market ratio.
RET(-1)	Stock returns in the past 1 month.
RET(-12,-2)	Cumulative stock returns between the past 2 to 12 months.
RET(-60,-13)	Cumulative stock returns between the past 13 to 60 months.
ROE	Return on Equity, Net profit / Total equity.
IVOL	Idiosyncratic volatility estimated by Fama-French 3 factor model.
VOLUME	Total volume in 100 million CNY.
PRC	Monthly average of share prices in CNY.

Panel B Data description

	Full sample				AB			
	A-shares		B-shares		A-shares		B-shares	
#Firms	3156		109		81		81	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
LnME	22.409	22.228	22.525	22.429	22.662	22.532	22.662	22.532
BtM	0.637	0.649	0.711	0.760	0.686	0.719	0.686	0.719
ROE	0.029	0.035	0.318	0.031	0.324	0.033	0.324	0.033
VOLUME	0.027	0.024	0.022	0.020	0.027	0.025	0.022	0.019
IVOL	0.017	0.015	0.015	0.013	0.018	0.016	0.015	0.013
PRC	14.930	10.880	3.946	1.730	12.999	10.150	3.844	1.577

Table 2 Time-series return seasonalities of both A-shares and B-shares

Table 2 presents regression coefficient and t statistics of following Fama-Macbeth regression:

$$r_{i,t} = \Sigma_{k\in\mathcal{K}}\mu_k Dummy_{t,k} + \varepsilon_{i,t}$$

$\mathcal{K}_{monthly} = \{Jan, Feb, \dots, Dec\}$, $\mathcal{K}_{daily} = \{Mon, Tues, \dots, Fri\}$, $\mathcal{K}_{intraday} = \{Overnight, 9{:}30 - 10{:}00, \dots, 14{:}30 - 15{:}00\}$

$Dummy_{t,k}$ is the dummy that equals 1 if interval t belongs to period k. The t-statistics is adjusted by Newey and West (1987) correction with 12 lags for monthly regression, 5 lags for daily regression and 10 lags for intraday regression. Panel A reports monthly level results. Panel B reports daily level results and Panel C reports intraday level results. We use the whole market sample for calculation, the sample period is from Jul. 1999 to Dec. 2019.

Panel A Monthly level regression

Calendar month		Jan.	Feb.	Mar.	Apr.	May.	Jun.
A-shares	Mean	0.79%	6.44%	4.17%	1.57%	1.00%	-4.56%
	T stat	0.5	9.03	3.19	1.3	0.98	-3.78
B-shares	Mean	3.57%	3.79%	0.42%	2.98%	-0.02%	-5.04%
	T stat	1.47	4.96	0.48	1.75	-0.03	-4.87
Calendar month		Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
A-shares	Mean	0.52%	0.38%	-0.36%	0.56%	2.35%	-1.49%
	T stat	0.34	0.42	-0.45	0.55	3.35	-0.79
B-shares	Mean	0.22%	-2.08%	2.80%	2.06%	1.98%	-1.61%
	T stat	0.16	-2.21	2.63	1.95	1.61	-1.19

Panel B Daily level regression

Weekday		Mon.	Tue.	Wed.	Thu.	Fri.
A-shares	Mean	0.256%	0.115%	0.048%	-0.125%	-0.041%
	T stat	2.36	1.55	0.73	-1.65	-0.34
B-shares	Mean	0.111%	0.015%	0.171%	0.005%	0.103%
	T stat	0.94	0.18	2.09	0.01	0.99

Panel C Intraday level regression

Period		Overnight	9:30-10:00	10:00-10:30	10:30-11:00	11:00-11:30
A-shares	Mean	-1.649%	1.702%	0.166%	-0.033%	-0.006%
	T stat	-8.35	8.39	1.01	-1.79	-0.84
B-shares	Mean	-0.395%	0.420%	-0.009%	-0.037%	-0.025%
	T stat	-7.61	12.28	-0.78	-5.07	-3.56
	Period	Noon break	13:00-13:30	13:30-14:00	14:00-14:30	14:30-15:00
A-shares	Mean	0.018%	0.018%	0.025%	0.004%	0.013%
	T stat	2.43	2.54	3.63	0.36	0.60
B-shares	Mean	0.394%	-0.010%	0.002%	-0.015%	0.071%
	T stat	7.57	-1.42	0.35	-2.23	0.62

Table 3 Different time-series return seasonalities between A-shares and B-shares

We conduct the following Fama-Macbeth regression on twin stocks to compare the difference in return seasonalities.

$$r_{i,t} = \alpha + \mu_k Dummy_{t,k} + \gamma_k Dummy_{t,k} * B_i + X_t + \varepsilon_{i,t} \quad (2)$$

$Dummy_{t,k}$ equals to 1 for the calendar period we are interested. At monthly level, we include dummies for February (**Feb**), June and December (**Jun** &**Dec**). At daily level, we include dummy for early-of-the-week (**Early**) that equals to 1 if the weekday is between Monday to Tuesday and dummy later-of-the-week (**Later**) that equals 1 if the weekday is between Thursday to Friday. At intraday level, we include dummies for overnight (**OV**), first half hour of trading at the morning open (**P1**). B is the dummy that equals to 1 if this is a B-share stock. Control variables $X_{i,t}$ include all other calendar dummies. The t-statistics is adjusted by Newey and West (1987) correction with 12 lags for monthly regression, 5 lags for daily regression and 10 lags for intraday regression.

Monthly level			Daily level			Intraday level		
	Coeff.	T stat.		Coeff.	T stat.		Coeff.	T stat.
Feb	6.452%	9.07	**Early**	1.515%	2.15	**OV**	-3.938%	-14.43
Jun&Dec	-4.150%	-1.98	**Later**	0.013%	0.02	**P1**	1.644%	11.62
Feb*B	-2.835%	-7.43	**Early*B**	-1.080%	-2.49	**OV*B**	4.182%	17.01
Jun&Dec*B	0.784%	1.82	**Later*B**	0.761%	2.75	**P1*B**	-0.057%	-0.47
Controls	√	√	**Controls**	√	√	**Controls**	√	√

Table 4 Cross-sectioanl return seasonalities of both A-shares and B-shares

Table 4 reports time-series mean and t value of regression coefficient in regression (4).

$$r_{i,t} = \alpha_t + b_t\, \bar{u}_{i,t} + X_{i,t} + \varepsilon_{k,t} \tag{4}$$

Where $\bar{u}_{i,t}$ denotes averages of past returns, it can be: (a) average of all past returns during formation period (**All**), (b) average of past same-calendar return (**Same-calendar**), (c) average of past other-calendar return (**Other-calendar**), (d) the difference of same-calendar average and other-calendar average (**Dif**). $X_{i,t}$ denotes firm characteristics including past 1-month return, past 2-to-12-month return, and past 13-to-60-month return, size and book-to-market ratio. To test the persistency of the predictability, multiple formation periods are aligned here, we take year (-1), year (-3,-2), year (-5,-3), year (-10,-5) and year(-10,-1) as time window, we separately run regression (4) in all A-shares and all B-shares. Panel A reports result of monthly regression, consider same and other calendar month, Panel B reports result of daily regression, consider same and other calendar day, Panel C reports result of intraday regression, consider same and other calendar intraday period. The sample period is July 1999 to December 2019.

Panel A Monthly return prediction

	A-shares								B-shares							
	(a) All		(b) Same-calendar		(c) Other-calendar		(d) Dif		(a) All		(b) Same-calendar		(c) Other-calendar		(d) Dif	
Year	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t
year (-1)	-0.064	-2.7	0.005	0.91	-0.073	-3.3	0.019	3.57	-0.042	-0.61	0.017	1.37	-0.026	-0.44	0.021	1.62
year (-3,-2)	-0.061	-2.52	0.005	0.63	-0.061	-2.79	0.014	1.87	-0.147	-2.08	-0.006	-0.30	-0.153	-2.32	0.017	1.10
year (-5,-3)	-0.056	-3.35	0.010	1.57	-0.058	-3.78	0.015	2.95	-0.066	-1.06	-0.008	-0.63	-0.076	-1.21	0.000	-0.02
year (-10,-5)	-0.037	-2.03	0.017	2.53	-0.047	-2.68	0.019	3.06	-0.151	-1.15	0.054	2.16	-0.181	-1.64	0.067	2.89
year(-10,-1)	-0.111	-3.37	0.023	2.06	-0.118	-4.15	0.036	3.64	-0.425	-2.24	0.040	1.54	-0.391	-3.02	0.070	2.64

Panel B Daily return prediction

Year	A-shares								B-shares							
	(a) All		(b) Same-calendar		(c) Other-calendar		(d) Dif		(a) All		(b) Same-calendar		(c) Other-calendar		(d) Dif	
	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t
year (-1)	-0.002	-0.13	0.026	4.03	-0.012	-0.94	0.023	5.10	-0.002	-0.06	0.025	1.60	-0.034	-0.75	0.033	2.56
year (-3,-2)	-0.025	-1.55	0.003	0.39	-0.022	-1.34	0.007	1.37	-0.121	-2.44	-0.015	-0.71	-0.098	-2.29	0.012	0.68
year (-5,-3)	-0.042	-3.01	0.006	0.95	-0.043	-4.20	0.015	3.55	-0.033	-0.71	-0.034	-1.44	-0.020	-0.38	-0.007	-0.35
year (-10,-5)	-0.013	-1.25	0.012	1.83	-0.031	-2.98	0.011	2.48	-0.048	-0.89	0.004	0.16	-0.040	-0.85	0.015	0.83
year(-10,-1)	-0.009	-0.43	0.029	3.43	-0.050	-2.94	0.040	4.56	-0.199	-1.90	-0.009	-0.24	-0.209	-3.20	0.047	1.60

Panel C Intraday return prediction

Year	A-shares								B-shares							
	(a) All		(b) Same-calendar		(c) Other-calendar		(d) Dif		(a) All		(b) Same-calendar		(c) Other-calendar		(d) Dif	
	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t	Mean	t
year (-1)	0.095	8.96	0.198	42.1	-1.541	-24.77	0.102	37.99	0.297	11.14	-0.270	-0.48	0.150	0.43	-0.206	-0.48
year (-3,-2)	0.921	14.91	0.151	32.01	-0.833	-5.14	0.129	18.83	0.165	5.65	0.257	16.32	-0.053	-1.47	0.192	13.95
year (-5,-3)	-0.395	-6.38	0.212	16.72	-0.674	-4.67	0.048	15.46	0.546	9.89	0.234	16.61	0.428	6.80	0.162	14.59
year (-10,-5)	-0.479	-5.25	0.019	12.47	-0.093	-3.19	0.097	7.28	0.008	3.44	0.196	13.11	0.292	7.80	0.018	2.63
year(-10,-1)	0.227	12.55	0.291	32.92	-0.084	-6.39	0.058	18.54	0.027	6.40	-0.263	-0.47	0.170	9.47	0.035	16.15

Table 5 Seasonality and seasonal reversals in dual-share stocks

Table 5 reports time-series mean and t value of regression coefficient b_t in regression (4).

$$r_{i,t} = \alpha_t + b_t \bar{u}_{i,t} + X_{i,t} + \varepsilon_{k,t} \quad (4)$$

Where $\bar{u}_{i,t}$ denotes averages of past returns, it can be: average of same-calendar returns during the past 120 month (Same-calendar) as shown in Panel A. Or the average of other-calendar returns during the past 120 month (Other-calendar) as shown in Panel B. "Monthly" consider same-calendar and other-calendar month in daily regression, "daily" consider same day-of-the-week and other-day-of-the-week in a daily regression, "intraday" consider same time-of-the-day and other time-of-the-day returns in an intraday regression. "A|AB" represents the A-shares of the dual-share firms, "B|AB" represents the B-shares of the dual-share firms, and the "A|AB-B|AB" and "A-B" represents difference of coefficients between A-shares and B-shares in dual-share firms and whole sample, respectively.

Panel A Same-calendar

Freq.	(a) A|AB		(b) B|AB		(c) A|AB-B|AB		(d) A-B	
	Mean	t	Mean	t	Mean	t	Mean	t
Monthly	0.0880	3.51	0.0176	0.65	0.0846	2.49	0.0266	0.95
Daily	0.1098	2.50	-0.0298	-0.64	0.1269	2.40	0.0741	1.57
Intraday	0.2808	14.20	0.1351	7.03	0.1406	5.21	0.0456	2.31

Panel B Other-calendar

Freq.	(a) A|AB		(b) B|AB		(c) A|AB-B|AB		(d) A-B	
	Mean	t	Mean	t	Mean	t	Mean	t
Month	-0.4356	-3.92	-0.3252	-2.69	-0.2268	-1.68	-0.0866	-0.77
Daily	-0.2712	-3.53	-0.2108	-2.78	-0.0854	-0.84	-0.2395	-2.78
Intraday	-0.1596	-7.34	0.1406	12.14	-0.3219	-12.27	-1.5250	-13.08

Table 6 Trading time preferences of institutions

Table 6 reports Fama-Macbeth regression result of regressing yearly institutional ownership changes(△ IO) on contemporaneous yearly cumulative returns. For example, in column "Feb", the dependent variable is annual change of institutional ownership and the independent variable is contemporaneous Feb return. In column "Non", the dependent variable is contemporaneous Non-Feb return. Column "Dif" reports the coefficient difference between "Feb" and "Non". "Early" represents cumulative returns from Monday to Wednesday, and "P1&P5" is the cumulative return of the first half hour of trading in both the morning and the afternoon. We sort portfolios according to Institutional ownership, "5" with the highest, "1" with the lowest. "5-1" reports the coefficient differences between high IO and low IO.

Panel A Monthly trading patterns

	A-shares			B-shares		
	Feb	Non	Dif	Feb	Non	Dif
1	-0.005	0.027	-0.033	-0.886	0.086	-0.973
	-0.42	3.58	-2.43	-1.47	0.51	-1.31
5	0.055	0.072	-0.017	-0.794	0.05	-0.844
	3.38	6.6	-1.42	-0.91	1.59	-0.93
"5-1"	0.06	0.045		0.043	-0.033	
	2.73	3.97		0.04	-0.19	
	Jun&Dec	Non	Dif	Jun&Dec	Non	Dif
1	0.027	0.022	0.005	-0.071	-0.068	-0.004
	3.18	3.17	0.5	-0.18	-0.66	-0.01
5	0.076	0.065	0.011	0.014	-0.005	0.019
	7.67	7.8	1.06	0.3	-0.21	0.52
"5-1"	0.048	0.043		0.086	0.062	
	3.96	4.02		0.22	0.56	

Panel B Daily trading patterns

	Early	Non	Dif	Early	Non	Dif
1	0.02	0.025	-0.005	-0.006	0.271	-0.278
	3.97	4.31	-1.26	-0.03	0.92	-0.75
5	0.053	0.059	-0.006	-0.098	0.245	-0.343
	6.66	4.6	-0.54	-1.06	0.98	-1.02
"5-1"	0.034	0.034		-0.097	-0.01	
	3.48	2.6		-0.35	-0.03	

Panel C Intraday trading patterns						
	Overnight	Non	Dif	Overnight	Non	Dif
1	0.003	0.008	-0.005	0.04	-0.036	0.076
	0.37	1.58	-0.08	0.97	-0.95	0.96
5	0.04	0.033	0.007	0.003	-0.001	0.003
	2.43	6.09	0.4	1.23	-0.34	0.82
"5-1"	0.037	0.025		-0.037	0.035	
	2.08	3.14		-0.89	0.92	
	First	Non	Dif	First	Non	Dif
1	0.005	0.008	-0.003	-0.019	0.052	0.138
	1.03	0.98	-0.46	-0.63	1.29	0.92
5	0.038	0.043	-0.004	-0.212	-0.35	-0.071
	5.64	5.49	-0.71	-1.48	-1.89	-1.39
"5-1"	0.033	0.035		0.191	0.405	
	3.51	2.6		1.23	2.1	

Table 7 Time-varying clientele effect on time-series return seasonalities

Table 7 reports the regression result includes 3rd order intersection of △NRA, A-share dummy and calendar time dummy. The △NRA represents difference between number of newly created retail accounts in A-shares and number of newly created retail accounts in B-shares. The regression is as follows:

$$r_{i,t} = \alpha + \gamma_{i,t} * Dummy_{k,t} * A_i * High_t + X_{i,t} + \varepsilon_{i,t} \tag{5}$$

Where $High_t$ is the dummy that equals 1 when △NRA is above the time-series median. The control variable $X_{i,t}$ includes the first order and second order intersections of $Dummy_{k,t}$, A_i and $High_t$.

(1)		(2)		(3)	
Feb*A*High	0.0586 5.46	Early*A*High	0.0015 7.38	OV*A*High	-0.0035 -13.60
Jun&Dec*A*High	-0.0138 -1.28	Later*A*High	-0.0075 -0.67	P1*A*High	0.0028 13.29
Control	√	Control	√	Control	√

Table 8 Time-varying clientele effect on cross-sectional return seasonalities

Table 8 reports the cross-sectional return seasonalities between high and low Δ NRA. We split the dual-share firm stock sample by the median of Δ NRA, and run the following regression respectively:

$$r_{i,t} = \alpha + \gamma_{i,t} * \bar{u}_{i,t} * A_i + X_{i,t} + \varepsilon_{i,t} \tag{6}$$

Where $\bar{u}_{i,t}$ represents R_{same} and R_{other}, respectively. The control variable $X_{i,t}$ includes $\bar{u}_{i,t}$, A_i.

	Monthly		Daily		Intraday	
	High	Low	High	Low	High	Low
R_{Same}*A	0.163	-0.027	-0.009	0.027	1.879	0.022
	2.04	-0.42	-0.16	0.47	3.38	0.67
R_{other}*A	-0.284	-0.181	-0.232	-0.081	-1.202	-0.120
	-1.19	-0.64	-2.03	-0.7	-4.31	-3.32

Table 9 Time-series seasonalities of return volatility and trading volume

We run regression (7) on all the dual-share stocks to compare the time-series volatility and volume seasonality difference.

$$Dep_{i,t} = \alpha + \beta_{i,t} * Dummy_{k,t} + \gamma_{i,t} * Dummy_{k,t} * B + \varepsilon_{i,t} \qquad (7)$$

Here $Dep_{i,t}$ can be return volatility or trading volume (in shares), $Dummy_{k,t}$ refers to a series of dummies that equals to 1 for the calendar period k. At monthly level, we include dummies for June and December (**Jun&Dec**). At daily level, we include dummy for late-of-the-week (**Later**) that equals to 1 if the weekday is between Thursday to Friday. At intraday level, we include dummies for overnight (**OV**). B is the dummy that equals to 1 if this is a B-share stock.

Panel A Monthly regression			
		Jun&Dec	**Jun&Dec*B**
Vol	coeff	3.398%	-3.797%
	T	4.79	-3.80
Volume	coeff	0.848	-1.872
	T	5.89	-8.91
Panel B Daily regression			
		Later	**Later*B**
Vol	coeff	0.740%	-1.540%
	T	34.16	-66.59
Volume	coeff	0.644	-1.230
	T	70.87	-126.97
Panel C Intraday regression			
		OV	**OV*B**
Vol	coeff	0.080%	-0.290%
	T	10.88	-27.78
Volume	coeff	0.001	-0.002
	T	151.67	-178.50

Table 10 Cross-sectional seasonalities of return volatility and trading volume

Table 10 reports time-series mean and t value of regression coefficients in regression (8) for all the dual-share stocks.

$$v_{i,t} = \alpha_t + b_t \bar{v}_{i,t} + c_t \bar{v}_{i,t} * B + \varepsilon_{k,t} \quad (8)$$

Where $v_{i,t}$ represents return volatility or trading volume (in 10 million shares), $\bar{v}_{i,t}$ denotes averages of past volatility or trading volume, it can be: average of same-calendar v during the past 120 month ($\boldsymbol{V_{same}}$). Or the average of other-calendar v during the past 120 month ($\boldsymbol{V_{other}}$).

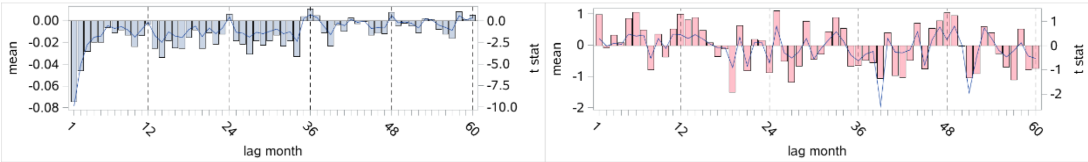
		Monthly		Daily		Intraday	
		Volatility	Volume	Volatility	Volume	Volatility	Volume
$\boldsymbol{V_{same}}$	coeff	0.195	1.030	1.361	2.845	0.037	0.030
	T	0.51	1.91	1.51	1.92	0.51	1.91
$\boldsymbol{V_{other}}$	coeff	-0.144	-1.448	-0.137	-1.457	-0.144	-0.973
	T	-0.29	-2.70	-2.28	-2.71	-3.29	-2.70
$\boldsymbol{V_{same}}$*B	coeff	0.010	0.097	-0.029	-0.101	0.010	0.097
	T	0.31	0.17	-0.32	-1.39	0.31	4.17
$\boldsymbol{V_{other}}$*B	coeff	0.103	0.832	0.115	1.229	0.103	0.832
	T	3.07	1.43	2.73	3.28	9.73	8.37

Figure 1 Cross-sectional seasonality in monthly, daily and intraday level

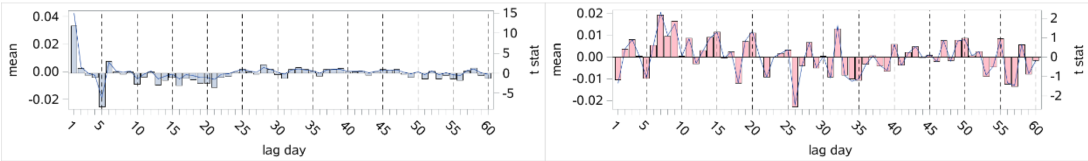
Figure 1 presents estimates of return responses in equation (3):

$$r_{i,t} = \alpha_{k,t} + \Sigma_{k=1}^{l} b_{k,t}\, r_{i,t-k} + \varepsilon_{k,t} \tag{3}$$

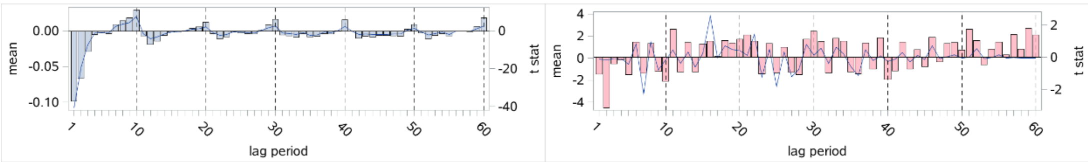
Where $r_{i,t}$ denotes current monthly, daily and intraday returns, $r_{i,t-k}$ is the lag k period return, coefficient $b_{k,t}$ is called return responses. We separately run regressions in all A-shares and all B-shares at all frequency, we take $k = 60$ simply to present the trend. Figure 1a/1b/1c reports monthly/daily/intraday level results, respectively. The solid lines represent time-series average of return responses, the corresponding y axis is “mean” on the left. The blue bar represents time-series t value of return responses, the corresponding y axis is “t stat” on the right. The sample period is from July 1999 to December 2019. The left panel results of A-shares and the right panel is results of B-shares.



1a Seasonality in monthly cross-sectional returns



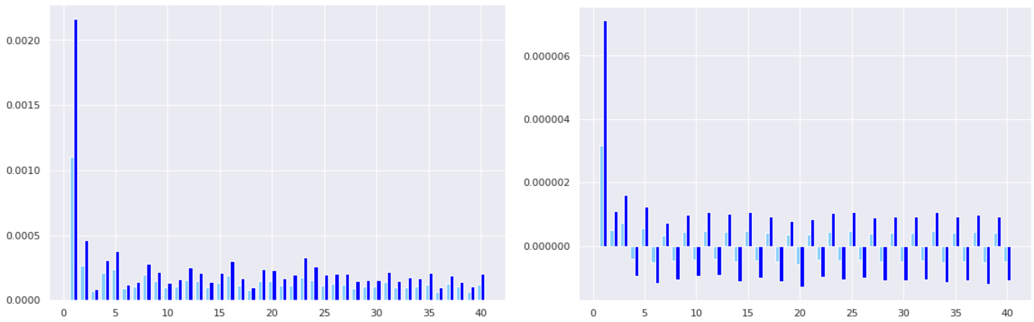
1b Seasonality in daily cross-sectional returns



1c Seasonality in intraday cross-sectional returns

Figure 2 Simulated autocorrelations of return volatility and trading volume

Figure 2 presents the autocorrelations of return volatility and trading volume in our model. Deep parameters we used for simulation are: α_1= 0.97, α_2 = 0.7, $\bar{d}$ = 1%, σ_d = 1%, $\bar{\mu}$ = 1%, σ_u = 2%, K = 2000. The x-axis in the following graphs is lag periods, the blue bar represents autocorrelations when N = 1 and the navy bar represents autocorrelations when N = 1.5.

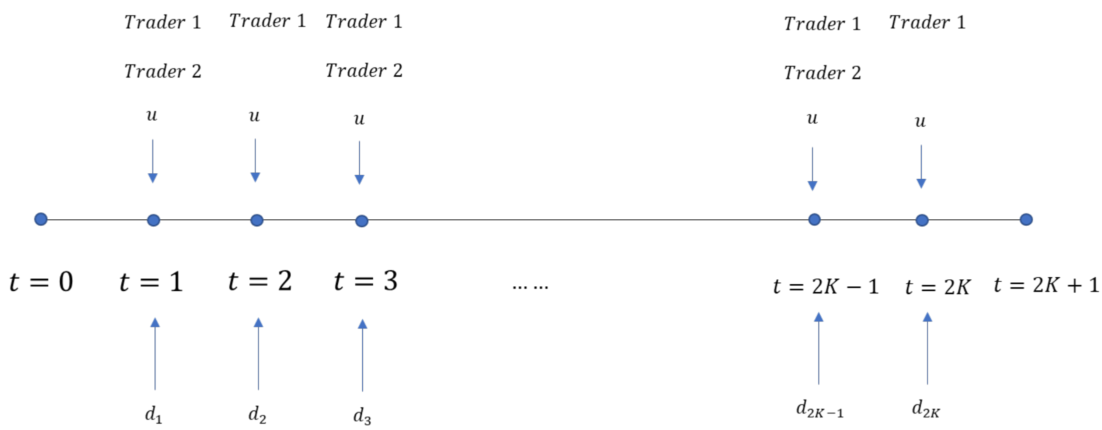


(a) Autocorrelations of return volatility (b) Autocorrelations of trading volume

Appendix A

1. The economy

Consider an economy with discrete time periods of $T = 0,1,2,\dots,2K+1$, $t = 0$ is the initial time period when no dividend is paid. $T = 2K+1$ is the terminal time period that all dividends are revealed. A risky asset exists in this economy, the asset pays off dividend $d_i = \bar{d}_i + \sigma\varepsilon_i$, $\varepsilon_i \sim N(0,1)$ and ε_i is i.i.d in each period i $(i \in \{1,2,\dots,2K+1\})$. $P_{2K+1} = d_1 + \cdots + d_{2K} + d_{2K+1}$. For simplicity, we assume risk-free return is 0. A frequent trader (trader 1) exists and would trade at each time period $T = 1,\dots,2K$. The frequent trader is analog to the market maker presents to clear the market demand. Trader 1 has CARA utility with risk-aversion of α_1, the asset demand of trader 1 is denoted by $x_{1,t}$. A seasonal trading investor (trader 2) exists and would trade only at odd time period $t = 2k-1$ $(1 \le k \le K+1)$. In this setting, "same-calendar" periods refer to odd vs. odd periods and even vs. even periods. "Other-calendar" periods refer to odd vs. even periods. The infrequent trader also has a CARA utility with risk-aversion of α_2, his demand of risky-asset is $x_{2,t}$. There are N noise traders exist in $T = 1,\dots,2K$, the net demand of each noise trader is u. Assume $E(u) = \bar{u} > 0$ due to short-selling constraint in China. N represent the dominance of noise traders.



At $t = 2(K-i)$, the frequent trader would face the problem of maximize his utility.

$$\max_{x_{1,2(K-i)}} \; E\left(W_{2(K-i)+1}\right) - \frac{\alpha_1}{2} Var\left(W_{2(K-i)+1}\right)$$

$$s.t.\, W_{2(K-i)+1} = W_{2(K-i)} + x_{1,2(K-i)}\left(P_{2(K-i)+1} - P_{2(K-i)}\right)$$

Plug the constraint equation into objective function, we have:

$$x_{1,2(K-i)} = \frac{E\left[P_{2(K-i)+1} \middle| \mathcal{F}_{2(K-i)}\right] - P_{2(K-i)}}{2 * \left(\frac{\alpha_1}{2}\right) * (2i+1)\sigma^2}$$

By market clearing

$$x_{1,2(K-i)} + N * u = 0$$

We get

$$P_{2(K-i)} = E[P_{2(K-i)+1}|\mathcal{F}_{2(K-i)}] - \alpha_1(2i+1)N\sigma^2 u \qquad \text{(A1)}$$

At $t = 2(K-i)-1$, similarly, the frequent trader's problem is:

$$\max_{x_{1,2(K-i)-1}} E\big(W_{2(K-i)}\big) - \frac{\alpha_1}{2} Var(W_{2(K-i)})$$

$$s.t. W_{2(K-i)} = W_{2(K-i)-1} + x_{1,2(K-i)-1}\big(P_{2(K-i)} - P_{2(K-i)-1}\big)$$

Plug the constraint equation into objective function, we have:

$$x_{1,2(K-i)-1} = \frac{E[P_{2(K-i)}|\mathcal{F}_{2(K-i)-1}] - P_{2(K-i)-1}}{2 * \left(\frac{\alpha_1}{2}\right) * (2i+2)\sigma^2}$$

And at this time, frequent trader joins the market, as he only cares his utility in $2(K-i)+1$, he faces the problem as follows:

$$\max_{x_{2,2(K-i)-1}} E\big(W_{2(K-i)+1}\big) - \frac{\alpha_2}{2} Var(W_{2(K-i)+1})$$

$$s.t. W_{2(K-i)+1} = W_{2(K-i)-1} + x_{2,2(K-i)-1}\big(P_{2(K-i)+1} - P_{2(K-i)-1}\big)$$

Plug the constraint equation into objective function, we have:

$$x_{2,2(K-i)-1} = \frac{E[P_{2(K-i)+1}|\mathcal{F}_{2(K-i)-1}] - P_{2(K-i)-1}}{2 * \left(\frac{\alpha_2}{2}\right) * (2i+2)\sigma^2}$$

The demand of both traders should satisfy the following market clearing equation:

$$x_{1,2(K-i)-1} + x_{2,2(K-i)-1} + N * u = 0 \qquad \text{(A2)}$$

2. Solution

Proposition 1: We solve the prices of risky asset P_t in this economy by guess and verification, we claim the prices are:

$$P_{2(K-i)-1} = \Sigma_{j=1}^{2(K-i)-1} d_j + \Sigma_{j=2(K-i)}^{2K+1} \bar{d}_j + (i+1)(2i+3)\alpha_2[\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2 u]$$

$$P_{2(K-i)} = \Sigma_{j=1}^{2(K-i)} d_j + \Sigma_{j=2(K-i)+1}^{2K+1} \bar{d}_j + (2i+1)[\alpha_1 + (i+1)\alpha_2][\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2 u]$$

$P_{2(K-i)}$, $P_{2(K-i)-1}$ satisfy equation (A1) and (A2).

Proof:

Equation (A1):

NTS: $P_{2(K-i)} = E[P_{2(K-i)+1}|\mathcal{F}_{2(K-i)}] - \alpha_1(2i+1)N\sigma^2 u$

The right-hand side is:

$$d_1 + \cdots + d_{2(K-i)} + \bar{d}_{2(K-i)+1} + \cdots + \bar{d}_{2K+1} + i(2i+1)\alpha_2[\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2 u] - (2i+1)\alpha_1 N\sigma^2 u$$

$= d_1 + \cdots + d_{2(K-i)} + \bar{d}_{2(K-i)+1} + \cdots + \bar{d}_{2K+1} + [i(2i+1)\alpha_2 + (2i+1)(\alpha_1 + \alpha_2)][\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2 u]$

$= d_1 + \cdots + d_{2(K-i)} + \bar{d}_{2(K-i)+1} + \cdots + \bar{d}_{2K+1} + (2i+1)[\alpha_1 + (i+1)\alpha_2][\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2 u]$

$= P_{2(K-i)}$

Equation (A2):

$$\text{NTS: } \frac{E[P_{2(K-i)}|\mathcal{F}_{2(K-i)-1}]-P_{2(K-i)-1}}{2*\left(\frac{\alpha_1}{2}\right)*(2i+2)\sigma^2}+\frac{E[P_{2(K-i)+1}|\mathcal{F}_{2(K-i)-1}]-P_{2(K-i)-1}}{2*\left(\frac{\alpha_2}{2}\right)*(2i+2)\sigma^2}+N*u=0$$

The left-hand side is:

$$\frac{E\left[P_{2(K-i)}\middle|\mathcal{F}_{2(K-i)-1}\right]-P_{2(K-i)-1}}{2*\left(\frac{\alpha_1}{2}\right)*(2i+2)\sigma^2}+\frac{E[P_{2(K-i)+1}|\mathcal{F}_{2(K-i)-1}]-P_{2(K-i)-1}}{2*\left(\frac{\alpha_2}{2}\right)*(2i+2)\sigma^2}$$

$$=\frac{(2i+1)[\alpha_1+(i+1)\alpha_2][\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2u]-(i+1)(2i+3)\alpha_2\left[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2u\right]}{2*\left(\frac{\alpha_1}{2}\right)*(2i+2)\sigma^2}+\frac{i(2i+1)\alpha_2[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2u]-(i+1)(2i+3)\alpha_2[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2u]}{2*\left(\frac{\alpha_2}{2}\right)*(i+2)\sigma^2}$$

$$=\left(\frac{[(2i+1)\alpha_1-(2i+2)\alpha_2]}{\alpha_1}+\frac{[-4i-3]\alpha_2}{\alpha_2}\right)\frac{\left[\frac{\alpha_1}{\alpha_1+\alpha_2}Nu\right]}{(2i+2)}$$

$$=-\frac{2(1+i)(\alpha_1+\alpha_2)}{\alpha_1}*\frac{\left[\frac{\alpha_1}{\alpha_1+\alpha_2}Nu\right]}{(2i+2)}$$

$$=-N*u$$

The equilibrium is unique as F.O.C.s for every period are linear.

3. Comparative analysis:

Returns:

$$r_{2(K-i)+2}=P_{2(K-i)+2}-P_{2(K-i)+1}=\varepsilon_{2(k-i)+2}+[(2i-1)\alpha_1-2i\alpha_2][\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2u]$$

$$r_{2(K-i)+1}=P_{2(K-i)+1}-P_{2(K-i)-1}=\varepsilon_{2(k-i)+1}-(2i+1)(\alpha_1+\alpha_2)[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2u]$$

$$r_{2(K-i)}=P_{2(K-i)}-P_{2(K-i)-1}=\varepsilon_{2(k-i)}+[(2i+1)\alpha_1-(2i+2)\alpha_2][\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2u]$$

$$r_{2(K-i)-1}=P_{2(K-i)-1}-P_{2(K-i)-2}=\varepsilon_{2(k-i)-1}-(2i+3)(\alpha_1+\alpha_2)[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2u]$$

Demands:

$$x_{1,2(K-i)+1}=\frac{(2i-1)\alpha_1-2i\alpha_2}{2*\left(\frac{\alpha_1}{2}\right)*2i\sigma^2}[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2u]$$

$$x_{2,2(K-i)+1}=\frac{(i-1)\alpha_1-(4i-1)\alpha_2}{2*\left(\frac{\alpha_2}{2}\right)*2i*\sigma^2}[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2u]$$

$$x_{1,2(K-i)}=\frac{-(1+i)\alpha_1-(1+2i)\alpha_2}{2*\left(\frac{\alpha_1}{2}\right)*(2i+1)\sigma^2}[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2u]$$

$$x_{1,2(K-i)-1}=\frac{(1+2i)\alpha_1-2(1+i)\alpha_2}{2*\left(\frac{\alpha_1}{2}\right)*(2i+2)\sigma^2}[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2u]$$

$$x_{2,2(K-i)-1}=\frac{i\alpha_1-(3+4i)\alpha_2}{2*\left(\frac{\alpha_2}{2}\right)*(2i+2)\sigma^2}[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2u]$$

Time-series seasonalities:

1. Expected returns w/o the presence of infrequently trading investor (even period) is as follows.

$$E\left[r_{2(K-i)}\right] = [(2i+1)\alpha_1 - (2i+2)\alpha_2][\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2\bar{u}]$$

2. Expected returns w/i the presences of infrequently trading investor (odd period) is as follows.

$$E\left[r_{2(K-i)+1}\right] = -(2i+1)(\alpha_1+\alpha_2)\left[\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2\bar{u}\right] < 0$$

Due to the presence of equity premium in China, average return (in excess of risk-free return) in real stock market data is positive. This need $E\left[r_{2(K-i)}\right] > 0$ for some periods to avoid deriving negative equity premium in the equilibrium, which is equal to $\exists I$, s.t. $\frac{\alpha_1}{\alpha_2} > \frac{2i+1}{2i}$ for $i > I$.

Seasonalities and seasonal reversals:

The autocovariances between same-calendar period returns are, for $\forall i, j \in \{1,2,\dots,K\}$ and $i < j$:

$$Cov\left(r_{2(K-i)+1}, r_{2(K-j)+1}\right) = (2i+1)(2j+1)(\alpha_1+\alpha_2)^2\left[\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2\sigma_\mu\right]^2 > 0$$

$$Cov\left(r_{2(K-i)}, r_{2(K-j)}\right) = [(2i+1)\alpha_1 - (2i+2)\alpha_2][(2j+1)\alpha_1 - (2j+2)\alpha_2]\left[\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2\sigma_\mu\right]^2 > 0$$

The autocovariances between other-calendar period returns are, for $\forall i, j \in \{1,2,\dots,K\}$ and $i < j$:

$$Cov\left(r_{2(K-i)+1}, r_{2(K-j)}\right) = -(2i+1)(\alpha_1+\alpha_2)[(2j+1)\alpha_1 - (2j+2)\alpha_2]\left[\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2\sigma_\mu\right]^2 < 0$$

$$Cov\left(r_{2(K-i)}, r_{2(K-j)+1}\right) = -(2j+1)(\alpha_1+\alpha_2)[(2i+1)\alpha_1 - (2i+2)\alpha_2]\left[\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2\sigma_\mu\right]^2 < 0$$

Predictions about investor clientele:

The number of noise traders N would affect the covariances. When N is close to 0, all these covariances would be close to 0. As N grows larger, we would see larger covariances in their absolute value. In this way, we predict that the dominance of noise traders would magnify return seasonalities.

Hypothesis 1: Time-series seasonality should be stronger in A-share market than B-share market, as A-share market is dominated by noise traders.

$$\frac{\partial E\left[r_{2(K-i)}\right]}{\partial N} = [(2i+1)\alpha_1 - (2i+2)\alpha_2]\left[\frac{\alpha_1}{\alpha_1+\alpha_2}\sigma^2\bar{u}\right] > 0,$$

$$\frac{\partial E\left[r_{2(K-i)+1}\right]}{\partial N} = -(2i+1)(\alpha_1+\alpha_2)\left[\frac{\alpha_1}{\alpha_1+\alpha_2}\sigma^2\bar{u}\right] < 0$$

Hypothesis 2: The seasonalities and seasonal reversals should be stronger in A-share market than in B-share market.

For $\forall i, j \in \{1,2,\dots,K\}$ and $i < j$:

$$\frac{\partial Cov\left(r_{2(K-i)+1}, r_{2(K-j)+1}\right)}{\partial N} = 2N(2i+1)(2j+1)(\alpha_1+\alpha_2)^2\left[\frac{\alpha_1}{\alpha_1+\alpha_2}\sigma^2\sigma_\mu\right]^2 > 0$$

$$\frac{\partial Cov\left(r_{2(K-i)}, r_{2(K-j)}\right)}{\partial N} = 2N[(2i+1)\alpha_1 - (2i+2)\alpha_2][(2j+1)\alpha_1 - (2j+2)\alpha_2]\left[\frac{\alpha_1}{\alpha_1+\alpha_2}\sigma^2\sigma_\mu\right]^2 > 0$$

$$\frac{\partial Cov\left(r_{2(K-i)+1}, r_{2(K-j)}\right)}{\partial N} = -2N(2i+1)(\alpha_1+\alpha_2)[(2j+1)\alpha_1 - (2j+2)\alpha_2]\left[\frac{\alpha_1}{\alpha_1+\alpha_2}\sigma^2\sigma_\mu\right]^2 < 0$$

$$\frac{\partial Cov\left(r_{2(K-i)}, r_{2(K-j)-1}\right)}{\partial N} = -2N(2j+1)(\alpha_1+\alpha_2)[(2i+1)\alpha_1-(2i+2)\alpha_2]\left[\frac{\alpha_1}{\alpha_1+\alpha_2}\sigma^2\sigma_\mu\right]^2 < 0$$

Hypothesis 3: Return volatility and trading volume should be higher when infrequent trader is active. And when the market has more noise traders, the presence of infrequent can cause greater return volatility and trading volume.

According to the expressions of returns, we can derive the following return volatility:

$$Var[r_{2(k-i)+1}] = \sigma^2 + (2i+1)^2(\alpha_1+\alpha_2)^2\left[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2\right]^2\sigma_u^2$$

$$Var\left[r_{2(k-i)}\right] = \sigma^2 + [(2i+1)\alpha_1-(2i+2)\alpha_2]^2\left[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2\right]^2\sigma_u^2$$

$$Var[r_{2(k-i)-1}] = \sigma^2 + (2i+3)^2(\alpha_1+\alpha_2)^2\left[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2\right]^2\sigma_u^2$$

We find return volatility in odd periods are greater than in even period:

$$Var\left[r_{2(k-i)-1}\right] - Var\left[r_{2(k-i)}\right] = (4(1+i)\alpha_1+\alpha_2)(2\alpha_1+(5+4i)\alpha_2)\left[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2\right]^2\sigma_u^2 > 0$$

$$Var\left[r_{2(k-i)+1}\right] - Var\left[r_{2(k-i)}\right] = (3+4i)\big((2+4i)\alpha_1-\alpha_2\big)\alpha_2\left[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2\right]^2\sigma_u^2 > 0$$

This return difference should be greater when the market has more noise traders.

$$\frac{\partial\left[Var\left[r_{2(k-i)-1}\right] - Var\left[r_{2(k-i)}\right]\right]}{\partial N} > 0$$

$$\frac{\partial\left[Var\left[r_{2(k-i)+1}\right] - Var\left[r_{2(k-i)}\right]\right]}{\partial N} > 0$$

And we derive the expected trading volume as follows: (Note if $u_t \sim N(0,1)$, $E[u_{t1}-u_{t2}] = \sqrt{\frac{2}{\pi}}$).

$$E\left[Volume_{2(k-i)}\right] = E\left[\frac{\left|x_{1,2(k-i)}-x_{1,2(k-i)-1}\right| + N*\left|u_{2(k-i)}-u_{2(k-i)-1}\right|}{2}\right]$$

$$= E\left[\frac{\left|\left[\frac{-(1+i)\alpha_1-(1+2i)\alpha_2}{2*\left(\frac{\alpha_1}{2}\right)*(2i+1)\sigma^2} - \frac{(1+2i)\alpha_1-2(1+i)\alpha_2}{2*\left(\frac{\alpha_1}{2}\right)*(2i+2)\sigma^2}\right]\left[\frac{\alpha_1}{\alpha_1+\alpha_2}N\sigma^2 u\right]\right|}{2}\right] + \frac{1}{2}\sqrt{\frac{2}{\pi}}$$

$$= \left\{\left|\frac{3+8i+6i^2}{2(1+i)(1+2i)\sigma^2}\right|\left[\frac{\alpha_1}{\alpha_1+\alpha_2}\sigma^2\right] + 1\right\}N\sqrt{\frac{2}{\pi}}$$

$$E[Volume_{2(k-i)+1}] = E\left[\frac{\left|x_{1,2(k-i)+1}-x_{1,2(k-i)}\right| + \left|x_{2,2(k-i)+1}-x_{2,2(k-i)-1}\right| + N*\left|u_{2(k-i)+1}-u_{2(k-i)-1}\right|}{2}\right]$$

$$= E\left[\frac{\left|\left[\frac{(2i-1)\alpha_1 - 2i\alpha_2}{2 * \left(\frac{\alpha_1}{2}\right) * 2i\sigma^2} - \frac{-(1+i)\alpha_1 - (1+2i)\alpha_2}{2 * \left(\frac{\alpha_1}{2}\right) * (2i+1)\sigma^2}\right]\left[\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2 u\right]\right|}{2}\right]$$

$$+ E\left[\frac{\left|\left[\frac{(i-1)\alpha_1 - (4i-1)\alpha_2}{2 * \left(\frac{\alpha_2}{2}\right) * 2i * \sigma^2} - \frac{i\alpha_1 - (3+4i)\alpha_2}{2 * \left(\frac{\alpha_2}{2}\right) * (2i+2)\sigma^2}\right]\left[\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2 u\right]\right|}{2}\right] + \frac{1}{2}\sqrt{\frac{2}{\pi}}$$

$$= E\left[\frac{\left|\left[\frac{-1+2i+6i^2}{2(i+2i^2)\sigma^2}\right]\left[\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2 u\right]\right|}{2}\right] + E\left[\frac{\left|\left[\frac{-\alpha_1+\alpha_2}{2(i+i^2)\sigma^2\alpha_2}\right]\left[\frac{\alpha_1}{\alpha_1+\alpha_2} N\sigma^2 u\right]\right|}{2}\right] + \frac{1}{2}\sqrt{\frac{2}{\pi}}$$

$$= \{\frac{-1+2i+6i^2}{2(i+2i^2)\sigma^2} + \frac{\alpha_1-\alpha_2}{2(i+i^2)\sigma^2\alpha_2}\}\left\{\left[\frac{\alpha_1}{\alpha_1+\alpha_2}\sigma^2\right] + 1\right\} N\sqrt{\frac{2}{\pi}}$$

$$= \frac{3 - \frac{2}{i} + \frac{1}{1+i} + \frac{1}{1+2i} + \frac{\alpha_1}{(i+i^2)\alpha_2}}{2\sigma^2}\left\{\left[\frac{\alpha_1}{\alpha_1+\alpha_2}\sigma^2\right] + 1\right\} N\sqrt{\frac{2}{\pi}}$$

$$E\left[Volume_{2(k-i)+1}\right] - E\left[Volume_{2(k-i)}\right] = \frac{6i + \frac{\alpha_1}{(1+i)\alpha_2} - 1 - \frac{1}{1+2i}}{2i\sigma^2}\left\{\left[\frac{\alpha_1}{\alpha_1+\alpha_2}\sigma^2\right] + 1\right\} N\sqrt{\frac{2}{\pi}} > 0$$

This volume difference between odd and even period should be greater when the market has more noise traders:

$$\frac{\partial\left[E\left[Volume_{2(k-i)+1}\right] - E\left[Volume_{2(k-i)}\right]\right]}{\partial N} > 0$$

(Please see Online appendix for the related Mathematica code).

Appendix B

Fig B1 Characteristics of current winner and loser by calendar month

Figure B1 presents the mean characteristic value of winner and loser stocks in each calendar month. We sort stocks into deciles each month, the decile with highest returns are called winners, the decile with the lowest returns are called losers. We plot the mean of lnME, BtM, ROE, VOLUME, IVOL and PRC of winners in navy solid line for all A-shares (in dark red for all B shares), and those of losers in light blue solid line for all A-shares (in pink for all B-shares). The mean difference between winners and losers are in black dotted line. The axis for these characteristics is named "characteristics" on the left. Also, we plot the t value of mean differences between winners and losers in bar, the axis is named "t stat for W-L" on the right. The blue bar highlight t value in February for A-shares (pink for all B-shares), navy bar highlights t value in June and December for all A-shares (Red for all B-shares).

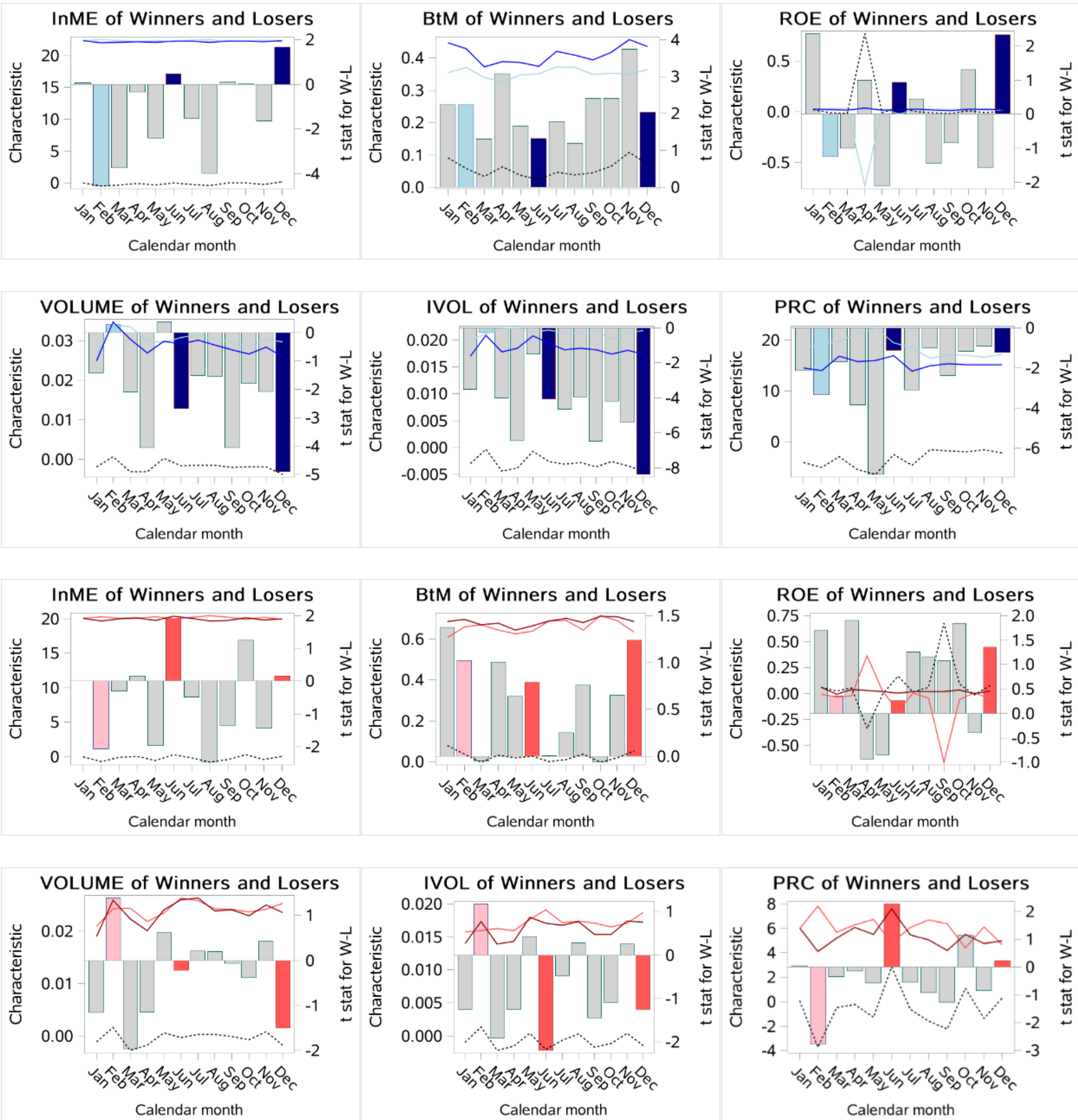


Fig B2 Characteristics of current winner and loser by day of the week

Figure B2 presents the mean characteristic value of winner and loser stocks in each day of the week. We sort stocks into deciles each trading day, the decile with highest returns are called winners, the decile with the lowest returns are called losers. We plot the mean of lnME, BtM, ROE, VOLUME, IVOL and PRC of winners in navy solid line for all A-shares (in dark red for all B shares), and those of losers in light blue solid line for all A-shares (in pink for all B-shares). The mean difference between winners and losers are in black dotted line. The axis for these characteristics is named "characteristics" on the left. Also, we plot the t value of mean differences between winners and losers in bar, the axis is named "t stat for W-L" on the right. The blue bar highlight t value in early-of-the-week for A-shares (pink for all B-shares), navy bar highlights t value in late-of-the-week for all A-shares (Red for all B-shares).

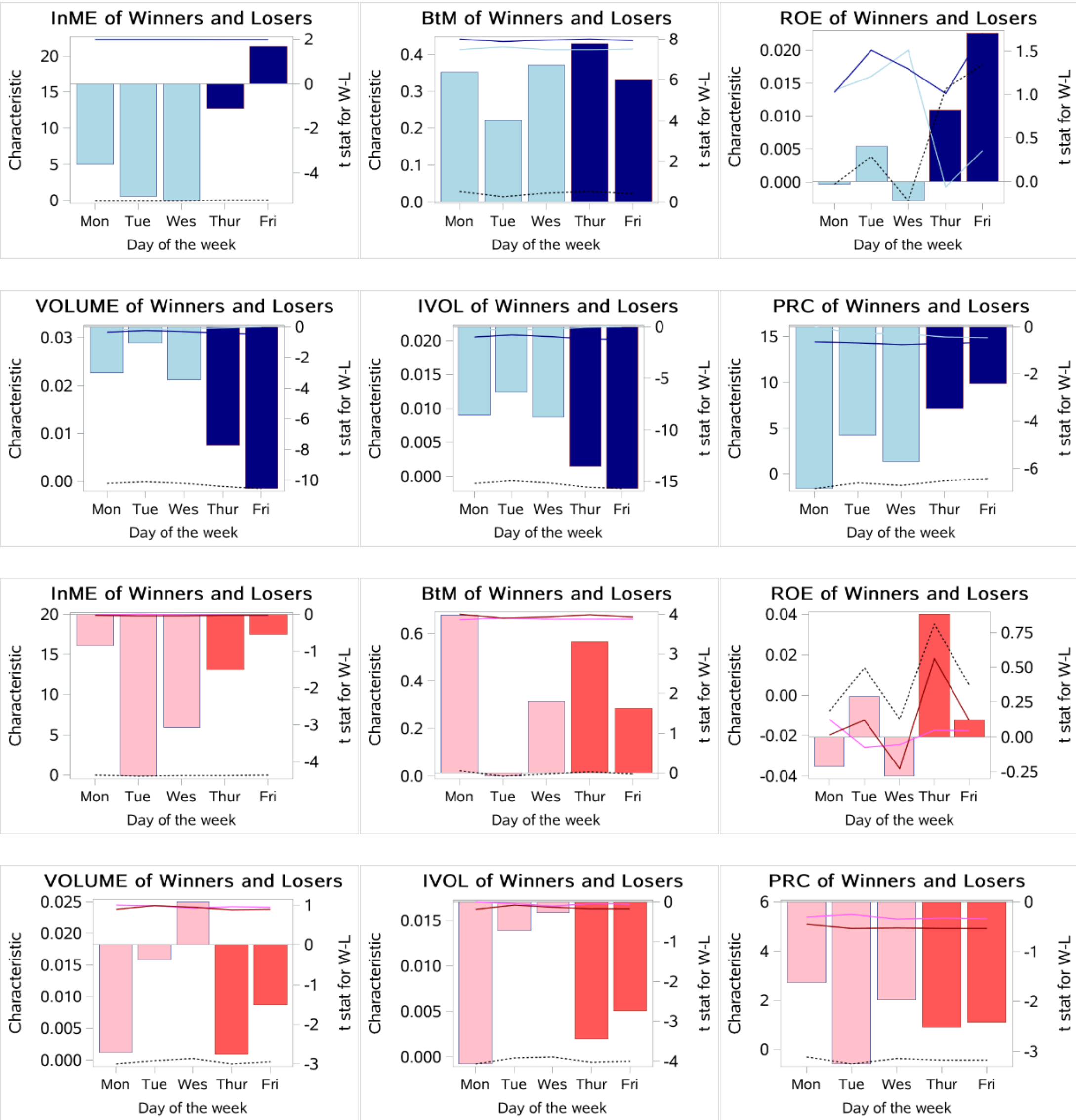


Fig B3 Characteristics of current winner and loser by time of the day

Figure B3 presents the mean characteristic value of winner and loser stocks in each time of the day. We sort stocks into deciles each trading period, the decile with highest returns are called winners, the decile with the lowest returns are called losers. We plot the mean of lnME, BtM, ROE, VOLUME, IVOL and PRC of winners in navy solid line for all A-shares (in dark red for all B shares), and those of losers in light blue solid line for all A-shares (in pink for all B-shares). The mean difference between winners and losers are in black dotted line. The axis for these characteristics is named “characteristics” on the left. Also, we plot the t value of mean differences between winners and losers in bar, the axis is named “t stat for W-L” on the right. The blue bar highlight t value during first half hour of trading in the morning and the afternoon for A-shares (pink for all B-shares), navy bar highlights t value during overnight for all A-shares (Red for all B-shares).