A Practical Model for Prediction of Intraday Volatility

Young Li
Bloomberg Enterprise Quants
xli528@bloomberg.net

Sunday 1st August, 2021

Abstract

Although intraday volatility has been studied extensively for many asset classes, there are still important questions to be answered: Is the unconditional mean diurnal profile time-invariant? Does statistical ergodicity hold for the profile? Is it possible to predict intraday volatility in absolute terms? In this study, we explore answers to these questions. Intraday bar data are collected for securities in Russell 3000 Index, FTSE 100 Index and CAC All-Tradable Index. For intraday volatility measure, we choose the one that makes use of open-high-low-close prices of each time bucket. We first propose a predictive model where the intraday volatility is decomposed into three multiplicative components: daily volatility, time-scaling factor, and normalized diurnal profile. We then estimate unconditional mean diurnal profiles of securities in four market capitalization groups: MEGA-, BIG-, MID-, and SMALL-CAP over time and observe that they are time invariant. We further compare time-averaged with ensemble-averaged diurnal profiles and find that strict statistical ergodicity doesn't hold but the two are not far apart. Finally we evaluate model performance using both timeseries and cross-sectional approaches and conclude that both approaches are unbiased.

1 Introduction

Intraday volatility has been investigated for many asset classes including equity, futures, foreign exchange, and fixed income. Andersen and Bollerslev (1997) observed pervasive intraday periodicity in the return volatility in foreign exchange and equity markets. They proposed a decomposition model where intraday volatility was a product of daily level and intraday periodic component. Engle and Sokalska (2012) used a similar decomposition model for forecasting of volatility of high-frequency returns. They found that forecasts from a pooled cross-section of companies outperformed the corresponding forecasts from a company-by-company estimation. Vatter et al. (2015) performed a non-parametric estimation of intraday spot volatility by disentangling instantaneous trend and seasonality. With a study of exchange rate returns, they found that failing to factor in the seasonality led to misestimation of the intraday spot volatility. Bollerslev et al. (2000) provided a detailed characterization of the return volatility in US Treasury bond futures contracts and found that public information in the form of regularly scheduled macroeconomic announcements was an important source of volatility at the intraday level. Zhang and Dufour (2019) studied the intraday volatility of European government bonds using the framework of the multiplicative component GARCH model.

A general finding of these studies is that the intraday volatility demonstrates a U-shaped diurnal profile. Although extensive studies have been conducted in this area, there are still important questions to be answered: Is the unconditional mean diurnal profile time-invariant? Does statistical ergodicity hold for the profile? Is it possible to predict intraday volatility in absolute terms? In this study, we will attempt to answer these questions.

2 Methodologies

2.1 Measurement of Intraday Volatility

Garman and Klass (1980) proposed the following volatility measure using open-high-low-close:

$$V_{ohlc} = 0.5\left[\log(H) - \log(L)\right]^2 - \left[2\log(2) - 1)\right]\left[\log(C) - \log(O)\right]^2$$

This model is based on the assumption that price returns follow a Wiener process with zero drift and constant infinitesimal variance. It's constructed by minimizing the variance of a quadratic estimator subject to the constraints of price and time symmetry and scale invariance of volatility. Chan and Lien (2003) shows that the above measure is unbiased and more efficient than the measure using only high/low proposed by Parkinson (1980). We will use $\sigma_{ohlc} = \sqrt{V_{ohlc}}$ as the volatility measure in this paper to be compatible with convention.

For each intraday time interval, the above equation can be applied to calculate volatility during that period.

2.2 Decomposition of Intraday Volatility

Andersen and Bollerslev (1997) studied the dynamics of intraday returns and built the following decomposition model:

$$R_{t,n} = E(R_{t,n}) + \frac{\sigma_t s_{t,n} Z_{t,n}}{\sqrt{N}}$$

where t refers to a day, n refers to a time interval within day t, $E(R_{t,n})$ denotes the unconditional mean, N refers to the number of return intervals per day, σ_t is the average volatility level of day t, $s_{t,n}$ is the periodic component for nth interval, and $Z_{t,n}$ is a i.i.d. random variable with mean zero and variance 1. The decomposition model has been used in predicting equity intraday volatilities (Engle and Sokalska 2012).

In this study, we build our intraday volatility prediction model using the decomposition as follows:

$$\begin{aligned} \sigma_{t,n} &= \sigma_t \rho_t s_{t,n} \\ \frac{\sum_n s_{t,n}}{N} &= 1 \end{aligned} \tag{2.1}$$

where σ_t is daily volatility estimate for day t, ρ_t is the estimate of ratio between average intraday volatility level and daily volatility level for day t, and $s_{t,n}$ is the intraday periodic volatility component for nth interval estimated for day t. The periodic components are scaled to have mean equal to 1. ρ_t is introduced to account for the effect of time scaling of volatility: it can be considered a factor of conversion from daily volatility σ_t to intraday volatility mean level. Henceforth, we will refer to $\{s_{t,n}$, n $\in [1..N]\}$ as the *intraday volatility profile*.

2.2.1 Daily Volatility Prediction Using EWMA

The following exponentially weighted moving average model (EWMA) is used for prediction of daily volatility:

$$\sigma_{t+1}^2 = \lambda \sigma_t^2 + (1 - \lambda) r_t^2 \tag{2.2}$$

where r_t is price return on day t, and λ is the decay factor. In the EWMA model, the latest observations carry the highest weight. J.P.Morgan/Reuters (1996) recommends the decay factor of 0.94 for daily timeseries based on prediction error analysis of over 480 financial timeseries. We will use this value in this study.

2.2.2 Intraday Volatility Profile Prediction

We will use two approaches to estimate time-scaling factor ρ_t and diurnal profile $\{s_{t,n}, n \in [1..N]\}$ in Equation 2.1.

- timeseries-based: the time-scaling factor and diurnal profile are estimated as averages over time for the same security,
- cross-section or ensemble-based: the time-scaling factor and diurnal profile are estimated as averages over securities in an ensemble.

We use market capitalization group as the ensemble in this study. The capitalization groups are defined as follows:

- MEGA-CAP: $\geq$ \$200 bln for US securities, or $\geq$ \$100 bln for non-US securities,
- BIG-CAP: $\geq$ \$10 bln,
- MID-CAP: $\geq$ \$2 bln,
- SMALL-CAP: $<$ \$2 bln

3 Data

Securities in the following indices are used for this study:

- Russell 3000 Index: an index composed of 3000 large U.S. companies, as determined by market capitalization. This portfolio of Securities represents approximately 98% of the investable U.S. equity market.
- FTSE 100 Index: a capitalization-weighted index of the 100 most highly capitalized companies traded on the London Stock Exchange.
- CAC All-Tradable Index: an index that contains all the stocks of the Euronext Paris market that have an annual Free Float Velocity over 20%.

Daily and intraday data used are as follows:

- Daily index member securities of the above indices from 2021-01-01 to 2021-07-23
- Daily closing price time series are obtained for each security in the above indices from 2020-04-01 to 2021-07-23
- Bloomberg intraday 1-minute frequency bar data (open/high/low/close) are obtained for each security in the above indices from 2021-01-01 to 2021-07-23

4 Results

4.1 Daily Volatility Prediction Using EWMA

Figure 1 shows plots of EWMA-predicted daily volatility and absolute daily return timeseries for four securities, respectively in MEGA-, BIG-, MID-, and SMALL-CAP market capitalization group of Russell 3000 Index. The absolute daily return is used here as a measure of true daily volatility. The predicted values capture the levels and trends fairly well. The measure fluctuates much more widely than the predicted; it's expected since the measure only uses two day prices. We use it here as it's a more timely measure than other measures such as 30-day realized volatility. Its time period matches the 1-day prediction horizon of EWMA.

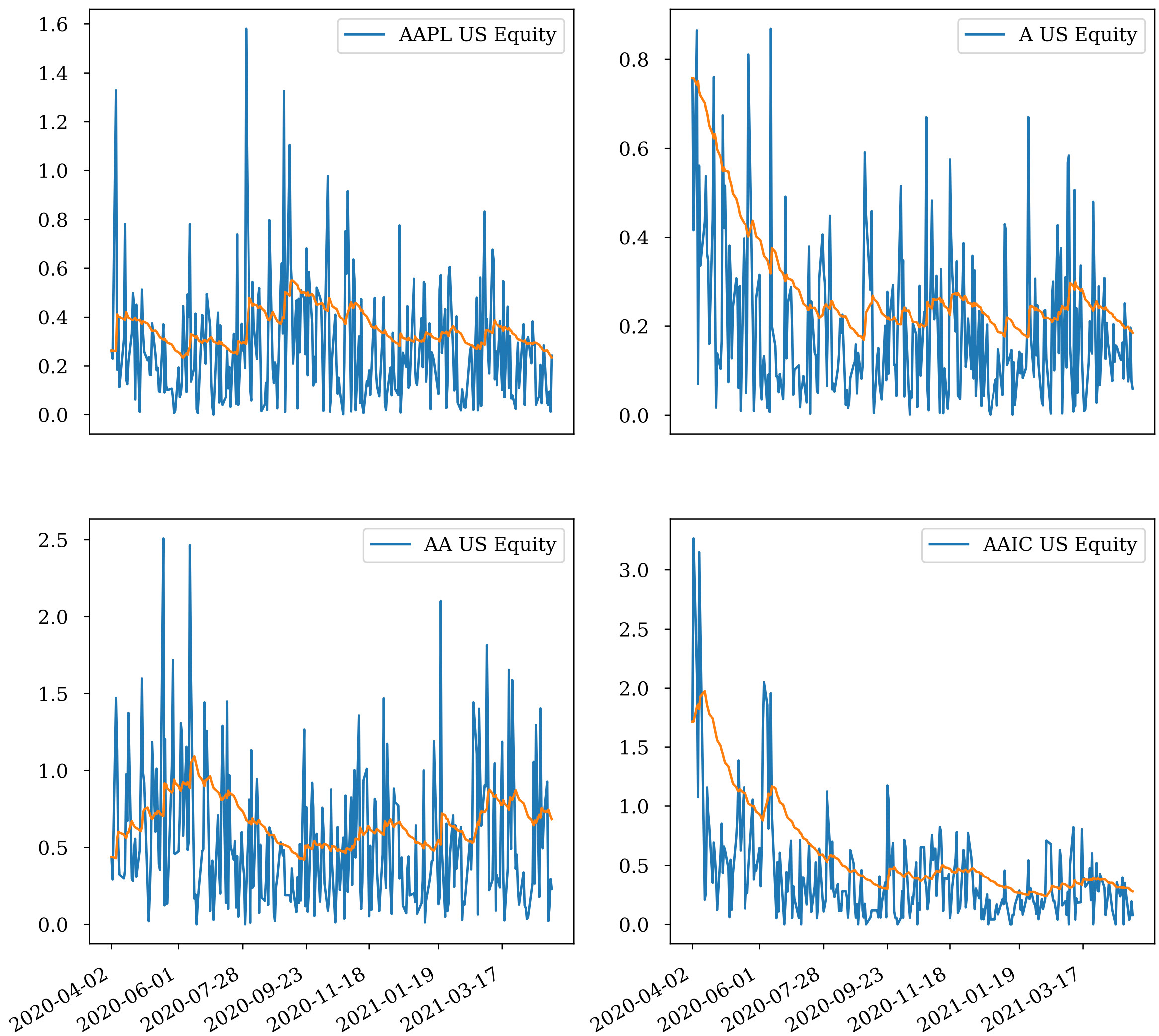


Figure 1: EWMA-Predicted (orange colored) vs. observed (blue colored) daily volatilities for four stocks in Russell 3000 Index that fall in the MEGA-CAP, BIG-CAP, MID-CAP, and SMALL-CAP groups respectively. The observed volatility here is taken as the absolute daily returns scaled by the annualization factor $\sqrt{252}$.

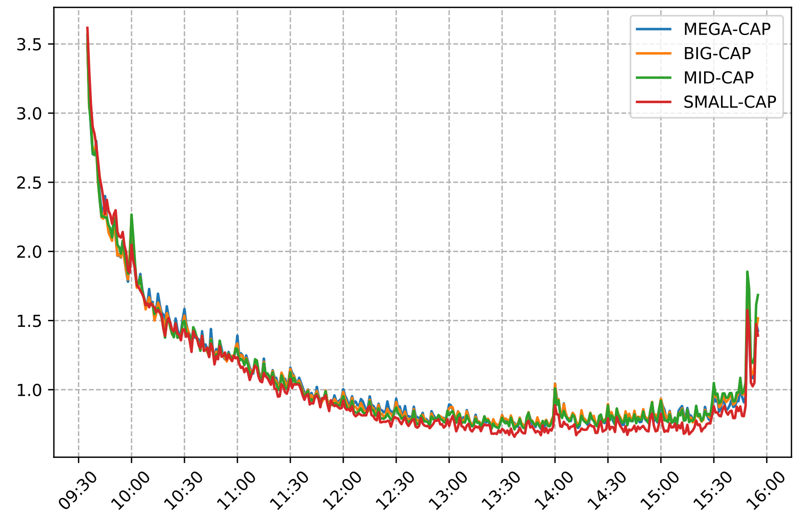
4.2 Average Intraday Volatility Profiles

Figure 2a shows intraday volatility profiles averaged over Russell 3000 Index member market capitalization groups. The differences between the top three groups are trivial, but SMALL-CAP shows slightly smaller values than the other three groups for the second half of the regular trading session.

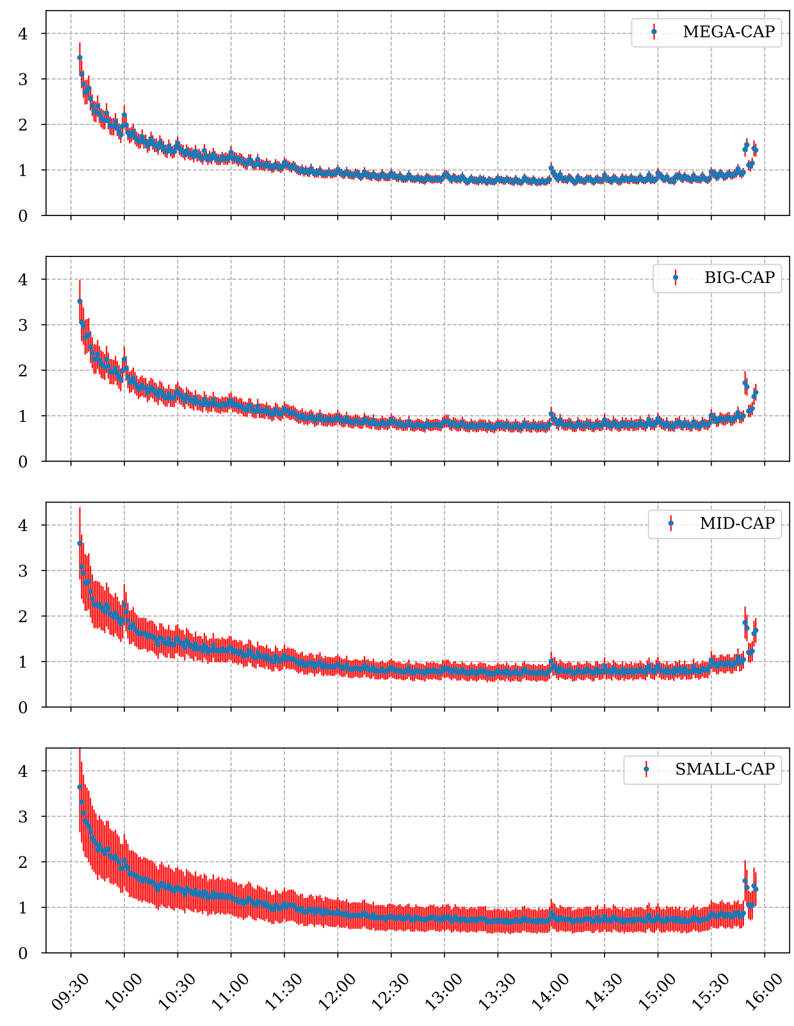
Figure 2b shows the same average profiles with error bars added, where the standard deviation of a bucket average is calculated using inter-quantile-range. As expected, SMALL-CAP demonstrates larger variability than the other three groups.

Figure 3 shows average intraday profiles for securities in FTSE 100 Index in three periods. The first period is before US daylight savings time starts (2021-03-14), the second between British summer time starts (2021-03-28) and US daylight savings time starts, and the last after British summer time starts. What's interesting is that the afternoon peaks occur respectively at 2:30 PM, 1:30 PM, and again 2:30 PM. These times correspond to 9:30 AM New York time when the New York Stock Exchange and the Nasdaq Stock Market open. We can thus infer with high confidence that the afternoon peaks are the result of US stock market open.

The same linkage of afternoon peaks to US stock market open is also seen for securities in CAC All-Tradable Index (Figure 4).

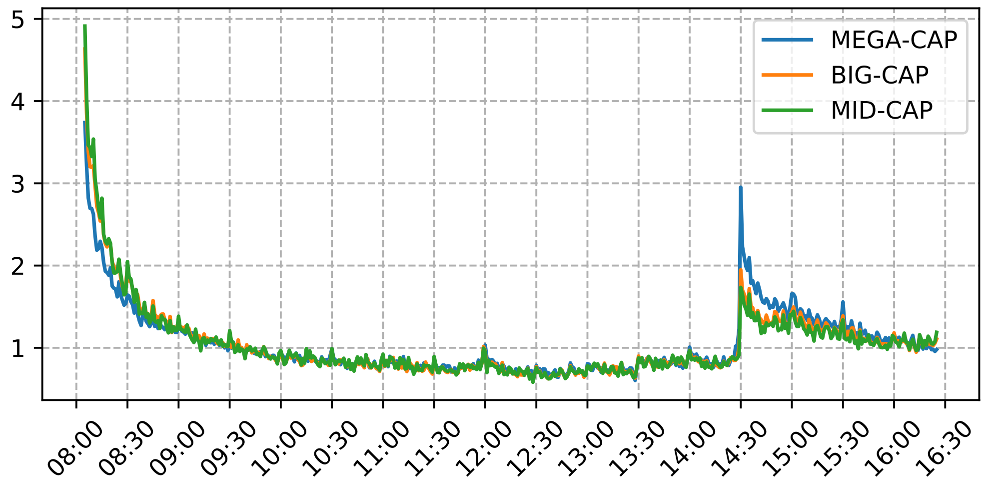


(a) average profiles

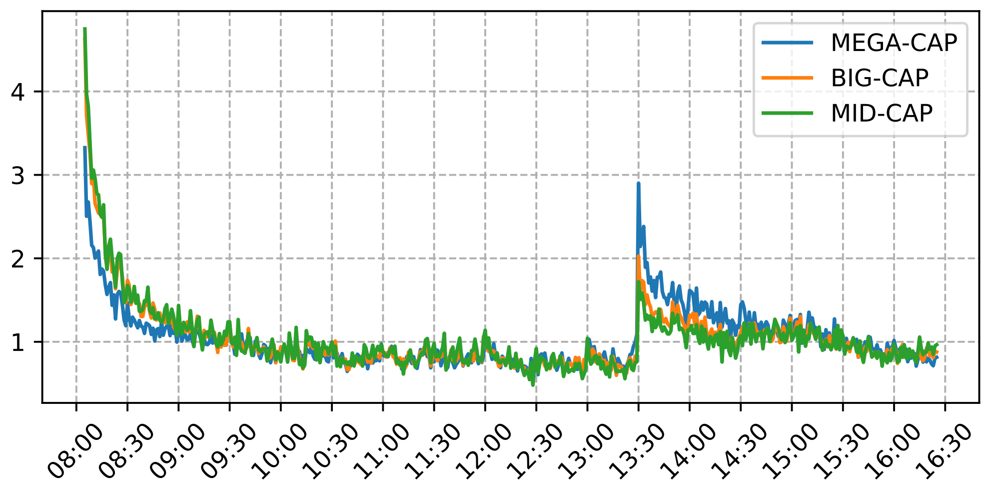


(b) profiles with error bars

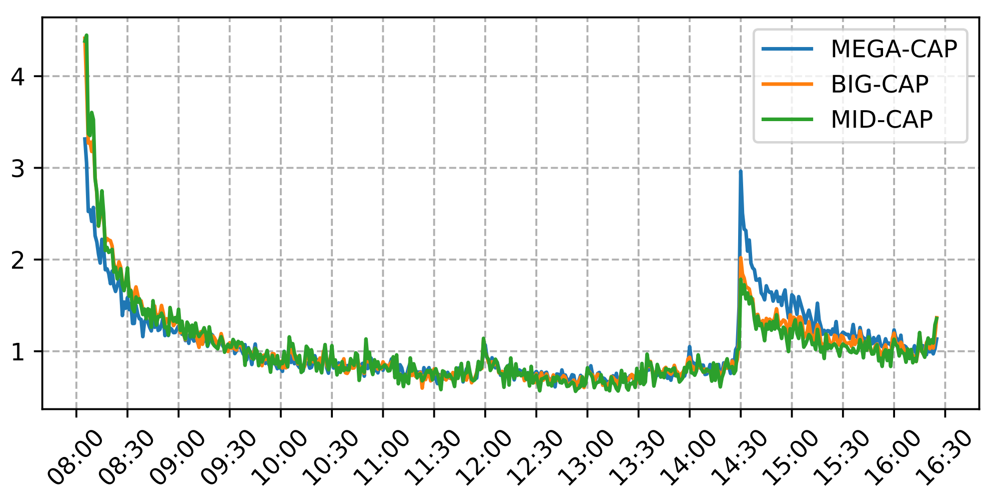
Figure 2: Average intraday volatility profiles of Russell 3000 Index members grouped by market capitalization during 2021-01-01 to 2021-04-30. The averaging is performed using quantiles to mitigate impacts of outliers: $(q_{25}+q_{50}+q_{75})/3$; the standard deviation is calculated using inter-quantile-range: $(q_{75}-q_{25})/1.35$.



(a) 2021-01-01 – 2021-03-13.

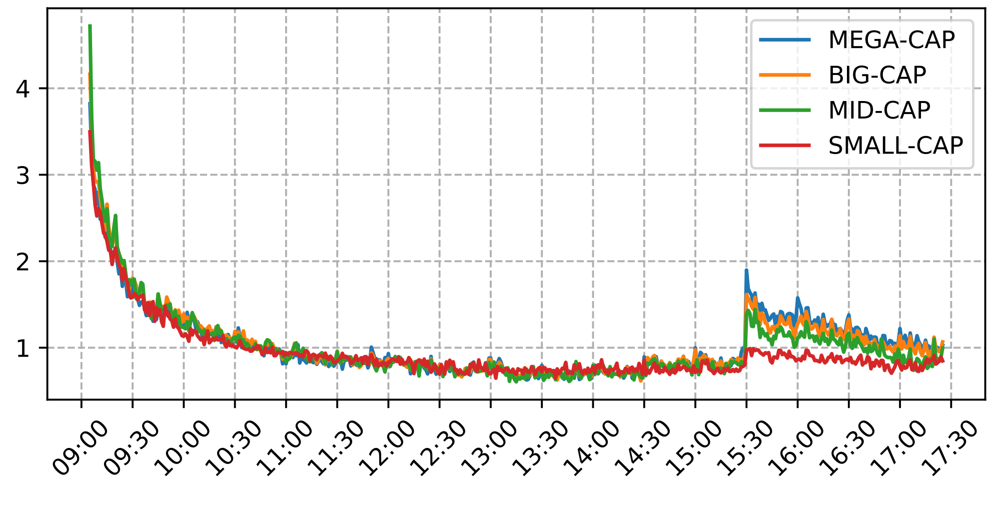


(b) 2021-03-14 (US Daylight Saving Time Starts) – 2021-03-27

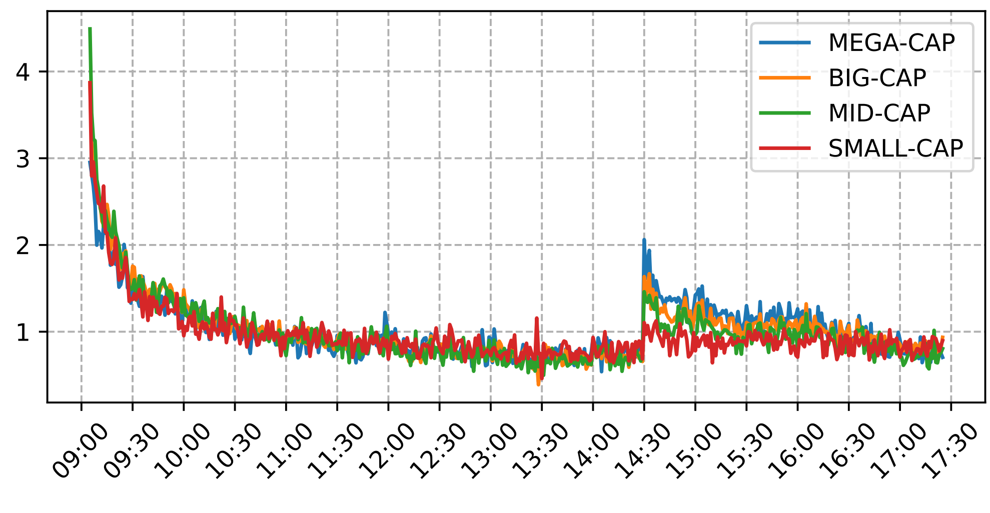


(c) 2021-03-28 (British Summer Time Starts) – 2021-04-30

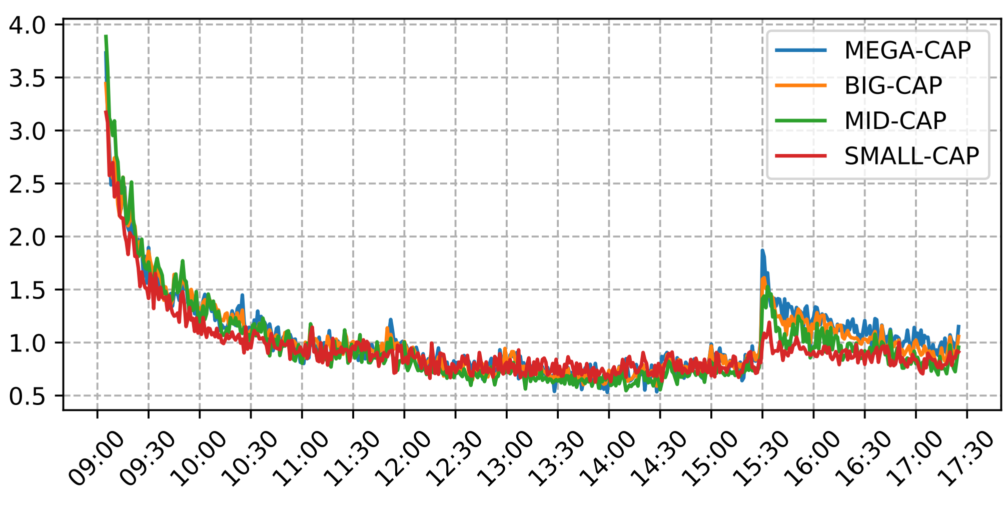
Figure 3: Intraday volatility profiles of FTSE 100 Index members for three periods.



(a) 2021-01-01 – 2021-03-13.



(b) 2021-03-14 (US Daylight Saving Time Starts) – 2021-03-27



(c) 2021-03-28 (France Summer Time Starts) – 2021-04-30

Figure 4: Intraday volatility profiles of CAC All-Tradable Index members for three periods.

4.3 Is Intraday Volatility Profile Time Invariant?

Figure 5 compares average intraday volatility profiles over three periods for four market capitalization groups of Russell 3000 Index. The differences over time are visually trivial, suggesting that the profile is approximately time invariant.

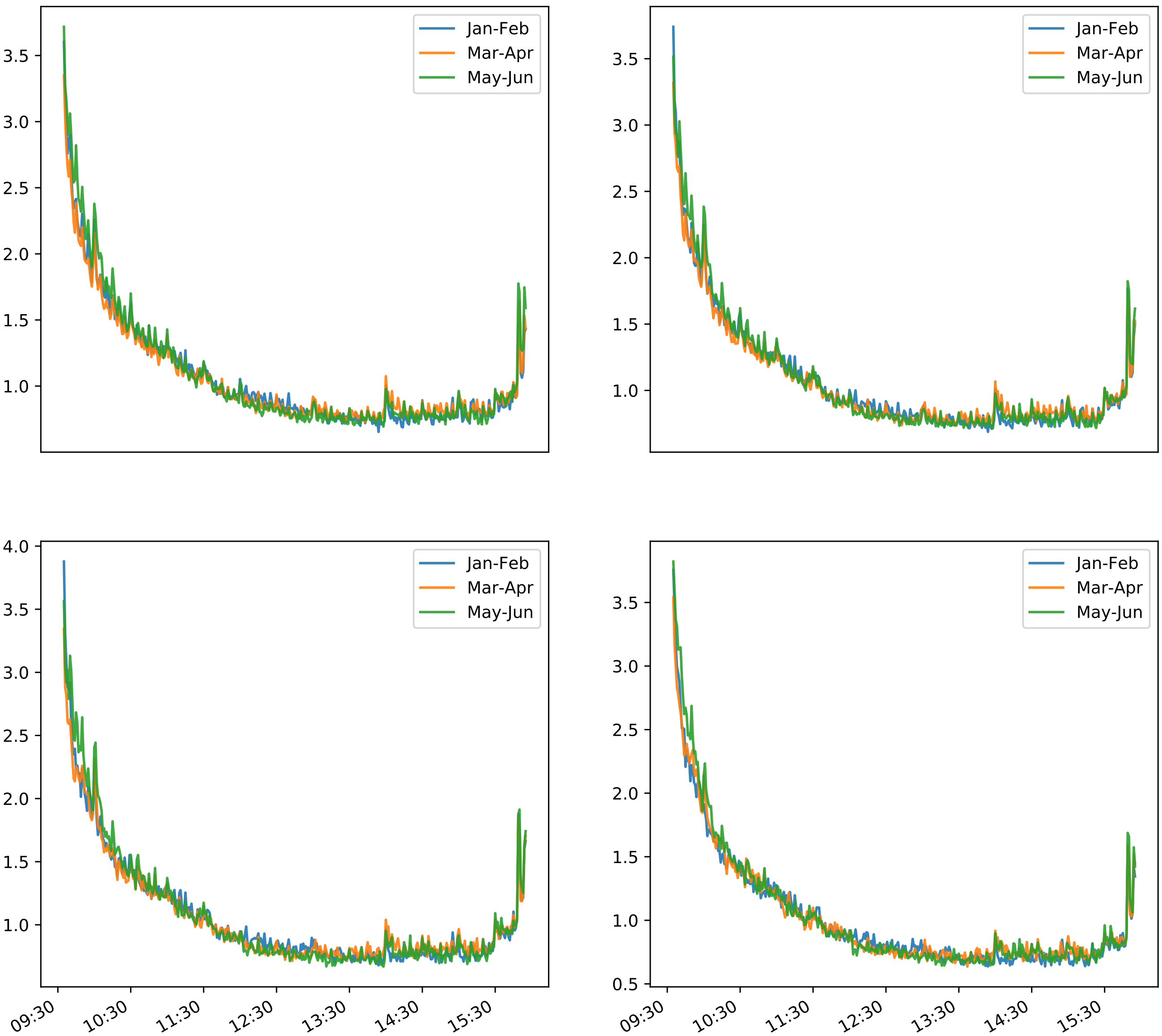


Figure 5: Intraday volatility profiles averaged over three periods of Jan-Feb, Mar-Apr, and May-Jun of 2021 for Russell 3000 Index members divided into four market capitalization groups, from left-top: MEGA-CAP, BIG-CAP, MID-CAP, and SMALL-CAP.

4.4 Ergodicity of Intraday Volatility Profiles

An ergodic process is one which conforms to the ergodic theorem(Birkhoff 1942 and Walters 1982). The theorem states that the time average of a conforming process is equal to the ensemble average. In practice this means that statistical sampling can be performed at one instant across a group of identical processes or sampled over time on a single process with no change in the measured result. We investigate here if we can treat the intraday volatility profile as an ergodic process, where the time average for a single security over a prolonged period is equal to the average over an ensemble of securities in the same market capitalization group during a very short period of time.

The ergodicity has practical implication. For some illiquid securities, there are few intraday ticks even over a prolonged time period, and in this case it would be extremely helpful if we can combine the ticks for a large ensemble of equivalent securities to compute the intraday profile.

Figure 6 compares the intraday volatility profile of a single security averaged over a prolonged period with that averaged over an ensemble and short time period. The AAPL/MEGA-CAP and AA/MID-CAP pairs show very small discrepancies, whereas the A/BIG-CAP and AAN/SMALL-CAP pairs show visually evident discrepancies. Due to illiquidity, for some small-cap securities, even in a prolonged time period, there are still not sufficiently large number of data points, which can lead to a widely fluctuating time-averaged profile. This is the case for AAN. On the other hand, the SMALL-CAP ensemble has a large number of data points available, which allows for a much more reliable estimation of ensemble-average profile. In general the ergodicity does not seem to hold strictly, but the two averages are not far apart.

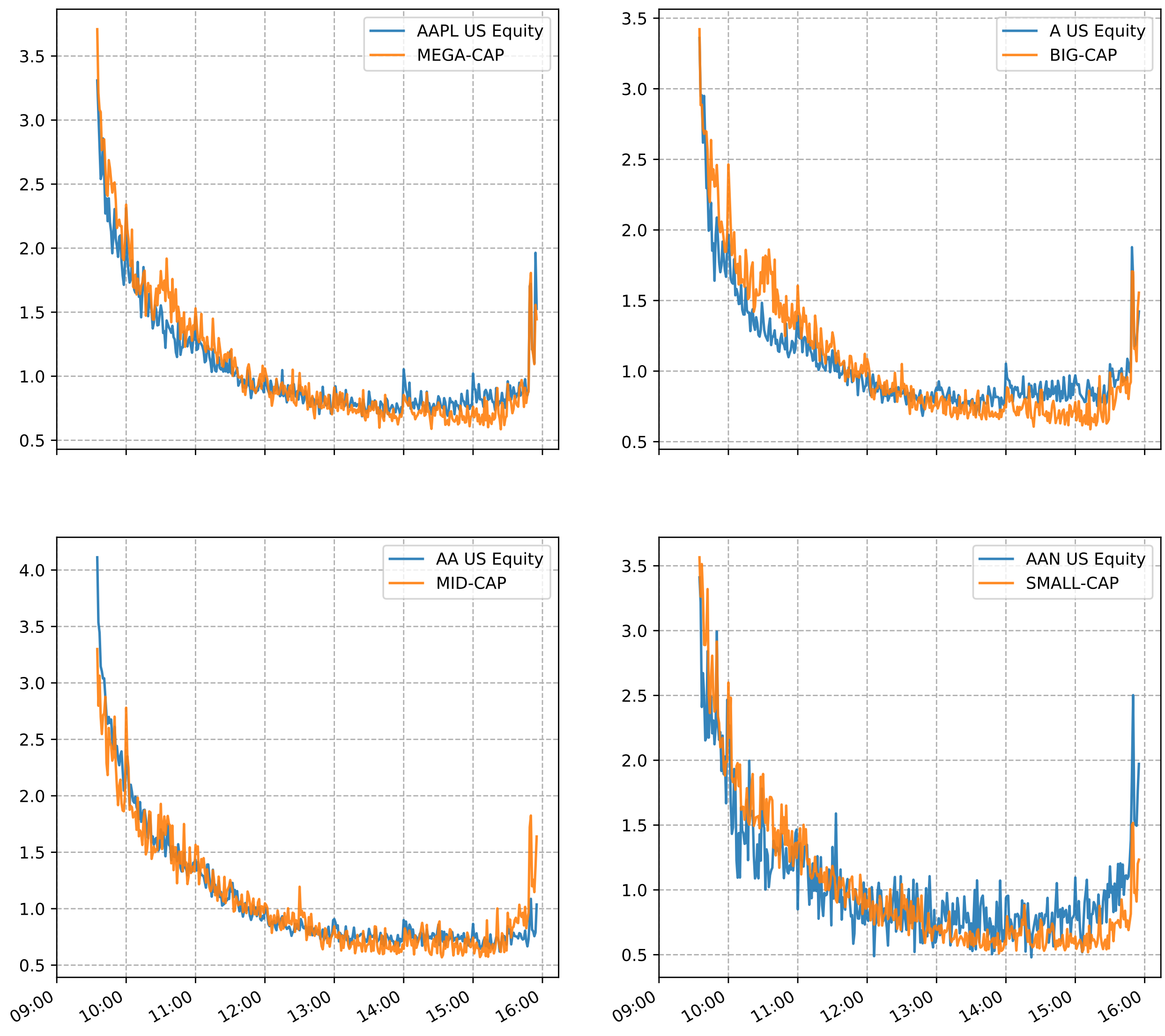


Figure 6: Intraday volatility profiles of a single security over a prolonged period vs. averaged over an ensemble of market capitalization group. The prolonged period is from 2021-01-01 to 2021-06-30, and the ensemble average is over the securities in the capitalization group and over the first week of July of 2021.

4.5 Intraday Volatility Mean Level Relative to Daily Volatility

Figure 7 shows intraday volatility mean levels relative to EWMA-estimated daily volatility for four securities in the MEGA-, BIG-, MID-, and SMALL-CAP group over time. The ratios for the first three groups fluctuate around 1, whereas the ratio for the SMALL-CAP security are below 1 most of the time. The Augmented Dickey-Fuller test using statsmodels.tsa.stattools.adfuller shows that the four ratio timeseries are all stationary.

Figure 8 shows cross-sectional distributions of the ratios for four market capitalization groups. The mean ratio decreases with decreasing market capitalization, with it as low as 0.73 for SMALL-CAP.

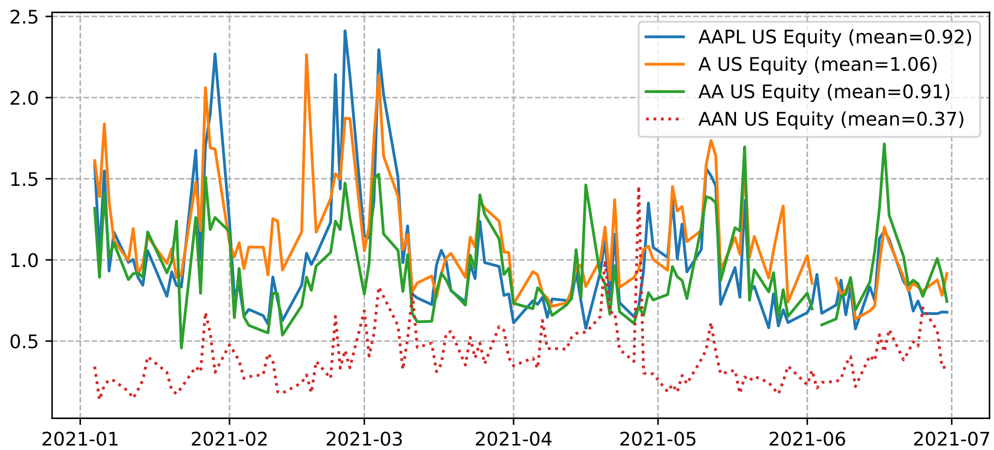


Figure 7: Timeseries of intraday volatility mean level divided by EWMA-estimated daily volatility for four securities in MEGA-, BIG-, MID-, and SMALL-CAP market capitalization group of Russell 3000 Index.

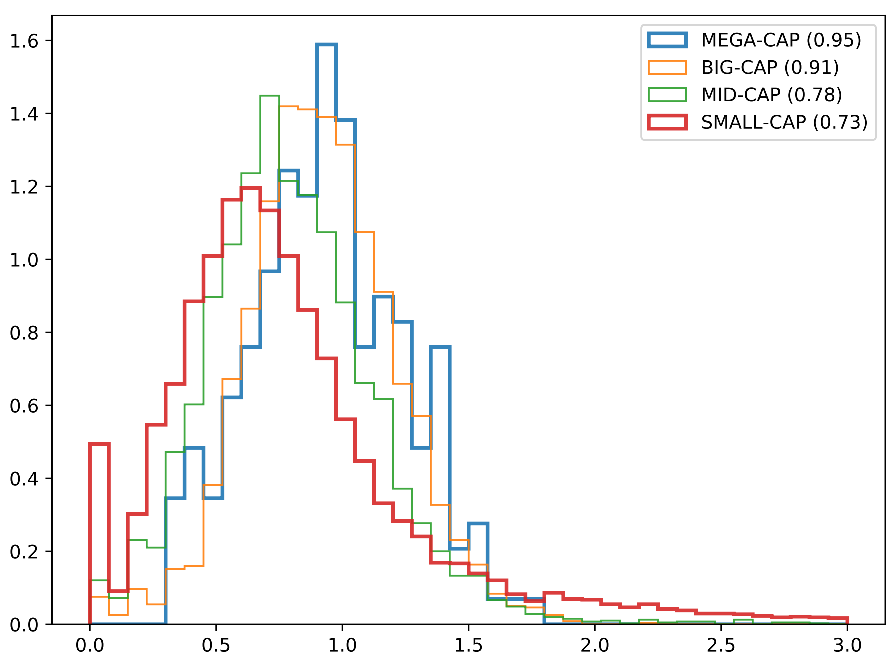


Figure 8: Normalized histograms of intraday volatility mean level divided by EWMA-estimated daily volatility for securities in Russell 3000 Index grouped by market capitalization. The numbers in parentheses are sample means. The data used is for date range between 2021-07-02 and 2021-07-09.

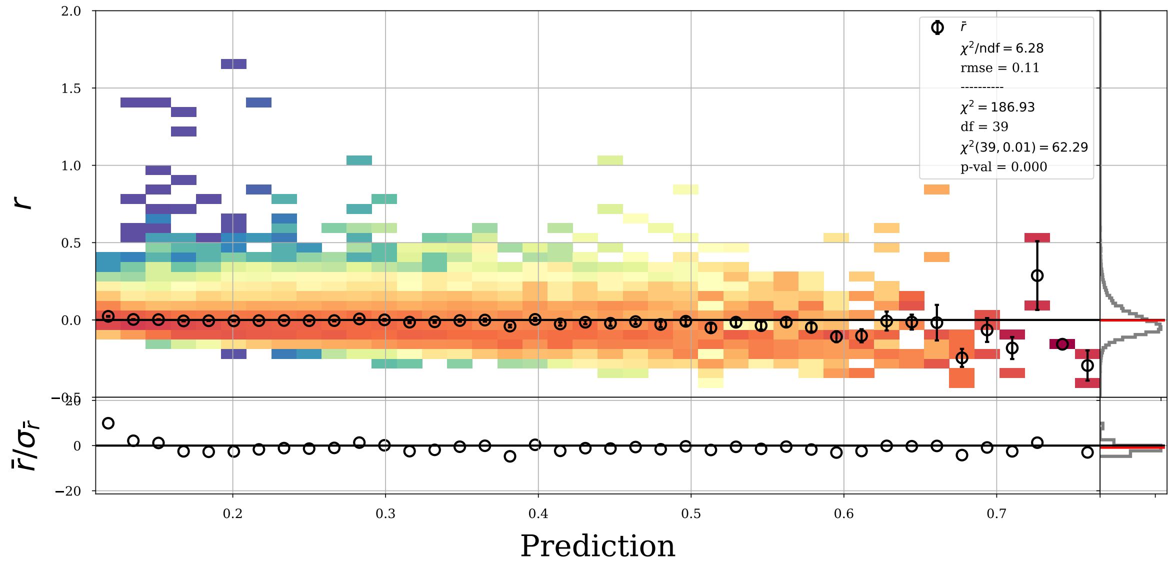
4.6 Intraday Volatility Model Performance

To assess performance of the intraday volatility model (Equation 2.1), we will use two securities, MSFT and KODK, the former being a MEGA-CAP security and the latter a SMALL-CAP security.

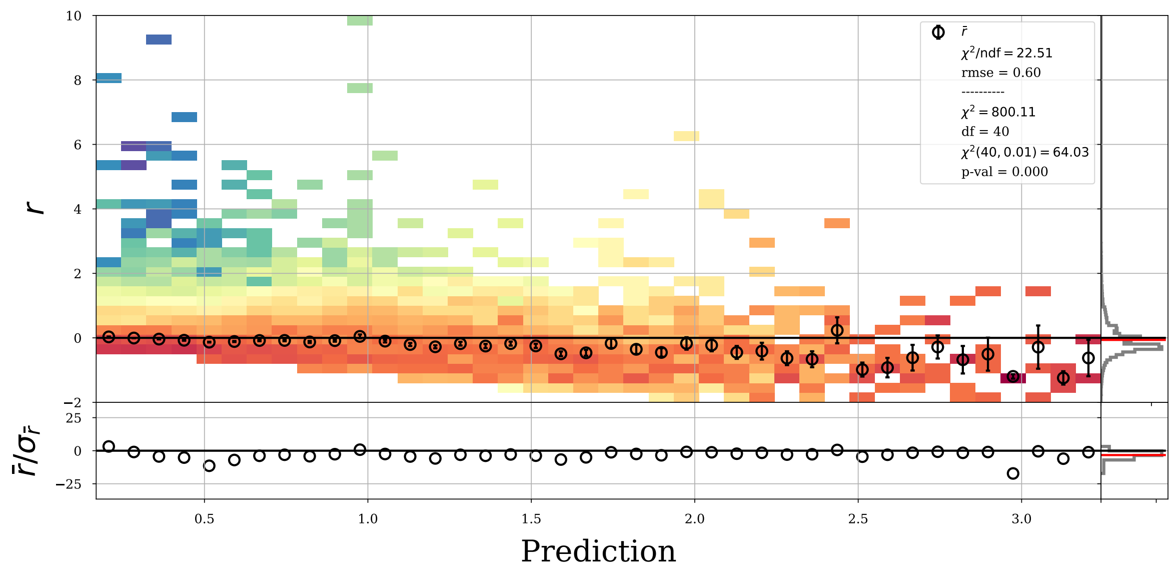
Figure 9 shows prediction residual distribution analysis for the intraday volatility model using timeseries approach, where the conversion factor ρ_t and profile $\{s_{t,n}, n \in [1..N]\}$ are estimated as their respective averages over all days prior to t for the same security, and σ_t is estimated using EWMA Equation 2.2. The plot shows that the model predictions match means of observations well for both securities.

Figure 10 shows prediction residual distribution analysis for the intraday volatility model using cross-section approach, where the conversion factor ρ_t and profile $\{s_{t,n}, n \in [1..N]\}$ are estimated as their respective *ensemble* means during the week prior to prediction date t, and σ_t is estimated using EWMA Equation 2.2. The *ensemble* here refers to the set of securities in the same market capitalization group as the target security. The plot shows that the model predictions match means of observations well for both securities.

Although both approaches provide unbiased estimation, the timeseries-based approach gives smaller root mean squared errors: 0.11 vs. 0.15 for MSFT and 0.60 vs. 1.01 for KODK.

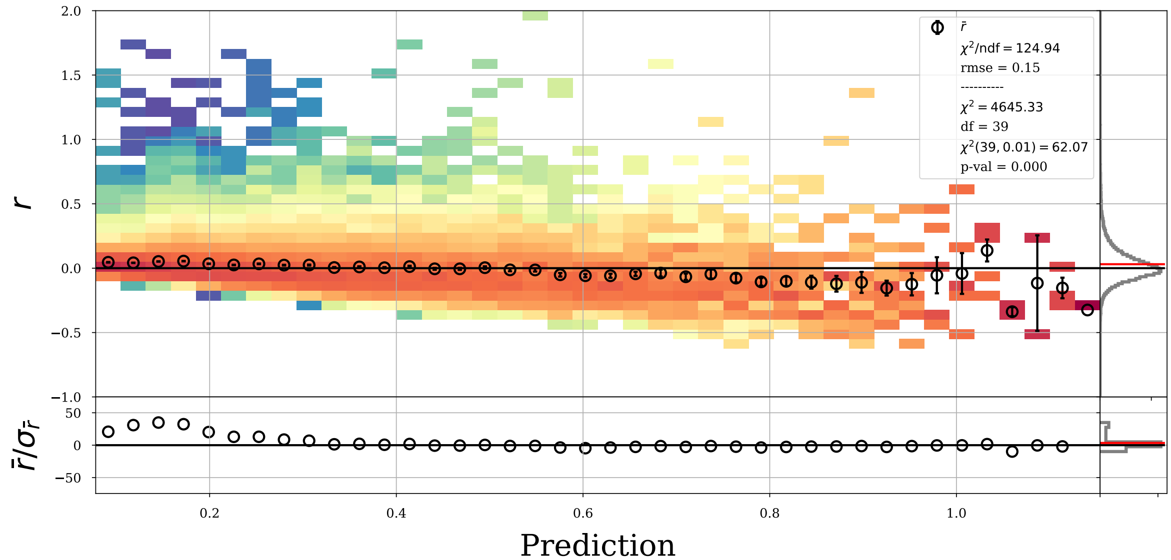


(a) MSFT

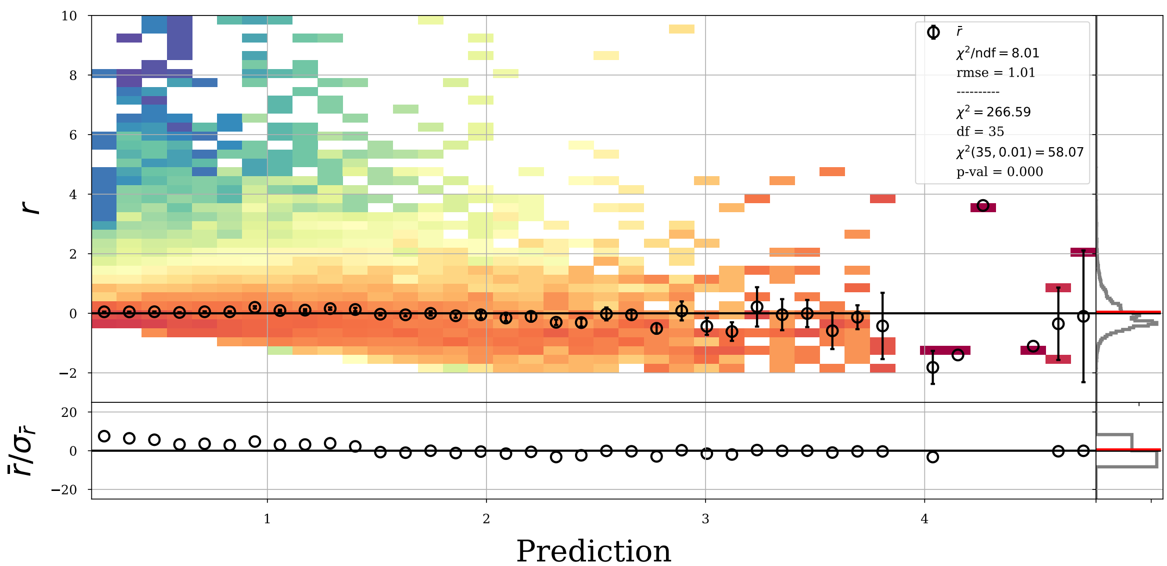


(b) KODK

Figure 9: Performance of intraday volatility model using timeseries approach. Data from 2021-01-01 to 2021-07-23 are used. Predictions are made for dates since 2021-05-03. For each prediction date, intraday volatility average profiles and mean relative levels are estimated using all data prior to that date.



(a) MSFT



(b) KODK

Figure 10: Performance of intraday volatility model using cross-section approach. Data from 2021-01-01 to 2021-07-23 are used. For each prediction date, intraday volatility average profiles and mean relative levels are estimated as the ensemble means of the same market capitalization groups in the previous week.

5 Discussion

In our prediction model Equation 2.1, the volatility time-scaling factor ρ is introduced to allow for the definition of a normalized profile whose average over time buckets is unity. This is important as we want to separate out intraday volatility level so that we have a pure diurnal profile. We show that the time-scaling factor for small-cap securities is significantly smaller than those for bigger market capitalization groups (Figure 7 and Figure 8).

It's well known that daily volatilities for small-cap securities are much greater than those for large-cap securities. A natural question is thus: does intraday volatility profile demonstrate the same level of difference? Our results with securities in Russell 3000, FTSE 100 and CAC All-Tradable Index members show that the discrepancies between different market capitalization groups are much less significant than with daily volatility (Figure 2a, Figure 3, and Figure 4).

The profiles used here are their respective averages over a long period of time and can be considered *unconditional mean profiles*. Although it's conceivable that a profile for a security on a particular day is influenced by relevant news/announcements (e.g., Torben et al. 2019), when these profile realizations are averaged over a sufficiently long period of time, those influences can be netted out and the unconditional mean can be obtained. Figure 5 shows that the unconditional mean profiles obtained for two-month periods change little over time for every market capitalization groups in the Russell 3000 Index. This suggests that the unconditional mean profile is time invariant.

The preliminary investigation of profile ergodicity shows that it doesn't hold strictly; at the same time, the time-averaged and the ensemble-averaged profiles are not far apart (Figure 6). Considering the often observed illiquidity of small-cap securities, the ensemble-averaged profile may still provide a good estimation of unconditional mean.

The performance of model Equation 2.1 is evaluated with one MEGA-CAP security MSFT and one SMALL-CAP security KODK. Estimation is performed using timeseries and cross-section approaches respectively. The results show that both approaches give unbiased estimates, with the timeseries approach gives smaller overall RMSE (Figure 9 and Figure 10).

We find that intraday volatility profiles of securities traded in London Stock Exchange and EURONEXT Paris show local peaks in the afternoons that correspond to New York Stock Exchange and NASDAQ Stock Market open time. This illustrate significant systematic inter-market impacts on volatility profiles. Ozenbas (2008) investigated the pattern of intra-day volume of trading in US and European equity markets and found that the pattern of volume in the European markets was significantly impacted by the opening of trading in the US markets.

References

Andersen, T. G. and T. Bollerslev (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance 4*(2), 115–158. High Frequency Data in Finance, Part 1.

Birkhoff, G. D. (1942). What is the ergodic theorem? *The American Mathematical Monthly 49*(4), 222–226.

Bollerslev, T., J. Cai, and F. M. Song (2000). Intraday periodicity, long memory volatility, and macroeconomic announcement effects in the us treasury bond market. *Journal of Empirical Finance 7*(1), 37–55.

Chan, L. and D. Lien (2003). Using high, low, open, and closing prices to estimate the effects of cash settlement on futures prices. *International Review of Financial Analysis 12*, 35–47.

Engle, R. F. and M. E. Sokalska (2012, 01). Forecasting intraday volatility in the US equity market. Multiplicative component GARCH. *Journal of Financial Econometrics 10*(1), 54–83.

Garman, M. B. and M. J. Klass (1980). On the estimation of security price volatilities from historical data. *Journal of Business 53*, 67–78.

J.P.Morgan/Reuters (1996). Riskmetricstechnical document. Technical report, J.P.Morgan/Reuters.

Ozenbas, D. (2008). Intra-day trading volume patterns of equity markets: A study of us and european stock markets. *International Business & Economics Research Journal 7*(8), 79–93.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business 53*, 61–65.

Torben, G. A., T. Martin, and T. Viktor (2019). Time-varying periodicity in intraday volatility. *Journal of the American Statistical Association 114*(528), 1695–1707.

Vatter, T., H.-T. Wu, V. Chavez-Demoulin, and B. Yu (2015). Non-parametric estimation of intraday spot volatility: Disentangling instantaneous trend and seasonality. *Econometrics 3*(4), 864–887.

Walters, P. (1982). *An introduction to ergodic theory*. USA: Springer-Verlag.

Zhang, H. and A. Dufour (2019). Modeling intraday volatility of european bond markets: A data filtering application. *International Review of Financial Analysis 63*, 131–146.