Predicting Intraday Return Patterns based on Overnight Returns for the US Stock Market

Hao Li[1], Cees Diks[1,2], and Valentyn Panchenko[3]

[1]CeNDEF, Amsterdam School of Economics, University of Amsterdam, Amsterdam, The Netherlands

[2]Tinbergen Institute, Amsterdam, The Netherlands

[3]School of Economics, Faculty of Business, University of New South Wales, Sydney, Australia

March 12, 2020

Abstract

This paper investigates predicting intraday return patterns conditional on observed overnight returns. Based on Trade and Quote data, we find evidence for dependence between overnight returns and subsequent intraday first and last half-hour return patterns for the S&P 500 Exchange-Traded Fund for the time period from 2003 to 2013 with both statistical and economic significance. Our methodology allows studying the return patterns documented in the existing theoretical and empirical literature in more detail. Moreover, we find that both the first and the last half-hours offer trading opportunities for day traders. Specifically, 20-minute after the market opens and the last 30-minute before the market closes seem to be the best holding periods for investors in terms of annualized returns, Sharpe ratios, and Certainty Equivalent Returns.

JEL Classification: C14, C22, C53

Keywords: Overnight returns; Intraday returns; High-frequency trading; Nonlinear dependence; Day trading

1 Introduction

Studies such as Wood *et al.* (1985), Smirlock and Starks (1986), Harris (1986), Jain and Joh (1988), Hong and Wang (2000) and Bogousslavsky (2016) show that average returns (mean returns, or market mean returns) vary over the day.[1] To explain this, Hong and Wang (2000) solve an equilibrium model for a competitive stock market in which investors trade for asset allocation as well as for informational reasons. The model indicates that time patterns in intraday stock returns are driven by time-varying hedging demand and time-varying information asymmetry. In addition, Hong and Wang (2000) find U-shaped and inverted U-shaped mean intraday stock return patterns during trading sessions.[2]

Here we propose a novel approach based on cumulative returns, referred to as cumulative regression (CumRe), to forecast intraday market return patterns conditional on overnight returns. By using this intraday pattern forecast, day traders can decide when to close out same-day trading positions to make profits. The literature already reports evidence of dependence between the overnight return and the subsequent intraday returns. For instance, Liu and Tse (2017) find that overnight returns significantly help predict the first half-hour market returns and the last half-hour market returns. Berkman *et al.* (2012) find a strong tendency for positive overnight returns followed by negative intraday returns. In practice, traders often view the close-to-open gap (i.e. overnight return in this paper) as an important marker; serving as either a support zone or resistance zone for that day's trading activity. Meanwhile, major macroeconomic announcements, such as GDP and CPI announcements, and the vast majority of earnings announcements are released either before the market opens or after the market closes (see e.g., Oldfield and Rogalski, 1980; Gao *et al.*, 2018; Bogousslavsky, 2018). As stated by Hong and Wang (2000), market prices cease to provide information to the uninformed investors when the market is closed, but continue to provide information after the market reopens. We propose predicting the

[1] A more precise statement is that market average returns vary during exchange trading sessions. We decompose returns into intraday and overnight returns based on exchange trading and non-trading periods for simplicity.

[2] See Subsection C.3 "Time Patterns of Returns" of Section V in Hong and Wang (2000) for details.

subsequent intraday return patterns on day t based on close-to-open returns (i.e. overnight returns) observed earlier on the same day.

We employ both a parametric CumRe method, based on predictive regressions, and a non-parametric CumRe method using kernel regressions, to predict intraday patterns in mean returns. The predictive regression is widely used in the return predictability literature (see e.g., Ang and Bekaert, 2007; Goyal and Welch, 2008; Rapach and Zhou, 2013) considering linear dependence between the predictor(s) and the response variable. Non-parametric methods provide flexibility in accommodating nonlinear dependence between the overnight returns and intraday returns. Guidolin *et al.* (2009) find that capturing nonlinear effects may be the key to improving forecasts and that non-linear dynamics should be modelled for the US stock market. Maasoumi and Racine (2002) conclude that the evidence of dependence-based on linear models is somewhat inconclusive and indicate the presence of nonlinear dependence in returns. Tsay (2010) also shows that nonlinearities exist in high-frequency financial time series. Abhyankar *et al.* (1997) find nonlinear dependence and chaos in US stock market returns. Therefore, we compare the intraday return patterns predicted by using both the parametric CumRe and the more flexible non-parametric CumRe. We evaluate the parametric and non-parametric versions of CumRe in terms of out-of-sample forecast performance as well as economic significance.

The main goal of our paper is to identify upward/downward trends in intraday cumulative returns in response to overnight returns. We utilize these time patterns in dependence between the overnight returns and the subsequent intraday returns to formulate promising trading strategies for day traders. We examine the intraday return patterns within the first and the last half-hours separately. These two half-hours during each trading day are most popular among traders. According to, among others, Wood *et al.* (1985), Andersen and Bollerslev (1997), and Heston *et al.* (2010), volatilities during these two half-hours are usually higher than during the other trading periods. In addition, the trading volumes during these time intervals are, on average, larger than during other time intervals (see e.g., Gao *et al.*, 2018). Moreover, overnight returns can significantly help predict returns within the first and last half-hour, compared with intraday returns in other time intervals.

Similar to studies such as Rogalski (1984), Jain and Joh (1988), Wood *et al.* (1985), we also use the return on the market index as a proxy of the market return. However, we use ETFs transaction-level data[3], instead of the underling non-tradable cash index. This is because the spot prices of the underling index may be distorted by non-synchronous trading (Cliff *et al.*, 2008; Yang *et al.*, 2010). We examine SPDR S&P 500 ETF (ticker symbol SPY), the largest ETF with the structure of Unit Investment Trust. Compared with other ETFs for the US market, the much larger trading volumes and higher daily turnover of SPY indicate that it is widely used as a trading vehicle for more active traders. Besides, SPDR S&P 500 ETF has the lowest and relatively most stable bid-ask spread.[4] According to Amihud and Mendelson (1987), the stochastic nature of stock returns is affected by two major factors: the arrival of new information and noise transactions, i.e. the smaller the bid-ask spread, the smaller the noise-induced variance. Therefore, we are likely to capture the impact of new information on returns when we examine SPY data.

Our results show that CumRe can identify the intraday patterns in returns documented in the previous literature (see e.g., Amihud and Mendelson, 1987; Hong and Wang, 2000; Liu and Tse, 2017). Specifically, the market overreaction at the beginning, the market under-reaction at the closing and the opposite trends during the first and the last half-hours of trading. Besides, non-parametric CumRe shows intraday return patterns in more detail and we find that the non-parametric CumRe improves the out-of-sample forecast performance in terms of the Diebold and Mariano (1995) test statistic and the leave-one-*year*-out cross-validation.

Overall, the overnight returns help predict the market directions within the first and the last half-hours of the trading sessions. The timing strategies based on the signs of

[3]We use data from NYSE Trade and Quote (TAQ) database of Wharton Research Data Services. This database contains intraday transactions data in two databases, Consolidated Quote database for intraday quotes and Consolidated Trade database for intraday quotes, for all securities listed on the New York Stock Exchange (NYSE) and American Stock Exchange (AMEX), as well as Nasdaq National Market System (NMS) and SmallCap issues.

[4]The average bid/ask spread is one penny ($0.01) even when market volatility spikes and during financial crisis. Ross, J.E. (2018, January 17). Steady Under Pressure: How SPY Performs in Times of Market Stress. Retrieved from https://global.spdrs.com/blog/post/2018/jan/steady-under-pressure-how-spy-performs-in-times-of-market-stress.html. Last access date 03-12-2018.

the overnight returns always generate higher annualized returns, and usually yield better Sharpe ratios. Besides, our results indicate that holding established positions for 15-minute or 20-minute after the market opens yields higher annualized returns and substantially higher Sharpe ratios, and always positive and higher utility gains than holding positions for the full last half-hour of trading time.

This paper is organized as follows. Section 2 introduces methodology. Section 3 describes data and presents evidence of predictability. Section 4 compares the out-of-sample forecast performance of non-parametric CumRe and parametric CumRe. Section 5 presents economic significance in terms of market timing strategies and utility gains of mean-variance investors. Section 6 concludes.

2 Parametric and Non-parametric CumRe

We consider the case of discrete-time stock prices. Let t indicate the trading day. The regular trading hours in the United States are from 9:30 a.m. to 4 p.m. Eastern time. Consequently, the trading period on day t lasts 6.5 hours (390 minutes).

We employ the cumulative returns similar to Tsay (2010, Ch. 5) and explore their pattern forecasts. We consider day traders, who open and close their positions on the same day, trying to maximize their profits based on cumulative returns.

Cumulative returns are formulated by expanding a 5-minute interval starting from the same initial price. We focus on 5-minute interval, because it strikes a balance between market microstructure effects and the blurring price reactions over larger time intervals (see e.g., Andersen *et al.*, 2007). There are 78 cumulative 5-minute return observations per trading day.

Formally, for each i and j, satisfying $0 \leq i < j \leq 78$, let $p_{t,i}$ be the log price at which a security is first traded on day t, i.e. $i \times 5$ minutes after the market opens to establish a position on that security, and $p_{t,j}$ be the log price at which the security is executed to close out the position established earlier on day t.

A return pattern on day t can be defined as $\mathbf{y}_t = \left(y_t^{i,i+1}, y_t^{i,i+2}, \ldots, y_t^{i,j}\right)$, where $y_t^{i,j}$ are the cumulative returns on day t, given by $y_t^{i,j} \equiv p_{t,j} - p_{t,i}$. We focus on trading within the first half-hour and the last half-hour per day, that is, patterns $\mathbf{y}_t^f = \left(y_t^{0,1}, y_t^{0,2}, \ldots, y_t^{0,6}\right)$ and $\mathrm{y}_t^l = \left(y_t^{72,73}, y_t^{72,74}, \ldots, y_t^{72,78}\right)$, respectively.

Let τ represent the τ^{th} element in the return pattern vector $\mathbf{y}_t$. There are 6 elements in the vector of the first or the last half-hour return pattern. The τ^{th} element of the return pattern on day t, $\mathbf{y}_t$, can be expressed as $y_{t,\tau}$, where $\tau = 1, \ldots, 6$, $t = 1, \ldots, T$, T is the total number of trading days in the sample.

In the setting of cumulative regression, we assume the vector-valued time-series process $\{\mathbf{Y}_t\}$ under consideration (e.g. the first or last half-hour return pattern) to be strictly stationary. Given an overnight return $X_t = x$, the conditional mean of the τ^{th} element in $\mathbf{Y}_t$ is denoted by

$$m(\tau, x) \equiv \mathbb{E}(Y_{t,\tau}|x). \tag{1}$$

Parametric predictive regression is commonly used in the return predictability literature (see e.g., Stambaugh,1999; Ang and Bekaert, 2007; Goyal and Welch, 2008; Rapach and Zhou, 2013). Gao *et al.* (2018) and Liu and Tse (2017) employ the predictive regressions to investigate the linear dependence between the predictor(s) and the response variables, namely the overnight returns and intraday returns.[5] We propose a parametric CumRe based on predictive regressions of intraday return patterns on the overnight returns. The parametric CumRe is specified as

$$y_{t,\tau} = \alpha_\tau + \beta_\tau x_t + \epsilon_{t,\tau}, \tag{2}$$

where α_τ is a constant term and β_τ is the coefficient of the predictor x_t, i.e. the observed overnight return on day t.

In nonparametric regression, Nadaraya-Watson estimators based on Gaussian kernels are among the most popular estimators. The corresponding nonparametric estimator of

[5]The overnight return is included in formulating the first half-hour return in Gao *et al.* (2018), resulting in a statistically linear dependence between the first half-hour returns and the last half-hour returns.

$m(\cdot)$ is given by

$$\hat{m}(\tau, x) \equiv \sum_{t=1}^{T} y_{t,\tau} w_t(x, h), \tag{3}$$

where $w_t(x,h) = K\left(\frac{x_t - x}{h}\right) / \sum_{t'=1}^{T} K\left(\frac{x_{t'} - x}{h}\right)$ with $K(u) = 1/\sqrt{2\pi}\exp(-u^2/2)$, and the bandwidth for the τ^{th} element in $\mathbf{y}_t$, h_τ, is to be selected from a pre-defined set of values $h_1, \cdots, h_q$ according to least-squares cross-validation (CV). Here we consider a range of bandwidths between 0.1 and 15.[6]

Also in time-series settings, CV can be used for bandwidth selection (Yao and Tong, 1998). For details regarding the corresponding mild regularity conditions for strictly stationary discrete-time time series, we refer to Yao and Tong (1998). According to Härdle *et al.* (2004), minimizing CV(h) is equivalent to minimizing the average squared error (ASE) while trading off bias and variance. For the purpose of prediction, we employ leave-one-out[7] CV to select a bandwidth which minimizes the mean squared prediction error (MSPE). The selected bandwidth for τ, h_τ is given by

$$h_\tau = \underset{h \in \{h_1, \ldots, h_q\}}{\arg\min} \frac{1}{T} \sum_{t=1}^{T} (y_{t,\tau} - \hat{y}_{-t,\tau}(\tau, x))^2, \tag{4}$$

where $\hat{y}_{-t,\tau}(x, h)$ is the estimator of $y_{t,\tau}$ formed by leaving out the t^{th} observation when generating the prediction for observation t.

In the setting of conditional mean forecasts of return patterns, the bandwidth h_τ herein

[6]The function for generating the bandwidth selection set is taken to be equidistant on a logarithmic scale, and given by $h_k = a \cdot \exp\left[(k-1)/(q-1) \cdot (\log b - \log a)\right]$, $k = 1, \ldots, q$, where a is the lower bound of the bandwidth selection set, b is the upper bound of the bandwidth selection set, q is the number of bandwidths in the selection set. We take $q = 80$ to account for the trade-off between computational intensity and precision of bandwidth selection. The upper bound here is determined by the extreme values of returns in percentage according to our data set.

[7]To take into account serial correlation in time series, we used the autocorrelation functions and the partial autocorrelation functions of elements in $\mathbf{y}_t^f$ and $\mathbf{y}_t^l$. The first order autocorrelation is not significantly different from 0 at the 90% significance level for all the elements in $\mathbf{y}_t^f$. Even if the first order autocorrelation for a couple of elements in $\mathbf{y}_t^f$ is significantly different from 0 at the 95% significance level, these autocorrelation values (less than $|0.18|$) are not large enough to make a dramatic difference in bandwidth selection. That is, given this level of autocorrelation, the bandwidths selected by cross-validation and a correlation-corrected method (CDPI) are comparable for a time series process. We refer to Opsomer *et al.* (2001) for details.

controls the amount of data which is effectively used. This approach gives more weight to the historical overnight return observations X_t such that $\|X_t - x\|$ is closer to 0, which is similar to some local weighted average methods such as the nearest-neighbors method in machine learning.

For each τ, the estimator of the conditional cumulative distribution function (CDF) $F_{Y_\tau|X}(y|x)$ of the cumulative returns $Y_{t,\tau}$ given an overnight return x is

$$\widehat{F}_{Y_\tau|X}(\mathbf{y}|x) = \sum_{t=1}^{T} I(Y_{t,\tau} \leq y) w_t(x, h_\tau), \tag{5}$$

where $I(\cdot)$ is a indicator function, which equals 1 if its argument is true, and 0 otherwise. The CDF of the non-parametric CumRe estimator will provide more detail in intraday return patterns, e.g., opposite trends in the first and the last half-hours, larger return volatility at the opening phase than at the closing, leverage effects in both the first and the last half-hours.

To avoid capturing spurious predictability due to in-sample overfitting, we examine the out-of-sample predictive performance. For the conditional mean forecasts of return patterns, we introduce leave-one-year-out cross-validation and the Diebold-Mariano test to compare the forecast accuracy of non-parametric CumRe with that of the parametric CumRe. We propose the squared prediction error (SPE) of a return pattern $\mathbf{y}_t$ during the first or the last half-hour to measure the expected squared distance between the pattern predicted by a specific value of the overnight return and the true intraday return pattern during the trading session. We define the distance between the realized return pattern and the predicted return pattern on day t by

$$\text{SPE}_t = \sum_{\tau=1}^{6} \left(\frac{\epsilon_{t,\tau}}{\sqrt{\tau}} \right)^2, \tag{6}$$

where $\epsilon_{t,\tau} = y_{t,\tau} - \hat{y}_{t,\tau}$ is the predition error when $\hat{y}_{t,\tau}$ is predicted by the overnight return. Herein the prediction error $\epsilon_{t,\tau}$ is standardized by the square root of τ, since each particular increment of this cumulative return pattern has a prediction error variance that is proportional to the time over which the return is realized.

We assume that returns on the market during the first and the last half-hours in one year are independent with the other years. We use leave-one-year-out cross-validation (LOYOCV). Let N denote the number of calendar years, then the data set is partitioned into N disjoint sets for N calendar years, say $I = \{I_1, \ldots, I_N\}$. Define $I_{n^-} = \cup_{m \neq n} I_m$, so that for a particular evaluation n in the cross-validation, I_n is used as test set, and the remaining sets, I_{n^-}, are used for model fitting. The mean squared prediction error (MSPE) for year n across all τs is then

$$\mathrm{MSPE}_n(I_{n^-}) = \frac{1}{|I_n|} \sum_{\forall t \in I_n} \mathrm{SPE}_t(I_{n^-}), \tag{7}$$

where $|\cdot|$ denotes the cardinality of a set and thus $|I_n|$ is the number of trading days in year n, and I_{n^-} denotes conditioning on information from all the other years for parameter estimation.The sum of squared prediction errors (SSPE) for all N available calendar years is then just $\mathrm{SSPE} = \sum_{n=1}^{N} \mathrm{MSPE}_n(I_{n^-})$. The smaller the SSPE, the better the predictive performance is.

We will use the SPEs to formulate a type of Diebold and Mariano (1995) test to compare the forecast accuracy of non-parametric and parametric CumRe for return pattern forecasts. The null hypothesis is that the two forecast methods have the same forecast accuracy. The test statistic and corresponding test results are provided in Section 4.

3 Data and Evidence of Predictability

We use SPDR S&P 500 ETF prices, sampled at the 5-minute frequency. The sample period ranges from January 1, 2003 to December 31, 2013. Following studies using the TAQ dataset on higher-frequency data such as Boffelli *et al.* (2016), Maneesoonthorn *et al.* (2017), and Perera and Koul (2017), the raw data has been cleaned using methods similar

Table 1: Comparison of the dispersion and deviations for cumulative returns within the first and the last half-hours (Jan 2003 - Dec 2013). The entries are based on 2714 observations for each return series, which did not contain missing values. This table shows the ratios of the dispersion and deviation measures. The superscripts f and l denote the first and last half-hours, respectively.

	$\tau = 1$	$\tau = 2$	$\tau = 3$	$\tau = 4$	$\tau = 5$	$\tau = 6$	Mean	S.E.	Ratio> 1
$\text{ABS}^f/\text{ABS}^l$	1.09	1.04	1.16	0.97	1.03	0.99	1.05*	0.02	4
$\text{IQR}^f/\text{IQR}^l$	1.37	1.34	1.26	1.27	1.24	1.24	1.29***	0.02	6
$\text{Min}^f/\text{Min}^l$	0.93	0.73	0.69	0.64	0.75	0.75	0.75***	0.04	0
$\text{Max}^f/\text{Max}^l$	1.25	1.30	1.81	1.37	1.37	1.32	1.40***	0.08	6
SD^f/SD^l	1.09	1.07	1.11	1.07	1.07	0.91	1.05*	0.03	5
$\text{Skew}^f/\text{Skew}^l$	-6.20	-2.65	-1.93	-2.62	-2.36	-1.13	-2.81***	0.72	0
$\text{Kurt}^f/\text{Kurt}^l$	0.80	1.09	1.26	1.02	1.03	0.83	1.00	0.07	4

Notes: ABS denotes range (the difference between the maximum and minimum observation). IQR is short for interquartile range. SD is standard deviation. Min and Max are the minimum (always negative) and maximum (always positive) values of the returns in the sample, respectively. Skew and Kurt are skewness and kurtotis, and all the values of skewness are different from 0 and all the values of Kurtotis are larger than 3. The cumulative returns during the first half-hour are always distributed skew to the right (positive skewness) while the corresponding returns during the last half-hour are always skew to the left (negative skewness), therefore all ratios are always negative. S.E. is short for standard error. *, ** and *** denote significance at the 1%, 5% and 10% significance level, respectively.

to those of Brownlees and Gallo (2006).[8]

In line with Oldfield and Rogalski (1980), Wood *et al.* (1985), and Liu and Tse (2017), the percentage overnight return is calculated as

$$\text{ovr}_t = 100 \times (p_{t,0} - p_{t-1,78}), \tag{8}$$

where ovr_t is the overnight return on day t, $p_{t,0}$ is the log price at 9:30 a.m.[9] on day t, and $p_{t-1,78}$ is the log price at 16:00 p.m. on day $t-1$.

We conducted the Augmented Dickey-Fuller (ADF) test, and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test to examine the stationarity of the return series; all test results suggest that these return series are stationary.

According to studies such as Harris (1986), Amihud and Mendelson (1987) and Gerety

[8]The corresponding R package is TAQMNGR. We use a choice of $k = 60$, $\gamma = 0.02$ for the data cleaning algorithm, as mentioned by Brownlees and Gallo (2006).

[9]We also examined $\text{ovr}_t = 100 \times (p_{t,09:35} - p_{t-1,16:00})$ and $\text{ovr}_t = 100 \times (p_{t,09:45} - p_{t-1,16:00})$ as in Bogousslavsky (2018), and we obtained correlation values between overnight returns and the subsequent intraday returns similar to those found in the correlation matrix of the intraday returns, provided in Table A1 in the Appendix.

and Mulherin (1994), the return volatility is larger during the opening than during the closing period. The results in Table 1 confirm that cumulative returns within the first half trading hours tend to be more volatile than those within the last half trading hours. Based on measures of dispersion such as the range of the return distributions (the difference between the maximum and minimum returns), and the interquartile range (the difference between 75th and 25th percentiles of return distribution, also called the midspread or middle 50%), we find that, on average, the range of the return distribution during the first half-hour is 5% larger than that during the last half-hour. The midspread is even 29% larger.

Despite its greater dispersion, we observe that the cumulative returns during the first half-hour have larger maximum values but smaller minimum values, than those returns during the last half-hour. This implies that it is possible for traders to realize higher profits and smaller losses during the first half-hour. Meanwhile, the ratios of skewness show different degrees of asymmetry in the return distributions. Notice that the ratios of skewness are all negative. This is because all the cumulative returns during the first half-hour appear to be skewed to the right (positive skewness), while the corresponding returns during the last half-hour appear to be skewed to the left (negative skewness).

To illustrate the behavior of the cumulative 5-minute market returns series within the first and the last half-hours, we present graphical results of the means of the market return across the trading days in Figure 1. The left panel shows the changing trends in the mean returns within the first half-hour. It is evident that all means conditional on negative overnight returns (the red dots), averaged at -0.46%, are positive and present a rough upward trend within the first half-hour. The mean of cumulative return peaks at around 0.02% when $\tau = 4$. After that, it declines slightly when $\tau = 5$, but goes up again to slightly above 0.02% when $\tau = 6$. In contrast, all means conditional on positive overnight returns (green dots), averaged at 0.42%, lie below 0%. It shows a downward trend until 25 minutes after the market opening, reaching a low at around -0.03% when $\tau = 5$. It is interesting to note that the means of returns when $\tau = 5$ are inflection points for both the unconditional mean returns and the conditional mean returns. The unconditional mean

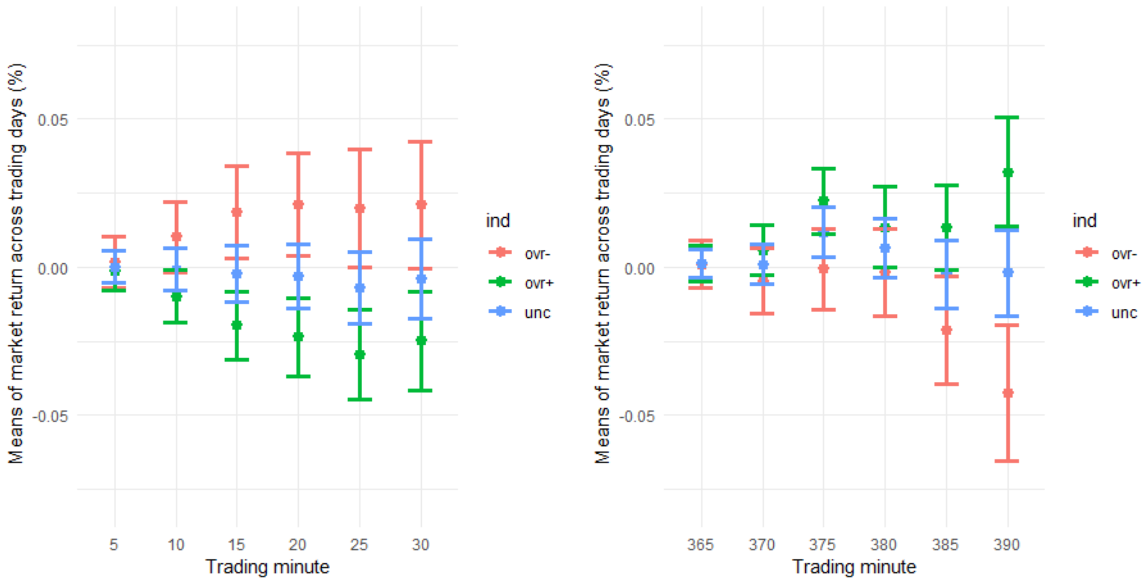

Figure 1: Cumulative market return (in %) within the first (left panel) and last (right panel) half trading hours conditional on positive returns, ovr+ (red dots), on negative returns, ovr− (green dots) and unconditional means, unc (blue dots). The error bars represent 95% confidence intervals.

(in blue), present a pattern similar to means conditional on positive overnight returns. The pattern shows a slight fall, ending at about -0.07% when $\tau = 5$. Then it increases, even though the unconditional mean is still below 0.

The right-hand side panel of Figure 1 reflects the changing trends in the means of market returns within the last half-hour of trading time. Notice that all the unconditional and conditional means of the market returns are positive when $\tau = 1$ during the last half trading hour. However, the mean conditional on positive overnight returns (represented by ovr+) and that conditional on negative overnight returns (ovr−) show a completely opposite trends afterwards; ovr+ experiences an increase until $\tau = 3$ (at slightly higher than 0.02%), and then a decrease at just above 0.01% when $\tau = 5$, but jumps up to over 0.03% when $\tau = 6$. In contrast, ovr− falls to a negative number when $\tau = 2$, then goes up just below 0 when $\tau = 3$. After that, it starts to decrease and reaches a number smaller than -0.04% when $\tau = 6$ during the last half trading hour. The unconditional mean illustrates a pattern similar to ovr+. It increases until $\tau = 3$, and then decreases until

$\tau = 5$ and slightly increases after, within the last half-hour.

Table 1 and Figure 1 provide evidence that the behavior of the cumulative 5-minute market mean returns within the first half-hour is different from that within the last half-hour. Besides, we find that the overnight returns are negatively (positively) correlated with the cumulative returns during the first (last) half-hour (see Table A). Therefore, we investigate the conditional pattern forecasts depending on overnight returns during the first half-hour and the last half-hour separately.

In light of Menkveld *et al.* (2007), a short-term market overreaction is consistent with liquidity suppliers who are compensated for their services through price reversals, while market under-reaction or positive serial correlation is consistent with strategic trading by informed investors who split their order across time to maximize profit. As observed by Amihud and Mendelson (1987), a short-term market overreaction to information is more likely during the opening phase. In that case, we may expect a negative trend pattern in market mean returns after the market opens conditional on a positive overnight return, because investors are more willing to earn profits by selling stocks after the market opens. For a negative overnight return, we may expect a positive trend pattern in market mean returns after the market opens. Amihud and Mendelson (1987) find evidence that a short-term market under-reaction to information is more likely during the closing phase of the market. Hence, we may expect a positive (negative) trend pattern in the market mean returns conditional on a positive (negative) overnight return in the closing phase.

When it comes to parametric CumRe, we find that the overnight returns are significant predictors for the cumulative returns during the first half-hour, according to Table 2.[10] In addition, overnight returns can also be used to predict the price movements from 15:30 to 15:55 and 15:30 to 16:00, i.e., the last half-hour before closing. Furthermore, we also find the negative relationship between the overnight returns and cumulative returns through the first 30 minutes, while the positive relationship between the overnight returns and returns

[10]We also examined the cumulative returns for the whole trading session i.e. $y_t^{0,1}, y_t^{0,2}, \ldots, y_t^{0,78}$. The overnight returns are significant predictors for the intraday returns from 9:30 to 14:55. However, a full discussion of longer trading periods is beyond the scope of this paper, since we focus on the first and and last half trading hours herein.

Table 2: Predictability during the first and the last half-hour. Panel A shows the results of regressing the cumulative returns during the first half-hour on the overnight returns, while Panel B shows the results of regressing the cumulative returns during the last half-hour on the overnight returns. The returns are in percentages and the coefficients are scaled up by 100. Newey and West (1987) robust t–statistics are in parentheses, and significance at the 1%, 5%, and 10% level is denoted by ***, **, and *, respectively.

	Dependent variable: $y^*_{t,\tau}$											
	Panel A: first half-hour						Panel B: last half-hour					
	$\tau = 1$	$\tau = 2$	$\tau = 3$	$\tau = 4$	$\tau = 5$	$\tau = 6$	$\tau = 1$	$\tau = 2$	$\tau = 3$	$\tau = 4$	$\tau = 5$	$\tau = 6$
Intercept	0.05	0.01	-0.09	-0.17	-0.51	-0.21	0.11	0.06	1.12^{***}	0.61	-0.34	-0.41
	(0.20)	(0.03)	(-0.21)	(-0.33)	(-0.87)	(-0.35)	(0.47)	(0.18)	(2.81)	(1.35)	(-0.71)	(-0.72)
ovr	-1.88^{**}	-4.14^{***}	-6.50^{***}	-7.21^{***}	-9.44^{***}	-9.43^{***}	0.01	1.72	3.29^{*}	2.55	5.37^{**}	10.95^{***}
	(-2.10)	(-2.90)	(-2.84)	(-3.01)	(-3.27)	(-2.67)	(0.01)	(1.05)	(1.84)	(1.35)	(2.25)	(2.71)
$R^2(\%)$	0.84	2.15	3.15	3.03	3.98	3.34	0.00	0.43	0.99	0.44	1.48	3.80
$R^2_{\text{ad}}(\%)$	0.80	2.11	3.12	2.99	3.95	3.31	-0.04	0.39	0.96	0.40	1.45	3.76

in the last half-hour. Figure 2 shows the predicted intraday return patterns depending on different values of overnight returns, namely 0%, ± 1%, ± 2.5%, during the first and the last half-hours. The opposite trends within the first and the last half-hours are consistent with existing literature such as Amihud and Mendelson (1987), Hong and Wang (2000), and Menkveld *et al.* (2007).

The parametric CumRe is able to reflect only linear dependence. The Ramsey RESET test (Ramsey, 1969) suggests that there are neglected non-linearities[11] in the setting of parametric CumRe, because quadratic or cubic terms of the fitted intraday return patterns are statistically significant in the conditional pattern forecasts (except for $y^l_{t,4}$, which is significant at the 10% significance level).

Next, we investigate non-parametric CumRe which may capture non-linearities. Figure 3 compares the forecasts provided by parametric CumRe and non-parametric CumRe for the first 20-mins returns and the last 30-mins returns, the periods of the highest profitability relative to risk, as we will show later. We observe substantial non-linearities in Figure 3. The predictions based on the non-parametric CumRe show a twisting shape,

[11] We also applied the Probability Integral Transform (PIT) on the overnight returns prior to non-parametric CumRe to account for skewness and excess kurtosis/fat tails in return distributions. We found that the areas of predicted intraday returns conditional on positive and negative values of the overnight returns are asymmetric, implying that non-linearities should not be neglected. The results are available upon request.

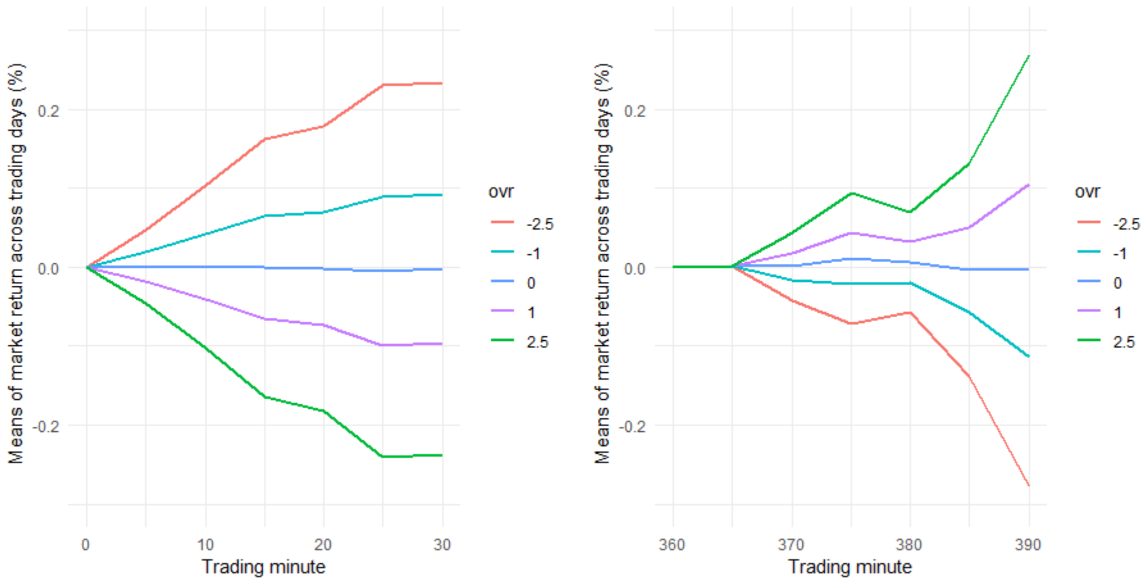

Figure 2: Parametric CumRe pattern forecasts depending on overnight returns within the first half trading hour (from 0 to 30-minute) and last half trading hour (from 360 to 390-minute). The symbol 'ovr' represents overnight returns. Returns are in percentages. The sample period is from January 1, 2003, through December 31, 2013.

varying with magnitude of the overnight returns. The linear prediction seems to be affected by large overnight returns and may not be accurate for relatively small overnight returns. There is also a substantial asymmetry; with the same magnitude of overnight returns, the degree of reaction to a positive overnight return is different from that to a negative overnight return. The returns for other time spans (τ-values) show similar qualitative asymmetries.

Figure 4 shows the predicted non-parametric return patterns during the first (top panel) and the last (bottom panel) half-hour trading periods conditional on negative and positive overnight returns of $\pm$ 2.5%.[12]. We show conditional means (left side) and conditional the distribution by plotting the conditional median as well as 80%, 90% and 95% predictive intervals (right side). As in the linear case, the mean returns predicted non-parametrically given a positive value of the overnight return (green line) shows a downward trend within

[12]In the Appendix we show patterns conditional on $\pm$ 1% overnight returns and the patterns of the standardized returns $\tilde{y}_{t,\tau} = \frac{y_{t,\tau}}{\hat{\sigma}_{t,\tau}}$ conditional on overnight returns of 0, $\pm$ 1%, and $\pm$ 2.5%.

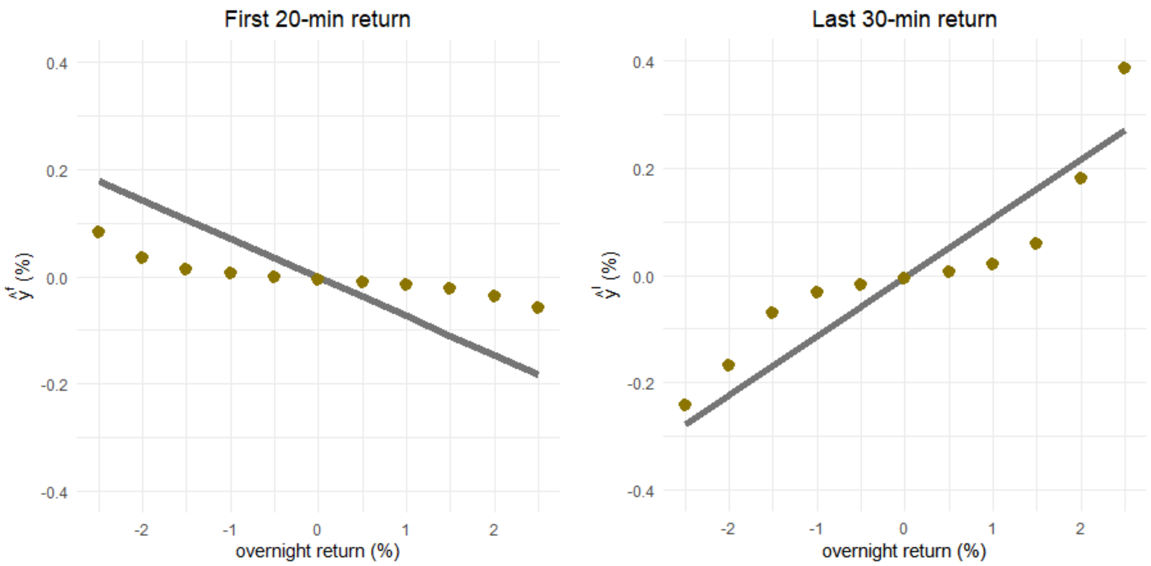

Figure 3: Comparison of parametric and non-parametric CumRe prediction. The left panel displays the predictions conditional on overnight returns during the first 20 trading minutes ($\tau = 4$). The right panel illustrates the predictions conditional on overnight returns for the last 30-min returns, that is, the returns from 360 to 390 trading minutes ($\tau = 6$). The overnight returns range from -2.5 to 2.5 on the figure which covers 99% of all observations. The solid line represents the predictions according to parametric CumRe. The dotted curve shows the predictions based on non-parametric CumRe. The returns are in percentages.

the first 30 trading minutes and an upward trend within the last 30 minutes; the pattern of mean returns conditional on a negative overnight return (red line) shows opposite trends within the same time intervals. Moreover, the forecast distribution during the first half-hour is skewed to the left (right) conditional on a positive (negative) overnight return; while the forecast distribution for the last half-hour is skewed to the right (left) for a positive (negative) overnight return.

4 Forecast Performance Evaluation

4.1 Leave-one-year-out cross-validation

Cross-validation (CV) is one of the most widely-used standard procedures for model evaluation in regression (Bergmeir *et al.*, 2018). Burman and Nolan (1992) propose bias cor-

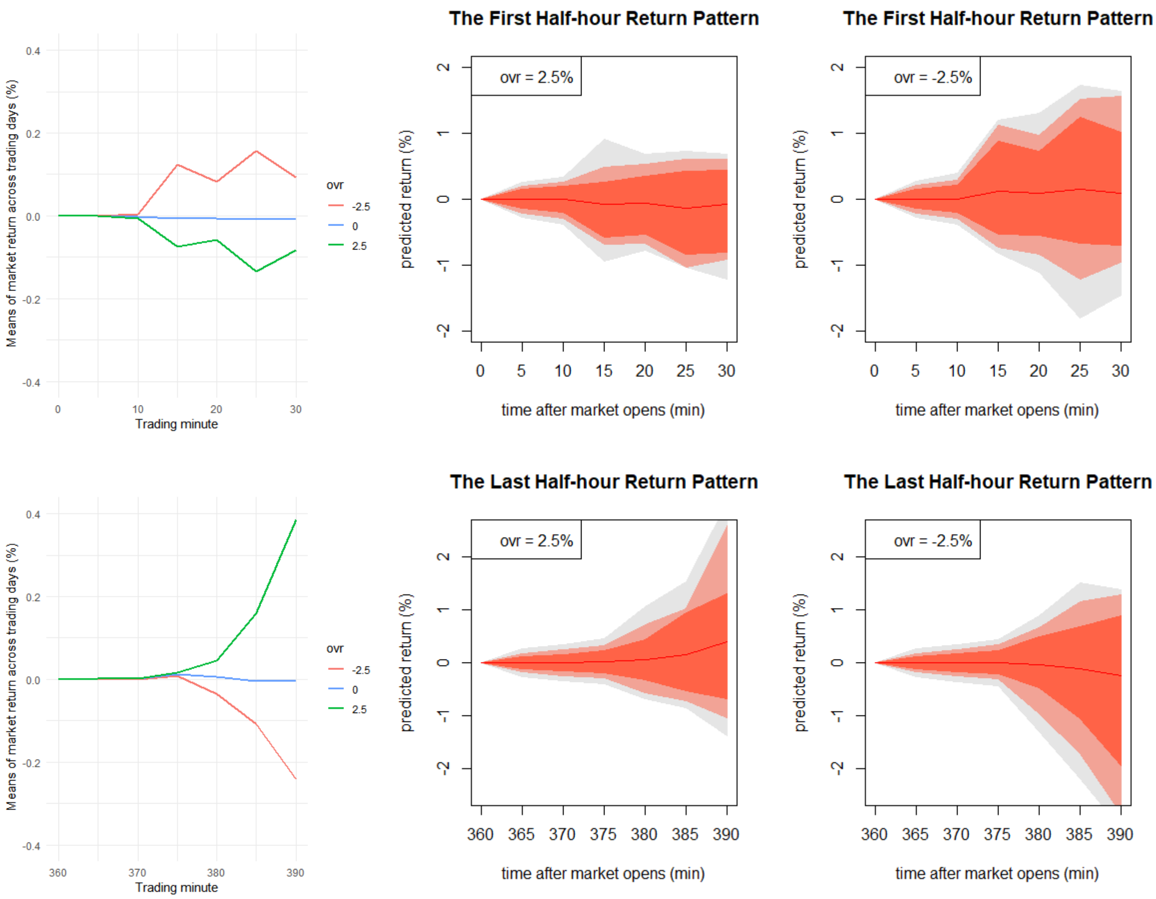

Figure 4: Predicted return pattern conditional on overnight return (-2.5%, 0%, and 2.5%). 80% prediction interval is shaded in red, 90% – light red and 90% – light grey. The predicted return is in percentage, and ovr is short for overnight return in the legend.

rection and Burman *et al.* (1994) propose η-block CV to account for serial dependence in time series data. However, these cross-validation procedures leave out the possibly dependent observations and use data insufficiently (Bergmeir *et al.*, 2018). To reflect on time-series dependence and efficiently use the available time series data set, we propose the leave-one-year-out cross-validation to evaluate the time series pattern prediction.

Table 3 shows MSPEs using the leave-one-year-out evaluation for parametric CumRe (column P) and the non-parametric CumRe (column NP). These values are benchmarked against the MSPE of a simple constant forecast of 0-returns (column $\emptyset$). While it is

Table 3: Mean squared prediction error using leave-one-year-out cross-validation, Eq. (7). '$\emptyset$' indicates MSPEs from a constant 0-returns forecast used as a benchmark, 'P' indicates MSPEs from the parametric approach, 'NP' indicates MSPEs from the non-parametric approach. All numbers are scaled up by 100.

	First half-hour			Last half-hour		
	$\emptyset$	P	NP	$\emptyset$	P	NP
2003	10.97	11.00	10.95	6.58	6.75	6.67
2004	3.84	3.92	3.89	3.75	3.77	3.74
2005	2.91	3.08	2.95	3.23	3.21	3.22
2006	3.67	3.75	3.72	2.82	2.82	2.81
2007	6.03	6.07	5.97	8.13	7.92	8.06
2008	46.37	45.13	46.38	54.10	53.47	53.60
2009	26.70	27.66	27.15	18.87	19.39	18.85
2010	10.61	10.92	10.90	7.30	7.40	7.25
2011	14.72	14.86	14.67	12.86	12.84	12.88
2012	5.82	6.05	5.81	3.37	3.43	3.36
2013	4.46	4.56	4.48	2.49	2.46	2.48
Sum	136.11	137.00	136.86	123.52	123.48	122.93

hard to outperform the benchmark for the first half-hour return pattern forecasts, the non-parametric CumRe overall shows smaller MSPE than the parametric CumRe in all of the years only except for year 2008. For the last half-hour return pattern forecasts, the conditional forecasts outperform the benchmark for most of the years, and overall the non-parametric CumRe shows better performance than the parametric CumRe again.

4.2 Comparison of predictive accuracy of patterns

Next we formally compare parametric and non-parametric CumRe-based forecasts using the Diebold and Mariano (1995) test statistic. Define the loss difference as the difference between the non-parameteric and parameteric squared prediction error on day t,

$$d_t = \mathrm{SPE}_{t,\mathrm{NP}} - \mathrm{SPE}_{t,\mathrm{P}},$$

with SPE_t as defined in Eq. (6). The Diebold-Mariano test statistic is defined as the average loss normalized by its standard error,

$$\text{t-stat}_{\text{DM}} = \frac{\bar{d}}{\sqrt{\hat{\sigma}_d^2/P}},$$

where the average is taken over the evaluation period of P days. A Newey and West (1987) HAC robust estimator of σ_d^2 is used to account for serial dependence in the sequence of loss differences $\{d_t\}$.[13] The null hypothesis is that the two methods have the same forecast accuracy against the alternative of unequal forecast accuracy.

We used the estimation period from January 1, 2003 till March 9, 2009 and evaluate performance for the period from March 10, 2009 to December 31, 2013, the period right after the financial crisis. We find that for the first half-hour return patterns, the null hypothesis is rejected at the 1% significance level ($\bar{d} = -0.0054$ and $\text{t-stat}_{\text{DM}} = -3.39$). For the last half-hour, the null is rejected at the 5% significance level ($\bar{d} = -0.0021$ and $\text{t-stat}_{\text{DM}} = -2.22$). This gives a strong indication that the non-parametric CumRe outperforms the parametric CumRe at the forecast accuracy during the evaluation period.

5 Economic Significance

Next, we develop market timing strategies and compute utility gains based on CumRe for mean-variance day traders. For easier comparison with existing results we annualize returns by multiplying mean daily returns for a given strategy by 252 trading days per year. We will also compute annualized Sharpe ratios, $\sqrt{252}(\bar{r}/\hat{\sigma})$, where $\bar{r}$ and $\hat{\sigma}$ are the sample mean and the sample standard deviation of the realized daily returns of a strategy, respectively. In addition to simple standard deviation, we also compute a HAC robust estimator. To measure utility gains of portfolios constructed based on different expected returns, we use the annualized certainty equivalent rate of return (CER) which we describe

[13] As in Diks *et al.* (2011), the HAC estimator of variance is calculated as $\hat{\sigma}^2 = \hat{\lambda}_0 + 2\sum_{z=1}^{Z-1} e_z \hat{\lambda}_z$, where $\hat{\lambda}_z$ denotes the lag-z sample covariance of the sequence of loss differences $\{d_t\}$ and e_z are the Bartlett weights $e_z = 1 - z/Z$ with $Z = \lfloor P^{1/4} \rfloor$, P is the number of days in the evaluation period.

in Section 5.2.

5.1 Market timing

Market timing is the practice of moving in and out of the market by attempting to predict the future direction of the market. To examine whether the overnight returns help predict the market direction in practice, we employ the overnight returns as market timing signals to trade in the market. Due to the negative relationship between the overnight returns and returns during first half-hour, on each trading day we will take a short position if the overnight return is positive at the beginning of the trading session; and then close it after $\tau \times 5$ min, $\tau \in \{1, \ldots, 6\}$; we take a long position otherwise. We examine different time intervals τ within the first half trading hour, for discovering when closing positions is most profitable for day traders.

Table 4 shows the results of these market timing strategies during the first and last half-hours. We compare the annualized returns based on the market time strategies I(ovr) with the returns based on simply holding long positions for the same periods.[14] When it comes to annualized returns, predicting the directions of the market during the first or the last half-hour based on the sign of overnight returns improves the performance of trading strategy. All of the annualized returns when we use the overnight return as the trading signal, on average, appear to be positive, while the simple long and short strategies yield smaller returns. The highest average return of 9.31% is obtained when we enter the market at 15:30, taking a position according to the sign of the overnight return on that day, and close the position at 16:00.

To take risk into account, we also present annualized Sharpe ratios of each timing strategy in Table 4. The higher the ratio, the better the performance. Conventionally, an annualized Sharpe ratio larger than 1 is considered acceptable to investors. During the first half-hour, holding a position for more than 10 minutes since 9:30 delivers a Sharpe ratio larger than 1. When it comes to trades during the last half-hour, the market timing

[14]Holding a short position would result in an opposite return sign.

Table 4: Market timing strategies during the first and the last half-hours and the passive buy-and-hold strategy on a daily basis. The timing strategy denoted by I(ovr) takes a short (long) position in the market when the observed overnight return is positive (negative) during the first half-hour on each trading day t, or a long (short) position in the market when the observed overnight return is positive (negative) during the last half-hour on each trading day t. Long indicates always taking long positions for different holding periods during the first half trading hour. For each timing strategy, the average return (Avg ret) and standard deviation (Std dev) are annualized by multiplying by 252 and $\sqrt{252}$, repectively, and are expressed in percentages. We also present annualized Sharpe ratio (SRatio), skewness (Skew.), and kurtosis (Kurt.) for each strategy. Sharpe ratios calculated based on Newey-West HAC estimator of variance are in parentheses.

	First half-hour					**Last half-hour**				
Timing	Avg ret	Std dev	SRatio	Skew.	Kurt.	Avg ret	Std dev	SRatio	Skew.	Kurt.
			$\tau = 1$					$\tau = 1$		
I(ovr)	0.37	2.25	0.16 (0.16)	0.74	12.14	0.06	2.07	0.03 (0.02)	0.27	15.21
Long	0.04	2.25	0.02 (0.02)	0.21	12.17	0.29	2.07	0.14 (0.14)	-0.03	15.22
			$\tau = 2$					$\tau = 2$		
I(ovr)	2.52	3.08	0.82 (0.80)	1.32	15.35	1.31	2.88	0.45 (0.40)	1.00	14.15
Long	-0.19	3.09	-0.06 (-0.06)	1.23	15.56	0.25	2.89	0.09 (0.09)	-0.47	14.25
			$\tau = 3$					$\tau = 3$		
I(ovr)	4.85	3.99	1.22 (1.22)	1.46	17.68	3.14	3.61	0.87 (0.67)	0.82	13.91
Long	-0.57	4.00	-0.14 (-0.14)	1.29	17.97	2.98	3.61	0.83 (0.89)	-0.67	14.22
			$\tau = 4$					$\tau = 4$		
I(ovr)	5.70	4.52	1.26 (1.28)	1.56	16.25	2.08	4.22	0.49 (0.41)	0.18	16.21
Long	-0.79	4.53	-0.18 (-0.18)	1.02	16.59	1.66	4.22	0.39 (0.43)	-0.39	16.26
			$\tau = 5$					$\tau = 5$		
I(ovr)	6.32	5.16	1.23 (1.19)	1.45	14.73	4.29	4.81	0.89 (0.71)	1.14	14.51
Long	-1.76	5.17	-0.34 (-0.36)	0.66	15.07	-0.59	4.82	-0.12 (-0.14)	-0.28	14.67
			$\tau = 6$					$\tau = 6$		
I(ovr)	5.84	5.63	1.04 (1.03)	1.88	18.14	9.31	6.11	1.52 (1.16)	2.16	21.84
Long	-1.01	5.64	-0.18 (-0.20)	0.51	18.50	-0.48	6.14	-0.08 (-0.10)	-0.45	22.25
	Buy-and-hold									
	Avg ret	Std dev	SRatio	Skew.	Kurt.					
	1.30	16.23	0.08 (0.09)	-0.26	8.84					

strategy conditional on the overnight return that holds the market for 30 minutes provides the highest Sharpe ratio of 1.52. All the timing strategies for the first and the last half-hours conditional on the overnight returns outperform the passive buy-and-hold strategy when we take risk into account. Moreover, the conditional timing strategies generate returns with positive skewness and much higher kurtosis, implying that these strategies often generate positive returns. We conclude that using the overnight return as the trading signal provides a good risk-return trade-off.

5.2 Utility Gains

Assume that a mean-variance day trader allocates funds between the market portfolio (SPY) with weight $\omega_{t,\tau}$ and a risk-free asset (3-month T-bill) with weight $(1 - \omega_{t,\tau})$. The portfolios are constructed on a daily basis conditional on overnight returns and held for $\tau \times 5$ minutes, either during the first or last half-hour. Therefore, the risk-free rate is assumed to be 0. The weight on the market portfolio, which incorporates the investor's optimal trade-off between the expected return and risk, can be obtained by solving the following expected utility maximization problem (see e.g., Brandt, 2010),

$$\omega_{t,\tau} = \arg\max_{\omega} \left(\mathbb{E}(\omega \cdot r_{t+1,\tau}) - \frac{\gamma}{2} \mathrm{Var}(\omega \cdot r_{t+1,\tau}) \right), \tag{9}$$

where γ measures the level of relative risk aversion of the day trader. The solution of this maximization problem is

$$\omega_{t,\tau} = \frac{\hat{r}_{t+1,\tau}}{\gamma \hat{\sigma}^2_{t+1,\tau}}, \tag{10}$$

where $\hat{r}_{t+1,\tau}$ and $\hat{\sigma}^2_{t+1,\tau}$ denote investor's estimate on the mean and variance of the returns of market portfolio for time interval τ, respectively. In Campbell and Thompson (2008), Ferreira and Santa-Clara (2011), and Gao *et al.* (2018), the expected returns of the market portfolios predicted by predictors are used to construct the optimal portfolios for mean-variance investors who allocate wealth between a risky asset (a stock market index or a ETF of that index as a proxy of a market portfolio) and a risk-free asset. Similarly, we predict expected returns on day $t + 1$ obtained from either parametric or non-parametric CumRe. The parameters of CumRe are estimated using an expanding window from January 1, 2003 to day t, where t ranges from March 9, 2009 to December 30, 2013. We also use the unconditional historical average returns, which is commonly used to compare the out-of-sample forecast performance of predictors (see e.g., Campbell and Thompson, 2008), as a benchmark calculated over the expanding window. The estimate of variance is obtained by calculating the sample variance using an expanding window as well.[15]

[15] We also employed the GARCH model to obtain the estimate of variance. The corresponding utility gains results showed us similar qualitative and quantitative portfolio performance. For easier comparison

There are different values of the risk aversion coefficient γ used in the literature: $\gamma = 2$ in Ferreira and Santa-Clara (2011), $\gamma = 3$ in Campbell and Thompson (2008), and $\gamma = 5$ in Gao *et al.* (2018). The higher the risk aversion coefficient, the lower the risk tolerance; and thus investors will invest less wealth on the market portfolio. In addition, we also consider $\gamma = 10$ to account for an even lower risk-tolerance case.

We impose the weight constraints for the market portfolios of different investment policy ranging from -4 to 4. This is because day traders in the US are allowed to use up to $4:1$ margin leverage according to the day-trading margin requirements.[16] In practice, we can short the SPY directly or buy the leveraged inverse ETFs (short ETFs or bear ETFs).[17]

We implement these portfolio policies according to the predetermined weights $\omega_{t,\tau}$ of Eq. (10) and thus the portfolio return on day $t+1$, $\rho_{t+1,\tau}$ is

$$\rho_{t+1,\tau} = \omega_{t,\tau} r_{t+1,\tau} \qquad \text{subject to} \qquad -4 \leq \omega_{t,\tau} \leq 4, \tag{11}$$

where $r_{t+1,\tau}$ is the realized return on the market portfolio during the time interval $\tau \times 5$ minutes on day $t+1$. We iterate this process over the out-of-sample period from March 10, 2009 to December 31, 2013 and get a time series of portfolio returns for each trading strategy. Since we employ the CER of each trading strategy to measure utility gains, we can compare the performance of the strategies that we use the unconditional historical mean (HM), and the parametric CumRe and the non-parametric CumRe both conditional on the overnight return to forecast the expected return. The CER of each portfolio is given

with existing results in the literature, we use sample variance, $\hat{\sigma}^2_{t+1,\tau} = \frac{\sum_{s=1}^{t}(r_{s,\tau}-\bar{r}_\tau)^2}{t-1}$, as investor's estimate on the variance of the market portfolio on day $t+1$.

[16] Day traders are allowed to have more leverage since their positions are short-term, and therefore each trade is likely to experience smaller price swings compared to positions held for days, weeks, or years. Financial Industry Regulatory Authority. Day-Trading Margin Requirements: Know the Rules. Retrieved from http://www.finra.org/investors/day-trading-margin-requirements-know-rules. Last access date 15-01-2019.

[17] For instances, Short S&P 500 (ticker symbol SH), UltraShort S&P 500 (ticker symbol SDS), and UltraPro Short S&P 500 (ticker symbol SPXU), or Daily S&P 500 Bear 1X ETF (ticker symbol SPDN), the Daily S&P 500 Bear 3X (ticker symbol SPXS) can deliver its $1\times$, $2\times$, or $3\times$ inverse exposure to the S&P 500 index. Likewise, we can go long by buying 2x and 3x Long ETFs or Bull ETFs to satisfy leverage ratios smaller than 4. We also examined the portfolio performance for different leverage ratios, which can be realized by trading these leveraged ETFs. We found that patterns in portfolio performance with lower leverage ratios are similar to the results in Table 5.

by

$$\text{CER} = \hat{\mu}_\rho - \frac{\gamma\hat{\sigma}_\rho^2}{2}, \tag{12}$$

where $\hat{\mu}_\rho$ and $\hat{\sigma}_\rho^2$ are the sample mean and sample variance of the realized portfolio returns over the out-of-sample period. The higher the CER, the larger the risk-adjusted return. Similar to Gao *et al.* (2018), we use the difference between the CERs of strategies using the overnight return as the predictor and the strategies using the historical mean forecast to measure the predictability significance. We annualize the CER by multiplying by 252 and multiply the result by 100. Thus, this measure (CER dif) can be interpreted as the percentage utility gain per year of the conditional mean forecast based on the overnight return instead of the unconditional historical mean forecast. A positive CER dif indicates that the conditional forecast based on the overnight return provides higher utility gains than an unconditional mean forecast for a mean-variance investor, ceteris paribus.

Table 5 reports the results of the out-of-sample portfolio performance for mean-variance investors with different levels of risk tolerance[18], who enter the market either during the first or the last half-hour to open a position and then close the position after $\tau \times 5$ minutes.

For the first half hour we find that portfolios constructed based on the expected returns predicted by parametric CumRe outperform those based on non-parametric CumRe or the unconditional historical mean forecasts for all holding periods. Among others, holding a position for 20 minutes ($\tau = 4$) from the beginning of each trading session delivers the highest positive annualized returns, which also yields a Sharpe ratio of over 1. In addition, the positive CERs in most cases imply that the conditional mean forecasts of expected returns depending on the overnight returns usually generate higher utility gains relative to the unconditional mean forecasts of expected returns for the mean-variance investors.

For the last half trading hour, while in the sense of constructing portfolios it is difficult for the conditional mean forecasts to consistently surpass the unconditional mean forecasts of expected returns for all the holding periods, there are two holding periods, 15-min ($\tau = 3$) and 30-min ($\tau = 6$), particularly notable. When $\tau = 3$, the portfolios depending

[18]We found results of portfolio performance corresponding to different levels of risk tolerance such as $\gamma = 3, 5$ similar to Table 5, which are available upon request.

Table 5: The results of portfolio performance for different levels of risk tolerance day traders. This table shows the economic value of recursively forecasting the market return based on the overnight return during the first and the last half-hours. We apply the predicted return as the expected return and then construct optimal portfolios for mean-variance day traders with risk aversion coefficients of 2 and 10. We impose portfolio constraints from -4 to 4 according to the day-trading margin requirements in the US. The average return (Avg ret), standard deviation (Std dev), and the difference of certainty equivalent rate of return (CER dif) between the optimal mean-variance strategies (P is short for parametric and NP for non-parametric CumRe) and the corresponding historical mean forecast strategy (HM) are annualized by multiplying by 252 and reported in percentages. We also present Sharpe ratio (SRatio), skewness (Skew.), and kurtosis (Kurt.) for each strategy. Sharpe ratios calculated based on Newey-West HAC estimators of variance are given in parentheses. The one-step ahead predicted returns are determined using the expanding window from March 10, 2009.

	$-4 \le w \le 4, \gamma = 2$											
	First half-hour						**Last half-hour**					
	Avg ret	Std dev	SRatio	Skew.	Kurt.	CER dif	Avg ret	Std dev	SRatio	Skew.	Kurt.	CER dif
	$\tau = 1$						$\tau = 1$					
P	6.14	8.35	0.74 (0.71)	-0.15	6.36	9.15	-3.60	5.26	-0.68 (-0.64)	-1.02	11.34	-3.22
NP	-2.03	3.80	-0.54 (-0.53)	-1.42	16.20	1.53	-0.44	4.69	-0.09 (-0.09)	-0.23	6.02	0.00
HM	-3.64	2.60	-1.40 (-1.50)	-4.06	65.18		-0.44	4.70	-0.09 (-0.09)	-0.23	6.02	
	$\tau = 2$						$\tau = 2$					
P	17.29	11.59	1.49 (1.51)	0.25	3.01	18.93	-1.26	8.61	-0.15 (-0.14)	-0.08	5.45	2.23
NP	2.14	9.66	0.22 (0.23)	0.24	3.75	4.18	-3.75	6.40	-0.59 (-0.58)	0.02	8.16	0.07
HM	-2.53	6.74	-0.38 (-0.42)	-0.05	4.62		-3.79	6.59	-0.58 (-0.57)	0.11	8.24	
	$\tau = 3$						$\tau = 3$					
P	**20.43**	14.75	**1.38 (1.44)**	0.23	3.39	**20.01**	8.13	11.63	0.70 (0.63)	-0.47	7.37	-12.08
NP	6.76	13.67	0.49 (0.52)	0.32	4.52	6.64	**19.24**	11.52	**1.67 (1.60)**	0.21	7.18	-0.94
HM	-1.58	4.12	-0.39 (-0.42)	-0.35	3.62		**20.28**	11.94	**1.70 (1.66)**	0.15	6.61	
	$\tau = 4$						$\tau = 4$					
P	**22.54**	15.95	**1.41 (1.50)**	0.12	2.23	**20.72**	-1.56	13.50	-0.12 (-0.11)	-0.85	10.85	-10.02
NP	7.63	14.54	0.53 (0.58)	0.17	3.27	6.24	3.41	12.63	0.27 (0.25)	-0.24	7.88	-4.82
HM	-0.02	8.36	0.00 (0.00)	-0.11	2.59		8.10	12.11	0.67 (0.69)	0.59	10.89	
	$\tau = 5$						$\tau = 5$					
P	18.47	18.21	1.01 (1.08)	0.49	4.28	14.93	3.38	15.49	0.22 (0.20)	-0.31	8.82	6.83
NP	5.88	17.07	0.34 (0.37)	0.53	5.55	2.73	1.01	12.89	0.08 (0.07)	-0.37	17.94	5.20
HM	2.79	16.00	0.17 (0.19)	0.00	2.95		-5.38	6.81	-0.79 (-0.91)	-0.68	18.19	
	$\tau = 6$						$\tau = 6$					
P	12.13	19.94	0.61 (0.66)	0.70	4.35	12.05	**20.33**	19.25	**1.06 (0.99)**	0.36	13.83	**22.16**
NP	1.72	17.20	0.10 (0.11)	1.03	8.26	2.67	**16.09**	17.20	**0.94 (0.86)**	0.63	22.02	**18.66**
HM	-2.68	11.04	-0.24 (-0.26)	0.00	3.86		-5.16	6.14	-0.84 (-1.02)	-1.48	28.43	
	$-4 \le w \le 4, \gamma = 10$											
	First half-hour						**Last half-hour**					
	Avg ret	Std dev	SRatio	Skew.	Kurt.	CER dif	Avg ret	Std dev	SRatio	Skew.	Kurt.	CER dif
	$\tau = 1$						$\tau = 1$					
P	4.74	7.46	0.64 (0.60)	-0.21	10.75	2.71	-2.01	1.98	-1.02 (-0.89)	-4.74	63.77	-2.01
NP	-0.41	0.76	-0.53 (-0.53)	-1.41	16.16	0.32	-0.15	1.01	-0.14 (-0.15)	-0.32	6.51	0.00
HM	-0.74	0.53	-1.39 (-1.50)	-4.75	80.00		-0.15	1.01	-0.15 (-0.15)	-0.32	6.51	
	$\tau = 2$						$\tau = 2$					
P	14.27	10.31	1.38 (1.37)	0.35	5.47	9.59	-2.57	7.03	-0.37 (-0.35)	-0.24	11.36	-3.99
NP	0.52	2.45	0.21 (0.22)	-0.11	7.73	0.86	-0.95	1.51	-0.63 (-0.64)	-0.58	14.21	-0.02
HM	-0.55	1.38	-0.40 (-0.45)	-0.06	5.29		-0.92	1.60	-0.57 (-0.58)	-0.09	11.28	
	$\tau = 3$						$\tau = 3$					
P	**16.77**	13.02	**1.29 (1.33)**	0.32	6.14	**8.64**	5.30	9.82	0.54 (0.47)	-0.70	13.97	-6.21
NP	2.01	6.56	0.31 (0.31)	-0.85	40.73	0.21	7.71	4.46	**1.73 (1.65)**	0.47	5.80	**0.02**
HM	-0.32	0.82	-0.38 (-0.42)	-0.35	3.62		7.80	4.71	**1.65 (1.58)**	0.59	7.31	
	$\tau = 4$						$\tau = 4$					
P	**17.05**	13.91	**1.23 (1.27)**	0.23	4.66	**7.53**	-2.12	10.21	-0.21 (-0.18)	-1.19	28.62	-8.53
NP	2.31	6.92	0.33 (0.34)	-1.73	36.83	0.06	1.18	5.42	0.22 (0.19)	1.51	75.09	-1.49
HM	0.00	1.67	0.00 (0.00)	-0.11	2.59		1.58	2.75	0.58 (0.59)	0.53	9.28	
	$\tau = 5$						$\tau = 5$					
P	12.70	16.17	0.79 (0.83)	0.69	7.80	-0.39	1.30	13.10	0.10 (0.09)	-0.37	16.76	-6.11
NP	2.04	9.23	0.22 (0.23)	0.13	27.85	-2.24	1.65	8.14	0.20 (0.18)	-0.11	41.83	-0.49
HM	0.54	3.21	0.17 (0.18)	0.00	2.90		-1.08	1.36	-0.79 (-0.91)	-0.68	18.19	
	$\tau = 6$						$\tau = 6$					
P	7.54	17.40	0.43 (0.47)	0.99	8.25	-6.81	**15.45**	16.99	**0.91 (0.84)**	0.56	23.35	**2.12**
NP	-1.65	7.44	-0.22 (-0.24)	-0.35	37.49	-3.64	**12.02**	12.15	**0.99 (0.87)**	2.22	63.76	**5.74**
HM	-0.54	2.21	-0.24 (-0.26)	0.00	3.86		-1.03	1.23	-0.84 (-1.02)	-1.48	28.43	

on the non-parametric CumRe perform similar to those depending on the unconditional mean forecasts; and these portfolios always yield the highest (HAC) Sharpe ratios (larger than 1) among portfolios across all holding periods, which is even higher than the most profitable case that the weights on the market portfolios depend on the expected returns predicted by parametric CumRe with a holding period of 30-minute.

Notice that given that a day trader with a risk aversion coefficient of 2 invests on the market portfolio for the last half hour, the annualized CER dif of the portfolio depending on the overnight return, 22.16% per year, is 15.55% higher than the CER dif depending on the predictors of the first and the twelfth half-hour returns in Gao *et al.* (2018) (6.61% in Table A.7 in the Appendix). This is because we impose the portfolio constraint that the weight on the market portfolio lies inside $[-4, 4]$, while the constraint in Gao *et al.* (2018) is $[-0.5, 1.5]$. If the constraint is the same, CER dif of the portfolio based on the overnight return is 7.57% and this portfolio yields a higher Sharpe ratio of 1.19. The unconstrained portfolio based on the overnight return generates an annualized return of 244.57% with a Sharpe ratio of 1.06 for the last half-hour.

We observe that the distributions of returns of portfolios depending on parametric and non-parametric CumRe have asymmetries and excess kurtosis. The Jarque-Bera test results suggest that these return series are not normally distributed with highly statistical significance. Autocorrelation and heteroskedasticity are present in these return series. While the measure of CER dif provides some qualitative evidence, it is not based on proper statistical inference procedure. In order to introduce proper inference, we use a related measure, which we refer to as Realized Utility (RU),

$$\mathrm{RU}_{t,\tau} = \omega_{t,\tau} \cdot r_{t+1,\tau} - \frac{\gamma}{2}\omega_{t,\tau}^2 \hat{\sigma}_{t+1,\tau}^2, \tag{13}$$

where $\hat{\sigma}_{t+1,\tau}^2$ is variance forecast for which we use the GARCH (1,1) model. For comparison with the CER dif results in Table 5, we use the same weights, $\omega_{t,\tau}$.

To compare the realized utilities depending on non-parametric ($\mathrm{RU_N}$) and parametric CumRe ($\mathrm{RU_P}$) for an investor who has a risk aversion coefficient of γ, we define the Diebold-

Mariano test statistic as $\text{t-stat}_{\text{DM}} = \frac{\bar{d}_\gamma}{\sqrt{\hat{\sigma}^2_{d_\gamma}/P}}$, where d_γ is the difference between RU_N and RU_P, $\hat{\sigma}^2_{d_\gamma}$ is a HAC robust estimator of variance to account for serial correlations and heteroskedasticity in portfolio return series over the evaluation period. A positive $\bar{d}_\gamma$ indicates that the portfolio constructed based on non-parametric CumRe on average yields higher utility gains relative to that based on the parametric CumRe to investors with the risk aversion coefficient γ. Similar to CER dif, we measure the realized utility gains in percentage and annualize d_γ by multiplying by 252.

Table 6: Constrained and unconstrained realized utility results according to Eq. (13). For easier comparison, we measure realized utility gains in percentage and annualize the difference between the realized utilities depending on non-parametric and parametric CumRe by multiplying by 252. The Newey-West (1987) robust t-statistic are given in parentheses.

				$-4 \leq w \leq 4$								
	First half-hour						**Last half-hour**					
	$\tau=1$	$\tau=2$	$\tau=3$	$\tau=4$	$\tau=5$	$\tau=6$	$\tau=1$	$\tau=2$	$\tau=3$	$\tau=4$	$\tau=5$	$\tau=6$
$\bar{d}_2$	-7.32^{**}	-14.65^{***}	-13.42^{**}	-14.42^{**}	-12.21^{*}	-9.05	3.20^{**}	-1.74	10.99	4.89	-0.57	-1.90
	(-1.99)	(-2.79)	(-2.14)	(-2.41)	(-1.87)	(-1.45)	(2.00)	(-0.36)	(1.60)	(1.08)	(-0.16)	(-0.59)
$\bar{d}_{10}$	-1.44	-7.88^{*}	-7.07	-5.66	0.08	4.08	1.99^{**}	5.20	7.21	7.48^{**}	7.67^{**}	9.14^{**}
	(-0.41)	(-1.81)	(-1.61)	(-1.24)	(0.02)	(0.72)	(2.22)	(1.55)	(1.51)	(2.11)	(2.15)	(2.34)
				Unconstrained Weights								
	First half-hour						**Last half-hour**					
	$\tau=1$	$\tau=2$	$\tau=3$	$\tau=4$	$\tau=5$	$\tau=6$	$\tau=1$	$\tau=2$	$\tau=3$	$\tau=4$	$\tau=5$	$\tau=6$
$\bar{d}_2$	40.23	103.29	176.16^{**}	161.49^{**}	260.48^{***}	276.76^{***}	11.31^{*}	81.15^{**}	57.57	33.21	80.78^{*}	113.15^{**}
	(0.79)	(1.29)	(2.29)	(2.23)	(2.99)	(3.17)	(2.48)	(2.14)	(1.07)	(1.33)	(1.95)	(2.09)
$\bar{d}_{10}$	8.05	20.66	35.23^{**}	32.30^{**}	52.10^{***}	55.35^{***}	2.26^{*}	16.23^{**}	11.51	6.64	16.16^{*}	22.63^{**}
	(0.79)	(1.29)	(2.29)	(2.23)	(2.99)	(3.17)	(2.48)	(2.14)	(1.07)	(1.33)	(1.95)	(2.09)

Table 6 reports the realized utility difference between the portfolios constructed depending on the non-parametric CumRe and the parametric CumRe forecasts. We find that, with the day-trading margin requirements in the US, trading based on parametric CumRe predictions during the first half hour can generate significantly higher realized utilities for mean-variance investors with lower level of risk aversion ($\gamma = 2$); however, non-parametric CumRe forecasts provide significantly higher realized utilities for higher level of risk aversion ($\gamma = 10$) investors during the last half hour. Recall the twisting shape of non-parametric CumRe predictions and the linear shape of parametric CumRe predictions in Figure 3, employing parametric CumRe predictions may lead to the magnitude of weights determined according to these predictions being higher than those determined according to non-parametric CumRe predictions, since 99% of the observations of overnight returns range between -2.5% and 2.5% over the evaluation period. In the absence of day-trading margin requirements for day traders, the trading strategies depending on non-parametric CumRe forecasts yield significantly higher realized utility gains relative to strategies depending on parametric CumRe forecasts.

6 Conclusion

Our results show that it is possible to forecast the subsequent intraday patterns in market returns based on observed market overnight returns (close-to-open returns) with both statistical and economic significance. Our forecasting approach CumRe confirms the intraday patterns in returns documented in the existing theoretical and empirical literature, such as the opposite trends during the beginning period and the closing period within the same trading day (Hong and Wang, 2000); a short-term market overreaction to information is possible in the opening, while the under-reaction effect is likely at the close (see e.g., Amihud and Mendelson, 1987); the positive or negative relationship depending on the overnight returns during the first or the last half trading hours (Liu and Tse, 2017; Gao *et al.*, 2018). Moreover, the overnight returns help predict the market directions within the first and the last half trading hours. In addition, provided that the dependence between the overnight returns and certain subsequent intraday returns is statistically significant, the magnitudes and the signs of the overnight returns can also help forecast the expected returns and thus yield higher annualized returns, Sharpe ratios, and certainty equivalent returns than the unconditional mean pattern forecasts of the expected returns—one-step-ahead (cumulative) forecasts based on a recursive approach—for mean-variance day traders.

We find non-linearities in return patterns. The non-parametric CumRe can accommodate nonlinear dependence between the overnight returns and the subsequent intraday returns, while the parametric CumRe only accounts for linear relationship. Based on transaction-level data of SPDR S&P 500 ETF from January 1, 2003 to December 31, 2013, which is an exchange-traded fund that closely tracks the S&P 500 index with the smallest bid-ask spread, we find that, statistically, non-parametric CumRe improves the out-of-sample forecast performance in terms of the Diebold and Mariano (1995) type test statistic and the leave-one-year-out cross-validation errors.

There are a number of possible extensions to predict intraday returns based on overnight returns. First, it might be helpful to incorporate ARMA and/or GARCH structure into the forecast methods or to take the asymmetry and fat-tails in returns into account when forecasting the intraday returns. Second, one might wonder whether overnight returns could be used as a predictor for other financial markets such as futures, currencies, and commodities. Third, we wonder whether this intraday market predictability can contribute to volatility forecasting and thus influences the performance of the portfolios constructed by both conditional mean and variance forecasts depending on the overnight returns. Fourth, as stated by Brandt (2010), the mean-variance problem ignores any preferences toward higher-order return moments, and thus we would be interested in the portfolio performance if the weights on the market portfolio are determined by a higher-order approximation of expected utility maximization. We leave these for future work.

Appendix A Correlation Matrix

This table reports the correlation coefficients between the overnight returns and the subsequent intraday returns during the first and the last half hours. The variable ovr is short for overnight returns. The variables $y^f_{t,\tau}$ and $y^l_{t,\tau}$ denote the τ^{th} element, where $\tau = 1, \ldots, 6$ of the first and last half hour return pattern, respectively. The sample period is from January 1, 2003, to December 31, 2013.

Table A.1: Correlation Matrix.

	ovr	$y^f_{t,1}$	$y^f_{t,2}$	$y^f_{t,3}$	$y^f_{t,4}$	$y^f_{t,5}$	$y^f_{t,6}$	$y^l_{t,1}$	$y^l_{t,2}$	$y^l_{t,3}$	$y^l_{t,4}$	$y^l_{t,5}$	$y^l_{t,6}$
$y^f_{t,1}$		1.00	0.65	0.47	0.46	0.45	0.45						
$y^f_{t,2}$			1.00	0.76	0.69	0.65	0.63						
$y^f_{t,3}$				1.00	0.87	0.81	0.74						
$y^f_{t,4}$					1.00	0.91	0.84						
$y^f_{t,5}$						1.00	0.93						
$y^f_{t,6}$							1.00						
$y^l_{t,1}$								1.00	0.72	0.54	0.40	0.38	0.36
$y^l_{t,2}$									1.00	0.75	0.62	0.59	0.54
$y^l_{t,3}$										1.00	0.87	0.79	0.70
$y^l_{t,4}$											1.00	0.88	0.79
$y^l_{t,5}$												1.00	0.91
$y^l_{t,6}$													1.00
ovr	1.00	-0.09	-0.15	-0.18	-0.17	-0.20	-0.18	0.00	0.07	0.10	0.07	0.12	0.19

Appendix B Intraday Return Patterns

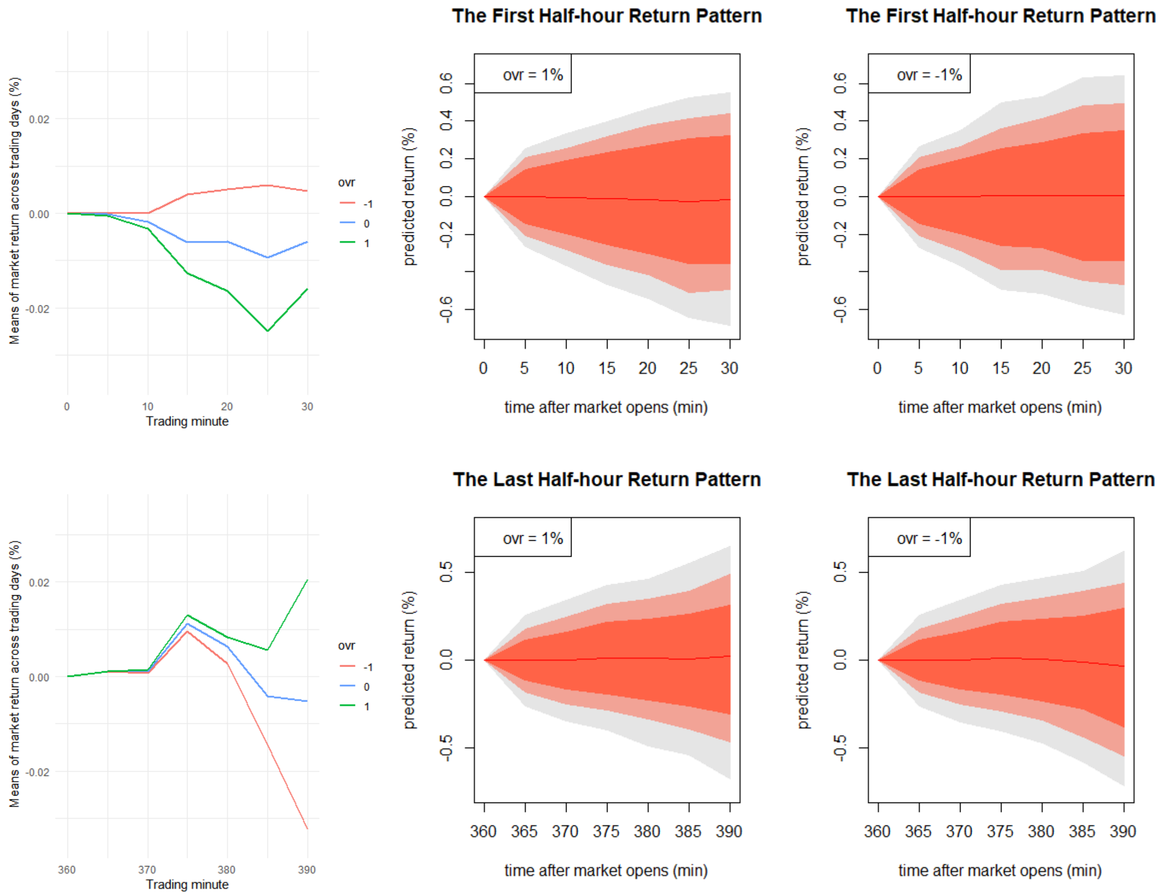

Figure B.1: Non-parametric CumRe: Predicted return pattern conditional on overnight returns ($-1\%, 0\%, +1\%$). The outer grey shade is 80% prediction interval, the inner shades are respectively 90% and 95% prediction intervals. The predicted return is in percentage, and ovr is short for overnight return in the legend. Note that we take a relatively conservative overnight return 1% as an example. These patterns are more striking based on a larger magnitude of overnight return.

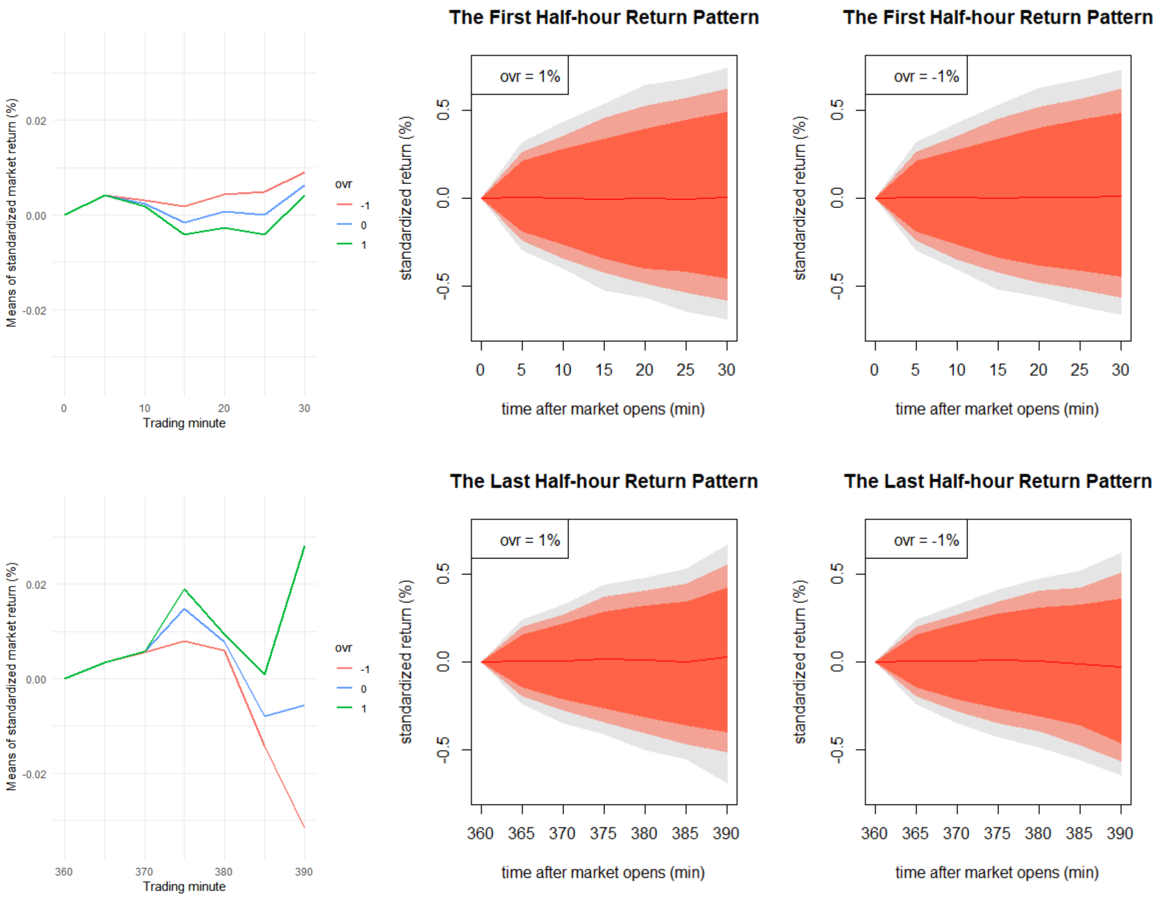

Figure B.2: Non-parametric CumRe: Predicted return pattern conditional on overnight returns $(-1\%, 0\%, +1\%)$. 80% prediction interval is shaded in red, 90% – light red and 90% – light grey. The intraday returns series are standardized by $\tilde{Y}_{t,\tau} = \frac{Y_{t,\tau}}{\hat{\sigma}_t}$, where σ_t is the conditional standard deviation of $Y_{t,\tau}$, σ_t is estimated by realized volatility (the square root of the realized variance). The predicted returns are in percentages, and ovr is short for overnight return in the legend.

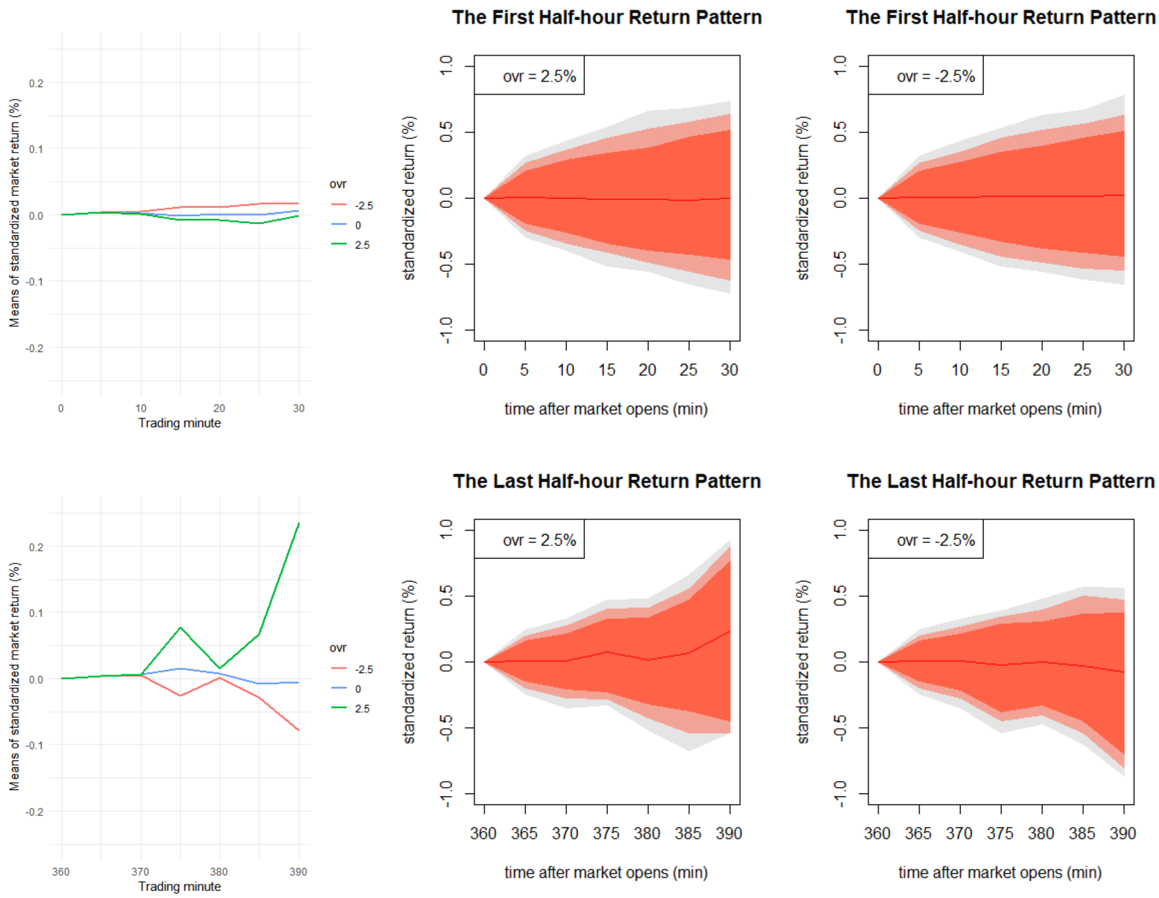

Figure B.3: Non-parametric CumRe: Predicted return pattern conditional on overnight returns $(-2.5\%, 0\%, +2.5\%)$. 80% prediction interval is shaded in red, 90% – light red and 90% – light grey. The intraday returns series are standardized by $\tilde{Y}_{t,\tau} = \frac{Y_{t,\tau}}{\hat{\sigma}_t}$, where σ_t is the conditional standard deviation of $Y_{t,\tau}$, σ_t is estimated by realized volatility (the square root of the realized variance). The predicted returns are in percentages, and ovr is short for overnight return in the legend.

Acknowledgments

The authors would like to thank Peter Boswijk, Pavel Cizek, Mike Cooper, Jan de Gooijer, Yi He, Bart Keijsers, Siem Jan Koopman, André Lucas, Albert Jan Menkveld, Anne Opschoor, Jeffrey S. Racine, Howell Tong, Kees Jan van Garderen, Qiwei Yao, and participants at 24$^{\text{th}}$ International Conference on Computing in Economics and Finance and the 2019 annual conference of the Netherlands Econometrics Study Group for providing useful comments and suggestions.

References

Abhyankar, A., Copeland, L. S. and Wong, W. (1997). Uncovering nonlinear structure in real-time stock-market indexes: The S&P 500, the DAX, the Nikkei 225, and the FTSE-100. *Journal of Business & Economic Statistics*, **15,** number 1, 1–14.

Amihud, Y. and Mendelson, H. (1987). Trading mechanisms and stock returns: An empirical investigation. *The Journal of Finance*, **42,** number 3, 533–553.

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

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Vega, C. (2007). Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics*, **73,** number 2, 251–277.

Ang, A. and Bekaert, G. (2007). Stock return predictability: Is it there? *The Review of Financial Studies*, **20,** number 3, 651–707.

Bergmeir, C., Hyndman, R. J. and Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics & Data Analysis*, **120,** 70–83.

Berkman, H., Koch, P. D., Tuttle, L. and Zhang, Y. J. (2012). Paying attention: overnight returns and the hidden cost of buying at the open. *Journal of Financial and Quantitative Analysis*, **47,** number 4, 715–741.

Boffelli, S., Skintzi, V. D. and Urga, G. (2016). High-and low-frequency correlations in European government bond spreads and their macroeconomic drivers. *Journal of Financial Econometrics*, **15,** number 1, 62–105.

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

Bogousslavsky, V. (2018). The cross-section of intraday and overnight returns. Working paper.

Brandt, M. W. (2010). Portfolio choice problems, chapter 5. In *Handbook of Financial Econometrics: Tools and Techniques*, *San Diego* (eds Yacine Aït-Sahalia and Lars Peter Hansen), Handbooks in Finance, volume 1, pp. 269 – 336. North-Holland.

Brownlees, C. T. and Gallo, G. M. (2006). Financial econometric analysis at ultra-high frequency: Data handling concerns. *Computational Statistics & Data Analysis*, **51,** number 4, 2232–2245.

Burman, P., Chow, E. and Nolan, D. (1994). A cross-validatory method for dependent data. *Biometrika*, **81,** number 2, 351–358.

Burman, P. and Nolan, D. (1992). Data-dependent estimation of prediction functions. *Journal of Time Series Analysis*, **13,** number 3, 189–207.

Campbell, J. Y. and Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, **21,** number 4, 1509–1531.

Cliff, M. T., Cooper, M. J. and Gulen, H. (2008). Return differences between trading and non-trading hours: Like night and day. Working paper.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, **13,** number 1, 253–263.

Diks, C., Panchenko, V. and Dijk, D. Van (2011). Likelihood-based scoring rules for comparing density forecasts in tails. *Journal of Econometrics*, **163,** number 2, 215–230.

Ferreira, M. A. and Santa-Clara, P. (2011). Forecasting stock market returns: The sum of the parts is more than the whole. *Journal of Financial Economics*, **100,** number 3, 514–537.

Gao, L., Han, Y., Li, S. Z. and Zhou, G. (2018). Market intraday momentum. *Journal of Financial Economics*, **129,** number 2, 394 – 414.

Gerety, M. S. and Mulherin, J. H. (1994). Price formation on stock exchanges: The evolution of trading within the day. *Review of Financial Studies*, **7,** number 3, 609–629.

Goyal, A. and Welch, I. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, **21,** 1455–1508.

Guidolin, M., Hyde, S., McMillan, D. and Ono, S. (2009). Non-linear predictability in stock and bond returns: When and where is it exploitable? *International Journal of Forecasting*, **25,** number 2, 373–399.

Härdle, W. K., Müller, M., Sperlich, S. and Werwatz, A. (2004). *Nonparametric and Semiparametric Models*. Springer.

Harris, L. (1986). A transaction data study of weekly and intradaily patterns in stock returns. *Journal of Financial Economics*, **16,** number 1, 99–117.

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

Hong, H. and Wang, J. (2000). Trading and returns under periodic market closures. *The Journal of Finance*, **55,** number 1, 297–354.

Jain, P. C. and Joh, G. (1988). The dependence between hourly prices and trading volume. *Journal of Financial and Quantitative Analysis*, **23,** number 3, 269–283.

Liu, Q. and Tse, Y. (2017). Overnight returns of stock indexes: Evidence from ETFs and futures. *International Review of Economics & Finance*, **48,** 440–451.

Maasoumi, E. and Racine, J. (2002). Entropy and predictability of stock market returns. *Journal of Econometrics*, **107,** number 1, 291–312.

Maneesoonthorn, W., Forbes, C. S. and Martin, G. M. (2017). Inference on self-exciting jumps in prices and volatility using high-frequency measures. *Journal of Applied Econometrics*, **32,** number 3, 504–532.

Menkveld, A. J., Koopman, S. J. and Lucas, A. (2007). Modeling around-the-clock price discovery for cross-listed stocks using state space methods. *Journal of Business & Economic Statistics*, **25,** number 2, 213–225.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, **55,** number 3, 703–708.

Oldfield, G. S. and Rogalski, R. J. (1980). A theory of common stock returns over trading and non-trading periods. *The Journal of Finance*, **35,** number 3, 729–751.

Opsomer, Jean, Wang, Yuedong and Yang, Yuhong (2001). Nonparametric regression with correlated errors. *Statistical Science*, **16,** 134–153.

Perera, I. and Koul, H. L. (2017). Fitting a two phase threshold multiplicative error model. *Journal of Econometrics*, **197,** number 2, 348–367.

Ramsey, J.B. (1969). Tests for specification errors in classical linear least squares regression analysis. *Journal of the Royal Statistical Society, Series* B, **31,** 350–371.

Rapach, D. and Zhou, G. (2013). Forecasting stock returns. In *Handbook of Economic Forecasting*, volume 2, pp. 328–383. Elsevier.

Rogalski, R. J. (1984). New findings regarding day-of-the-week returns over trading and non-trading periods: A note. *Journal of Finance*, **39,** number 5, 1603–14.

Smirlock, M. and Starks, L. (1986). Day-of-the-week and intraday effects in stock returns. *Journal of Financial Economics*, **17,** number 1, 197–210.

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics*, **54,** number 3, 375–421.

Tsay, R. S. (2010). *Analysis of Financial Time Series.* John Wiley & Sons.

Wood, R. A., McInish, T. H. and Ord, J. K. (1985). An investigation of transactions data for NYSE stocks. *The Journal of Finance*, **40,** number 3, 723–739.

Yang, J., Cabrera, J. and Wang, T. (2010). Nonlinearity, data-snooping, and stock index ETF return predictability. *European Journal of Operational Research*, **200,** number 2, 498–507.

Yao, Q. and Tong, H. (1998). A bootstrap detection for operational determinism. *Physica D: Nonlinear Phenomena*, **115,** number 1/2, 49–58.