Index-based Investing and Intraday Stock Dynamics*

Yiwen Shen† Meiqi Shi‡

This version: June 2020

Abstract

We investigate how the growth of index-based investing impacts the intraday stock dynamics using a large high-frequency dataset, which consists of 1-second level trade data for all S&P 500 constituents from 2004 to 2018. We estimate intraday trading volume, volatility, correlation, and beta using estimators that are statistically efficient under market microstructure noise and observation asynchronicity. We find the intraday patterns indeed change substantially over time. For example, in the recent decade, the trading volume and correlation significantly increase at the end of trading session; the betas of different stocks start dispersed in the morning, but generally move towards one during the day. Besides, the daily dispersion in trading volume is high at the market open and low near the market close. These intraday patterns demonstrate the implication of the growth of index-based strategies and the active-open, passive-close intraday trading profile. We theoretically support our interpretation via a market impact model with time-varying liquidity provision from both single-stock and index-fund investors.

Keywords: index-based investing, intraday stock dynamics, high-frequency data, efficient estimation

1 Introduction

Understanding the pattern of intraday stock dynamics is an important topic with various practical applications. Portfolio managers executing large orders can reduce the transaction costs by trading in the hours with abundant market liquidity; intraday traders can better exploit price comovement of different stocks during the periods when correlation is high; risk managers can reduce intraday risk by avoiding times with large price fluctuations. In this paper, we will show that the intraday

---

*We would like to thank Paul Glasserman and Harry Mamaysky for their valuable comments. The research of Yiwen Shen is supported by the Deming Center and Chazen Institute at Columbia Business School.
†Address: Graduate School of Business, Columbia University, 3022 Broadway, New York, NY 10027. E-mail: YShen21@gsb.columbia.edu.
‡Address: Haas School of Business, University of California Berkeley, 2220 Piedmont Ave, Berkeley, CA 94720. Email: mqshi@berkeley.edu.
patterns of US stocks have changed in important ways since 2004. For example, we find in the recent
decade, the trading volume and correlation increase significantly near the market close; the betas
of different stocks are dispersed in the morning, but generally move towards one throughout the
day. These patterns demonstrate the substantial implications from passive investment, and more
specifically, the index-based strategies.

In the recent decade, the growth of passive investment and index-based strategies have drawn
great attention from both industry and academia (Appel et al., 2016). The index-based strategies
tend to make investment decisions based on portfolio-level approaches instead of selecting individual
stocks in a discretionary way. For example, index-based strategies include buying or selling all S&P
500 constituents according to their market capitalization; investing in stocks with high or low beta;
buying stocks in specific sectors, etc. These strategies trade multiple stocks in a systematic manner,
and make stocks more likely to move in same directions. On the other hand, it has been widely
noticed by financial press that the index-based strategies from passive investors tend to concentrate
their trading near the market close (Strumpf (2015) and Driebusch et al. (2018)). This behavior
can be explained by, among other reasons, minimizing the tracking errors of orders benchmarked
to the closing price, efficiently buying or selling large number of stocks with market-on-close orders,
and reducing the inventory risk for redemption and creation settlement (see relevant discussion in
Cushing and Madhavan (2000), Foucault et al. (2005), and Wu (2019)).

In this paper, we first use two empirical studies to provide evidence for the growth of passive
investment and the intraday trading pattern of index-based strategies. In the first one, we find the
stocks with high (resp. low) passive ownership have higher trading volume near the market close
(resp. open). In the second one, we show the trading volume tends to drop dramatically near the
market close after a stock is removed from the S&P 500 Index, thus less tracked by the index-based
strategies. These findings suggest an active-open, passive-close intraday trading profile, i.e., more
discretionary (resp. index-based) trading at the market open (resp. close). Such trading profile can
have substantial impact on the intraday patterns of stock dynamics. Accordingly, we propose three
hypotheses that motivate our subsequent studies. First, we expect the correlation between stocks to
be low at the market open, and high at the end of trading session. This is because the index-based
strategies tend to drive multiple stocks to move in same directions. Similarly, we expect the betas
of different stocks to be more dispersed in the morning, but move towards one near the market close
due to index-based orders. Finally, the daily dispersion in trading volume is supposed to be lower
at the end of trading session, as the trading from institutional investors, who execute most of the
index-based strategies, is shown to be highly persistent across days (Campbell et al., 2009).

The impact of institutional and passive investment on stock trading, at both the intraday and
overnight level, has been a popular topic in recent years. Karolyi et al. (2012) and Koch et al. (2016)
provide empirical evidence showing that trading from ETFs and index funds contributes to the
commonality in daily trading volume. Heston et al. (2010) suggest that institutional fund flows and trading algorithms can generate periodicity in intraday returns and volumes. Subsequent work by Bogousslavsky (2016) and Gao et al. (2018) shows, both theoretically and empirically, that delayed portfolio rebalancing from institutional investors leads to positive correlation between the returns in the last and first half hours of adjacent trading sessions. Our work complements and extends this line of research by revealing the implication of passive investment from other important aspects in high-frequency setting, including intraday correlation, beta, and volume dispersion. We show the intraday patterns indeed changed substantially over years. In particular, the three hypotheses hold in our large dataset, especially in the recent decade during which the passive investment has become more prevalent.

To estimate intraday patterns, high-frequency data plays an indispensable role. With the development of financial technologies, there are growing applications of high-frequency data in various fields of finance. We list some examples below among many others in this rich area. For intraday volatility, Andersen and Bollerslev (1997) analyze the intraday periodicity in volatility and its impact on return dynamics; Andersen et al. (2001) test potential pattern shift in intraday volatility with variance-ratio statistics. Some literature on volatility forecasting with high-frequency data and its applications can be found in Andersen et al. (2003), Hansen et al. (2012), Stroud and Johannes (2014), and Liu et al. (2018). Intraday trading volume is also widely studied. For example, Kappou et al. (2010) analyze how the addition of a stock to an index impacts the trading volume and return on adjacent days; Min et al. (2018) develop a time-varying liquidity model based on intraday trading volume pattern and study its impact on optimal portfolio execution. Some recent work studies the estimation of covariance with high-frequency data and demonstrates its benefit in portfolio allocation. Boudt and Zhang (2015) show that an equal-risk portfolio constructed from jump-robust intraday covariance estimation delivers higher return and lower risk than traditional equal-weight portfolio; Bibinger et al. (2019) reveal that intraday covariances follow periodicity patterns, and increase strongly with the arrival of new information; Bollerslev et al. (2019) show the factor-based covariance estimates can improve the performance of risk minimization portfolio in the high-dimensional setting.

While high-frequency data can provide valuable information, the estimators based on high-frequency data are often contaminated by two undesired issues, i.e., market microstructure noise and asynchronicity in price observations. Ait-Sahalia et al. (2005) show that as sampling frequency decreases to zero, the return variance becomes fully induced by microstructure noise instead of the underlying price process. The well-known “Epps” effect (Epps, 1979) states that the asynchronicity in price observations tends to attenuate the correlation between stocks. To handle these difficulties in high-frequency setting, there is a vast literature on efficient high-frequency estimators. Consistent estimators for realized variance in the presence of market microstructure noise include the multi-scale sub-sampling method of Zhang et al. (2005) and Aït-Sahalia et al. (2011), the realized kernel esti-
mator of Barndorff-Nielsen et al. (2008), and the pre-averaging approach of Jacod et al. (2009). For realized covariance, estimators accounting for both microstructure noise and asynchronicity include, among others, the quasi maximum likelihood estimator of Aït-Sahalia et al. (2010), the multivariate realized kernel approach of Barndorff-Nielsen et al. (2011), the two-scale method of Zhang (2011), and the factor-based method in Bollerslev et al. (2019).

In this paper we estimate and analyze the intraday patterns of S&P 500 constituents with a large high-frequency dataset. The dataset consists of 1-second level trade data from the Trade and Quote (TAQ) database for all S&P 500 constituents over 15 years (2004 – 2018). This large sample, both cross-sectional and over time, allows us to obtain robust and general patterns for S&P 500 constituents and examine how the patterns change over years. Specifically, we estimate the intraday correlation, beta, volatility, and trading volume for all stocks or stock pairs in the S&P 500 Index. This establishes a comprehensive picture of various aspects of intraday stock dynamics. We employ the two-scale based estimators developed in Zhang et al. (2005) and Zhang (2011), which are unbiased under market microstructure noise and asynchronicity. Besides, the estimators make full use of the large sample, thus avoid the information loss suffered by traditional estimators based on sparse sampling. Furthermore, the nonparametric two-scale based estimators can be efficiently implemented on a large group of stocks, which is essential for our study.

We find informative intraday patterns for S&P 500 stocks. For realized correlation, We show it exhibits certain intraday patterns that evolve over time. First, in the recent decade, the realized correlation starts low and increases in the morning, stays flat in the middle of the day, and further increases near the market close. The magnitude of the intraday increase in realized correlation is on average larger than 0.2, which is relatively substantial. Second, in 2016 to 2018, the realized correlation for the stock pairs with low daily correlations starts even lower at the market open, and increases rapidly throughout the entire trading session. For example, in 2018, the average realized correlation for the stock pairs with bottom 1% daily correlations (1140 pairs) increases from $-0.1$ at the market open to 0.2 at the end of trading session. Similar patterns are also observed for realized correlation between sector pairs. On the other hand, we find the realized beta of different stocks are more dispersed in the morning, but moves towards one near the market close. In 2018, the average realized beta for the high-beta stocks (top 10% daily betas) decreases from 1.85 to 1.23 during the day, while that for the low-beta stocks (bottom 10% daily betas) increases from 0.17 to 0.65. These patterns confirm our hypotheses on the implication of index-based strategies, and reveal the substantial impact of the active-open, passive-close trading profile on various aspects of intraday stock dynamics. As an additional theoretical support, we develop a market impact model with single-stock and index-fund investors, and show the time-varying liquidity provision indeed produces the observed intraday patterns of realized correlation and beta.

Next, we find the intraday trading volume shows a U-shape pattern, with higher volume near the
market open and close. Furthermore, the U-shape pattern becomes more skewed to the right in the recent decade, as the trading volume near the market close increases significantly. In particular, the proportion of trading volume in the last half hour of trading session increases from 15% in 2004 to 22% in 2018. Such shift in trading volume, as a consequence of the growth of passive investing, has been widely noticed in recent research (see, e.g., Min et al. (2018) and Wu (2019)). Moreover, we show the daily variation in trading volume is high in the morning, but low at the end of trading session. This can be attributed, in part, to the persistent trading from institutional investors who execute most of the index-based strategies near the market close. Finally, we find the intraday realized volatility shows a U-shape pattern skewed to the left, i.e., starts relatively high at the market open and drops subsequently. Besides, we observe the realized volatility near the market close further decreases after 2012, making the intraday curves flatter at the end. While the intraday volatility and volume have been studied in the literature (see, e.g., Wood et al. (1985) and Pagano et al. (2008)), our large dataset and estimators that are efficient under market microstructure noise enable us to obtain robust intraday patterns and examine how they evolve over time.

The rest of this paper is organized as follows. In Section 2, we show the implication of index-based investment via two empirical studies, which motivate our estimation and analysis of intraday stock dynamics. Section 3 establishes the estimators used in our high-frequency setting. In Section 4, we describe the data and implementation details. We provide the main estimation results of intraday patterns in Section 5, including realized correlation, beta, volume, and volatility. In Section 6, we develop a theoretical market impact model with time-varying liquidity provision. Section 7 concludes the paper and provides further discussions.

2 Implication of Index-based Investment on Intraday Trading

In this section, we use two empirical studies to show the growth of index-based investment indeed impacts the intraday trading activities. We propose three hypotheses on the intraday patterns of stock dynamics, which motivate our study with high-frequency data in subsequent sections.

2.1 Evidence from Passive Fund Ownership

First, we demonstrate the growth of index-based investment strategies using the degree of passive fund ownership of S&P 500 constituents. We calculate the passive and active mutual fund ownership following the classification method in Appel et al. (2016). Specifically, a fund is classified as either passive or active by searching for certain strings in its name that identify index funds and the supplementary information on the index fund indicator from CRSP. Figure 1 plots the equal-weight average of the fractions of shares owned by either passive or active funds (left), as well as the ratio

1A detailed description can also be found in Appendix A.3 of Glasserman et al. (2019).
of shares owned by passive funds to the total shares owned by both types of mutual funds2 (right). Note the passive and active shares on the left do not sum to 1 as not all shares are owned by mutual funds. We see the average proportion owned by passive mutual funds significantly increases in the recent decade. Thus, we expect more trading from index-based strategies for stocks in the S&P 500 Index.

Figure 1: The left panel shows the fractions of shares owned by either active and passive funds. The right panel shows the ratio of shares owned by passive fund to that owned by mutual funds.

The growth of index-based investment strategies leads to substantial change in the intraday pattern of trading activities. To show this, we check how the intraday distribution of trading volume differs across stocks with high and low passive ownership. We define the scaled trading volume for stock $i$ in time interval $t$ of day $d$ as

$$S\text{Volm}_{idt} = \frac{\text{Volm}_{idt}}{\text{TotVolm}_{id}},$$

(1)

where $\text{Volm}_{idt}$ is the trading volume in the interval, i.e., the number of shares traded, and $\text{TotVolm}_{id}$ denotes the total trading volume of stock $i$ on day $d$. The scaled trading volume allows us to compare across different stocks, which can have very different shares outstanding and trading volume.

We construct two bins of stocks with low and high passive ownership in each year as follows. For each year, we select the stocks that are in the S&P 500 Index for the entire year. We define their degree of passive ownership as the percent of shares outstanding held by passive mutual funds (averaged over the four quarters). The two bins consist of the stocks with degree of passive ownership below the fifth percentile and above the 95th percentile, respectively. Thus, each bin has

---

2The patterns in Figure 1 match well with the results in Figure 2 of Glasserman et al. (2019); see also Figure 2.8 in 2018 Investment Company Institute Fact Book.
approximately 24 stocks in a given year. The average passive ownership for the two bins is reported in Table 1 in Appendix B.

In Figure 2, we plot the scaled trading volume of the two bins in the first and last half hours of the trading session, i.e., 9:30 – 10:00 and 15:30 – 16:00. The results are computed as the equal weight average across the corresponding stocks and trading days in the given year. By the left panel, we find the scaled trading volume at the market open is higher for the low passive ownership bin than that for the high passive one since 2008. Moreover, the gap keeps increasing after 2015, and reaches approximately two percentage points in 2018. On the other hand, from the right panel, we see the scaled trading volume near the market close increases significantly since 2008 for both bins, but the magnitude of increase is much larger for the high passive one. In 2018, the scaled trading volume in the last half hour is approximately three percentage points higher for the high passive ownership bin than that for the low passive one. As stocks with low (resp. high) passive ownership are more likely to be traded by active (resp. index-based) strategies, the results in Figure 2 suggest an active-open, passive-close profile for intraday trading, i.e., there is more discretionary trading in the morning, and the index-based strategies tend to concentrate their trading near the market close.

The surge in trading volume near the market close, as a consequence of concentrated trading from passive investors, has been widely noticed by financial press (see, e.g., Driebusch et al. (2018) and Strumpf (2015)). The motivations for such behavior include, among others, to minimize the tracking errors of orders benchmarked to the closing price, to efficiently deploy capital to hundreds of underlying stocks using market-on-close order, and to reduce the inventory risk for redemption and creation settlement (see e.g., Cushing and Madhavan (2000), Foucault et al. (2005), and Wu (2019)). Besides, the active-open, passive-close trading profile is also obtained in Min et al. (2018) using the intraday trading volume data of S&P 500 constituents in 2017.

2.2 Evidence from Index Removal Effect

We have shown in the previous section that the trading from index-based strategies drives up the trading volume near the market close. In this section, we provide consistent evidence for such effect using an impact analysis for the stocks that are removed from the S&P 500 Index. Specifically, we find that after a stock is removed from the index, thus less tracked by index-based strategies, its scaled trading volume tends to drop dramatically near the market close.

We select the stocks that are removed from the S&P 500 Index between 2011 and 2018, during which the index-based strategies have become more prevalent. Moreover, to mitigate the potential impact of delisting, we restrict to the stocks that have at least 60 days of observations in the Trade and Quote (TAQ) database after removal. This leaves us with 54 stocks between 2011 and 2018. We then compute the average scaled trading volume for these stocks in the 30 trading days before and after their removal.
Figure 2: Scaled trading volume in 9:30 – 10:00 (left) and 15:30 – 16:00 (right) for the low and high passive ownership bins.

Figure 3: The left panel shows the intraday scaled trading volume before and after removal. The right panel shows the absolute change.
The left panel of Figure 3 shows the intraday pattern of scaled trading volume before (red) and after (blue) the removal, which is computed for each half hour interval with a moving step of five minutes from 9:30 to 16:00. As shown in the left panel, the intraday curve of scaled volume is almost unchanged after the removal, except for the last point representing the time interval 15:30 – 16:00. This can be seen more clearly from the absolute change in the right panel, where the shaded area represents the 95% confidence interval of the estimates. We see that after the removal from the S&P 500 Index, the scaled trading volume significantly drops near market close. On average, the scaled trading volume in 15:30 – 16:00 drops by 2.5 percentage points, which translates to a relative drop of more than 11%. This decrease is offset by the increases in scaled trading volume before 14:00. However, such increases are much smaller in magnitude and more scattered across the day.

After a stock is removed from the S&P 500 Index, it will be less traded by index-based strategies. This analysis shows that the major impact of the removal is the drop in the proportion of trading volume near the market close. Similar effect is also observed in the literature from other perspectives. For example, Grynkiv and Russell (2015) use the data from 2012 to 2015 and find that the increase in trading volume at the end of trading session is more significant for S&P 500 constituents than for stocks in the less liquid exchange-traded products (ETPs). These results provide consistent evidence for the conclusion that index-based strategies tend to concentrate their trading at the end of trading session.

2.3 Hypothesis on Intraday Stock Patterns

From the two empirical studies, we see the growth of index-based strategies and their trading activities substantially impact the intraday stock dynamics. However, unlike the trading volume, the implication on other important aspects, e.g., correlation and beta, remain less studied. In the rest of this paper, we aim to explore these implications with a large high-frequency dataset and efficient estimation methods.

To begin with, we propose three hypotheses on the intraday patterns of stock dynamics based on the active-open, passive-close trading profile. We expect the patterns described in the hypotheses to be more significant in the recent decade, during which the index-based strategies have become more prevalent.

First, we expect the correlation between different stocks to be lower in the morning, but higher near the market close. This is because the index-based strategies tend to trade multiple stocks in the same direction simultaneously, thus driving up the correlation at the end of trading session. Some examples include buying all S&P 500 constituents and investing in all stocks in target sectors. On the other hand, more discretionary trading from active strategies in the morning tends to result in lower correlation, as active strategies focus more on the specific shocks related to individual stocks. The above analysis leads to our following hypothesis on the intraday pattern of correlation.
Hypothesis 1 (H1). The intraday correlation is low at the market open and high near the market close.

Next, we propose a hypothesis on the intraday pattern of beta, which measures the level of systematic risk in individual stocks. With more index-based trading near the market close, we expect the betas of different stocks to move towards one at the end of trading session, as individual stock returns are more driven by index-level orders. By contrast, we expect betas of different stocks to be more dispersed at the market open, as discretionary trading captures the heterogeneity in their levels of systematic risk. This leads to our second hypothesis as below.

Hypothesis 2 (H2). The intraday betas of different stocks are more dispersed in the morning, but move towards one near the market close.

Finally, we analyze the intraday pattern of daily dispersion in trading volume. It has been empirically observed that the trading from institutional investors is highly persistent across days (Campbell et al., 2009). As the index-based strategies are mostly executed by institutional investors (via passive vehicles), we expect lower daily volume dispersion near the market close, i.e., the corresponding trading volume varies less across days. On the other hand, the active strategies at the market open focus more on the short-term price fluctuations and incoming news flows, leading to trading activities that vary much across days. Thus, we expect the daily volume dispersion to be higher at the market open. To summarize, we have following hypothesis on daily dispersion in trading volume.

Hypothesis 3 (H3). The daily dispersion in trading volume is high at the market open and low near the market close.

In subsequent sections, we use a large high-frequency dataset to show the intraday patterns discussed in the three hypotheses indeed hold for S&P 500 constituents, especially in the recent decade. This reveals, from multiple aspects, the substantial implications of index-based investment on intraday stock dynamics.

3 Estimation Methodologies in High-Frequency Setting

In this section, we introduce the estimation methods in our high-frequency setting. Specifically, we define the estimators for realized variance, covariance, correlation, and beta. The estimators account for both market microstructure noise and observation asynchronicity, and can be efficiently implemented on our large dataset.

3.1 Estimators for Realized Variance and Covariance

We first introduce the estimators for realized variance and covariance with high-frequency data, which serve as an indispensable foundation for the estimation of realized correlation and beta. We employ
the Two-Scale Realized Variance (TSRV) and Two-Scale Realized Covariance (TSRCV) estimators
developed in Zhang et al. (2005), Aït-Sahalia et al. (2011), and Zhang (2011). The two estimators
are unbiased under market microstructure noise and asynchronicity, and avoid information loss by
using all price observations.

The TSRV estimator is established as follows. Suppose we estimate the realized variance for stock
$Y$ over a target time interval. We observe the log price $Y_i$ at a series of time points $i = 0, 1 \ldots, n$. Here
we consider a fixed grid of sampling intervals, e.g., every five seconds. We select the last observation
in each interval, or use the most recent one if there is no observation in the current interval. This
is analogous to the previous-tick interpolation commonly used in high-frequency literature (see, e.g.,
 Gençay et al. (2001)). As we focus on the S&P 500 constituents in this study, the stocks considered
are generally highly liquid with frequent price observations.

The observed price $Y_i$ can be viewed as a sum of the true underlying price and the market
microstructure noise. This introduces an essential challenge for estimating realized variance with
high-frequency data. The most naive way to estimate realized variance is to sum all the squared
returns in the time interval, i.e., $RV^{(nv)} = \sum_{i=0}^{n-1} (Y_{i+1} - Y_i)^2$. However, as shown in Aït-Sahalia
et al. (2005), this naive estimator is biased by market microstructure noise, and the bias increases
in the number of observations $n$. Thus, this estimator can be severely contaminated when sampling
frequency is high. The most straightforward remedy for this is to sample sparsely. For instance, the
estimator with observations sampled every $J$ steps can be constructed as $RV^{(sp)} = \sum_{i=0}^{n/J-1} (Y_{(i+1)J} - Y_{iJ})^2$. This sparse estimator is widely employed in the literature, with the sampling interval chosen
in an ad hoc way from 5 to 30 minutes (see, e.g., Gençay et al. (2002) and Barndorff-Nielsen and
Shephard (2002)). While the sparse estimator reduces bias, it inevitably leads to information loss.
Such loss can be significant when sampling frequency is high: if we sample every minute for 1-second
level price observations, we implicitly discard 59/60 of the original data as only the last observation
of each minute is used. An explicit analysis of the naive and sparse estimators can be found in Zhang
et al. (2005).

To overcome the above dilemma, we estimate realized variance by the TSRV estimator proposed
in Zhang et al. (2005) and further developed in Aït-Sahalia et al. (2011). The TSRV estimator
circumvents the two challenges discussed above: it uses all price observations, but yields an unbiased
and consistent estimator of the underlying integrated volatility\footnote{Aït-Sahalia et al. (2011) further show the bias-corrected and consistent properties of the TSRV estimator hold even when microstructure noise exhibits time series dependence.}. The spirit of the TSRV estimator
is to correct the bias by combining the returns from two time scales, i.e., a fast and a slow one. For
a time scale $J$, define the following sum of squared returns

$$[Y, Y]^{(J)} = \frac{1}{J} \sum_{i=0}^{n-J} (Y_{i+J} - Y_i)^2.$$
Similar to the sparse estimator, the term \([Y, Y]^{(J)}\) is also based on \(J\)-step returns. However, it moves by one step each time and thus uses all the observations. This avoids any loss in the price information. Then, the TSRV estimator with fast scale \(J\) and slow scale \(K\) is given by

\[
RV^{(J,K)} = \frac{n}{(K-J)n_K} \left( [Y, Y]^{(K)} - \frac{n_K}{n_J} [Y, Y]^{(J)} \right) \quad \text{for } J < K, \tag{2}
\]

where \(n_J = (n - J + 1)/J\) and \(n_K\) is defined similarly. Thus, the TSRV estimator is a linear combination of the squared terms of two time scales. As a nonparametric estimator, it can be efficiently implemented on a large set of stocks over a long period. This advantage is essential for our study, which involves estimation for all S&P 500 constituents across 15 years. The realized volatility is simply computed as the square root of the TSRV estimator in (2).

Next, we briefly introduce the TSRCV estimator for realized covariance. As we have discussed, the estimation of covariance under high-frequency setting is biased due to asynchronicity and microstructure noise. To cope with these two challenges, we employ the TSRCV estimator proposed in Zhang (2011), which can eliminate the two types of bias simultaneously. The TSRCV estimator follows the same spirit of the TSRV estimator in (2), i.e., correcting the bias by combining the returns from two time scales. Besides, similar to the TSRV estimator, the TSRCV estimator is nonparametric and can be efficiently implemented on a large set of stocks.

The TSRCV estimator is constructed as follows. Suppose we estimate the realized covariance for two stocks \(X\) and \(Y\) over a target time interval. We observe the log prices \(X_i\) and \(Y_i\) at a series of time points \(i = 0, 1, \ldots, n\). Same as for the TSRV estimator, here we consider a fixed time grid and apply previous-tick interpolation to handle missing observation. For a time scale \(J\), define the term \([X, Y]^{(J)}\) as

\[
[X, Y]^{(J)} = \frac{1}{J} \sum_{i=0}^{n-J} (X_{i+J} - X_i)(Y_{i+J} - Y_i).
\]

Then the TSRCV estimator is given by

\[
RCV^{(J,K)} = \frac{n}{(K-J)n_K} \left( [X, Y]^{(K)} - \frac{n_K}{n_J} [X, Y]^{(J)} \right) \quad \text{for } J < K, \tag{3}
\]

where \(n_J = (n - J + 1)/J\) and \(n_K\) is defined similarly. The TSRCV estimator is unbiased under both observation asynchronicity and market microstructure noise. More detailed analysis of its properties can be found in Section 8 of Zhang (2011).

### 3.2 Estimators for Realized Correlation and Beta

In this section, we develop the estimators for realized correlation and beta, which are based on the TSRV and TSRCV estimators in the previous section. Specifically, we develop two methods to estimate realized correlation. The first one estimates the realized correlation between stock pairs, while the second one estimates the portfolio-implied realized correlation between two sets of stocks.
3.2.1 Pairwise Realized Correlation

The estimator for pairwise realized correlation is simply the high-frequency counterpart of the traditional correlation, which is defined by

$$\rho_{X,Y} = \frac{\text{Cov}(R_X, R_Y)}{\sqrt{\text{Var}(R_X)} \cdot \sqrt{\text{Var}(R_Y)}}$$

for stocks $X$ and $Y$. To estimate their realized correlation, we just plug in the high-frequency counterparts of the covariance and variance, i.e.,

$$\text{RCorr}_{X,Y}^{(J,K)} = \frac{\text{RCV}_{X,Y}^{(J,K)}}{\sqrt{\text{RV}_{X}^{(J,K)} \cdot \sqrt{\text{RV}_{Y}^{(J,K)}}}},$$

(4)

where $J$ and $K$ denote the two time scales employed in the TSRV and TSRCV estimators.

3.2.2 Portfolio-implied Realized Correlation

Besides, we propose an estimator for the realized correlation between two mutually exclusive sets of stocks. Unlike the pairwise estimator, the new estimator is based on the realized variances of suitably constructed portfolios. Consider two mutually exclusive stock sets $A$ and $B$. Denote by $w_i > 0$ the weight of stock $i$ (e.g., market-capitalization). Note that we do not require $\sum_i w_i = 1$ as long as the weights are fixed. Define the average return correlation between the two sets of stocks as

$$\bar{\rho}_{A,B} = \sum_{i \in A, j \in B} w_{i,j} \rho_{i,j},$$

(5)

where $\rho_{i,j}$ is the return correlation between stocks $i$ and $j$; $w_{i,j}'$ is defined by

$$w_{i,j}' = \frac{w_i w_j \sigma_i \sigma_j}{\sum_{k \in A, l \in B} w_k w_l \sigma_k \sigma_l},$$

where $\sigma_i$ denotes the standard deviation of the return of stock $i$. Thus, $\bar{\rho}_{A,B}$ in (5) is a weighted average of the pairwise correlations $\rho_{i,j}$ for $i \in A$ and $j \in B$. It puts more weights on the stock pairs with larger portfolio weights ($w_i$ and $w_j$) or more volatile returns ($\sigma_i^2$ and $\sigma_j^2$).

We now propose the estimator for (5) under high-frequency setting. We construct three portfolios using the stock weights $w_i$: the first two include stocks in $A$ and $B$ respectively, and the third one combines stocks from both $A$ and $B$. Then, the average correlation (5) in high-frequency setting can be estimated by

$$\text{RCorr}_{A,B}^{(J,K)} = \frac{\text{RV}_S^{(J,K)} - \text{RV}_A^{(J,K)} - \text{RV}_B^{(J,K)}}{2 \sum_{i \in A, j \in B} w_i w_j \sqrt{\text{RV}_i^{(J,K)} \cdot \sqrt{\text{RV}_j^{(J,K)}}}},$$

(6)

Here $\text{RV}_A^{(J,K)}$, $\text{RV}_B^{(J,K)}$, and $\text{RV}_S^{(J,K)}$ denote the realized variances of the three portfolios respectively; $\text{RV}_i^{(J,K)}$ denotes the realized variance of stock $i$. They are estimated by the TSRV estimator with
time scales $J$ and $K$. Equation (6) defines the portfolio-implied estimator for realized correlation between two sets of stocks. We document its explicit derivation in Appendix A.

When $A$ and $B$ only contain one stock each, we can show by simple algebraic calculation that the portfolio-implied estimator $\text{RCorr}_{A,B}^{(J,K)}$ coincides with the pairwise estimator in (4). However, when the two sets have multiple stocks, the portfolio-implied estimator (6) significantly reduces the computational burden. In particular, if both sets have $N$ stocks, the portfolio-implied estimator only needs to estimate $2N + 3$ realized variances, while the average correlation based on the pairwise estimator needs to estimate $N^2$ realized covariances (for each pair) and $2N$ realized variances (for each stock).

### 3.2.3 Estimation of Realized Beta

Finally, we propose the estimator for realized beta under high-frequency setting. A stock’s beta measures the level of systematic risk in its return. However, the study of intraday beta with high-frequency data, to our best knowledge, is relatively rare.

Denote the market return by $R_{M,t}$. The traditional beta of stock $i$ is estimated by

$$\beta_i = \frac{\text{Cov}(R_i, R_M)}{\text{Var}(R_M)},$$

(7)

where $\text{Cov}(R_i, R_M)$ is the covariance between individual and market returns; $\text{Var}(R_M)$ is the variance of market return. In high-frequency setting, we estimate the realized beta by plugging in the realized variance and covariance into above equation, i.e.,

$$\text{RBeta}_i^{(J,K)} = \frac{\text{RCV}_{M,i}^{(J,K)}}{\text{RV}_M^{(J,K)}},$$

(8)

where $\text{RCV}_{M,i}^{(J,K)}$ and $\text{RV}_M^{(J,K)}$ denote the TSRCV and TSRV estimators with time scales $J$ and $K$. In this study, we use the S&P 500 ETF from SPDR (ticker SPY) to compute the market return, as its high-frequency data are conveniently available in the Trade and Quote (TAQ) database.

### 4 Data and Implementation Details

#### 4.1 Data

In this study, we use all the stocks in the S&P 500 Index from 2004 to 2018. The universe is adjusted dynamically to reflect the quarterly rebalancing of the index. The high-frequency data is obtained from the Trade and Quote (TAQ) Daily Product database. The database contains intraday transaction data (both trades and quotes) for all securities listed on US equity exchanges since September 2003. The original data provided is in millisecond level. In our study, we use the trade data and sample every five seconds. By the previous-tick interpolation, we use the last price
observation for each interval, or the most recent observation if no trade happens in the interval. Accordingly, trade sizes are summed within each five-second interval to measure the trade volume in the interval. Days with only morning trading hours are discarded, including the days before Independence Day, Thanksgiving, and Christmas.

We get other information on stocks from the Center for Research in Security Prices (CRSP) and Compustat databases. This includes daily prices, market capitalization, and sector information. The TAQ and CRSP databases are linked via the ticker and permno code of stocks. We discard stocks with multiple share classes as well as preferred stocks. Besides, we obtain fund holding data used in Section 2 from Thomson Reuters Mutual Fund Holdings and mutual fund classification from CRSP.

Before estimation, we first identify and handle errors and outliers in the high-frequency data by the following two filters. In the first one, we handle “bounce-backs” where price moves by a large amount but then returns to almost the same level immediately. This filter is also employed in previous literature on realized variance (see e.g., Aıt-Sahalia et al. (2011)). Denote three consecutive prices as $p_1$, $p_2$, and $p_3$ (each for a five second interval). We regard $p_2$ as a “bounce-back” if both conditions below are satisfied

$$|r_2| = \left| \ln \left( \frac{p_2}{p_1} \right) \right| > 0.001 \text{ and } |p_3 - p_1| < 0.001.$$ 

That is, the first five second return is larger than 0.1%, and the difference between the first and last prices is smaller than 0.001.

In the second filter, we handle those consecutive outliers which cannot be captured by the first one. At each time point $t$, we first compute the one-minute moving average of the prices, i.e.,

$$MA_t = \frac{1}{12} \sum_{i=0}^{11} p_{t-i}.$$ 

We regard $p_t$ as an outlier if

$$\left| \ln \left( \frac{p_t}{MA_t} \right) \right| > 0.01,$$

i.e., if it is 1% away from the one-minute moving average. For the identified outliers, we set their price levels using the most recent observation and set the corresponding trading volume to be zero. Through experiments, we find the two filters identify fewer than 0.3% of the observations as outliers in normal years, and fewer than 1% in 2008 and 2009. For robustness checks, we find our results are not impacted when using other thresholds for the two filters.

### 4.2 Implementation Details

In all our empirical studies, we set the length of estimation interval to be 30 minutes. This choice of time interval balances two considerations. First, to estimate the intraday patterns, we prefer short estimation interval to enhance granularity. Second, there need to be enough observations in each
time interval to obtain reliable estimates. We regard 30 minutes as a good balance between the two. With a length of 30 minutes, each interval contains 360 price and volume observations sampled every five seconds. Besides, we apply a moving step of five minutes to obtain smooth intraday patterns. Consequently, there are in total 73 time intervals for each day, corresponding to 9:30 to 10:00, 9:35 to 10:05, . . . , and 15:30 to 16:00.

For all two-scale based estimators in Section 3, we set the fast and slow scales to be

\[ J = 2 \text{ and } K = 12. \]

As the prices are sampled every five seconds, the fast and slow scales correspond to the returns over ten seconds and one minute respectively. It has been shown empirically that the two scale estimators are robust to the specific choice of time scales (Aït-Sahalia et al., 2011).\(^4\)

To obtain the intraday patterns, we average the estimates from individual stocks and trading days in each year. Before computing the average, we winsorize the individual estimates between their first and 99th percentiles to mitigate the impact from outliers. Besides, in very rare cases, the estimated realized variance can be negative and the realized correlation can fall outside of \([-1, 1]\). Such anomalies are probably due to large price jumps. When the estimated realized variance is negative, we replace it with the average of the estimates in that day, and we truncate the realized correlation to \([-1, 1]\). The main results are not impacted when we use different thresholds for winsorizing or discarding the anomalies from the estimates entirely.

5 Empirical Results of Intraday Stock Patterns

In this section, we provide the empirical results of intraday stock patterns estimated from our high-frequency data, including realized correlation, beta, trading volume, and volatility.

5.1 Intraday Realized Correlation Between Stock Pairs

In this section, we report the estimation results for intraday realized correlation between stock pairs. While there has been some literature on the comovement in the trading volume of different stocks (see, e.g., Karolyi et al. (2012), Koch et al. (2016), and Min et al. (2018)), the intraday correlation of stock returns is much less studied. We shed light on this topic by estimating realized correlation from a large high-frequency dataset with robust estimators. Our results reveal that the intraday realized correlation indeed shows specific patterns that change over years. The findings support our statement in hypothesis H1 on the implication of index-based investment.

\(^4\)As an additional robustness check, we select several examples and compare the estimated realized variances by the TSRV method with that by the parametric MLE method in Ait-Sahalia et al. (2005). The results match well in most cases. Note that the MLE method requires separate optimization in each estimation. Thus it can not be practically implemented on our large dataset.
The estimation proceeds as follows. Using the pairwise estimator (4), we estimate the realized correlation for each stock pair in the S&P 500 Index. Then, the most convenient way to obtain the general intraday pattern is to average across all stock pairs. However, the correlation between two stocks can be very different across pairs. For example, correlation may be positive or negative depending on the fundamental similarity of the two stocks. Such variation can have substantial impact on the intraday realized correlation as well. To capture the potential heterogeneity in realized correlation, we divide stock pairs into bins based on the correlation of their daily returns, and compute the average realized correlation for each bin separately.

We construct twelve bins of stock pairs as follows. For each year, we pick those stocks that are in the S&P 500 Index for the entire year. This leaves us with a set of 475 stocks on average in each year. For each stock pair constructed from this set, we compute its daily correlation using the daily returns of the two stocks in the year. We then divide the stock pairs into different bins by the levels of their daily correlations, which can be regarded as a measurement of the fundamental similarity between the two stocks. Denote by $p_\alpha$ the $\alpha$-th percentile of the daily correlations across all stock pairs in the given year. The first three bins include stock pairs with daily correlations within $[p_0, p_1]$, $(p_1, p_5]$, and $(p_5, p_{10}]$, respectively, i.e., the stock pairs with the bottom 10% daily correlations. The other nine bins contain the stock pairs with daily correlations within $(p_{10n}, p_{10(n+1)}]$ for $n = 1, 2, ..., 9$. As most stocks in the S&P 500 Index are positively correlated, we use a more granular partition via the first three bins for the stock pairs with low, potentially negative, daily correlations. The average daily correlation for each bin is reported in Table 3 in Appendix B.

For each pair bin denoted by $B_j$ ($j = 1, 2, ..., 12$), we compute its realized correlation in time interval $t$ as the equal weight average of all the stock pairs and trading days, i.e.,

$$\text{RCorr}_{jt} = \frac{1}{N \times |B_j|} \sum_{d=1}^{N} \sum_{(i_1, i_2) \in B_j} \text{RCorr}_{i_1,i_2}^{dt},$$

where $\text{RCorr}_{i_1,i_2}^{dt}$ denotes the realized correlation between stocks $i_1$ and $i_2$ in time interval $t$ of day $d$; $|B_j|$ is the number of pairs in bin $j$ and $N$ is the number of trading days. Note that even the smallest bin contains a large number of stock pairs. For example, with 475 stocks in a year, we would have in total $475 \times 474/2 = 112,575$ stock pairs. Thus even the smallest bin (below the first percentile) includes over 1,000 stock pairs, which translates to more than $1,000 \times 250 = 250,000$ samples every year for a given time interval $t$. Such large sample size improves the robustness of the estimated intraday pattern.

The results for intraday realized correlation of different bins are shown in Figure 4. Each panel represents a given year between 2004 and 2018. The horizontal axis corresponds to the trading hours in a day, where the first (resp. last) point represents the half hour time interval 9:30 – 10:00 (resp. 5:00 – 5:30). The numbers of such stocks for each year are reported in Table 2 in Appendix B.
Figure 4: Intraday realized correlation for different stock pair bins
15:30 – 16:00). For each year, we plot the estimated intraday realized correlation for six selected bins: the first three bins for low-correlated pairs \( \{p_0, p_1\}, \{p_1, p_5\}, \text{ and } \{p_5, p_{10}\} \), the bin with median correlation level \( \{p_{40}, p_{50}\} \), and the two bins for high-correlated pairs \( \{p_{80}, p_{90}\} \text{ and } \{p_{90}, p_{100}\} \). The six bins with daily correlation from high to low are represented by the red, purple, orange, green, cyan, and dark blue lines, respectively. The results for other bins are qualitatively similar. The standard deviations for the intraday curves are also estimated. They are generally very small thanks to the large sample size. Indeed, the standard deviation for the estimates in Figure 4 is smaller than 0.006 in all cases.

By Figure 4, we have two direct observations for the intraday realized correlation. First, in most cases, the realized correlation is positive. This is not surprising as most stocks in the S&P 500 Index are positively correlated. Besides, the relative ranking in daily correlation is mostly preserved in realized correlation. That is, the bin with higher daily correlation also has higher realized correlation in a given time interval.

We then take a closer look at the intraday pattern of realized correlation. Comparing the intraday curves for different bins and over years, we can see the realized correlation indeed demonstrates specific patterns that change over time. We summarize our findings in the following three points.

First, in the period 2004 – 2007 (the first four panels), the intraday realized correlation generally shows an M-shape pattern for all the six bins: it starts lowest at the market open and increases to a peak around 11:00, stays flat around noon, further increases to the highest level around 15:00, and finally drops near the market close.

Second, after 2009, the drop in realized correlation near the market close vanishes. Instead, the realized correlation increases in the entire afternoon, and reaches the highest level near the market close for all bins. The market open (9:30 to 10:00) still witnesses the lowest level of realized correlation. The curve in the middle of the day (e.g., 11:00 to 14:00) becomes flatter, indicating stable periods for realized correlation. This holds except for the three low correlation bins in 2016 – 2018, which is further discussed below.

Finally, in recent years from 2016 to 2018, the intraday pattern for the low correlation bins (green, cyan, and blue lines) becomes different from that in previous years as well as that for the high correlation bins. For these three bins, their realized correlation at the market open further decreases, even to the negative regime, which is not seen in previous years. Besides, the middle part of their intraday curves become steeper, showing their realized correlation increases rapidly during the day. This change is not seen in the intraday pattern for the high correlation bins, which still stays flat during the middle of the day. For all bins, the realized correlation reaches the highest level near the market close. This suggests the stock prices are more likely to move in the same direction.

---

6 The estimates of standard deviations for these as well as other intraday curves in the paper are available from authors upon request.
at the end of trading session.

The above observations are illustrated more directly in Figure 5, where we plot the intraday realized correlation for the stock pairs with the top 10% (left) and bottom 10% (right) daily correlations. Each line in the panel represents the average over different years in a given period (2004–2006, 2007–2010, 2011–2015, and 2016–2018). In 2004–2006, the realized correlation for both high and low correlated pairs show an M-shape (purple lines). After that, the realized correlation significantly increases near the market close. Moreover, in 2016–2018, the realized correlation for the bottom 10% pairs starts below zero at the market open, and monotonically increases during the day to above 0.2 at the end of trading session (red line in the right panel).

Figure 5: Realized correlation for the stock pairs with daily correlations above the 90th percentile (left) and below the 10th percentile (right), averaged over the yearly results in different periods.

The intraday patterns of realized correlation in the recent decade confirms our hypothesis H1 in Section 2.3, i.e., correlation is low in the morning and high near the market close\(^7\). This demonstrates the implication of the growth of index-based strategies and the active-open, passive-close trading profile. In particular, more discretionary (resp. index-based) trading tends to decrease (resp. increase) the correlation between different stocks at the market open (resp. close). Besides this hypothesis, our estimation results reveal other significant pattern shifts for intraday realized correlation in recent years, especially the even lower starting level and monotonically increasing shape for the low correlated pairs.

In the recent work of Buccheri et al. (2020), they propose a score-driven model to estimate the covariance dynamics under high-frequency setting. With a much smaller dataset (transaction data of ten stocks in 2014), they show the opening hours are dominated by idiosyncratic risk and a common

\(^7\)The difference between the first and last half hour intervals is statistically significant at 0.1% level.
market factor emerges in the afternoon. This result is consistent with the intraday patterns we obtained for realized correlation. Moreover, in Section 6, we develop a market impact model with time-varying liquidity provision from different types of investors, and use it to show the active-open, passive-close trading profile indeed generates the intraday correlation pattern that is qualitatively similar to the one observed in recent years. This provides additional theoretical support for our interpretation.

The intraday pattern of realized correlation has various applications, especially in intraday trading and portfolio execution. For example, if the traders want to exploit the low (resp. high) correlation in the intraday price movement of different stocks, they may set up their positions around the market open (resp. market close). On the other hand, if the portfolio managers prefer a period with stable correlation for order execution, the middle of the day appears as a better choice. These applications, among others, can be topics for future research.

5.2 Intraday Realized Correlation Between Sectors

In this section, we estimate the intraday realized correlation between different sectors. While we can compute the average between-sector correlation using the pairwise correlations of corresponding stocks, here we employ the portfolio-implied approach developed in Section 3.2.2. This allows us to estimate the realized correlation from the portfolio level that incorporates different weights of stocks. We reveal that the intraday realized correlation between sectors demonstrates similar patterns to that for stock pairs observed in Section 5.1.

We implement the portfolio-implied estimator in (6) as follows. As for pairwise correlations, we choose the stocks that are in the S&P 500 Index for the entire year, and divide them to eleven mutually exclusive sectors based on the first two digits of their GICS codes. The mapping from GICS codes to sector names and the numbers of stocks in each sector are summarized in Tables 4 and 5 in Appendix B. We exclude the real estate sector as it is not formally included as a sector before 2016, and the number of stocks in this sector is small in early years. We set the weight of each stock proportional to its average market capitalization throughout the year. The sector and market capitalization data are obtained from CRSP.

With ten remaining sectors, there are in total 45 sector pairs. For each year, we focus on the sector pairs with high and low average daily correlations\(^8\). Specifically, we report the estimation results for six sector pairs, three with the highest average daily correlation and three with the lowest. The sector names for the six selected pairs in each year are reported in Table 6 in Appendix B. The results for between-sector intraday realized correlation are shown in Figure 6. The six sector pairs, with average daily correlations from high to low, are represented by the red, purple, orange, green,

\(^8\)The average daily correlation between two sectors are simply computed as the average of the pairwise daily return correlations between their constituents.
cyan, and dark blue lines respectively. At first glance, we see the ranking of sector pairs is generally
preserved in intraday realized correlation: the three pairs with higher daily correlations also have
higher realized correlation than the other three for most of the time during the day.

A more interesting finding is revealed by comparing the intraday patterns in Figures 4 and 6. Recall Figure 4 plots the intraday realized correlation for stock pairs estimated by the pairwise estimator (4). Thus the realized correlation in the two figures are very different with respect to both the estimated object and the estimation method. Surprisingly, however, we see the patterns in the two figures are similar, especially in the recent decade. Thus the discussion in previous section regarding the shapes of intraday patterns generally applies here. In the period 2004 – 2007, the intraday realized correlation demonstrates an upward M-shape for most sector pairs. After 2009, the realized correlation starts low in the morning, stays relatively flat in the middle of the day, and increases significantly near the market close. Moreover, in 2016 to 2018, the realized correlation of the three low correlated sector pairs starts negative at the market open and increases quickly during the day, which is similar to the patterns of the three low correlation bins in Figure 4.

The consistent results for sector pairs highlight the generality and robustness of the intraday correlation patterns observed in Figures 4 and 6. It demonstrates, from the portfolio-level, the implication of the growth of index-based strategies and the active-open, passive-close trading profile. Specifically, concentrated trading from index-based strategies drives up the correlation between different sectors near the market close, while more discretionary trading tends to lower the correlation in the morning. Such patterns have important applications in sector-based intraday trading and portfolio execution, which are deferred to future research.

5.3 Intraday Pattern of Realized Beta

In this section, we analyze the estimation results for intraday realized beta of S&P 500 constituents. We find that in recent years, the realized betas of different stocks start dispersed at the market open, but generally move towards one near the market close. This confirms our second hypothesis H2 in Section 2.3. The convergence pattern of realized beta echos our results for realized correlation discussed in previous sections, and shows the impact of the index-based strategies and the active-open, passive-close trading profile.

We estimate intraday realized beta using the estimator (8) as follows. Similar to the study of pairwise realized correlation, we select the stocks that are in the S&P 500 Index for the entire year, and divide stocks into bins based on their daily betas, which are computed by (7) using daily returns. We construct eleven bins for each year. Denote by $p_\alpha$ the $\alpha$-th percentile of daily betas among all stocks. The first two bins include the stocks with daily betas in $[p_0, p_5]$ and $(p_5, p_{10}]$, i.e., stocks with the bottom 10% daily betas. The other nine bins consist of the stocks with daily betas in $(p_{10n}, p_{10(n+1]}$ for $n = 1, 2, ..., 9$. The average daily beta for each bin is reported in Table 7 in
Figure 6: Intraday realized correlation for sector pairs
Appendix B. For each bin $B_j$ ($j = 1, 2, ..., 11$), we estimate its realized beta over time interval $t$ as

$$RBeta_{jt} = \frac{1}{N \times |B_j|} \sum_{d=1}^{N} \sum_{i \in B_j} RBeta_{it}^d,$$

where $RBeta_{it}^d$ denotes the realized beta of stock $i$ in time interval $t$ of day $d$; $N$ and $|B_j|$ are the number of trading days and number of stocks in bin $j$, respectively. With 475 stocks in a year, the smallest bin would have 24 stocks, which translates to approximately $250 \times 24 = 6,000$ samples every year for a given time interval $t$.

The estimation results of intraday realized beta are shown in Figure 7. For each year, we plot the estimated results for six selected bins: the first three bins with low daily betas ($[p_0, p_5]$, $(p_5, p_{10}]$, and $(p_{10}, p_{20}]$), and the three with high daily betas ($(p_{70}, p_{80}]$, $(p_{80}, p_{90}]$, and $(p_{90}, p_{100}]$). The six bins with daily betas from high to low are represented by the red, purple, orange, green, cyan, and dark blue lines, respectively. The horizontal dashed black line denotes the level of beta equal to one. The standard deviations of the estimated intraday curves in Figure 7 are below 0.015 in all cases, which are relatively small compared with the estimated levels.

By Figure 7, we have the following observations of the intraday pattern of realized beta. First, the ranking of daily beta across different bins is mostly preserved in the intraday pattern: the bin with higher daily beta also has higher realized beta across the day. Next, we see the intraday patterns of realized beta indeed show specific shapes that evolve over time. For the high beta bins (red, pink, and yellow lines), their intraday realized beta exhibits a smirk pattern in 2004 – 2007, and a monotonically decreasing pattern in the years after. For the low beta bins (green, cyan, and dark blue lines), their intraday realized beta stays relatively flat before 2013, but significantly increases during the day in the years after. Consequently, in the recent decade, we see the realized betas of different bins are more dispersed in the morning, but generally move towards one near the market close. This confirms our second hypothesis H2 in Section 2.3, and shows the implication of the active-open, passive-close trading profile on intraday beta.

To demonstrate the changes in the intraday patterns more directly, we plot in Figure 8 the realized beta for the stocks with top (left) and bottom (right) 10% daily betas, averaged over the years in four different periods. By Figure 8, we see significant changes in the intraday patterns in recent years (2016 – 2018), as shown by the red lines in the two panels. First, at the market open, the realized betas of the high beta and low beta stocks become more dispersed than previous years, as shown by the even higher (resp. lower) red line in the left (resp. right) panel. On the other hand, the magnitude of the intraday movement is larger, i.e., the realized beta drops (resp. increases) more during the day for the high (resp. low) beta stocks. Such changes are especially noticeable for the low beta stocks: their average realized beta starts below 0.2 at the market open, but rises dramatically to above 0.6 at the end of trading session.
Figure 7: Intraday realized beta for different stock bins
As a consequence of above changes, we see a more significant convergence pattern of intraday realized beta in recent years. Specifically, the divergence in realized betas of different stocks shrinks during the day. At the end of trading session, the realized betas of all bins move towards one, suggesting the individual stock returns are more similar to the market return. This can be attributed to the growth of index-based investment, and the active-open, passive-close trading profile. In particular, more discretionary trading in the morning make realized betas to be more dispersed, while more index-based strategies drive realized betas towards one near the market close. In Section 6, we provide theoretical support for such interpretation using a market impact model with time-varying liquidity provision from single-stock and index-fund investors. The intraday pattern of realized beta has potential applications for intraday trading strategies that exploit the levels of systematic risk in stocks’ returns. For example, it may be better to execute strategies that hinge on the heterogeneity in stocks’ systematic risk levels in the morning rather than in the afternoon.

5.4 Intraday Patterns of Trading Volume and Realized Volatility

In this section, we look into the intraday patterns of trading volume and realized volatility. First, we show in Figure 9 the intraday pattern of scaled trading volume defined in (1). The results are computed as the equal weight average of all S&P 500 constituents in a given year. The four panels plot the estimation results for 2004 – 2007, 2008 – 2010, 2011 – 2014, and 2015 – 2018 respectively. Comparing across the four panels, we see the intraday pattern of scaled trading volume indeed changes over time. In 2004 – 2007, the scaled trading volume demonstrates a symmetric U-shape pattern that is relatively stable across years. The trading volume near the market close is quite close
to that at the market open. However, from 2008, the trading volume near the market close increases dramatically, and the symmetric U-shape pattern becomes skewed to the right. This trend becomes more significant in recent years, as the trading volume near the market close keeps increasing. In 2018, the final half hour 15:30 – 16:00 (the last point) consists of more than 20% of the total trading volume in the day.

The change in the intraday pattern of trading volume can be seen more clearly from the left panel of Figure 10, where we fix four time intervals (9:30 – 10:00, 11:30 – 12:00, 13:30 – 14:00, 15:30 – 16:00) and plot the scaled trading volume across different years. By the red line, we see the scaled trading volume in 15:30 – 16:00 increases significantly, especially during 2007 to 2010 and after 2016. Such increase is offset by the drop in scaled trading volume in the middle of the day, although the magnitude is much smaller. The results observed here indicate the increase in end-of-day trading volume holds for our large stock universe, thus generalize the finding in Figure 2 for the stocks with low and high passive ownership. Such increase can be attributed, in part, to the
growth of index-based strategies and their concentrated trading near the market close. For instance, Cushing and Madhavan (2000) and Foucault et al. (2005) point out the importance of closing price for institutional investors, who execute most of the index-based strategies. The recent work in Wu (2019) finds ETF flows make increased usage of market-on-close orders.

Figure 10: The left panel shows the scaled trading volume in four time intervals. The right panel shows the intraday pattern of daily trading volume dispersion for five selected years.

Next, motivated by hypothesis H3 in Section 2.3, we study the intraday pattern of daily dispersion in trading volume. For stock $i$, its daily volume dispersion in time interval $t$ is defined as

$$\text{VolmDisp}_{it} = \frac{\text{StdVolm}_{it}}{\text{AvgVolm}_{it}}.$$  

Here $\text{AvgVolm}_{it}$ and $\text{StdVolm}_{it}$ denote the mean and standard deviation of the corresponding trading volume across different days, i.e.,

$$\text{AvgVolm}_{it} = \frac{1}{N} \sum_{d=1}^{N} \text{Volm}_{idt},$$

and

$$\text{StdVolm}_{it} = \sqrt{\frac{1}{N-1} \sum_{d=1}^{N} (\text{Volm}_{idt} - \text{AvgVolm}_{it})^2},$$

where $N$ is the number of trading days for stock $i$ (e.g., roughly 250 trading days in a year). The normalization by $\text{AvgVolm}_{it}$ in the denominator of (9) allows us to average and compare across stocks with very different trading volume levels.

We show the intraday pattern of daily volume dispersion for five selected years (2004, 2006, 2009, 2013, and 2018) in the right panel of Figure 10, which are computed as the equal weight average of all the stocks in each year. The results for other years are qualitatively similar and available from
authors upon request. We have the following observations for the intraday pattern of daily volume dispersion. First, we see the daily volume dispersion demonstrates a similar intraday pattern in all the five years: starts high at the market open, drops and stays flat in the middle of the day, and further decreases near the market close. This pattern suggests the trading volume is more volatile across days in the morning, but much less near the market close. The standard deviations of the estimates in Figure 10 are smaller than 0.018 in all cases, and the difference between market open and close is statistically significant at 0.1% level for all five years. It confirms our hypothesis H3 on daily volume dispersion and shows the impact of the growth of index-based strategies and the active-open, passive-close trading profile. Specifically, the trading from discretionary investors in the morning varies more across days, while the trading from index-based strategies is more persistent.

Besides the general decreasing pattern of daily volume dispersion, more interesting results are revealed by comparing the intraday curves of the five different years. First, the overall dispersion level in 2009 (purple) is the lowest among the five years. This is likely a consequence of the financial crisis, during which trading volume was consistently high. More importantly, we see the magnitude of the intraday drop from market open to close increases significantly in 2013 and 2018 compared with that in previous years\(^9\). The increased drop is mainly driven by the lower dispersion levels at the end of trading session. This shift can be attributed to the prevalence of index-based strategies in the recent decade, and particularly, their concentrated trading near market close.

We further support this interpretation by comparing the intraday pattern of daily volume dispersion for the high and low passive ownership bins defined in Section 2.1, i.e., stocks with passive ownership below the fifth percentile and above the 95th percentile each year. The left and right panels of Figure 11 plot the average of yearly results for the two bins in the first (2004 – 2010) and second periods (2011 – 2018), respectively. By the left panel, we see the daily volume dispersion is quite close for the two bins in 2004 – 2010. However, in 2011 – 2018, the daily volume dispersion for high passive ownership bin becomes significantly lower than that for the low passive one. The increased gap in daily volume dispersion can be explained by the prevalence of passive mutual funds in the recent decade, which leads to larger difference in the degree of passive ownership between the two bins. Indeed, the average difference in passive ownership between the two bins increases from 7% for the first period to 20% for the second period, which can be computed from Table 1.

Finally, we look into the intraday pattern of realized volatility. While the intraday volatility pattern has been widely studied in previous literature, here we use a large dataset with all S&P 500 constituents over 15 years, instead of just the market index or a few selected stocks. The results are shown in Figure 12, with each curve computed as the equal weight average of all stocks in a given year. We have following observations for the intraday pattern of realized volatility. First, the intraday volatility shows a U-shape pattern skewed to the left: it starts relatively high in the

\(^9\)The increase in drop magnitude is also observed for other years after 2010.
morning and decreases during the day. In the majority of years, the realized volatility reaches the lowest level around noon, and slightly increases near the market close. This U-shape pattern is widely observed in the literature (see, e.g., Wood et al. (1985) and Pagano et al. (2008)). Not surprisingly, the financial crisis period (2008 and 2009) is associated with extremely high realized volatility in the entire day, as seen from the magnitude of vertical axis in the upper-right panel. Finally, after 2012, we observe the realized volatility generally decreases near the market close, thus the “tail” of the intraday curve tends to flatten at the end. This can be potentially attributed to the concentrated trading from index-based strategies at the end of trading session.

6 A Market Impact Model with Time-varying Liquidity Provision

In this section, we provide additional theoretical support for the relation between the growth of index-based strategies and the intraday patterns of stock dynamics. In particular, we use a market impact model to show that the active-open, passive-close trading profile indeed generates the intraday pattern of realized correlation and beta observed in Sections 5.1 and 5.3.

Suppose the market has $N$ individual stocks, indexed by $i = 1, 2, ..., N$, and a fund consisting of these stocks. The fund can be an ETF or a mutual index fund. The weight of each stock in the fund is given by $w_i$, which is assumed to be positive with $\sum_{i=1}^{N} w_i = 1$. For stock $i$, denote its actual price and investors’ reservation value by $p_i$ and $r_i$, respectively. We use the vector representations $w = (w_1, w_2, ..., w_N)^\top$, $r = (r_1, r_2, ..., r_N)^\top$, and $p = (p_1, p_2, ..., p_N)^\top$.

There are both single-stock and index-fund investors in the market. The single-stock investors buy or sell individual stocks in response to the change in the gap between the actual price and
Figure 12: Intraday realized volatility (annualized)
the reservation value. Specifically, active investors will buy (resp. sell) $\psi_{i,t}$ shares if the gap $r_{i} - p_{i}$ increases (resp. decreases) by one dollar in time interval $t$. This linear assumption is often assumed in microstructure literature (Kyle, 1985), and can be justified under the case of CARA utility investors and normally distributed beliefs. In contrast, the index-fund investors only trade the fund based on its price and reservation value implied by the stocks. They will buy (resp. sell) $\psi_{f}$ share of the fund if the gap $w^\top r - w^\top p$ increases (resp. decreases) by one dollar. Such index-based strategy trades stocks on a portfolio-level, translating to $w\psi_{f}$ position change for each stock.

The parameters $\psi_{i,t}$ and $\psi_{f}$ measure the liquidity provided by investors, which are allowed to be time-varying during the day. This set-up is similar to the liquidity provision model in Min et al. (2018). The trading from both types of investors impacts the stock prices. For illustration purpose, we use a simple linear function to model the price impact. The linear price impact model is widely used in microstructure literature (see, e.g., Huberman and Stanzl (2004) and Alfonsi et al. (2012)). In particular, we assume every share bought (resp. sold) of stock $i$ would increase (resp. decrease) its price by $\phi_{i}$, which measures the sensitivity of price to trading demand.

The market evolves as follows. In time interval $t$, a random shock $\epsilon_{i,t}$ happens to the reservation value of stock $i$. Natural sources for the random shocks can be news flow or unpredictable announcements that impact the stock valuation. We assume the shocks are i.i.d. over different time intervals, but allow them to be correlated across stocks. This captures the fact that different stocks may be driven by some common factors (e.g., news for sectors). Denote the vector representation of the shocks by $\epsilon_{t} = (\epsilon_{1,t}, \epsilon_{2,t}, \ldots, \epsilon_{N,t})^\top$. After the random shocks, both types of investors trade the stocks (or index fund) based on the new valuation. This leads to a new equilibrium price level that accounts for both the shock in reservation value and the market impact from trading.

Denote the equilibrium price change by an $N$-dimensional vector $\Delta p_{t}$. Then, the change in the gap between price and reservation value is equal to $\epsilon_{t} - \Delta p_{t}$. The trading demand comes from both single-stock and index-fund investors. First, single-stock investors will buy $\Psi_{t}^a(\epsilon_{t} - \Delta p_{t})$ shares of each stock, where the matrix $\Psi_{t}^a = \text{diag}_{i=1}^{N}(\psi_{i,t}^a)$. Next, index-fund investors will buy $\psi_{f}^f w^\top(\epsilon_{t} - \Delta p_{t})$ shares of the fund, which translates to $w\psi_{f}^f w^\top(\epsilon_{t} - \Delta p_{t})$ shares for individual stocks. Summing up the two sources of demand, the market impact model implies

$$\Phi \cdot \left[ \Psi_{t}^a(\epsilon_{t} - \Delta p_{t}) + w\psi_{f}^f w^\top(\epsilon_{t} - \Delta p_{t}) \right] = \Delta p_{t},$$

where $\Phi = \text{diag}_{i=1}^{N}(\phi_{i})$ measures the market impact. From this linear equation, we can solve $\Delta p_{t}$ as

$$\Delta p_{t} = M_{t}\epsilon_{t},$$

which is proportional to the shock $\epsilon_{t}$. The matrix $M_{t}$ is given by

$$M_{t} = (I_{N} + \Phi \Gamma_{t})^{-1} \Phi \Gamma_{t},$$

32
where $I_N$ is the $N$-dimensional identity matrix; $\Gamma_t$ is defined by

$$\Gamma_t = \Psi_t^a + w\psi_t^f w^\top.$$  

The matrix $\Gamma_t$ measures the total market liquidity from both single-stock and index-fund investors, which is allowed to be time-varying during the day. As the sum of two symmetric and strictly positive-definite matrices, it is also symmetric and strictly positive-definite.

By (10), the covariance matrix of $\Delta p_t$ can be computed as

$$\Sigma_p^p = M_t \Sigma r M_t^\top,$$ (11)

where $\Sigma_r$ is the covariance matrix of the random shocks $\varepsilon_t$ in stocks’ reservation values. When there are only single-stock investors, i.e., $\psi_f^t \equiv 0$, it is easy to verify the matrix $\Gamma_t$ (and $M_t$) becomes diagonal, and the price change of stock $i$ follows by

$$\Delta p_{i,t} = \phi_i \psi_{i,t}^a + \phi_i \psi_{i,t}^a \varepsilon_{i,t}.$$  

In this case, the price change of a stock only depends on the shock to its own reservation value. With deterministic liquidity parameters $\psi_{i,t}^a$, the correlation between price changes equals to that between the corresponding random shocks. As we assume the distribution of random shocks does not change over time, the pairwise correlation between $\Delta p_{i,t}$ is constant even if the liquidity parameters $\psi_{i,t}^a$ are time-varying. However, this does not hold when there are index-fund investors, as positive $\psi_f^t$ leads to non-zero off-diagonal entries in $\Gamma_t$ (and $M_t$). By (10), the change in the fund price is given by

$$\Delta p_f^t = w^\top \Delta p_t = w^\top M_t \varepsilon_t.$$  

We compute the beta of stock $i$ using the price changes as

$$\beta_{i,t} = \frac{\text{Cov}(\Delta p_{i,t}, \Delta p_f^t)}{\text{Var}(\Delta p_f^t)}.$$ (12)

Accordingly, its average beta over all intraday intervals is defined by

$$\bar{\beta}_i = \frac{1}{T} \sum_{t=1}^{T} \beta_{i,t}.$$ (13)

We conduct following numerical experiment to show the impact of time-varying liquidity provision on intraday stock dynamics. We model the intraday liquidity provision in line with the active-open, passive-close pattern. We divide the trading hours of each day (9:30 to 16:00) to 78 five-minute intervals, indexed by $t = 1, 2, ..., 78$. We assume the liquidity parameters $\psi_{i,t}^a$ and $\psi_f^t$ vary parametrically as

$$\psi_{i,t}^a = \alpha_t \psi_i^a \quad \text{and} \quad \psi_f^t = \beta_t \psi_f^f,$$
with $\sum_{t=1}^{78} \alpha_t = \sum_{t=1}^{78} \beta_t = 1$ and constants $\bar{\psi}^a_i$ and $\bar{\psi}^f_i$. The time-varying parameters $\alpha_t$ and $\beta_t$ determine the liquidity profile of single-stock and index-fund investors throughout the day. For illustration purpose, we consider following simplified liquidity profile:

$$
\alpha_t = \begin{cases} 
0.0256, & \text{for } t = 1, 2, \ldots, 6 \\
0.0105, & \text{for } t = 7, 8, \ldots, 72 \\
0.0256, & \text{for } t = 73, 74, \ldots, 78
\end{cases}
$$

and

$$
\beta_t = \begin{cases} 
0.0075, & \text{for } t = 1, 2, \ldots, 72 \\
0.0769, & \text{for } t = 73, 74, \ldots, 78
\end{cases}
$$

That is, the liquidity from single-stock investors is higher at the market open (9:30 – 10:00) and near the market close (15:30 – 16:00), while the liquidity from index-fund investors concentrates near the market close (15:30 – 16:00). Thus, there is more discretionary trading in the morning, and the end of trading session is dominated by index-based strategies. This liquidity profile is qualitatively analogous to the result in Min et al. (2018) (Figure 2 therein), which is calibrated from the intraday trading volume data of S&P 500 constituents in 2017.

We assume the market has $N = 478$ stocks, which is the number of stocks that are in the S&P 500 Index throughout the entire 2018. The index fund is benchmarked to the average price of the stocks, i.e., $w = (1/478, 1/478, \ldots, 1/478)^\top$. For illustration purpose, we assume the random shocks to stocks' reservation values have standard deviation of 1, and are correlated between the stocks following the daily return correlation matrix in 2018. We set the liquidity and market impact parameters as $\bar{\psi}^a = 10$, $\bar{\psi}^f = 843.5$, and $\phi_i = 0.9$ for $i = 1, 2, \ldots, 478$. Thus, the index-fund investors provide on average $\bar{\psi}^f / (\bar{\psi}^f + \sum_i \bar{\psi}^a_i) = 15\%$ of total market liquidity, which is consistent with the passive ownership degree in Figure 1. Besides, these parameters imply that, if there are only single-stock investors, the equilibrium price will increase by $0.9$ dollar for every dollar increase in the reservation value.

Similar to previous studies in the paper, we employ half-hour intervals with a moving step of five minutes to estimate the intraday curves for correlation and beta. For each half an hour, we compute the average correlation in $\Delta p_t$ and stock beta $\beta_{t,i}$ over the six 5-minute intervals, which can be explicitly obtained from (11) and (12). Denote by $p^c_{\alpha}$ the $\alpha$-th percentile for the pairwise correlations between the random shocks $\varepsilon_{t,i}$, i.e., daily return correlation in our study. We focus on the intraday correlation pattern for three bins of stock pairs, which have underlying correlations (between corresponding random shocks) below $p^c_{5}$, between $[p^c_{45}, p^c_{55}]$, and above $p^c_{95}$ respectively. They correspond to the stock pairs with high, median, and low correlations. Besides, denote by $p^b_{\alpha}$ the $\alpha$-th percentile of the average stock beta $\bar{\beta}_i$ given in (13), we study the intraday beta pattern for the stocks with $\bar{\beta}_i$ below $p^b_{5}$ and above $p^b_{95}$, i.e., the low and high-beta stocks of the bottom and top 5%.

Figure 13 plots the intraday patterns of pairwise correlation in $\Delta p_t$ (left) and the stock beta $\beta_{t,i}$ (right). By the left panel, we see the correlations indeed change substantially during the day, as a consequence of the index-fund investors and time-varying liquidation provision. At the market open,
The correlations are lower for all three bins, and are closer to the underlying levels when there is no index-fund investor. This shows the implication of more discretionary trading from single-stock investors in the morning. On the other hand, the correlation for all bins increase near the market close. The effect is large, and even more significant for the low correlation bin, which increases from around 0 in the morning to above 0.2 at the end of trading session. This increase can be explained by the concentrated trading from the index-fund investors, which drives stocks to move in the same direction via index-based orders. Besides, by two panels on the right, we see the intraday beta starts relatively high (resp. low) for the high- (resp. low-) beta stocks, but drops (resp. increases) during the trading session. This generates the convergence pattern of intraday realized beta observed in Section 5.3.

Figure 13: The left panel plots the model-implied intraday correlation for three pair bins. The right panels plot the intraday beta for high- (upper) and low- (lower) beta stocks.

The intraday patterns in Figure 13, which are solved explicitly from the model, are in line with the observations for realized correlation and beta in Sections 5.1 and 5.3, especially in the recent decade: The realized correlation starts low at the market open, stays relatively flat in the middle of the day, and increases significantly near the market close; the realized beta starts dispersed in the morning, but generally moves towards one near the market close. This theoretical study confirms the time-varying liquidity provision, in particular, the active-open, passive-close trading profile, indeed contributes to the observed intraday patterns of realized correlation and beta.
7 Conclusion

The rapid growth of passive investment and index-based strategies has drawn much attention in the recent years. In this paper, we demonstrate its implication on various aspects of intraday stock dynamics. In particular, we estimate intraday correlation, beta, volatility, and trading volume with a large high-frequency dataset, i.e., 1-second level trade data for all S&P 500 constituents from 2004 to 2018. We find the intraday patterns indeed change over time. In the recent decade, the realized correlation starts low in the morning and increases near the market close; the realized betas of different stocks start dispersed and generally move towards one at the end of trading session. Besides, we find the trading volume is more volatile across days in the morning than that in the afternoon. These patterns confirm our hypotheses on the implication of index-based strategies, which have become more prevalent in the recent decade.

With the development of financial technologies, high-frequency data has been widely used in various fields of finance. However, estimators under high-frequency setting are often biased by market microstructure noise and observation asynchronicity. Besides, large-scale application of high-frequency data is inevitably hindered by heavy computational burden. Due to these reasons, most previous studies use a limited dataset (e.g., several stocks or indices over a short period) or employ ad hoc estimators that do not take market noise and asynchronicity into account. In this paper, we overcome these challenges with estimators that can be efficiently implemented on large set of stocks and account for both market microstructure noise and observation asynchronicity. Furthermore, the size of the dataset allow us to obtain the general intraday patterns of US stocks and examine how they evolve over time.

The estimation of intraday patterns have various practical applications, including trading strategies, portfolio execution, and risk management. Besides, the intraday patterns facilitate the development of market impact and liquidity provision models that incorporate stylish and realistic features of intraday stock dynamics, e.g., time-varying correlation between stocks. Moreover, with the richness of the dataset, the cross-sectional dimension of the intraday estimates can be leveraged to study the determinants of intraday returns across stocks. Another interesting direction is to examine the intraday patterns for other markets, especially those emerging markets where institutional investors play a less important role. These potential directions can be explored in future research.

References


Aït-Sahalia, Y., Mykland, P. A., Zhang, L., 2005. How often to sample a continuous-time process in

Aït-Sahalia, Y., Mykland, P. A., Zhang, L., 2011. Ultra high frequency volatility estimation with


Andersen, T. G., Bollerslev, T., 1997. Intraday periodicity and volatility persistence in financial

Andersen, T. G., Bollerslev, T., Das, A., 2001. Variance-ratio statistics and high-frequency data:


Financial Economics 121 (1), 111–141.

to measure the ex post variation of equity prices in the presence of noise. Econometrica 76 (6),
1481–1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2011. Multivariate realised kernels:
consistent positive semi-definite estimators of the covariation of equity prices with noise and non-

(Statistical Methodology) 64 (2), 253–280.

419–435.

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

Bollerslev, T., Meddah, N., Nyawa, S., 2019. High-dimensional multivariate realized volatility esti-

Boudt, K., Zhang, J., 2015. Jump robust two time scale covariance estimation and realized volatility
budgets. Quantitative Finance 15 (6), 1041–1054.


Appendix A  Derivation of Portfolio-implied Realized Correlation

In this section, we provide the derivation for the portfolio-implied realized correlation (6). Denote the two mutually exclusive sets of stocks by $A$ and $B$. We construct three portfolios. The first two include stocks in $A$ and $B$, respectively. Their returns can be computed by

$$R_A = \sum_{i \in A} w_i R_i \quad \text{and} \quad R_B = \sum_{i \in B} w_i R_i,$$

where $w_i$ and $R_i$ denote the weight and return of stock $i$, respectively. For ease of representing the portfolio return, here we use $R_i$ to denote the simple return instead of log return of individual stocks. In our high-frequency setting, these two are generally very close. Note that we do not require $\sum_i w_i = 1$ as long as $w_i$ are fixed. The third portfolio $S$ consists of stocks from both $A$ and $B$, with return given by

$$R_S = \sum_{i \in A \cup B} w_i R_i.$$

The return variances of the three portfolios can be computed by

$$\sigma^2_A = \sum_{i,j \in A} w_i w_j \rho_{i,j} \sigma_i \sigma_j, \quad \sigma^2_B = \sum_{i \in B} w_i w_j \rho_{i,j} \sigma_i \sigma_j,$$

and

$$\sigma^2_S = \sigma^2_A + \sigma^2_B + 2 \sum_{i \in A, j \in B} w_i w_j \rho_{i,j} \sigma_i \sigma_j, \quad (A.1)$$

where $\sigma_i$ denotes the standard deviation of the returns of stock $i$. The last summation term in $\sigma^2_S$ captures the covariance between the stocks from the two sets. With average correlation $\bar{\rho}_{A,B}$ defined in (5), the last term can be expressed by

$$\sum_{i \in A, j \in B} w_i w_j \rho_{i,j} \sigma_i \sigma_j = \bar{\rho}_{A,B} \sum_{i \in A, j \in B} w_i w_j \sigma_i \sigma_j.$$

Combining this equation with (A.1), we can solve for $\bar{\rho}_{A,B}$ from the return variances as

$$\bar{\rho}_{A,B} = \frac{\sigma^2_S - \sigma^2_A - \sigma^2_B}{2 \sum_{i \in A, j \in B} w_i w_j \sigma_i \sigma_j}. \quad (A.2)$$

To estimate the realized version of $\bar{\rho}_{A,B}$ with high-frequency data, we plug the realized variances into (A.2) to obtain

$$\text{RCorr}_{A,B}^{(J,K)} = \frac{\text{RV}_{S}^{(J,K)} - \text{RV}_{A}^{(J,K)} - \text{RV}_{B}^{(J,K)}}{2 \sum_{i \in A, j \in B} w_i w_j \sqrt{\text{RV}_{i}^{(J,K)}} \cdot \sqrt{\text{RV}_{j}^{(J,K)}}} \quad (A.3)$$

This leads to the portfolio-implied realized correlation estimator in (6).

Appendix B  Supplementary Tables
| Table 1: Average passive ownership for the highest and lowest passive ownership bins |
|---------------------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Low   | 0.04 | 0.04 | 0.03 | 0.04 | 0.05 | 0.05 | 0.05 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.05 |
| High  | 0.09 | 0.09 | 0.09 | 0.10 | 0.13 | 0.14 | 0.15 | 0.17 | 0.19 | 0.20 | 0.22 | 0.23 | 0.26 | 0.26 |

<p>| Table 2: Number of stocks in the S&amp;P 500 Index in each entire year |</p>
<table>
<thead>
<tr>
<th>----------------</th>
<th>---</th>
<th>---</th>
<th>---</th>
<th>---</th>
<th>---</th>
<th>---</th>
<th>---</th>
<th>---</th>
<th>---</th>
<th>---</th>
<th>---</th>
<th>---</th>
<th>---</th>
<th>---</th>
<th>---</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>475</td>
<td>478</td>
<td>458</td>
<td>458</td>
<td>461</td>
<td>469</td>
<td>483</td>
<td>474</td>
<td>481</td>
<td>476</td>
<td>483</td>
<td>472</td>
<td>471</td>
<td>472</td>
<td>478</td>
</tr>
</tbody>
</table>

<p>| Table 3: Average daily correlation of stock pairs in each daily correlation bin |
|--------------------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|</p>
<table>
<thead>
<tr>
<th>Years</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>-0.01</td>
<td>0.09</td>
<td>0.15</td>
<td>0.20</td>
<td>0.24</td>
<td>0.28</td>
<td>0.31</td>
<td>0.34</td>
<td>0.38</td>
<td>0.42</td>
<td>0.47</td>
<td>0.57</td>
</tr>
<tr>
<td>2004</td>
<td>-0.02</td>
<td>0.05</td>
<td>0.09</td>
<td>0.13</td>
<td>0.17</td>
<td>0.20</td>
<td>0.22</td>
<td>0.25</td>
<td>0.28</td>
<td>0.31</td>
<td>0.35</td>
<td>0.45</td>
</tr>
<tr>
<td>2005</td>
<td>-0.02</td>
<td>0.04</td>
<td>0.09</td>
<td>0.12</td>
<td>0.16</td>
<td>0.19</td>
<td>0.22</td>
<td>0.24</td>
<td>0.27</td>
<td>0.30</td>
<td>0.34</td>
<td>0.45</td>
</tr>
<tr>
<td>2006</td>
<td>-0.04</td>
<td>0.02</td>
<td>0.06</td>
<td>0.10</td>
<td>0.13</td>
<td>0.16</td>
<td>0.19</td>
<td>0.22</td>
<td>0.25</td>
<td>0.28</td>
<td>0.32</td>
<td>0.42</td>
</tr>
<tr>
<td>2007</td>
<td>0.06</td>
<td>0.14</td>
<td>0.19</td>
<td>0.23</td>
<td>0.27</td>
<td>0.30</td>
<td>0.33</td>
<td>0.36</td>
<td>0.39</td>
<td>0.42</td>
<td>0.47</td>
<td>0.57</td>
</tr>
<tr>
<td>2008</td>
<td>0.18</td>
<td>0.28</td>
<td>0.34</td>
<td>0.40</td>
<td>0.44</td>
<td>0.48</td>
<td>0.51</td>
<td>0.53</td>
<td>0.56</td>
<td>0.59</td>
<td>0.63</td>
<td>0.70</td>
</tr>
<tr>
<td>2009</td>
<td>0.04</td>
<td>0.16</td>
<td>0.23</td>
<td>0.29</td>
<td>0.34</td>
<td>0.38</td>
<td>0.42</td>
<td>0.46</td>
<td>0.49</td>
<td>0.53</td>
<td>0.58</td>
<td>0.67</td>
</tr>
<tr>
<td>2010</td>
<td>0.08</td>
<td>0.19</td>
<td>0.27</td>
<td>0.33</td>
<td>0.38</td>
<td>0.42</td>
<td>0.45</td>
<td>0.49</td>
<td>0.52</td>
<td>0.56</td>
<td>0.60</td>
<td>0.68</td>
</tr>
<tr>
<td>2011</td>
<td>0.23</td>
<td>0.33</td>
<td>0.39</td>
<td>0.44</td>
<td>0.49</td>
<td>0.53</td>
<td>0.56</td>
<td>0.59</td>
<td>0.62</td>
<td>0.66</td>
<td>0.69</td>
<td>0.76</td>
</tr>
<tr>
<td>2012</td>
<td>0.01</td>
<td>0.09</td>
<td>0.14</td>
<td>0.19</td>
<td>0.23</td>
<td>0.27</td>
<td>0.30</td>
<td>0.34</td>
<td>0.37</td>
<td>0.41</td>
<td>0.47</td>
<td>0.56</td>
</tr>
<tr>
<td>2013</td>
<td>0.01</td>
<td>0.09</td>
<td>0.14</td>
<td>0.19</td>
<td>0.23</td>
<td>0.27</td>
<td>0.30</td>
<td>0.33</td>
<td>0.36</td>
<td>0.40</td>
<td>0.44</td>
<td>0.54</td>
</tr>
<tr>
<td>2014</td>
<td>-0.02</td>
<td>0.05</td>
<td>0.11</td>
<td>0.16</td>
<td>0.20</td>
<td>0.24</td>
<td>0.28</td>
<td>0.32</td>
<td>0.35</td>
<td>0.39</td>
<td>0.45</td>
<td>0.56</td>
</tr>
<tr>
<td>2015</td>
<td>-0.01</td>
<td>0.11</td>
<td>0.18</td>
<td>0.23</td>
<td>0.28</td>
<td>0.32</td>
<td>0.36</td>
<td>0.39</td>
<td>0.43</td>
<td>0.47</td>
<td>0.52</td>
<td>0.63</td>
</tr>
<tr>
<td>2016</td>
<td>-0.14</td>
<td>-0.01</td>
<td>0.07</td>
<td>0.13</td>
<td>0.19</td>
<td>0.23</td>
<td>0.27</td>
<td>0.31</td>
<td>0.35</td>
<td>0.40</td>
<td>0.46</td>
<td>0.58</td>
</tr>
<tr>
<td>2017</td>
<td>-0.20</td>
<td>-0.09</td>
<td>-0.04</td>
<td>0.00</td>
<td>0.04</td>
<td>0.07</td>
<td>0.10</td>
<td>0.13</td>
<td>0.16</td>
<td>0.20</td>
<td>0.25</td>
<td>0.39</td>
</tr>
<tr>
<td>2018</td>
<td>-0.06</td>
<td>0.03</td>
<td>0.10</td>
<td>0.17</td>
<td>0.23</td>
<td>0.27</td>
<td>0.31</td>
<td>0.35</td>
<td>0.39</td>
<td>0.43</td>
<td>0.48</td>
<td>0.60</td>
</tr>
<tr>
<td>GICS code</td>
<td>Sector</td>
<td>Abbreviation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Energy</td>
<td>Eng</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Materials</td>
<td>Mat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Industrials</td>
<td>Ind</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Consumer discretionary</td>
<td>Disc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Consumer staples</td>
<td>Stap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Healthcare</td>
<td>Heal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Financial</td>
<td>Fin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Information technology</td>
<td>Tech</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Communication services</td>
<td>Comm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>Utilities</td>
<td>Util</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>Real estate</td>
<td>Estat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Number of stocks in different sectors

<table>
<thead>
<tr>
<th>Years</th>
<th>Comm</th>
<th>Disc</th>
<th>Stap</th>
<th>Eng</th>
<th>Fin</th>
<th>Heal</th>
<th>Ind</th>
<th>Tech</th>
<th>Mat</th>
<th>Util</th>
<th>Estat</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>23</td>
<td>71</td>
<td>36</td>
<td>25</td>
<td>78</td>
<td>50</td>
<td>55</td>
<td>78</td>
<td>35</td>
<td>33</td>
<td>3</td>
</tr>
<tr>
<td>2004</td>
<td>21</td>
<td>70</td>
<td>35</td>
<td>25</td>
<td>71</td>
<td>51</td>
<td>52</td>
<td>78</td>
<td>34</td>
<td>33</td>
<td>5</td>
</tr>
<tr>
<td>2005</td>
<td>18</td>
<td>67</td>
<td>34</td>
<td>26</td>
<td>76</td>
<td>58</td>
<td>53</td>
<td>77</td>
<td>32</td>
<td>32</td>
<td>5</td>
</tr>
<tr>
<td>2006</td>
<td>19</td>
<td>63</td>
<td>35</td>
<td>26</td>
<td>72</td>
<td>56</td>
<td>51</td>
<td>68</td>
<td>30</td>
<td>31</td>
<td>7</td>
</tr>
<tr>
<td>2007</td>
<td>23</td>
<td>64</td>
<td>38</td>
<td>30</td>
<td>73</td>
<td>52</td>
<td>50</td>
<td>61</td>
<td>28</td>
<td>29</td>
<td>10</td>
</tr>
<tr>
<td>2008</td>
<td>18</td>
<td>63</td>
<td>37</td>
<td>34</td>
<td>68</td>
<td>52</td>
<td>53</td>
<td>64</td>
<td>27</td>
<td>31</td>
<td>14</td>
</tr>
<tr>
<td>2009</td>
<td>19</td>
<td>61</td>
<td>40</td>
<td>37</td>
<td>64</td>
<td>54</td>
<td>54</td>
<td>67</td>
<td>28</td>
<td>31</td>
<td>14</td>
</tr>
<tr>
<td>2010</td>
<td>19</td>
<td>64</td>
<td>41</td>
<td>37</td>
<td>65</td>
<td>52</td>
<td>56</td>
<td>70</td>
<td>31</td>
<td>31</td>
<td>17</td>
</tr>
<tr>
<td>2011</td>
<td>18</td>
<td>65</td>
<td>39</td>
<td>40</td>
<td>65</td>
<td>51</td>
<td>57</td>
<td>62</td>
<td>30</td>
<td>31</td>
<td>16</td>
</tr>
<tr>
<td>2012</td>
<td>21</td>
<td>66</td>
<td>38</td>
<td>40</td>
<td>66</td>
<td>53</td>
<td>57</td>
<td>63</td>
<td>30</td>
<td>30</td>
<td>17</td>
</tr>
<tr>
<td>2013</td>
<td>18</td>
<td>65</td>
<td>39</td>
<td>44</td>
<td>62</td>
<td>52</td>
<td>59</td>
<td>58</td>
<td>31</td>
<td>30</td>
<td>18</td>
</tr>
<tr>
<td>2014</td>
<td>20</td>
<td>69</td>
<td>37</td>
<td>42</td>
<td>62</td>
<td>54</td>
<td>64</td>
<td>57</td>
<td>28</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>2015</td>
<td>20</td>
<td>69</td>
<td>35</td>
<td>39</td>
<td>62</td>
<td>51</td>
<td>63</td>
<td>56</td>
<td>26</td>
<td>28</td>
<td>21</td>
</tr>
<tr>
<td>2016</td>
<td>21</td>
<td>67</td>
<td>35</td>
<td>35</td>
<td>58</td>
<td>56</td>
<td>63</td>
<td>58</td>
<td>24</td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td>2017</td>
<td>19</td>
<td>63</td>
<td>33</td>
<td>29</td>
<td>63</td>
<td>57</td>
<td>67</td>
<td>56</td>
<td>23</td>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td>2018</td>
<td>20</td>
<td>63</td>
<td>32</td>
<td>28</td>
<td>64</td>
<td>58</td>
<td>66</td>
<td>60</td>
<td>23</td>
<td>28</td>
<td>31</td>
</tr>
</tbody>
</table>
Table 6: Selected sector pairs with high and low daily correlations in each year

<table>
<thead>
<tr>
<th>Years</th>
<th>Lowest</th>
<th>2nd lowest</th>
<th>3rd lowest</th>
<th>3rd highest</th>
<th>2nd highest</th>
<th>Highest</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>(Eng, Stap)</td>
<td>(Eng, Tech)</td>
<td>(Eng, Heal)</td>
<td>(Ind, Fin)</td>
<td>(Fin, Mat)</td>
<td>(Mat, Ind)</td>
</tr>
<tr>
<td>2005</td>
<td>(Heal, Eng)</td>
<td>(Eng, Stap)</td>
<td>(Eng, Tech)</td>
<td>(Ind, Mat)</td>
<td>(Fin, Util)</td>
<td>(Eng, Util)</td>
</tr>
<tr>
<td>2006</td>
<td>(Stap, Eng)</td>
<td>(Heal, Eng)</td>
<td>(Disc, Eng)</td>
<td>(Mat, Eng)</td>
<td>(Fin, Ind)</td>
<td>(Ind, Mat)</td>
</tr>
<tr>
<td>2007</td>
<td>(Heal, Tech)</td>
<td>(Heal, Disc)</td>
<td>(Heal, Eng)</td>
<td>(Eng, Mat)</td>
<td>(Fin, Mat)</td>
<td>(Mat, Ind)</td>
</tr>
<tr>
<td>2008</td>
<td>(Eng, Disc)</td>
<td>(Eng, Fin)</td>
<td>(Fin, Util)</td>
<td>(Util, Eng)</td>
<td>(Eng, Mat)</td>
<td>(Mat, Ind)</td>
</tr>
<tr>
<td>2009</td>
<td>(Disc, Heal)</td>
<td>(Stap, Heal)</td>
<td>(Tech, Heal)</td>
<td>(Ind, Mat)</td>
<td>(Eng, Ind)</td>
<td>(Eng, Mat)</td>
</tr>
<tr>
<td>2010</td>
<td>(Stap, Disc)</td>
<td>(Tech, Stap)</td>
<td>(Heal, Stap)</td>
<td>(Ind, Mat)</td>
<td>(Ind, Fin)</td>
<td>(Eng, Ind)</td>
</tr>
<tr>
<td>2011</td>
<td>(Tech, Stap)</td>
<td>(Disc, Stap)</td>
<td>(Comm, Stap)</td>
<td>(Mat, Fin)</td>
<td>(Ind, Mat)</td>
<td>(Fin, Ind)</td>
</tr>
<tr>
<td>2012</td>
<td>(Disc, Util)</td>
<td>(Util, Tech)</td>
<td>(Util, Comm)</td>
<td>(Mat, Ind)</td>
<td>(Fin, Mat)</td>
<td>(Fin, Ind)</td>
</tr>
<tr>
<td>2013</td>
<td>(Tech, Util)</td>
<td>(Tech, Stap)</td>
<td>(Eng, Heal)</td>
<td>(Mat, Ind)</td>
<td>(Fin, Mat)</td>
<td>(Ind, Fin)</td>
</tr>
<tr>
<td>2014</td>
<td>(Util, Disc)</td>
<td>(Util, Tech)</td>
<td>(Comm, Util)</td>
<td>(Fin, Mat)</td>
<td>(Tech, Fin)</td>
<td>(Fin, Ind)</td>
</tr>
<tr>
<td>2015</td>
<td>(Util, Eng)</td>
<td>(Real, Eng)</td>
<td>(Disc, Util)</td>
<td>(Fin, Tech)</td>
<td>(Util, Real)</td>
<td>(Fin, Ind)</td>
</tr>
<tr>
<td>2016</td>
<td>(Fin, Util)</td>
<td>(Disc, Util)</td>
<td>(Util, Eng)</td>
<td>(Fin, Tech)</td>
<td>(Mat, Fin)</td>
<td>(Fin, Ind)</td>
</tr>
<tr>
<td>2017</td>
<td>(Fin, Util)</td>
<td>(Eng, Util)</td>
<td>(Util, Comm)</td>
<td>(Mat, Ind)</td>
<td>(Mat, Fin)</td>
<td>(Ind, Fin)</td>
</tr>
<tr>
<td>2018</td>
<td>(Disc, Util)</td>
<td>(Eng, Util)</td>
<td>(Fin, Util)</td>
<td>(Fin, Tech)</td>
<td>(Mat, Ind)</td>
<td>(Fin, Ind)</td>
</tr>
</tbody>
</table>

Table 7: Average daily beta of stocks in each daily beta bin

<table>
<thead>
<tr>
<th>Years</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>0.38</td>
<td>0.52</td>
<td>0.63</td>
<td>0.77</td>
<td>0.86</td>
<td>0.93</td>
<td>1.02</td>
<td>1.13</td>
<td>1.25</td>
<td>1.45</td>
<td>1.88</td>
</tr>
<tr>
<td>2004</td>
<td>0.42</td>
<td>0.53</td>
<td>0.64</td>
<td>0.73</td>
<td>0.82</td>
<td>0.90</td>
<td>1.00</td>
<td>1.10</td>
<td>1.25</td>
<td>1.50</td>
<td>2.03</td>
</tr>
<tr>
<td>2005</td>
<td>0.50</td>
<td>0.63</td>
<td>0.73</td>
<td>0.82</td>
<td>0.91</td>
<td>0.99</td>
<td>1.06</td>
<td>1.12</td>
<td>1.20</td>
<td>1.33</td>
<td>1.62</td>
</tr>
<tr>
<td>2006</td>
<td>0.38</td>
<td>0.51</td>
<td>0.60</td>
<td>0.71</td>
<td>0.80</td>
<td>0.89</td>
<td>1.00</td>
<td>1.11</td>
<td>1.25</td>
<td>1.46</td>
<td>1.94</td>
</tr>
<tr>
<td>2007</td>
<td>0.48</td>
<td>0.61</td>
<td>0.69</td>
<td>0.79</td>
<td>0.87</td>
<td>0.94</td>
<td>1.01</td>
<td>1.09</td>
<td>1.21</td>
<td>1.36</td>
<td>1.71</td>
</tr>
<tr>
<td>2008</td>
<td>0.50</td>
<td>0.62</td>
<td>0.71</td>
<td>0.80</td>
<td>0.87</td>
<td>0.95</td>
<td>1.03</td>
<td>1.13</td>
<td>1.29</td>
<td>1.48</td>
<td>1.87</td>
</tr>
<tr>
<td>2009</td>
<td>0.32</td>
<td>0.47</td>
<td>0.61</td>
<td>0.75</td>
<td>0.89</td>
<td>1.03</td>
<td>1.21</td>
<td>1.36</td>
<td>1.60</td>
<td>1.96</td>
<td>2.74</td>
</tr>
<tr>
<td>2010</td>
<td>0.46</td>
<td>0.56</td>
<td>0.69</td>
<td>0.81</td>
<td>0.93</td>
<td>1.03</td>
<td>1.14</td>
<td>1.26</td>
<td>1.38</td>
<td>1.53</td>
<td>1.82</td>
</tr>
<tr>
<td>2011</td>
<td>0.46</td>
<td>0.58</td>
<td>0.73</td>
<td>0.86</td>
<td>0.97</td>
<td>1.05</td>
<td>1.15</td>
<td>1.25</td>
<td>1.37</td>
<td>1.53</td>
<td>1.80</td>
</tr>
<tr>
<td>2012</td>
<td>0.36</td>
<td>0.48</td>
<td>0.62</td>
<td>0.78</td>
<td>0.90</td>
<td>1.01</td>
<td>1.11</td>
<td>1.23</td>
<td>1.38</td>
<td>1.54</td>
<td>1.88</td>
</tr>
<tr>
<td>2013</td>
<td>0.58</td>
<td>0.72</td>
<td>0.78</td>
<td>0.84</td>
<td>0.92</td>
<td>0.99</td>
<td>1.07</td>
<td>1.15</td>
<td>1.25</td>
<td>1.37</td>
<td>1.60</td>
</tr>
<tr>
<td>2014</td>
<td>0.38</td>
<td>0.55</td>
<td>0.68</td>
<td>0.80</td>
<td>0.90</td>
<td>0.99</td>
<td>1.07</td>
<td>1.14</td>
<td>1.22</td>
<td>1.34</td>
<td>1.61</td>
</tr>
<tr>
<td>2015</td>
<td>0.53</td>
<td>0.67</td>
<td>0.75</td>
<td>0.85</td>
<td>0.91</td>
<td>0.96</td>
<td>1.01</td>
<td>1.05</td>
<td>1.12</td>
<td>1.20</td>
<td>1.38</td>
</tr>
<tr>
<td>2016</td>
<td>0.30</td>
<td>0.51</td>
<td>0.67</td>
<td>0.82</td>
<td>0.91</td>
<td>1.01</td>
<td>1.14</td>
<td>1.24</td>
<td>1.35</td>
<td>1.53</td>
<td>1.96</td>
</tr>
<tr>
<td>2017</td>
<td>0.06</td>
<td>0.27</td>
<td>0.50</td>
<td>0.66</td>
<td>0.77</td>
<td>0.90</td>
<td>1.03</td>
<td>1.17</td>
<td>1.29</td>
<td>1.46</td>
<td>1.80</td>
</tr>
<tr>
<td>2018</td>
<td>0.19</td>
<td>0.40</td>
<td>0.57</td>
<td>0.72</td>
<td>0.82</td>
<td>0.89</td>
<td>0.94</td>
<td>1.01</td>
<td>1.08</td>
<td>1.19</td>
<td>1.44</td>
</tr>
</tbody>
</table>