

EXTREME RETURNS AND INTENSITY OF TRADING ¹

WEI LIN

*International School of Economics and Management
Capital University of Economics and Business, China*

GLORIA GONZÁLEZ-RIVERA

*Department of Economics
University of California, Riverside, USA*

March 21, 2016

¹We are grateful to the participants at the International Symposium in Forecasting, Riverside, 2015, Tsinghua International Conference on Econometrics, Beijing, 2015, and the 1st International Symposium on Interval Data Modeling, Beijing, 2015; and to the seminar participants at Guanghua School of Management at Peking University, Universidad Carlos III de Madrid (UC3M), CEMFI, Universidad Autonoma de Madrid, Universidad de Alcalá de Henares, Universidade do Minho, and Universidad Complutense de Madrid for useful comments. Gloria González-Rivera wishes to thank the Department of Statistics at UC3M, where this manuscript was written, for their hospitality and for the financial support of the 2015 Chair of Excellence UC3M/Banco de Santander. Wei Lin appreciates the financial support from the Research and Innovation Center of Metropolis Economic and Social Development, CUEB.

Abstract

Asymmetric information models of market microstructure claim that variables like trading intensity are proxies for latent information on the value of financial assets. We consider the interval-valued time series (ITS) of low/high returns and explore the relationship between these extreme returns and the intensity of trading. We assume that the returns (or prices) are generated by a latent process with some unknown conditional density. At each period of time, from this density, we have some random draws (trades) and the lowest and highest returns are the realized extreme observations of the latent process over the sample of draws. In this context, we propose a semiparametric model of extreme returns that exploits the results provided by extreme value theory. If properly centered and standardized extremes have well defined limiting distributions, the conditional mean of extreme returns is a nonlinear function of conditional moments of the latent process and of the conditional intensity of the process that governs the number of draws. We implement a two-step estimation procedure. First, we estimate parametrically the regressors that will enter into the nonlinear function, and in a second step, given the generated regressors, we estimate nonparametrically the conditional mean of extreme returns. Unlike current models for ITS, the proposed semiparametric model is robust to misspecification of the conditional density of the latent process. We fit several nonlinear and linear models to the 5-min low/high returns to three major bank stocks, Wells Fargo, Bank of America, and J.P. Morgan, and find that, either in-sample or out-of-sample, the nonlinear specification is superior to the current linear models and that the conditional standard deviation of the latent process and the conditional intensity of the trading process are major drivers of the dynamics of extreme returns. We find that there is an asymmetric relationship between extreme returns and trading intensity. While the lowest returns are sensitive to any large or small volume of trading (the largest the volume, the lower the lowest returns), the highest returns are responding mostly to extremely large trading volumes.

Key Words: Trading intensity, Interval-valued Time Series, Generalized Extreme Value Distribution, Nonparametric regression, Generated Regressor.

JEL Classification: C01, C14, C32, C51, G12.

1 Introduction

We explore the modeling and forecasting of interval-valued time series (ITS) of extreme returns, which are defined as the interval formed by the highest and the lowest returns in a given period of time. In contrast to the modeling of a classical time series of returns, in which it is very difficult to find any time-dependence, the intervals formed by extreme returns have statistical properties that can be exploited. For instance, in González-Rivera and Lin (2013), the authors estimate a constrained bivariate linear system for the daily lowest/highest returns of the SP500 index and find that there is statistically significant dependence with adjusted R-squared (in-sample) of about 50%. Though this work generalizes specifications of previous regression models on lower/upper bounds or center/radius of intervals (see the references herein), it relies on the assumption of bivariate normality. In a subsequent analysis, unlike the regression-type models just mentioned, Lin and González-Rivera (2015) propose an alternative modeling approach by pondering how interval-valued data is generated. They consider the lower and upper bounds of the interval as the realizations of minimal and maximal order statistics coming from a sample of N_t random draws from the conditional density of a latent random process $\{Y_t\}$. Through the statistical implementation of this approach to prices of agricultural commodities, they also find that their models provide a very good fit of extreme returns to livestock commodities with average coverage rates (percentage overlap of the actual low/high interval with the fitted interval) of 83%. However, there are also some disadvantages of this approach. First, the joint probability density function of minimal and maximal order statistics degenerates as the number of random draws goes to infinity. Second, the normality assumption on the latent random process $\{Y_t\}$ may be too restrictive.

To overcome these drawbacks and, in particular, the restrictions imposed by the distributional assumptions, we propose a new two-step semiparametric model that exploits the extreme property of the lower and upper bounds of the interval. We maintain the general setup of Lin and González-Rivera (2015) by assuming that there is a latent process $\{Y_t\}$ with conditional density $f_{Y_t}(\cdot)$, from which, at every moment of time, there are N_t random draws and the lower and upper bounds of the interval are the *realized extreme* observations of Y_t over the sample of draws. However, we will not assume any particular functional form of f_{Y_t} , so that the estimation procedure is robust to density

misspecification of the underlying stochastic process. We will only need conditional moments of the latent process and we will rely on limiting results provided by *extreme value theory* to estimate the conditional mean of the lower and upper bounds of the interval. The proposed estimation procedure consists of two steps. First, we obtain parametric estimates of the conditional mean and conditional variance of the latent process $\{Y_t\}$ and estimates of the conditional trading intensity of the process $\{N_t\}$. Second, with the generated conditional moments of the first step as the regressors, we specify a nonparametric model for the conditional means of the lower and upper bounds. We propose a nonparametric function because, according to extreme value theory, the conditional mean of an extreme value is often a nonlinear function that is difficult to estimate parametrically.

The proposed semiparametric model, in the context of the financial econometrics literature, is a natural vehicle to analyze the role of trading intensity on the generation of extreme returns. Asymmetric information models of market microstructure claim that variables like trading volume (or trading intensity) are proxies for latent information on the value of financial assets (see Easley and O'Hara, 1992). We have numerous empirical studies that link volume with volatility but, in contrast, the link between trading volume and extreme returns has not been analyzed in much detail. A sample of most representative results on volume and volatility follows in historical order. Lamoureux and Lastrapes (1990) find that identical latent factors drive trade volume and return volatility. Anderson (1996) propose a model in which informational asymmetries and liquidity needs motivate trading, which in turn, drives the dynamics of a stochastic volatility model. Engle (2000) analyzes an Autoregressive Conditional Duration model and a GARCH model to conclude that the absence of trading means either bad news or no news and translates into low volatility regimes. With high frequency data (5-min intraday data), Darrat, Rahman, and Zhong (2003) find evidence of significant lead-lag relations between volume and volatility in agreement with the sequential information arrival hypothesis. Fleming and Kirby (2011) analyze the joint dynamics of trading volume and realized volatility and find that there is a strong correlation between the innovations to volume and volatility. Sita and Westerholm (2011) find that trade durations (inversely related to trade volume) have forecasting power for returns but only within the trading day. One can argue that the range of the interval of extreme returns is a very good volatility estimator (Parkinson, 1980) and in this sense, the result of the aforementioned studies may apply. However, the dynamics

of the low/high interval are richer than those of the range itself because the modeling of the interval captures not only variability but also the dynamics of the bounds themselves. For instance, Ning and Wirjanto (2009) find that for East-Asian stocks, there is a significant and asymmetric return-volume dependence at the extremes. The largest returns tend to be associated with extremely large trading volumes but the lowest returns tend not to be related to either large or small volumes. We fit the proposed semiparametric model to the 5-min low/high returns to three major bank stocks, Wells Fargo, Bank of America, and J.P. Morgan, and find that, either in-sample or out-of-sample, extreme returns are driven by the conditional mean and conditional standard deviation of the latent process and the conditional intensity of the trading process, and that a nonlinear model is superior to the current linear specifications. We also find that the extreme returns respond asymmetrically to trading volume. The lowest returns are sensitive to any volume, large or small, while the highest returns are responsive to extremely large volume. Overall, the largest volume is associated with the most extreme returns.

The organization of this paper is the following. In section 2, we provide the basic assumptions for estimation of the model. In section 3, we present the two-step estimation procedure and establish the asymptotic properties of the second-step nonparametric regression with generated regressors. In section 4, we analyze several models to explain the relationship between extreme returns and the intensity of trading, and finally, we conclude in section 5.

2 Basic Assumptions

We describe the data generating process of the interval-valued time series. We need several assumptions, which are not too restrictive, and they accommodate many of the processes frequently encountered in financial data.

Assumption 1 (Data Generating Process). *Let $\{Y_t : t = 1, \dots, T\}$ be an underlying stationary stochastic process. The continuous random variable Y_t at time t has conditional density $f(y_t|\mathcal{F}_{t-1})$, where \mathcal{F}_{t-1} is the information set available at time t . At each time t , there are N_t independent draws from $f(y_t|\mathcal{F}_{t-1})$ collected in a set $\mathcal{S}_t \equiv \{y_{it} : i = 1, \dots, N_t\}$ with random sample size N_t , which it is assumed to follow a conditional distribution $H(n_t|\mathcal{F}_{t-1})$. Conditioning on \mathcal{F}_{t-1} , N_t and*

Y_t are independent for all t .

Let y_{lt} and y_{ut} denote the smallest and largest values in the sample \mathcal{S}_t at time t :

$$y_{lt} \equiv \min_i \mathcal{S}_t = \min_{1 \leq i \leq N_t} \{y_{it}\},$$

$$y_{ut} \equiv \max_i \mathcal{S}_t = \max_{1 \leq i \leq N_t} \{y_{it}\}.$$

Then, $\{(y_{lt}, y_{ut}) : t = 1, \dots, T\}$ is the observed interval time series (ITS) of lower and upper bounds.

The intuition behind Assumption 1 is straightforward. For instance, suppose that we have financial data and we choose a frequency, say, every five minutes. During these five minutes, trading takes place and, for every transaction, we observe a return (price). Then, in each block of five minutes, we will observe the lowest return, the highest return, and the number of trades. Our assumption means that the conditional density of returns $f(y_t | \mathcal{F}_{t-1})$ is updated every five minutes according to some dynamic specification. The number of trades during the five-minute time interval represents the number of random draws n_t from the conditional distribution of returns. Then, the lowest and the highest returns (y_{lt} and y_{ut}) are the two extremal (maximal and minimal) observations in the sample \mathcal{S}_t of size n_t .

Given this data generating mechanism, our analysis of ITS data proceeds with the analysis of extremal observations $\{(y_{lt}, y_{ut})\}$ based on the results of the extreme value theory. The asymptotic theory for maxima (and minima) is very different from the theory applied to averages. Central limit theorems provide a normal limiting distribution once the average is properly centered and standardized by its mean and standard deviation, respectively. In contrast, the centering and standardizing terms in the limit theorems for maxima (minima) are more difficult to derive because they depend on the tail characteristics of the assumed underlying density. The key result in extreme value theory is the Fisher-Tippett theorem that provides the limiting distributions of properly centered and standardized maxima (minima)¹. The three limiting distributions are *Fréchet*, *Weibull*, and *Gumbel*, which can be nested into a one-parameter generalized extreme value distribution (GEV)

¹We only consider continuous random variables, therefore the existence of a non-degenerate limiting distribution should always hold.

H_ξ defined as

$$H_\xi(x) = \begin{cases} \exp\{-(1 + \xi x)^{-1/\xi}\} & \text{if } \xi \neq 0, \\ \exp\{-\exp\{-x\}\} & \text{if } \xi = 0, \end{cases}$$

where $1 + \xi x > 0$ and ξ is a shape parameter. Then, (i) $\xi = \alpha^{-1} > 0$ corresponds to the Fréchet distribution, (ii) $\xi = 0$ corresponds to the Gumbel distribution, and (iii) $\xi = \alpha^{-1} < 0$ corresponds to the Weibull distribution.

It is said that the random variable Y_t belongs to the maximum domain of attraction (MDA) of the extreme value distribution H_ξ ($Y_t \in \text{MDA}(H_\xi)$) if the limiting distribution of standardized extremes, i.e. $c_u(t)^{-1}(Y_{ut} - d_u(t))$, is the extreme value distribution H_ξ . The standardizing and centering terms, $c_u(t)$ and $d_u(t)$, depend on t through the conditional distribution of Y_t and the number of random draws N_t . Explicitly, we write $c_u(t) \equiv c_u(N_t, f(y_t|\mathcal{F}_{t-1}))$ and $d_u(t) \equiv d_u(N_t, f(y_t|\mathcal{F}_{t-1}))$. The same argument holds for the minima process $\{Y_{lt}\}$. Formally, we have the following lemma.

Lemma 1. *Given Assumption 1, for all t , there exist*

$$\begin{aligned} c_l(t) &\equiv c_l(N_t, f(y_t|\mathcal{F}_{t-1})), \quad c_u(t) \equiv c_u(N_t, f(y_t|\mathcal{F}_{t-1})) \\ d_l(t) &\equiv d_l(N_t, f(y_t|\mathcal{F}_{t-1})), \quad d_u(t) \equiv d_u(N_t, f(y_t|\mathcal{F}_{t-1})) \end{aligned}$$

such that, conditioning on \mathcal{F}_{t-1} and for $N_t \xrightarrow{p} \infty$, the standardized maxima and minima have limiting GEV distributions H_{ξ_l} and H_{ξ_u} respectively, i.e.,

$$\begin{aligned} c_l(t)^{-1}[Y_{lt} - d_l(t)]|\mathcal{F}_{t-1} &\xrightarrow{d} H_{\xi_l}, \quad \forall t = 1, \dots, T \\ c_u(t)^{-1}[Y_{ut} - d_u(t)]|\mathcal{F}_{t-1} &\xrightarrow{d} H_{\xi_u}, \quad \forall t = 1, \dots, T. \end{aligned}$$

Since we would like to build conditional mean models for the extremes, the above convergence in distribution is too weak. We need to impose restrictions on the first moments of Y_{lt} and Y_{ut} to achieve stronger convergence. Thus, the following assumption and lemma:

Assumption 2. *For all t , there exists $\delta > 0$, such that*

$$\begin{aligned} \sup_{n_t} E \left(|c_l(t)^{-1}[Y_{lt} - d_l(t)]|^{1+\delta} | \mathcal{F}_{t-1} \right) &= M_l < \infty, \\ \sup_{n_t} E \left(|c_u(t)^{-1}[Y_{ut} - d_u(t)]|^{1+\delta} | \mathcal{F}_{t-1} \right) &= M_u < \infty. \end{aligned}$$

Lemma 2. Let Y_{ξ_l} and Y_{ξ_u} be the random variables with GEV distribution H_{ξ_l} and H_{ξ_u} , respectively.

If Assumptions 1 – 2 hold, then for all t , and for any $r \in (0, 1]$,

$$\begin{aligned}\lim_{n_t \rightarrow \infty} E(|c_l(t)^{-1}[Y_{lt} - d_l(t)]|^r | \mathcal{F}_{t-1}) &= E(|Y_{\xi_l}|^r), \\ \lim_{n_t \rightarrow \infty} E(|c_u(t)^{-1}[Y_{lt} - d_u(t)]|^r | \mathcal{F}_{t-1}) &= E(|Y_{\xi_u}|^r).\end{aligned}$$

Since the conditional expectation of the GEV random variable Y_{ξ_u} is $E(Y_{\xi_u}) = [\Gamma(1 - \xi_u) - 1]/\xi_u$, where $\Gamma(\cdot)$ is the Gamma function, the conditional expectations of the extrema are

$$E(Y_{ut}|N_t; \mathcal{F}_{t-1}) = d_u(N_t, f(y_t|\mathcal{F}_{t-1})) + c_u(N_t, f(y_t|\mathcal{F}_{t-1})) \frac{\Gamma(1 - \xi_u) - 1}{\xi_u} + o(c_u(N_t, \theta_t))$$

$$E(Y_{lt}|N_t; \mathcal{F}_{t-1}) = d_l(N_t, f(y_t|\mathcal{F}_{t-1})) + c_l(N_t, f(y_t|\mathcal{F}_{t-1})) \frac{\Gamma(1 - \xi_l) - 1}{\xi_l} + o(c_l(N_t, \theta_t)).$$

The conditional mean functions of the upper and lower bounds depend on the centering and standardizing terms associated with the assumed conditional density $f(y_t|\mathcal{F}_{t-1})$. Even for some common densities like normal or Student's t , these terms are highly nonlinear on the moments of interest.² Therefore, we propose to estimate the conditional mean functions nonparametrically so that they are robust to density misspecification of the underlying stochastic processes. In doing so, we also avoid the difficult task of calculating the associated standardizing and centering terms. We write:

$$E(Y_{ut}|N_t, \mathcal{F}_{t-1}) = m_u(N_t, f(y_t|\mathcal{F}_{t-1}), \xi_u), \quad (2.1)$$

$$E(Y_{lt}|N_t, \mathcal{F}_{t-1}) = m_l(N_t, f(y_t|\mathcal{F}_{t-1}), \xi_l). \quad (2.2)$$

²If y_t is normally distributed as $N(\mu_t, \sigma_t^2)$, we have

$$c_u(n_t, \mu_t, \sigma_t) = \frac{1}{\sigma_t \sqrt{2 \ln n_t}}; \quad d_u(n_t, \mu_t, \sigma_t) = \mu_t + \sigma_t \sqrt{2 \ln n_t} - \sigma_t \frac{\ln(4\pi) + \ln \ln n_t}{2(2 \ln n_t)^{1/2}}.$$

If y_t has t -distribution with mean μ_t and degrees of freedom ν_t , we have $d_u(n_t, \mu_t, \sigma_t) = 0$ and $c_u(n_t, \mu_t, \nu_t)$ is the solution to the following reduced form model,

$$\frac{1}{n} = \frac{1}{2} - (c - \mu_t) \Gamma\left(\frac{\nu_t + 1}{2}\right) \cdot \frac{{}_2F_1\left(\frac{1}{2}; \frac{\nu_t + 1}{2}; \frac{3}{2}; -\frac{(c - \mu_t)^2}{\nu_t}\right)}{\sqrt{\pi \nu_t} \Gamma\left(\frac{\nu_t}{2}\right)}$$

where ${}_2F_1$ is the hypergeometric function.

where $m_l(\cdot)$ and $m_u(\cdot)$ are the conditional mean functions depending on the conditional density of the underlying process $f(y_t|\mathcal{F}_{t-1})$, the number of random draws N_t , and the shape parameters of the limiting GEV distribution ξ_l and ξ_u . Note that ξ_l and ξ_u are constant, and therefore can be innocuously excluded from the functions.

We also assume that the conditional density $f(y_t|\mathcal{F}_{t-1})$ is indexed by a finite-dimensional parameter vector. We will include the first two moments of the underlying random process Y_t in a parameter vector θ_t , i.e. $\theta_t = (\mu_t, \sigma_t)$ to capture the location and the scale of the conditional distribution of Y_t at time t . Similarly, for the number of random draws N_t , we assume that the conditional distribution $H(n_t|\mathcal{F}_{t-1})$ is indexed by the first moment of N_t . Formally,

Assumption 3. (i) For any time period t , the conditional density $f(y_t|\mathcal{F}_{t-1})$ is indexed by the first two order conditional moments $\theta_t \equiv \theta(\mathcal{F}_{t-1}) \in \Theta \subset \mathbb{R}^2$, where Θ is a compact subset of Euclidean space, i.e., $f(y_t|\mathcal{F}_{t-1}) = f(y_t; \theta_t)$ for all t .

(ii) For any time period t , the conditional distribution $H(n_t|\mathcal{F}_{t-1})$ is indexed by the first order conditional moment $\lambda_t \equiv \lambda(\mathcal{F}_{t-1}) \in \Theta \subset \mathbb{R}$, where Θ is a compact subset of Euclidean space, i.e., $H(n_t|\mathcal{F}_{t-1}) = H(n_t; \lambda_t)$ for all t .

(iii) Let Ψ_1 and Ψ_2 be compact subsets on some finite k -dimensional Euclidean space \mathbb{R}^k . The expectational models $\mathcal{M}_1(\Psi_1)$ and $\mathcal{M}_2(\Psi_2)$ are correctly specified for $\theta_t \equiv (\mu_t, \sigma_t^2)$ and λ_t , respectively, i.e.,

$$\begin{aligned}\mu_t &\equiv E(Y_t|\mathcal{F}_{t-1}) = \mu(\mathcal{F}_{t-1}, \psi_1^o) \\ \sigma_t^2 &\equiv E[(Y_t - \mu_t)^2|\mathcal{F}_{t-1}] = \sigma^2(\mathcal{F}_{t-1}, \psi_1^o) \\ \lambda_t &\equiv E(n_t|\mathcal{F}_{t-1}) = \lambda(\mathcal{F}_{t-1}, \psi_2^o)\end{aligned}$$

almost surely for each time t with some $\psi_1^o \in \Psi_1$ and $\psi_2^o \in \Psi_2$. In addition, the point-valued time series $\{Y_t\}_{t=1}^T$ (e.g. returns based on closing prices), and $\{n_t\}_{t=1}^T$, used to estimate the parameters in the specifications \mathcal{M}_1 and \mathcal{M}_2 , satisfy regularity conditions such that the estimates $\hat{\psi}_1$ and $\hat{\psi}_2$ are \sqrt{T} -consistent.

Given assumption 3(i), the conditional expectations of maxima and minima in (2.1) and (2.2)

can be further simplified as

$$E(Y_{lt}|N_t, \mathcal{F}_{t-1}) = m_l(N_t, \theta_t), \quad (2.3)$$

$$E(Y_{ut}|N_t, \mathcal{F}_{t-1}) = m_u(N_t, \theta_t). \quad (2.4)$$

With assumption 3(ii), we can calculate the marginal expectations of the extremes as

$$E(Y_{lt}|\mathcal{F}_{t-1}) = M_l(\theta_t, \lambda_t) \equiv \int m_l(s, \theta_t) dH(s; \lambda_t), \quad (2.5)$$

$$E(Y_{ut}|\mathcal{F}_{t-1}) = M_u(\theta_t, \lambda_t) \equiv \int m_u(s, \theta_t) dH(s; \lambda_t). \quad (2.6)$$

Assumption 3(iii) is a high level assumption. In the framework of QMLE, it requires that the quasi log-likelihood function obeys the strong uniform law of large numbers (SULLN). Primitive conditions are available in the literature, see Domowitz and White (1982), among others.

3 Estimation

We propose to estimate (2.5) and (2.6) in two steps. First, we will generate the regressors θ_t and λ_t , and secondly we will estimate non-parametrically the conditional mean functions.

If the parameter λ_t and θ_t were known, we could directly use nonparametric methods to estimate the following two conditional mean models:

$$Y_{lt} = M_l(\theta_t, \lambda_t) + \varepsilon_{lt}, \quad (3.1)$$

$$Y_{ut} = M_u(\theta_t, \lambda_t) + \varepsilon_{ut}. \quad (3.2)$$

However, in most situations the regressors λ_t and θ_t are unknown. We will estimate them by proposing some parametric models that, according to assumption 3(iii), must be well specified. Consequently, our objective is the estimation of nonparametric conditional mean functions of generated regressors:

$$Y_{lt} = M_l(\widehat{\theta}_t, \widehat{\lambda}_t) + v_{lt}, \quad (3.3)$$

$$Y_{ut} = M_u(\widehat{\theta}_t, \widehat{\lambda}_t) + v_{ut}. \quad (3.4)$$

In order to estimate $\theta_t \equiv (\mu_t, \sigma_t^2)$, we work with a point-valued time series. If we are modeling returns, we can follow the standard practice of choosing the series of returns calculated at the end of each time period. Alternatively, we could also choose the series of the centers of the intervals as a realized sample path of the underlying process $\{Y_t\}$ and specify the dynamics of the conditional mean. The specification of the dynamics of the variance could be based on the time series of ranges of the intervals. Similarly, we work with the realized sample path $\{n_t\}$ and specify the dynamics of the conditional intensity $\lambda_t = E(N_t|\mathcal{F}_{t-1})$ to obtain their estimates.

It is possible to avoid these generated regressors by directly inserting into the nonparametric functions those observed regressors, in the information set \mathcal{F}_{t-1} , that drive the conditional moments μ_t, σ_t^2 and λ_t . The drawback of this approach is that the number of regressors could be very large so we face the curse of dimensionality of nonparametric models. The generated regressor approach offers a more parsimonious model, though we need to take into account the extra uncertainty generated by the estimation of the regressors.

There are two important differences with the approach in Lin and González-Rivera (2015). There, the estimation methodology is maximum likelihood where the log-likelihood function is built based on the joint density of the lowest and the highest rank order statistics of the random sample $\mathcal{S}_t \equiv \{y_{it} : i = 1, \dots, N_t\}$. The assumption of conditional normality for the underlying process $\{Y_t\}$ may seem to provide a QML estimator but as we discussed there, even with normality, the joint density of the ordinal statistics does not belong to the quadratic exponential family and the consistency of the QML estimator cannot be guaranteed. The approach that we propose here is robust to distributional assumptions:

- With the realized sample paths of point-valued time series, i.e., $\{y_t\}$ and $\{n_t\}$, associated with the underlying stochastic processes $\{Y_t\}$ and $\{N_t\}$, estimate consistently the conditional moments (θ_t, λ_t) .
- With the maxima and minima of the interval-valued time series $(\{y_t\}, \{y_{ut}\})$ and the parametrically generated covariates $(\hat{\theta}_t, \hat{\lambda}_t)$, estimate nonparametrically the two conditional mean functions (3.3) and (3.4).

The second difference is related to the feasibility of the order statistics approach when the

number of draws N_t is very large. A quick look at the log-likelihood function reveals that, for large number of trades, the function will explode, and the optimization exercise will not have a solution. Hence, the extreme value approach proposed here is general enough to provide feasibility and robustness.

Under Assumption 3(iii) and within a QMLE framework, the first step estimators $(\widehat{\theta}_t, \widehat{\lambda}_t)$ enjoy standard asymptotic properties. Now, we focus on the asymptotics of the second- step nonparametric estimator. Mammen, Rothe and Schienle (2012) provide a general theory on the statistical properties of two-stage nonparametric regression estimators. The first-step nonparametric estimates are the regressors in the second-step estimation of a nonparametric model. Our model is simpler than theirs in two aspects. First, the latent regressors in our model are generated parametrically, while their first-step estimators have parametric/nonparametric or just nonparametric components. Second, their latent regressors depend on the model itself, as in Conrad and Mammen (2009) who estimate a semiparametric GARCH-in-mean model where they need an iterative estimation procedure. Therefore, we only need a mild adaptation of their theorem to show that the oracle property of kernel-based nonparametric estimators also apply to our two step estimator.

Let $h_t^o \equiv (\theta_t^o, \lambda_t^o)$ be the true moments and $\widehat{h}_t \equiv (\widehat{\theta}_t, \widehat{\lambda}_t)$ be the estimates of those true moments.

Theorem 1 (Asymptotic properties of the two-step local linear estimator).

Let $j = l$ and u , $\sigma_j^2(x) = \mathbb{E}(\varepsilon_{jt}^2 | h_t = x)$ be the conditional variance of ε_{jt} , $f(x)$ be the pdf of h_t . Assume that $m(x)$, $f(x)$, and $\sigma_j^2(x)$ are twice differentiable. $K(x)$ is a bounded second order kernel, $\kappa = \int K(v)^2 dv$, $\kappa_2 = \int k(v)v^2 dv$, and $b = (b_1, \dots, b_q)$ are the bandwidths for covariates in h_t with $Tb_1 \dots b_q \rightarrow \infty$. Given assumptions 1 - 3 and some regularity conditions, we have

(1) (Asymptotic Equivalence) For $j = l$ and u , the two-step local linear estimator $\widehat{M}_j(x; \widehat{h}_t)$ with generated covariates \widehat{h}_t is asymptotic equivalent to the infeasible estimator $\widehat{M}_j(x; h_t^o)$ in the sense that

$$\sup_{x \in I} |\widehat{M}_j(x; \widehat{h}_t) - \widehat{M}_j(x; h_t^o)| = o_p(T^{-\alpha})$$

where α is the optimal convergence rate of the infeasible estimator $\widehat{M}_j(x; h_t^o)$;

(2) (Asymptotic Normality) The limiting distribution of the feasible two-step local linear estimator

is the same as that of the infeasible estimator, i.e.,

$$\sqrt{Tb_1 \cdots b_q} \left(\widehat{M}_j(x; \widehat{h}_t) - m(x) - \frac{\kappa_2}{2} \sum_{s=1}^q b_s^2 m''(x) \right) \xrightarrow{d} N \left(0, \frac{\kappa^q \sigma^2(x)}{f(x)} \right).$$

Essentially, the regularity conditions require that (a) the process from which the true moments h_t^o will be estimated be stationary and β -mixing; (b) the estimates \widehat{h}_t converge to their true values at the \sqrt{T} -rate, which is fast enough; (c) the dynamic specifications in Assumption 3(iii) can be well approximated by linear functions of parameters ψ in a neighborhood of their true values ψ^o ; (d) the error terms ε_{lt} and ε_{ut} have conditional subexponential tails; (e) some other regularity conditions on the kernel functions controlling the bias terms of the local linear smoothing. See also Assumption 1-11 in Conrad and Mammen (2009), and Assumption 1-6 in Mammen, Rothe and Schienle (2012).

4 Extreme Returns and Intensity of Trading

We model the interval-valued time series of the five-minute low/high returns to three bank stocks: Wells Fargo (WFC), Bank of America (BAC), and JP Morgan Chase (JPM), which are traded in the New York Stock Exchange, from January 3rd to January 31st, 2011 for a total of 20 trading days. We record the stock prices and the number of trades at the 5-minute frequency. Log returns (in per ten-thousand unit: ‰) are calculated with respect to the last price of the previous five-minute period:

$$y_{ut} \equiv \log(P_{high,t}/P_{close,t-1}) \times 10,000\text{‰}, \text{ (the highest return)}$$

$$y_{lt} \equiv \log(P_{low,t}/P_{close,t-1}) \times 10,000\text{‰}, \text{ (the lowest return)}$$

$$y_{ct} \equiv \log(P_{close,t}/P_{close,t-1}) \times 10,000\text{‰}, \text{ (close-to-close return)}$$

where $P_{high,t}$ and $P_{low,t}$ are the highest and lowest price in the five-minute period t , and $P_{close,t-1}$ is the last price in the previous five-minute period ($t-1$). Given the interval-valued return $[y_{lt}, y_{ut}]$, the center c_t and range r_t are defined as

$$c_t \equiv \frac{y_{lt} + y_{ut}}{2} = \log(\sqrt{P_{high,t}P_{low,t}}/P_{close,t-1}) \times 10,000\text{‰},$$

$$r_t \equiv y_{ut} - y_{lt} = \log(P_{high,t}/P_{low,t}) \times 10,000\%_{000}.$$

We remove any trades before 9:30 am (opening time) and after 4:00 pm (closing time) for a total number of observations: 20 days \times 78 observations per day = 1560.

[Table 1][Figure 1]

In Figure 1, we plot the 5-minute stock returns calculated from high, low, and close prices, and the numbers of trades for the three U.S. banks. In the first two columns of Table 1, we show the descriptive statistics of low and high returns. Both low and high returns exhibit large volatilities, skewness, and heavy tails. Low returns seem to be more volatile than high returns. BAC returns are more volatile than those of WFC and JPM. The correlations of low and high returns are around 0.27. The descriptive statistics for the center and range of the low/high interval returns are reported in the fourth and fifth column of the table. There is a mild negative correlation between the centers and the ranges. For the series of number of trades, in the last column of the table, we observe that the variances are more than 1000 times larger than their corresponding means, indicating that over-dispersion is present in the data. Therefore, in modeling the numbers of trades, a negative binomial distribution is a sensible choice.

4.1 Conditional moments of the latent process

We proceed to model the conditional moments of the underlying latent process Y_t and the series of the number of trades N_t . We implement two different approaches to model the dynamics of Y_t . First, we consider the center and range of the low/high intervals as good proxies for the location and scale of the conditional probability density function of Y_t . We will call this approach the “interval value approach for Y_t ”. Second, we simply use the close-to-close returns series $\{y_{ct}\}$ as a realized sample path of the underlying latent processes Y_t and model its conditional moments. We will call this approach the “point value approach for Y_t ”.

4.1.1 Interval value approach for modeling Y_t

We model the dynamics of the center and range series separately. For the center series $\{c_t\}$, we fit simple ARMA models and select the optimal number of lags by AIC. The preferred specifications are:

- WFC: AR(1), $c_t = \rho_1 c_{t-1} + u_t$
- BAC: AR(4), $c_t = \rho_1 c_{t-1} + \rho_2 c_{t-2} + \rho_3 c_{t-3} + \rho_4 c_{t-4} + u_t$
- JPM: AR(2), $c_t = \rho_1 c_{t-1} + \rho_2 c_{t-2} + u_t$,

We call \widehat{c}_t the fitted value for the center that proxies the estimated conditional mean of the latent process. In Table 2, we report the estimations results of the three ARMA models. We observe that, as expected, the time dependence in the conditional mean is very weak for the three bank stocks.

[Table 2]

For the range series $\{r_t\}$, we specify a Conditional AutoRegressive Range model with Weibull distribution (WCARR). Since the original range series $\{r_t\}$ exhibit a strong diurnal pattern (see Figures 2b-4b), we first remove the intraday seasonality by spline smoothing, that is,

$$r_t^* = r_t / \widehat{f}(i_t), \quad \widehat{f}(i_t) = \exp[\widehat{d}(i_t)], \quad \widehat{d}(i_t) = \widehat{\beta}_0 + \sum_{j=1}^L \widehat{\beta}_j f_j(i_t), \quad (4.1)$$

where i_t is the fraction of time in the trading day for the t -th observation such that

$$i_t = \begin{cases} 1, & \text{if } t \bmod 78 = 0 \\ t/78 - \lfloor t/78 \rfloor, & \text{otherwise.} \end{cases}$$

The coefficient estimates $\widehat{\beta}_j$ are obtained by ordinary least squares and the number of splines L is selected by least squares cross-validation. Then, we specify the conditional autoregressive range model with Weibull distribution, WCARR(p, s), for the adjusted range r_t^* ,

$$\begin{aligned} r_t^* &= \psi_t \varepsilon_t, \\ \psi_t &= \omega + \sum_{i=1}^p \gamma_i r_{t-i}^* + \sum_{j=1}^s \kappa_j \psi_{t-j}, \\ \varepsilon_t | \mathcal{F}_{t-1} &\sim g_{\text{weibull}}(\cdot; \theta), \end{aligned}$$

where ψ_t is the conditional mean of the adjusted range based on the information set available at time t . The normalized adjusted range $\varepsilon_t = r_t^*/\psi_t$ is assumed to be standardized Weibull distributed with density function $g_{\text{weibull}}(\cdot; \theta)$ with unit mean and shape parameter θ . The coefficients ω , γ_i , and κ_j are restricted to be positive and $\sum_{i=1}^r \gamma_i + \sum_{j=1}^s \kappa_j < 1$ to ensure that the series r_t^* is positive and stationary. The optimal lags p and s in $\text{WCARR}(p, s)$ are selected by AIC and the adequacy of the models is assessed with standard diagnostics. The selected models are:

- WFC: $\text{WCARR}(1,1)$,
- BAC: $\text{WCARR}(1,1)$,
- JPM: $\text{WCARR}(1,1)$.

We denote \widehat{r}_t^* as the fitted range with diurnal pattern adjustment and $\widehat{r}_t = \widehat{r}_t^* \widehat{f}(i_t)$ the fitted range for the original range. In Table 3, we report the estimation results of the three WCARR models. The dynamics of WFC and JPM are very similar with a persistence of about 0.90 ($\gamma_1 + \kappa_1$), while that of BAC is 0.76. The shape parameter θ is about 2 for all three banks indicating that the empirical conditional standardized probability density is far from exponential. Standard diagnostic tests indicate that the fitting is adequate. The standardized residuals have zero mean and unit standard deviation. The Q-statistics of the standardized residuals show that there is no residual dependence left in the data.

[Table 3]

In Figures 2–4 we plot the center and range series with their corresponding fitted values from the estimated AR and WCARR models.

[Figures 2–4]

4.1.2 Point value approach for modelling Y_t

We choose the close price return series $\{y_{ct}\}$ as a realized sample path of the underlying latent processes Y_t . The conditional mean is zero and we model the volatility of the process as follows,

$$y_{ct} = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i y_{c,t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad \varepsilon_t \sim \text{GED}(0, 1, \nu),$$

where the Generalized Error Distribution (GED) captures the heavy tail behavior of the data. The optimal number of lags is selected by AIC and the resulting specifications are the following,

- WFC: GARCH(1,2)-GED,
- BAC: GARCH(1,1)-GED,
- JPM: GARCH(1,1)-GED.

We denote $\hat{\sigma}_t$ the estimate of the conditional standard deviation. In Table 4, we report the estimation results for the three GARCH models. In Figures 5a – 7a, we plot the estimated conditional standard deviation for the three banks. The persistence from the three model is very similar with a value of 0.9. The shape parameter ν is about 1 for the three banks. The Q-statistics of the standardized squared residuals indicate that there is not dependence left in the data.

[Table 4]

4.2 Conditional trading intensity

We specify autoregressive dynamics in the conditional trading intensity to account for the temporal dependence in the number of trades series $\{n_t\}$. As in the range series $\{r_t\}$, the number of trades series exhibits a clear diurnal pattern (see Figure 1). We remove the intra-day seasonality by spline smoothing on the original series, that is,

$$n_t^* = n_t / \hat{f}(i_t), \tag{4.2}$$

where $\hat{f}(i_t)$ and i_t are defined as in the case of the range series explained in the previous section.

Then, for the adjusted numbers of trades n_t^* , we specify the following model

$$\psi_t = \omega + \sum_{i=1}^p \gamma_i n_{t-i}^* + \sum_{j=1}^s \kappa_j \psi_{t-j}. \tag{4.3}$$

Combining (4.2) and (4.3), we propose the following Autoregressive Conditional Intensity (ACI) model,

$$\lambda_t = \hat{f}(i_t) \psi_t, \quad \text{and} \quad n_t \sim g_{\text{NB}}(\cdot; \lambda_t, d). \tag{4.4}$$

where λ_t is the conditional trading intensity based the information set available at time t . The number of trades N_t is assumed to be negative binomial distributed with density function $g_{\text{NB}}(\cdot; \lambda_t, d)$ with mean λ_t and dispersion parameter d . We restrict $\sum_{i=1}^p \gamma_i + \sum_{j=1}^s \kappa_j < 1$ to ensure that the series $\{n_t\}$ stationary. We select the optimal number of lags p and s by AIC and we test the specification of the model with standard diagnostic statistics. The estimated conditional trading intensity is denoted as $\hat{\lambda}_t$.

We report the estimation results in Table 5 and we plot the estimated conditional intensity against the actual number of trades in Figures 5b – 7b. For WFC and JPM, the preferred model is ACI(2,1) and for BAC, the model is ACI(3,1). The persistence of trading for the three banks is about 0.95. The pseudo-Pearson residuals have mean zero and variance one and their Q-statistics do not show any dependence, indicating that these specifications are adequate.

[Table 5] [Figures 5–7]

4.3 Nonparametric conditional mean for lower and upper bounds

From the modeling of the latent process Y_t , we gather the regressors, i.e., conditional mean, standard deviation, and intensity, that will be fed into a nonparametric regression in order to obtain the conditional mean of the lower and upper bounds, i.e., y_{lt} and y_{ut} , of the return interval. Since the latent process has been modelled by two alternative approaches, we will propose and evaluate alternative nonparametric regressions.

The first model has as regressors the estimated conditional centers \hat{c}_t (ARMA models), conditional ranges \hat{r}_t (WCARR models), and conditional intensity $\hat{\lambda}_t$ (ACI models). It is the most general and parsimonious model.

- **Model 1:** $y_{jt} = m_j(\hat{c}_t, \hat{r}_t, \hat{\lambda}_t) + v_{jt}$, for $j = l, u$.

If the models for centers, ranges, and intensity have short dynamics, we could avoid the first estimation step and include directly original regressors such as c_{t-1} , r_{t-1} , n_{t-1} , etc. into the nonparametric regressions. The drawback of this approach is that, if the number of lags is very large, we run into the curse of dimensionality. With the current data, we experiment with the

following model,

- **Model 2:** $y_{jt} = m_j(c_{t-1}, r_{t-1}, n_{t-1}) + v_{jt}$, for $j = l, u$.

The next model considers regressors from the point value approach for modeling Y_t such as the estimated conditional standard deviation $\hat{\sigma}_t$ (GARCH models) in addition to the conditional intensity $\hat{\lambda}_t$ (ACI models), $\hat{\lambda}_t$, i.e.,

- **Model 3:** $y_{jt} = m_j(\hat{\sigma}_t, \hat{\lambda}_t) + v_{jt}$, for $j = l, u$.

We propose the next two nonparametric models in order to assess the importance of the intensity of trading in the modelling of the conditional mean of the upper and lower bounds,

- **Model 4:** $y_{jt} = m_l(\hat{c}_t, \hat{r}_t) + v_{jt}$, for $j = l, u$.
- **Model 5:** $y_{jt} = m_j(c_{t-1}, r_{t-1}) + v_{jt}$, for $j = l, u$,

Finally, we also propose Model 6 to assess whether the conditional trading intensity $\hat{\lambda}_t$ could be sufficient to predict the expected lower and upper bounds without information about the latent process Y_t ,

- **Model 6:** $y_{jt} = m_j(\hat{\lambda}_t) + v_{jt}$, for $j = l, u$.

4.3.1 In-sample model evaluation

We evaluate the performance of the proposed models by comparing several measures of fit for interval-valued data. In addition to the six specifications of the previous section, we include two additional models proposed in González-Rivera and Lin (2013), which are constrained VAR-type specifications satisfying the inequality $y_{lt} \leq y_{ut}$ for all t . These are Interval Autoregressive-Two Step (IAR-TS) and Interval Autoregressive-Modified Two Step (IAR-MTS). These models have been proven to be superior to the existing interval-valued regression approaches (see González-Rivera and Lin, 2013, and Lin and González-Rivera, 2015).

For a sample of size T , let $[\hat{y}_{lt}, \hat{y}_{ut}]$ be the fitted values of the corresponding interval $\mathbf{y}_t = [y_{lt}, y_{ut}]$ provided by each model. We consider the following criteria:

(i) Mean Squared Error (MSE) for upper and lower bounds separately.

$$\text{MSE}_{\text{lower}} = \sum_{t=1}^T (\hat{y}_{lt} - y_{lt})^2 / T \quad \text{MSE}_{\text{upper}} = \sum_{t=1}^T (\hat{y}_{ut} - y_{ut})^2 / T;$$

(ii) Multivariate Loss Functions (MLF) for the vector of lower and upper bounds (Komunjer and Owyang, 2011): $L_p(\boldsymbol{\tau}, \mathbf{e}) \equiv (\|\mathbf{e}\|_p + \boldsymbol{\tau}'\mathbf{e}) \|\mathbf{e}\|_p^{p-1}$ where $\|\cdot\|_p$ is the l_p -norm, $\boldsymbol{\tau}$ is a two-dimensional parameter vector that determines the asymmetry of the loss function (if $\boldsymbol{\tau} = \mathbf{0}$, the bivariate loss is symmetric), and $\mathbf{e} = (e_l, e_u)$ is the bivariate residual interval $(\hat{y}_{lt} - y_{lt}, \hat{y}_{ut} - y_{ut})$. We consider two norms, $p = 1$ and $p = 2$ and their corresponding $\boldsymbol{\tau}$ parameter vectors within the unit balls $\mathcal{B}_\infty \equiv \{(\tau_1, \tau_2) \in \mathbb{R}^2 : |\tau_1| \leq 1 \text{ and } |\tau_2| \leq 1\}$ and $\mathcal{B}_2 \equiv \{(\tau_1, \tau_2) \in \mathbb{R}^2 : \tau_1^2 + \tau_2^2 \leq 1\}$, respectively.

Then, the Multivariate Loss Functions (MLF) are defined by their sample averages:

$$MLF_1 = \sum_{t=1}^T L_1(\boldsymbol{\tau}_t^*, \mathbf{e}_t) / T, \quad MLF_2 = \sum_{t=1}^T L_2(\boldsymbol{\tau}_t^*, \mathbf{e}_t) / T;$$

where $\boldsymbol{\tau}_t^*$ is the optimal vector that defines the asymmetry of the loss.

(iii) Mean Distance Error (MDE) between the fitted and actual intervals (Arroyo *et al.*, 2011).

Let $D_q(\hat{y}_t, y_t)$ be a distance measure of order q between the fitted and the actual intervals, the mean distance error is defined as $MDE_q(\{\hat{y}_t\}, \{y_t\}) = \sum_{t=1}^T D_q^q(\hat{y}_t, y_t) / T$. We consider $q = 1$ and $q = 2$, with a distance measure such as,

$$\begin{aligned} D_1(\hat{y}_t, y_t) &= \frac{1}{2}(|\hat{y}_{lt} - y_{lt}| + |\hat{y}_{ut} - y_{ut}|), \\ D_2(\hat{y}_t, y_t) &= \frac{1}{\sqrt{2}}[(\hat{y}_{lt} - y_{lt})^2 + (\hat{y}_{ut} - y_{ut})^2]^{1/2}. \end{aligned}$$

Note that MDE_1 and MDE_2 are equal to a half of MLF_1 and MLF_2 respectively if $\boldsymbol{\tau} = \mathbf{0}$.

In Table 6, we report the in-sample evaluation of the models according to the aforementioned three loss functions. The first finding is that the nonparametric regressions are superior to the IAR-TS and IAR-MTS specifications as they deliver in most cases the smallest losses across loss functions and for all three banks. Within the six nonparametric models, the preferred specification is Model 1 across loss functions and for the three banks. Competing specifications are Models 2, 4, and 5 indicating that we could skip the first step estimation (Models 2 and 5) and that omitting trading

intensity (Models 4 and 5) is not very detrimental to the performance of the model. It is interesting to observe that trading intensity alone (Model 6) is far from being an optimal specification, it is only the interaction of the three regressors that are most helpful to estimate the means of the extreme bounds of the interval.

[Table 6]

In addition, in Table 7 we report Diebold-Mariano tests to formally test the superiority of Model 1 *versus* the parametric linear model and the remaining five non-parametric models. For WFC, there is overwhelming evidence that Model 1 is superior across loss functions; all p-values but one are practically zero so that we reject the null hypothesis of equally predictive accuracy (in-sample). For BAC, the evidence is mixed. While the linear model IAR-TS and Models 4 and 6 are easily rejected, Model 5 emerges as a competitor to Model 1, and to a lesser extent Models 2 and 3. For JPM, Models 1, 2 and 4 seem to be equivalent but IAR-TS and Models 3, 5, and 6 are inferior to Model 1. Overall, these results confirm that the joint inclusion of the three regressors, i.e. center, range, and intensity, are needed to produce the best fit for the bounds of the interval.

[Table 7]

We plot the nonparametric functions of Model 1 in Figures 8– 10 for the three bank stocks. Overall, we find that the relationship between extreme returns and trading intensity goes in the expected direction, i.e. the more trading, the larger the extreme returns are. We observe an asymmetric response of the extreme returns to trading intensity. The lowest returns are sensitive to small and large trading volumes but the highest returns are responsive mainly when the trading intensity is very large. For the WFC extreme returns with very large trading volume, we find a bend in the functions but the uncertainty around these returns is also very high because we do not have many observations in that area of the function. In general, the nonlinear behavior of the functions shows up for very large trading intensity.

[Figures 8-10]

4.3.2 Out-of-sample model evaluation

We evaluate the proposed models in an out-of-sample exercise. For each model, we calculate the one-step-ahead forecast and assess their predictive accuracy according to the aforementioned loss functions. We maintain the same estimation sample as before, i.e. January 3rd to January 31st, 2011 for a total of 20 trading days with 1560 observations. The prediction sample runs for the subsequent five trading days, from February 1st to February 7th, 2011 with 390 observations (78 observations per day). We implement a fixed scheme. All models are estimated only once with $t = 1, \dots, T = 1560$ observations. Over the prediction sample, we calculate the one-step-ahead forecast of the lower and upper bounds at time $T + h$ for $h = 1, \dots, P = 390$ as follows,

$$\begin{aligned}
 \text{Model 1 : } & \hat{y}_{j,T+h|T+h-1} = \hat{m}_j(\hat{c}_{T+h|T+h-1}, \hat{r}_{T+h|T+h-1}, \hat{\lambda}_{T+h|T+h-1}), \quad \text{for } j = l, u, \\
 \text{Model 2 : } & \hat{y}_{j,T+h|T+h-1} = \hat{m}_j(c_{T+h-1}, r_{T+h-1}, n_{T+h-1}), \quad \text{for } j = l, u, \\
 \text{Model 3 : } & \hat{y}_{j,T+h|T+h-1} = \hat{m}_j(\hat{\sigma}_{T+h|T+h-1}, \hat{\lambda}_{T+h|T+h-1}), \quad \text{for } j = l, u, \\
 \text{Model 4 : } & \hat{y}_{j,T+h|T+h-1} = \hat{m}_j(\hat{c}_{T+h|T+h-1}, \hat{r}_{T+h|T+h-1}), \quad \text{for } j = l, u, \\
 \text{Model 5 : } & \hat{y}_{j,T+h|T+h-1} = \hat{m}_j(c_{T+h-1}, r_{T+h-1}), \quad \text{for } j = l, u, \\
 \text{Model 6 : } & \hat{y}_{j,T+h|T+h-1} = \hat{m}_j(\hat{\lambda}_{T+h|T+h-1}), \quad \text{for } j = l, u,
 \end{aligned}$$

where the $\hat{m}_j(\cdot)$ functions are estimated nonparametrically within the estimation sample, and either the regressors are directly observable in the updated information set at time $T + h - 1$, i.e., lagged values $(c_{T+h-1}, r_{T+h-1}, n_{T+h-1})$ (Models 2 and 5), or they are the one-step-ahead forecasts coming from their respective models (Models 1, 3, 4, and 6)

In Table 8 we report the results. Overall, the nonparametric models deliver similar losses, which in turn are generally smaller than those from the linear models. For WFC, Models 2 and 3 are the best performers, for BAC Models 3 and 6, and for JPM, Models 1 and 4. The intensity of trading is key in all these models with the exception of Model 4. The results for WFC and JPM are roughly in agreement with the results of the in-sample evaluation. In these cases, Models 1, 2, and 4 seem to be the preferred specifications.

[Table 8]

In Table 9, for each bank, we compare the best non-parametric model against the linear specifications IAR-TS/MTS by computing the Diebol-Mariano test. For WFC, Model 2, and for BAC, Model 6 are superior to the linear models except when the loss function is the MSE (lower bound). For JPM, Model 4 is roughly equivalent to the linear model IAR-TS but it is superior to IAR-MTS except when the loss function is MSE. The overall conclusion is that non-parametric specifications that include the intensity of trading are superior specifications, either in-sample or out-of-sample, compared to autoregressive linear models of lower and upper bounds.

[Table 9]

5 Conclusion

In contrast to existing regression-type models for interval-valued data, we have exploited the extreme nature of the lower and upper bounds of intervals to propose a semiparametric model for interval-valued time series (ITS) data that is rooted in the limiting results provided by the extreme value theory. We have assumed that there are two stochastic processes that generated the interval-valued data. The first process $\{Y_t\}$ is latent, e.g. the process of financial returns, and follows some unknown conditional density. The second process $\{N_t\}$ consists of a collection of random draws, e.g. the process of number of trades, extracted from the conditional density of the latent process at each period of time. In this framework, the upper and lower bounds of the interval, e.g. the highest and the lowest returns at time t , are the realized extreme observations (maximum and minimum) within the sample of random draws at time t . We have shown that the conditional mean of extreme returns is a nonlinear function of the conditional moments of the latent process and of the conditional intensity of the process that governs the number of draws. This specification provides a natural context to test the relationship between intensity of trading and extreme returns, which is important according to information models of market microstructure. These models claim that trading volume is a proxy for latent information on the value of a financial asset. With an ITS of 5-minute returns to three major bank stocks, we have found that indeed there is a nonlinear relationship between extreme returns and intensity of trading, which is superior to linear specifications either in-sample or out-of-sample. In addition, we have found that the response of the extreme

returns to trading intensity is asymmetric with the lowest returns being more responsive to small and large volumes of trading than the highest returns; these are responsive mainly to very large trading intensity.

From an econometric point of view, the proposed semiparametric model has advantages over the existing models. It is general enough to accommodate linear specifications when these are granted, but the most important advantage is that the model is robust to misspecification of the conditional density of the latent process. We have estimated the conditional mean of the extremes, which is nonlinear on the conditional moments of the latent process, with nonparametric methods. In doing so, we have avoided to choose a specific conditional density, whose functional form is the driver of the nonlinearity. However, the nonparametric function depends on regressors that are generated in a first step. We have shown that the effect of the first-step parameter uncertainty into the second-step nonparametric estimator is asymptotically negligible, and therefore, our two-step estimator has typical nonparametric convergence rate and it is asymptotically normal.

References

- [1] Andersen, T. (1996), “Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility,” *Journal of Finance*, 51, pp. 169-204.
- [2] Arroyo, J., González-Rivera, G, and Maté, C. (2011), “Forecasting with Interval and Histogram Data. Some Financial Applications,” in *Handbook of Empirical Economics and Finance*, A. Ullah and D. Giles (eds.). Chapman and Hall, pp. 247-280.
- [3] Conrad, C., and Mammen, E. (2009), “Nonparametric Regression on Latent Covariates with an Application to Semiparametric GARCH-in-Mean Models,” University of Heidelberg, Department of Economics, Discussion Paper No. 473.
- [4] Darrat, A.F., Rahman, S., and Zhong, M. (2003). “Intraday Trading Volume and Return Volatility of the DJIA Stocks: A Note,” *Journal of Banking and Finance*, 27, pp. 2035-2043.
- [5] de Haan, L., and Ferreira, A. (2000), *Extreme Value Theory: An Introduction*. Springer, New York, NY.
- [6] Domowitz, I., and White, H. (1982), “Maximum Likelihood Estimation of Misspecified Models,” *Journal of Econometrics*, 20, pp. 35-58.
- [7] Easley, D., and O’Hara, M. (1992), “Time and the Process of Security Price Adjustment,” *Journal of Finance*, 47, pp. 905-927.
- [8] Embrechts, P., Klüppelberg, C., and Mikosch, T. (2003), *Modelling Extremal Events: for Insurance and Finance*. Springer, New York, NY.
- [9] Engle, R.F. (2000), “The Econometrics of Ultra-High-Frequency Data,” *Econometrica*, 68, pp. 1-22.
- [10] Fleming, J., and Kirby, C.(2011), “Long Memory in Volatility and Trading Volume,” *Journal of Banking & Finance*, 35, pp. 1714-1726.
- [11] González-Rivera, G., and Lin, W. (2013), “Constrained Regression for Interval-valued Data,” *Journal of Business and Economic Statistics*, Vol. 31, No. 4, pp. 473–490.

- [12] Komunjer, I., and Owyang, M. (2012), “Multivariate Forecast Evaluation and Rationality Testing,” *The Review of Economics and Statistics*. Vol. 94, pp. 1066-1080.
- [13] Lamoureux, C., and Lastrapes, W. (1990), “Heteroskedasticity in Stock Return Data: Volume Versus GARCH Effect,” *Journal of Finance*, 45, pp. 221-229.
- [14] Leadbetter, M.R., Lindgren, G., and Rootzén, H. (1988), “Extreme Theory for Stochastic Processes,” *Annals of Probability*. Vol. 16, pp. 431-478.
- [15] Li, Q., and Racine, J. (2007), *Nonparametric Econometrics*. Princeton University Press, Princeton, NJ.
- [16] Lin, W., and González-Rivera, G. (2015), “Interval-valued Time Series Models: Estimation based on Order Statistics. Exploring the Agriculture Marketing Service Data” *Computational Statistics and Data Analysis*, available on line July 2015.
- [17] Mammen, E., Rothe, C. and Schienle, M. (2012), “Nonparametric Regression with Nonparametrically Generated Covariates,” *The Annals of Statistics*. Vol. 40, No. 2, 1132–1170.
- [18] Ning, C., and Wirjanto T.S. (2009), “Extreme Return Volume Dependence in East-Asian Stock Markets: A Copula Approach,” *Finance Research Letters*, 6, pp. 202-209.
- [19] Parkinson, M. (1980). “The Extreme Value Method for Estimating the Variance of the Rate of Return,” *Journal of Business*, 53, pp. 61-65.
- [20] Rodrigues., P., and Salish, N. (2015), “Modeling and Forecasting Interval Time Series with Threshold Models,” *Advances in Data Analysis and Classification*, 9, pp. 41-57.
- [21] Sita, B.B., and Westerholm, P.J. (2011), “The Role of Trading Intensity Estimating the Implicit BidAsk Spread and Determining Transitory Effects,” *International Review of Financial Analysis*, 20, pp. 306-310.

Table 1: Descriptive Statistics for WFC, BAC, and JPM

Wells Fargo (2011/1/3 – 2011/1/31)						
Statistics	Returns in [†]			Center	Range	# of trades
	low price	high price	close price			
Minimum	-200	-79.37	-142.9	-122.2	3.068	201
1st Quartile	-18.16	4.611	-9.217	-5.596	15.430	895
Median	-9.297	9.311	0	0	21.510	1252
3rd Quartile	-3.198	15.86	9.231	4.757	30.790	1910
Maximum	92.95	165.5	165.5	125	185.300	13790
Mean	-13.33	12.81	0.2457	-0.2625	26.140	1589
Variance	278.593	221.041	363.666	158.996	364.013	1500781
Correlation	0.275			-0.12		
Skewness	-3.421	3.730	0.374	0.065	3.081	3.636
Kurtosis	25.609	26.585	12.351	25.063	14.258	21.569
Bank of America (2011/1/3 – 2011/1/31)						
Statistics	Returns in [†]			Center	Range	# of trades
	low price	high price	close price			
Minimum	-319.400	-82.53	-237.8	-188.4	6.625	522
1st Quartile	-20.960	6.669	-10.11	-6.912	16.780	1744
Median	-10.95	10.37	0.000	0.000	26.560	2608
3rd Quartile	-6.692	20.41	10.05	6.560	36.290	4159
Maximum	136.40	219.5	187.5	163.7	352.400	25410
Mean	-16.00	15.15	-0.07676	-0.4269	31.160	3302
Variance	497.032	408.582	483.34	285.800	671.3791	5733710
Correlation	0.264			-0.10		
Skewness	-5.454	4.022	-0.042	-0.6383	4.7474	2.473
Kurtosis	58.94	30.980	19.54	34.551	37.0742	10.304
J.P. Morgan Chase (2011/1/3 – 2011/1/31)						
Statistics	Returns in [†]			Center	Range	# of trades
	low price	high price	close price			
Minimum	-231.800	-77.910	-132.9	-112.3	3.409	372
1st Quartile	-17.560	4.428	-8.902	-5.581	14.410	1199
Median	-9.046	8.913	0.000	-0.245	20.170	1711
3rd Quartile	-4.445	15.610	8.874	5.0840	29.200	2532
Maximum	153.70	244.0	234.0	195.0	244.600	19380
Mean	-12.580	12.280	0.2465	-0.1491	24.860	2119
Variance	244.1	234.8	309.79	153.02	345.8719	2338652
Correlation	0.278			-0.03		
Skewness	-3.433	5.849	1.5094	2.5592	4.666	3.548
Kurtosis	47.73	70.02	24.42	57.452	39.433	24.78

[†] Returns are in the unit of per ten-thousand ($\%_{100}$), and sample size is 1559.

Table 2: ARMA models for center series of low/high return interval

ARMA Models [†]			
	WFC: AR(1)	BAC: AR(4)	JPM: AR(2)
ρ_1	-0.0583 (0.0253)	0.0726 (0.0253)	-0.0207 (0.0253)
ρ_2		-0.0501 (0.0254)	-0.0428 (0.0253)
ρ_3		0.0157 (0.0254)	
ρ_4		-0.0572 (0.0253)	
LogLike	-6160.51	-6612.45	-6131.36
AIC	12325.02	13234.89	12268.73

[†] Standard errors are reported in the parentheses.

Table 3: WCARR Model for range of low/high return interval

	WFC-WCARR(1,1) Model				BAC-WCARR(1,1) Model				JPM-WCARR(1,1) Model			
	estimate	s.e.	<i>t</i> -value	<i>p</i> -value	estimate	s.e.	<i>t</i> -value	<i>p</i> -value	estimate	s.e.	<i>t</i> -value	<i>p</i> -value
ω	0.0724	0.0156	4.6349	0.0000	0.2769	0.0533	5.1999	0.0000	0.1180	0.0390	3.0238	0.0000
γ_1	0.2219	0.0207	10.7263	0.0000	0.3076	0.0302	10.1747	0.0000	0.1863	0.0306	6.0883	0.0000
κ_1	0.7133	0.0285	25.0618	0.0000	0.4547	0.0662	6.8685	0.0000	0.7096	0.0594	11.9547	0.0000
θ	2.3369	0.0401	58.2897	0.0000	1.9226	0.0322	59.7791	0.0000	1.9380	0.0303	63.8635	0.0000
LogLikelihood: -981.6314				LogLikelihood: -1260.8521				LogLikelihood: -1192.2117				
Akaike Information Criteria: 1971.2628				Akaike Information Criteria: 2529.7042				Akaike Information Criteria: 2392.4233				
Ljung-Box Test for standardized residuals [†]				Ljung-Box Test for standardized residuals [‡]				Ljung-Box Test for standardized residuals [§]				
	statistic	<i>p</i> -value		statistic	<i>p</i> -value		statistic	<i>p</i> -value		statistic	<i>p</i> -value	
	Q(50)	40.1339	0.8395	Q(50)	56.1300	0.2560	Q(50)	60.4859	0.1471			
	Q(100)	65.2376	0.9972	Q(100)	97.0816	0.5640	Q(100)	87.9208	0.8005			
	Q(200)	148.4100	0.9975	Q(200)	172.5804	0.9201	Q(200)	159.0520	0.9851			

[†] For the standardized residuals, the mean and standard deviation are 0.0029 and 0.97 respectively.

[‡] For the standardized residuals, the mean and standard deviation are -0.005 and 1.03 respectively.

[§] For the standardized residuals, the mean and standard deviation are -0.002 and 1.01 respectively.

Table 4: Volatility Models for WFC, BAC, and JPM

	WFC-GARCH(1,2)-GED Model				BAC-GARCH(1,1)-GED Model				JPM-GARCH(1,1)-GED Model			
	estimate	robust s.e.	<i>t</i> -value	<i>p</i> -value	estimate	robust s.e.	<i>t</i> -value	<i>p</i> -value	estimate	robust s.e.	<i>t</i> -value	<i>p</i> -value
ω	29.5123	10.1286	2.9138	0.0036	35.1059	9.9250	3.5371	0.0004	36.1753	12.5979	2.8715	0.0041
α_1	0.3650	0.0726	5.0257	0.0000	0.1324	0.0317	4.1802	0.0000	0.2085	0.0472	4.4211	0.0000
β_1	0.2383	0.1412	1.6873	0.0915	0.7948	0.0356	22.2981	0.0000	0.6857	0.0599	11.4374	0.0000
β_2	0.3584	0.1190	3.0127	0.0026								
ν	1.1075	0.0849	13.0490	0.0000	0.9481	0.0828	11.4442	0.0000	1.0803	0.1112	9.7185	0.0000
LogLikelihood: -6463.12				LogLikelihood: -6705.238				LogLikelihood: -6432.971				
Ljung-Box test on standardized squared residuals												
	statistic		<i>p</i> -value		statistic		<i>p</i> -value		statistic		<i>p</i> -value	
Q(1)		0.0011		0.9738	Q(1)	0.0469		0.8285	Q(1)	0.0437		0.8344
Q(8)		0.4660		0.9970	Q(5)	0.2927		0.9845	Q(5)	0.1205		0.9973
Q(14)		0.9552		0.9997	Q(9)	0.5444		0.9980	Q(9)	0.1879		0.9999

Table 5: ACI Models for WFC, BAC, and JPM

WFC-ACI(2,1) Model					BAC-ACI(3,1) Model				JPM-ACI(2,1) Model			
	estimate	robust s.e.	t -value	p -value	estimate	robust s.e.	t -value	p -value	estimate	robust s.e.	t -value	p -value
ω	0.4481	0.1617	2.7720	0.0056	0.2499	0.0873	2.8614	0.0042	0.4667	0.2133	2.1876	0.0287
γ_1	0.3422	0.0333	10.2849	0.0000	0.3168	0.0292	10.8674	0.0000	0.4347	0.0326	13.3470	0.0000
γ_2	-0.2013	0.0482	-4.1727	0.0000	-0.1167	0.0432	-2.7030	0.0069	-0.2359	0.0738	-3.1979	0.0014
γ_3					-0.0812	0.0305	-2.6576	0.0079				
κ_1	0.8180	0.0463	17.6777	0.0000	0.8584	0.0288	29.8280	0.0000	0.7587	0.0788	9.6237	0.0000
$1/d$	0.1289	0.0055	23.2921	0.0000	0.1290	0.0053	24.3793	0.0000	0.1117	0.0049	22.7421	0.0000
LogLikelihood: -11839.9					LogLikelihood: -12949.54				LogLikelihood: -12190.78			
Ljung-Box test on standardized residuals [†]												
	statistic		p -value		statistic		p -value		statistic		p -value	
Q(50)	52.2080		0.3882		Q(50)		49.0242 0.5125		Q(50)		34.3940 0.9547	
Q(100)	115.7841		0.1338		Q(100)		90.8216 0.7332		Q(100)		85.7186 0.8449	
Q(200)	198.5963		0.5148		Q(200)		187.3835 0.7294		Q(200)		168.5562 0.9484	

[†] For the pseudo-Pearson residuals, the mean and standard deviation are $3.54e-05$ and 1.08 respectively for WFC; the mean and standard deviation are $-4.98e-05$ and 1.08 respectively for BAC; and the mean and standard deviation are $-1.71e-04$ and 1.08 respectively for JPM.

Table 6: In-sample model evaluation for 5-minute low/high stock returns

Wells Fargo Corporation						
Models	<i>MSE</i>		<i>MLF</i>		<i>MDE</i>	
	<i>Lower</i>	<i>Upper</i>	<i>p</i> = 1	<i>p</i> = 2	<i>q</i> = 1	<i>q</i> = 2
IAR-TS	225.4843	196.1827	35.1482	840.9812	8.7871	210.8335
IAR-MTS	225.5024	198.9469	35.2948	846.4363	8.8237	212.2246
Model 1	208.6861	120.8959	32.9122	657.4045	8.2280	164.7911
Model 2	184.6784	178.3050	34.1240	723.8444	8.5310	181.4921
Model 3	224.6382	188.5376	35.2679	824.1540	8.8170	206.5889
Model 4	218.6153	134.2691	33.6605	703.9149	8.4151	176.4434
Model 5	234.0900	194.4786	35.6289	854.6348	8.9072	214.2857
Model 6	239.2713	197.9835	36.6756	872.2196	9.1689	218.6288
Bank of America						
Models	<i>MSE</i>		<i>MLF</i>		<i>MDE</i>	
	<i>Lower</i>	<i>Upper</i>	<i>p</i> = 1	<i>p</i> = 2	<i>q</i> = 1	<i>q</i> = 2
IAR-TS	451.7409	369.5716	44.9750	1638.2019	11.2438	410.6563
IAR-MTS	458.2550	375.0649	45.5453	1662.1910	11.3863	416.6599
Model 1	413.6199	258.2997	41.6851	1340.324	10.4213	335.9598
Model 2	344.1869	340.7477	42.7650	1366.713	10.6912	342.4683
Model 3	363.0650	345.6019	42.8349	1414.036	10.7087	354.3317
Model 4	427.6282	334.3466	43.4654	1520.334	10.8663	380.9874
Model 5	395.9245	309.4655	41.2292	1407.391	10.3073	352.6959
Model 6	449.9118	359.3084	44.3920	1614.758	11.0980	404.6092
J.P. Morgan Chase						
Models	<i>MSE</i>		<i>MLF</i>		<i>MDE</i>	
	<i>Lower</i>	<i>Upper</i>	<i>p</i> = 1	<i>p</i> = 2	<i>q</i> = 1	<i>q</i> = 2
IAR-TS	209.8808	218.9624	33.9770	855.7951	8.4942	214.4216
IAR-MTS	209.9269	223.7333	34.1107	865.4129	8.5277	216.8301
Model 1	186.3672	162.8760	31.6307	696.9394	7.9077	174.6216
Model 2	153.9036	199.1422	31.9905	704.6221	7.9976	176.5229
Model 3	199.9346	204.6015	33.4886	807.2904	8.3722	202.2680
Model 4	197.0991	163.8791	31.6052	720.2441	7.9013	180.4891
Model 5	182.0219	215.1001	33.1919	792.5513	8.2980	198.5610
Model 6	213.4358	206.3306	34.0122	837.5538	8.5031	209.8832

The numbers in boldface correspond to the two lowest values for a given loss function.

Table 7: In-sample Diebold-Mariano tests of Model 1 v.s. other models

		WFC		BAC		JPM	
		statistic	p -value [†]	statistic	p -value [†]	statistic	p -value [†]
Model 1 v.s. IAR-TS	MSE: lower	-1.7706	0.0383	-1.7115	0.0435	-2.2009	0.0139
	MSE: upper	-3.5722	0.0002	-3.3316	0.0004	-3.5849	0.0002
	MLF: $p = 1$	-5.5251	0.0000	-6.6855	0.0000	-5.6545	0.0000
	MLF: $p = 2$	-4.1470	0.0000	-4.3175	0.0000	-3.6006	0.0002
	MDE: $q = 1$	-5.5251	0.0000	-6.6855	0.0000	-5.6545	0.0000
	MDE: $q = 2$	-4.1403	0.0000	-4.3167	0.0000	-3.6093	0.0002
Model 1 v.s. 2	MSE: lower	2.0622	0.9804	1.8070	0.9646	0.9329	0.8246
	MSE: upper	-3.1605	0.0008	-3.0428	0.0012	-1.1083	0.1339
	MLF: $p = 1$	-3.0931	0.0010	-2.0822	0.0187	-1.7549	0.0396
	MLF: $p = 2$	-1.6865	0.0459	-0.2793	0.3900	-1.0452	0.1480
	MDE: $q = 1$	-3.0931	0.0010	-2.0822	0.0187	-1.7549	0.0396
	MDE: $q = 2$	-1.6869	0.0458	-0.2751	0.3916	-1.0451	0.1480
Model 1 v.s. 3	MSE: lower	-2.2814	0.0113	1.3114	0.9051	-1.3829	0.0834
	MSE: upper	-3.3436	0.0004	-2.6705	0.0038	-3.1908	0.0007
	MLF: $p = 1$	-5.6205	0.0000	-2.3274	0.0100	-5.4948	0.0000
	MLF: $p = 2$	-3.8956	0.0000	-0.7705	0.2205	-2.8638	0.0021
	MDE: $q = 1$	-5.6205	0.0000	-2.3274	0.0100	-5.4948	0.0000
	MDE: $q = 2$	-3.8933	0.0000	-0.7679	0.2213	-2.8711	0.0020
Model 1 v.s. 4	MSE: lower	-2.3590	0.0092	-2.0840	0.0186	-2.1493	0.0158
	MSE: upper	-3.1899	0.0007	-3.2952	0.0005	-0.2922	0.3851
	MLF: $p = 1$	-4.2000	0.0000	-4.7639	0.0000	0.2000	0.5793
	MLF: $p = 2$	-3.5445	0.0002	-3.6268	0.0001	-1.9030	0.0285
	MDE: $q = 1$	-4.2000	0.0000	-4.7639	0.0000	0.2000	0.5793
	MDE: $q = 2$	-3.5429	0.0002	-3.6254	0.0001	-1.9025	0.0286
Model 1 v.s. 5	MSE: lower	-2.6425	0.0041	0.5477	0.7081	-0.0774	0.4691
	MSE: upper	-3.5136	0.0002	-1.4738	0.0703	-3.7016	0.0001
	MLF: $p = 1$	-6.5968	0.0000	0.7885	0.7848	-4.4300	0.0000
	MLF: $p = 2$	-4.4198	0.0000	-0.8195	0.2063	-2.0457	0.0204
	MDE: $q = 1$	-6.5968	0.0000	0.7885	0.7848	-4.4300	0.0000
	MDE: $q = 2$	-4.4192	0.0000	-0.8160	0.2072	-2.0479	0.0203
Model 1 v.s. 6	MSE: lower	-4.5786	0.0000	-1.5907	0.0558	-2.3501	0.0094
	MSE: upper	-3.7608	0.0001	-3.0339	0.0012	-3.2804	0.0005
	MLF: $p = 1$	-8.0081	0.0000	-5.9897	0.0000	-6.7402	0.0000
	MLF: $p = 2$	-4.9236	0.0000	-3.5191	0.0002	-3.3733	0.0004
	MDE: $q = 1$	-8.0081	0.0000	-5.9897	0.0000	-6.7402	0.0000
	MDE: $q = 2$	-4.9218	0.0000	-3.5222	0.0002	-3.3838	0.0004

[†] The p -values are calculated under the alternative hypothesis $H_a : L_p(e) < L_c(e)$, i.e., our proposed Model 1 has higher predicative accuracy than the other competing models.

Table 8: Out-of-sample evaluation of one-step-ahead forecast for 5-min low/high returns

Wells Fargo Corporation							
		MSE _{lower}	MSE _{upper}	MLF ₁	MLF ₂	MDE ₁	MDE ₂
	IAR-TS	360.1965	123.1417	31.7849	964.8324	7.9462	241.6691
	IAR-MTS	360.4025	129.3673	32.5674	977.6035	8.1418	244.8849
5 days [†]	Model 1	341.1291	164.1051	31.1969	1008.4783	7.7992	252.6171
	Model 2	336.9076	114.9612	30.5583	902.2852	7.6396	225.9344
	Model 3	340.2408	110.7288	30.6441	900.6928	7.6610	225.4848
	Model 4	345.0843	171.0736	32.0008	1030.3954	8.0002	258.0789
	Model 5	364.9090	126.9973	32.4253	981.7280	8.1063	245.9532
	Model 6	337.9289	114.1284	30.6995	902.7200	7.6749	226.0287
Bank of America							
		MSE _{lower}	MSE _{upper}	MLF ₁	MLF ₂	MDE ₁	MDE ₂
	IAR-TS	219.9246	293.1244	38.2963	1024.6959	9.5741	256.5245
	IAR-MTS	222.9597	292.2842	38.7431	1028.8959	9.6858	257.6220
5 days [†]	Model 1	249.8208	267.7880	37.0215	1032.6893	9.2554	258.8044
	Model 2	218.2605	280.8948	37.0770	996.7381	9.2693	249.5777
	Model 3	230.3074	268.1896	36.6976	994.8546	9.1744	249.2485
	Model 4	249.7301	249.9335	36.8214	996.3532	9.2053	249.8318
	Model 5	217.3751	285.9479	37.8088	1005.2017	9.4522	251.6615
	Model 6	223.4179	266.3878	36.3261	977.2970	9.0815	244.9029
J.P. Morgan Chase							
		MSE _{lower}	MSE _{upper}	MLF ₁	MLF ₂	MDE ₁	MDE ₂
	IAR-TS	130.7837	136.7175	28.2646	533.7610	7.0661	133.7506
	IAR-MTS	130.9228	142.1662	28.9521	544.8586	7.2380	136.5445
5 days [†]	Model 1	120.6567	136.5684	27.6903	513.2196	6.9226	128.6125
	Model 2	128.7980	137.1945	28.0210	530.9230	7.0053	132.9963
	Model 3	125.4281	138.0706	28.2265	525.6244	7.0566	131.7493
	Model 4	117.0483	134.9645	27.8764	502.7195	6.9691	126.0064
	Model 5	131.3306	139.2919	28.5458	540.1583	7.1365	135.3113
	Model 6	123.6573	138.6216	28.0895	523.5328	7.0224	131.1395

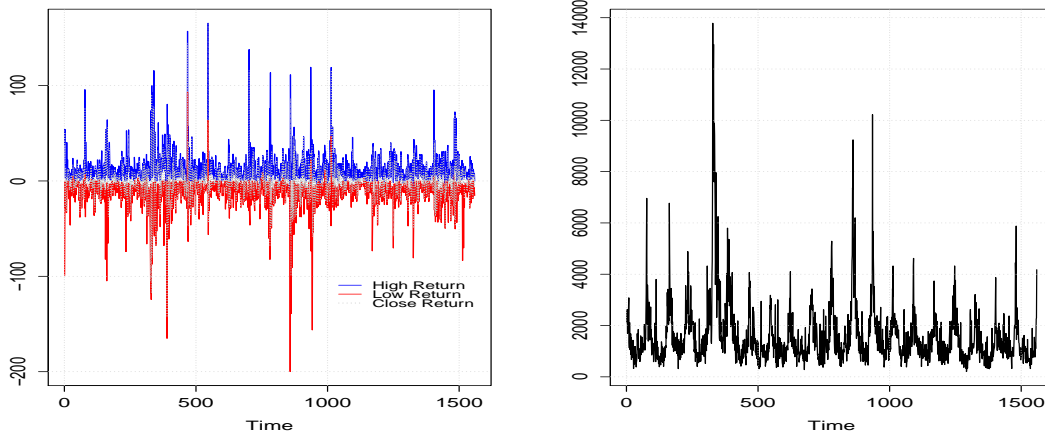
[†] The number of observations in each day is 78. Hence, the number of observations in a 5-day forecast sample is 390.

The numbers in boldface correspond to the two lowest values for a given loss function.

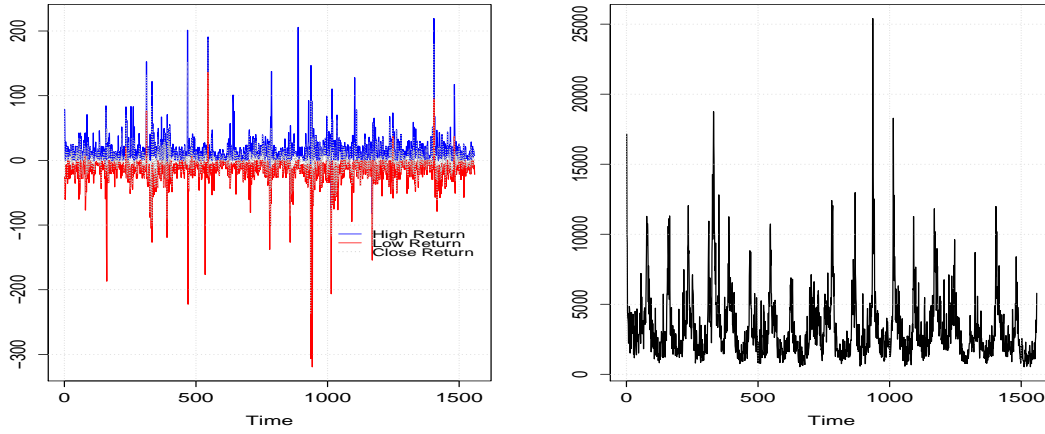
Table 9: Out-of-sample Diebold-Mariano tests of nonparametric models v.s. IAR-TS/MTS models

	WFC		BAC		JPM	
	(Model 2 v.s. IAR-TS)		(Model 6 v.s. IAR-TS)		(Model 4 v.s. IAR-TS)	
	statistic	p -value [†]	statistic	p -value [†]	statistic	p -value [†]
MSE: lower	-1.1255	0.1302	0.2624	0.6035	-1.3543	0.0878
MSE: upper	-1.2994	0.0969	-1.6995	0.0446	-0.1125	0.4552
MLF: $p = 1$	-2.6982	0.0035	-3.7893	0.0001	-1.1109	0.1333
MLF: $p = 2$	-1.7903	0.0367	-2.5713	0.0051	-1.3672	0.0858
MDE: $q = 1$	-2.6982	0.0035	-3.7893	0.0001	-1.1109	0.1333
MDE: $q = 2$	-1.7838	0.0372	-2.5342	0.0056	-1.3684	0.0856
	WFC		BAC		JPM	
	(Model 2 v.s. IAR-MTS)		(Model 6 v.s. IAR-MTS)		(Model 4 v.s. IAR-MTS)	
	statistic	p -value [†]	statistic	p -value [†]	statistic	p -value [†]
MSE: lower	-1.1353	0.1281	0.0331	0.5132	-1.3663	0.0859
MSE: upper	-2.0470	0.0203	-1.6320	0.0513	-0.4592	0.3230
MLF: $p = 1$	-4.2194	0.0000	-4.7888	0.0000	-2.8754	0.0020
MLF: $p = 2$	-2.0621	0.0196	-3.0183	0.0013	-1.8057	0.0355
MDE: $q = 1$	-4.2194	0.0000	-4.7888	0.0000	-2.8754	0.0020
MDE: $q = 2$	-2.0564	0.0199	-2.9970	0.0014	-1.8106	0.0351

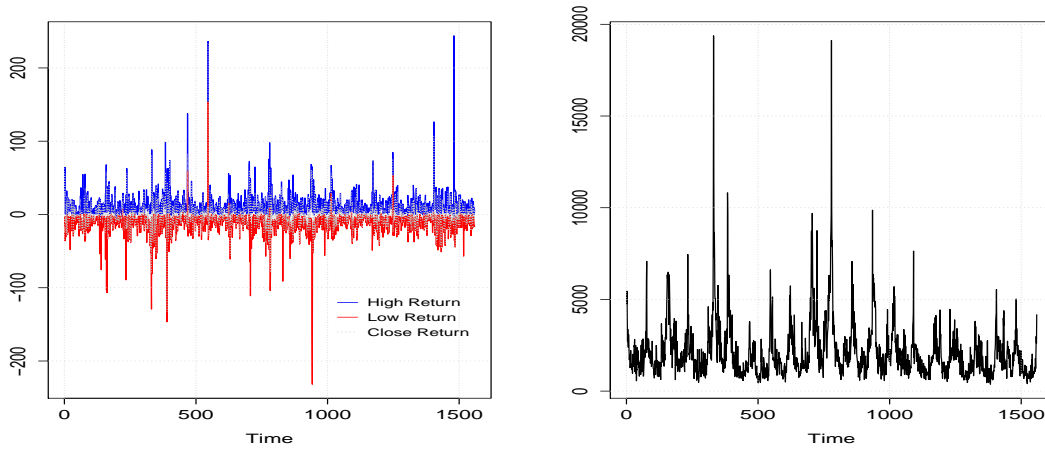
[†] The p -values are calculated under the alternative hypothesis $H_a : L_p(e) < L_c(e)$, i.e., our proposed Model 1 has higher predicative accuracy than the competing model.



(a) WFC: 5-min high, low, and close returns (in %) and numbers of trades



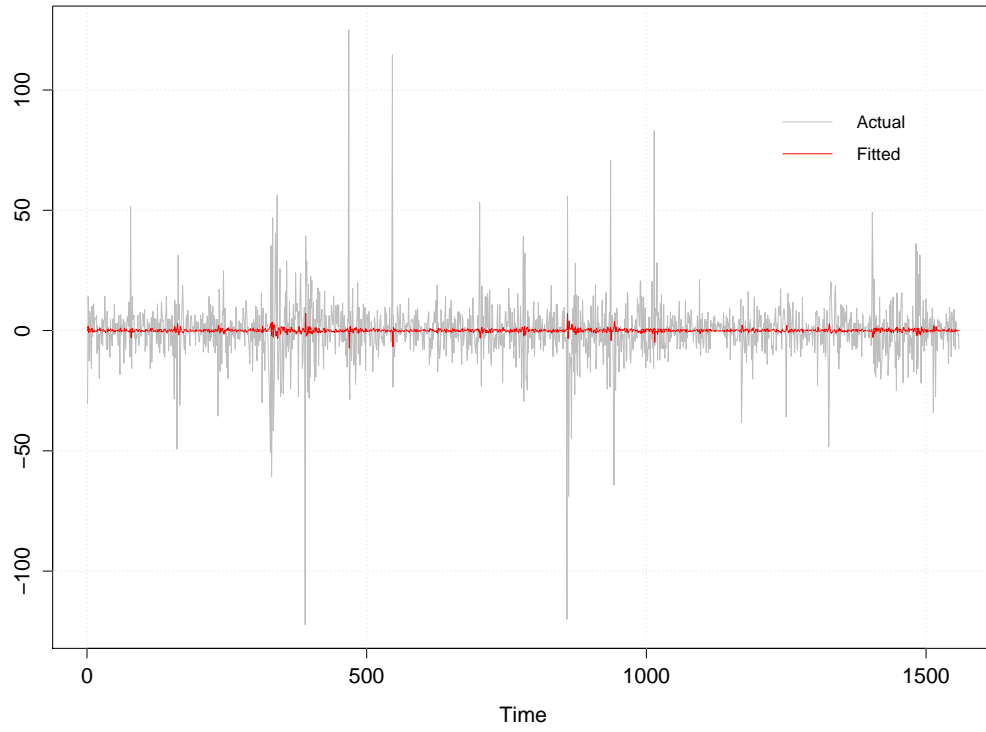
(b) BAC: 5-min high, low, and close returns (in %) and numbers of trades



(c) JPM: 5-min high, low, and close returns (in %) and numbers of trades

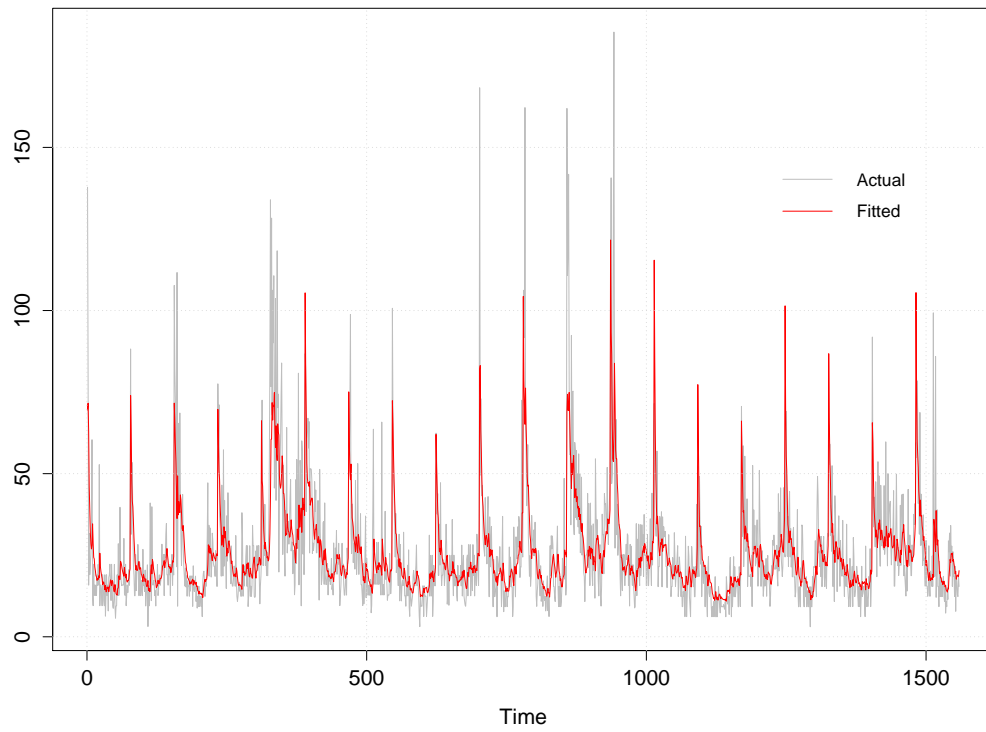
Figure 1: Time series plots

Center series of WFC and AR(1) fit



(a) Center series

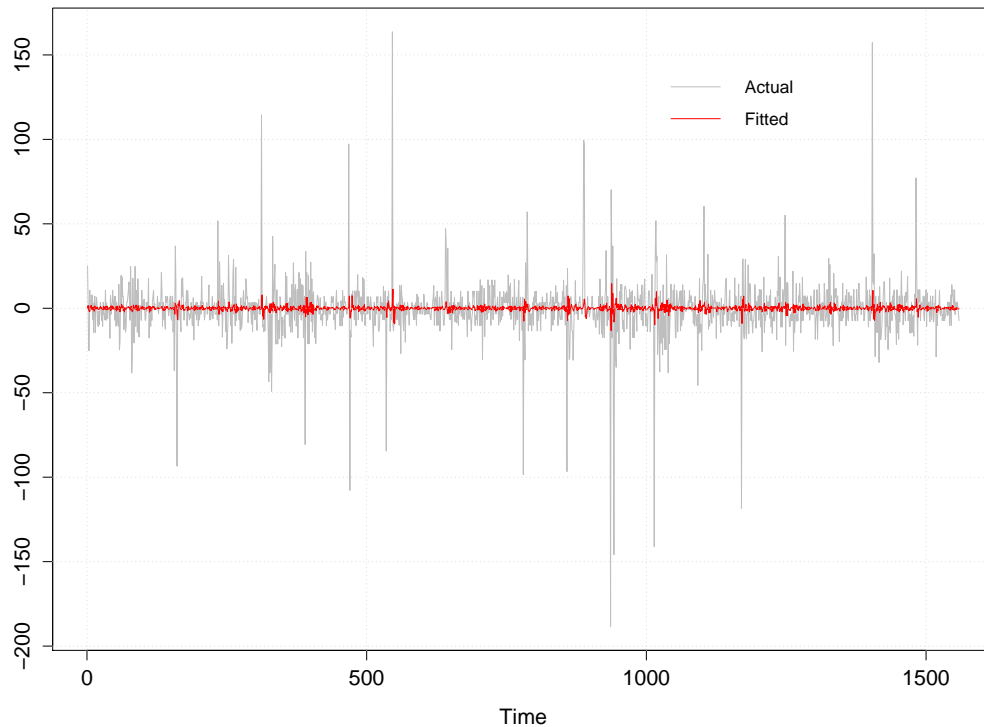
Range series of WFC and WCARR(1,1) fit



(b) Range series

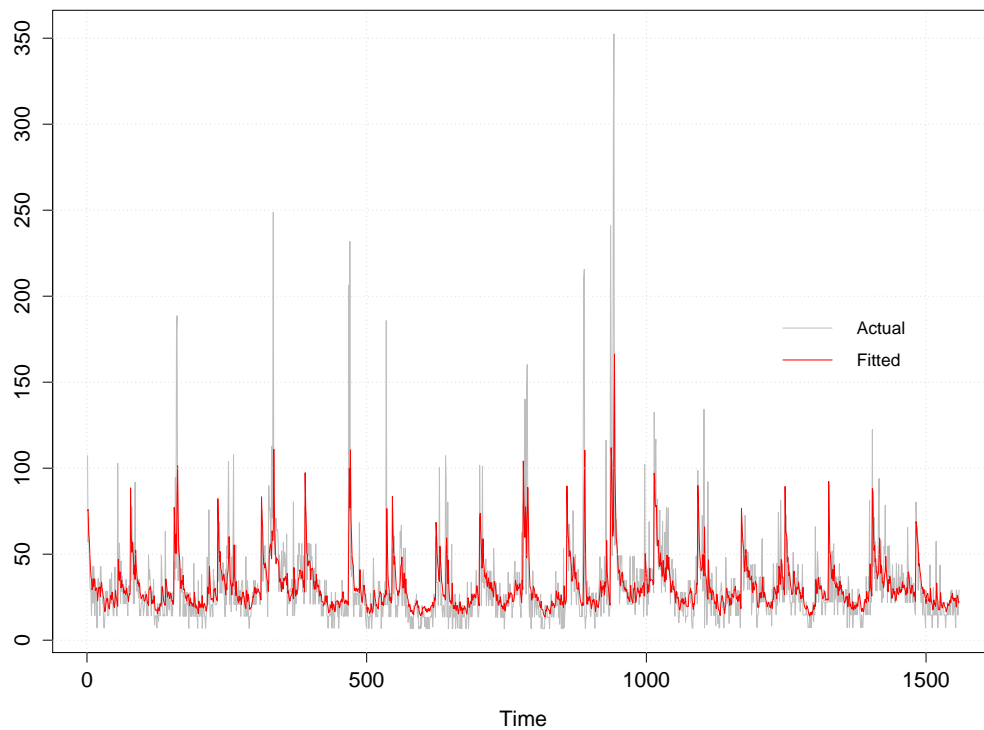
Figure 2: WFC: actual and fitted values of center and range return

Center series of BAC and AR(4) fit



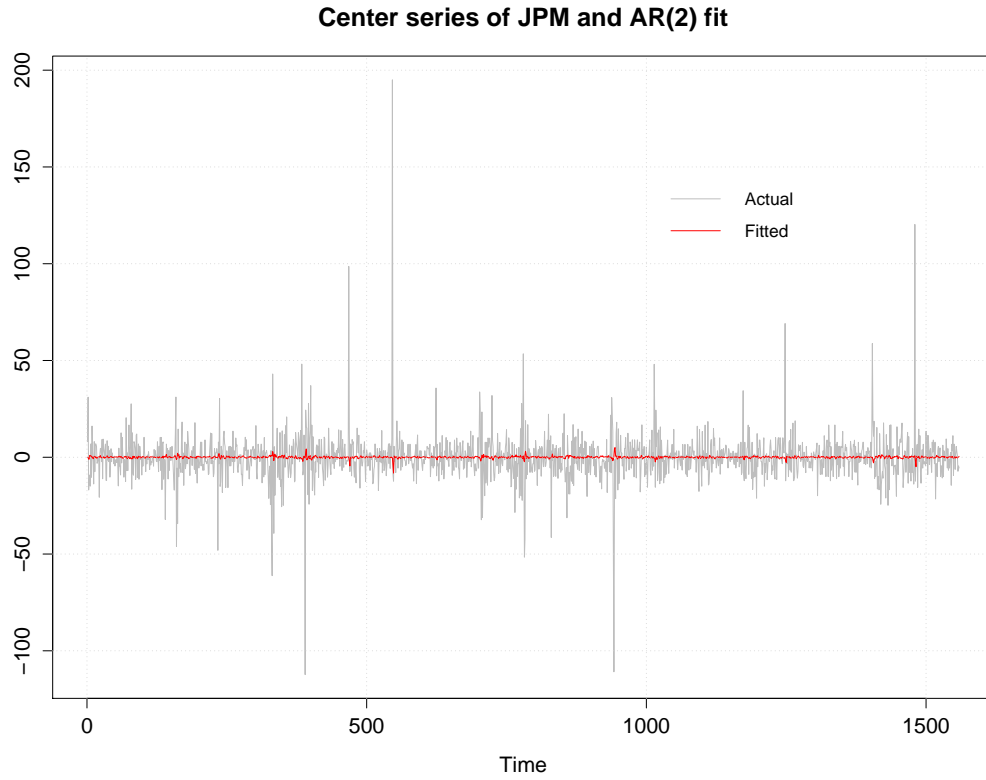
(a) Center series

Range series of BAC and WCARR(1,1) fit

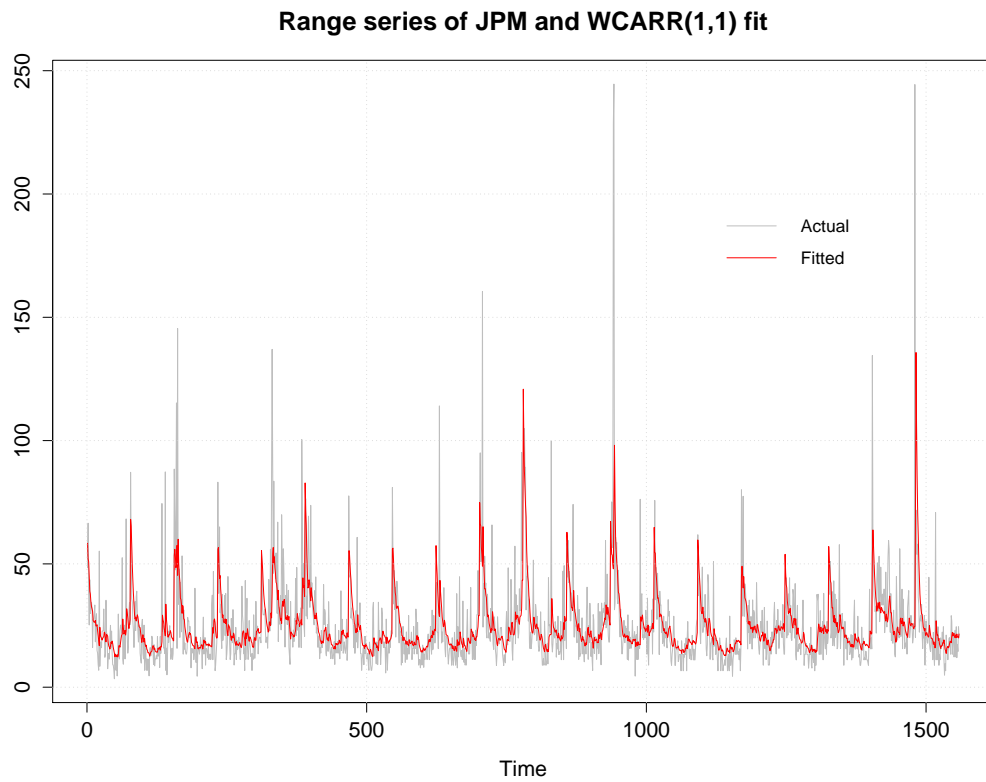


(b) Range series

Figure 3: BAC: actual and fitted values of center and range return



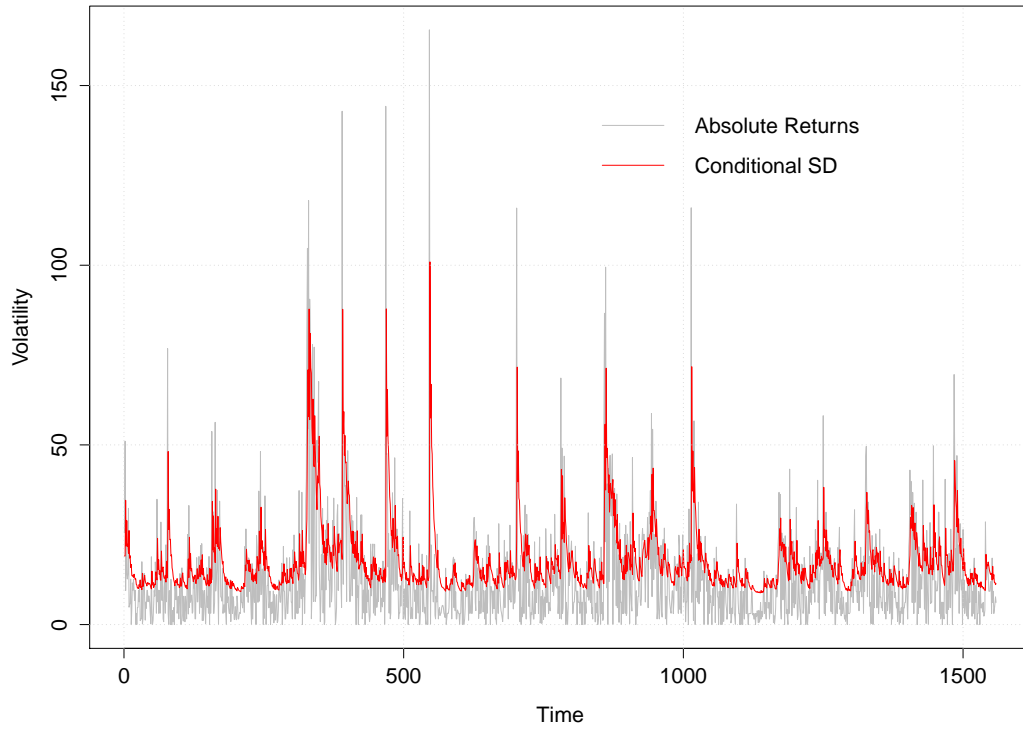
(a) Center series



(b) Range series

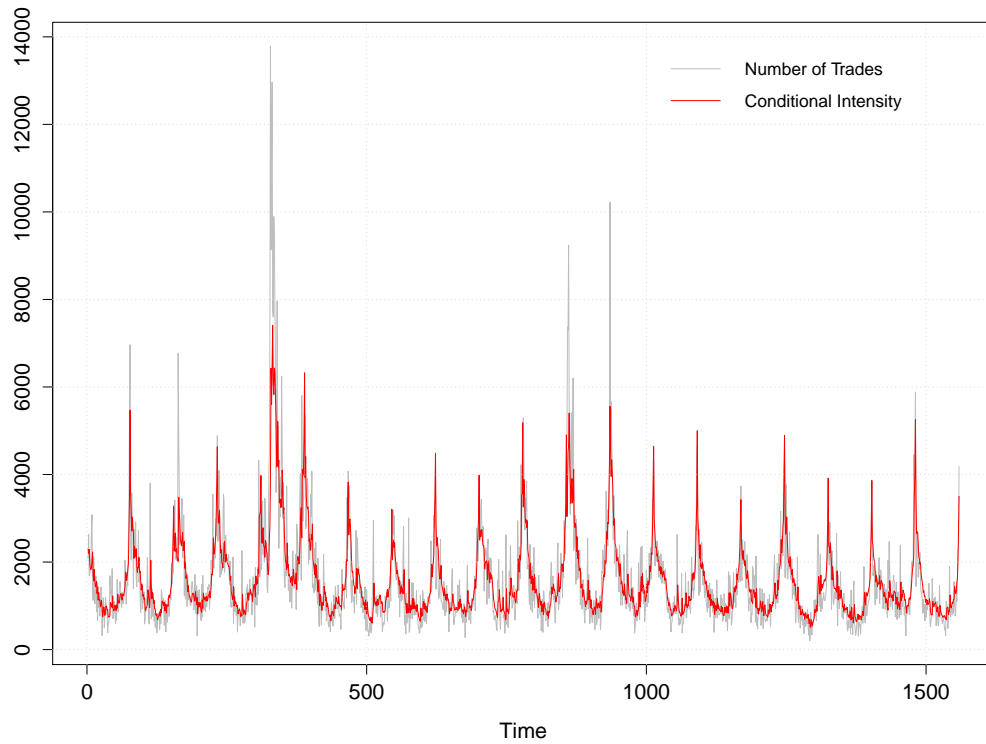
Figure 4: JPM: actual and fitted values of center and range return

Absolute returns of WFC and GARCH(1,2)-GED fit



(a) Absolute return and estimated conditional standard deviation

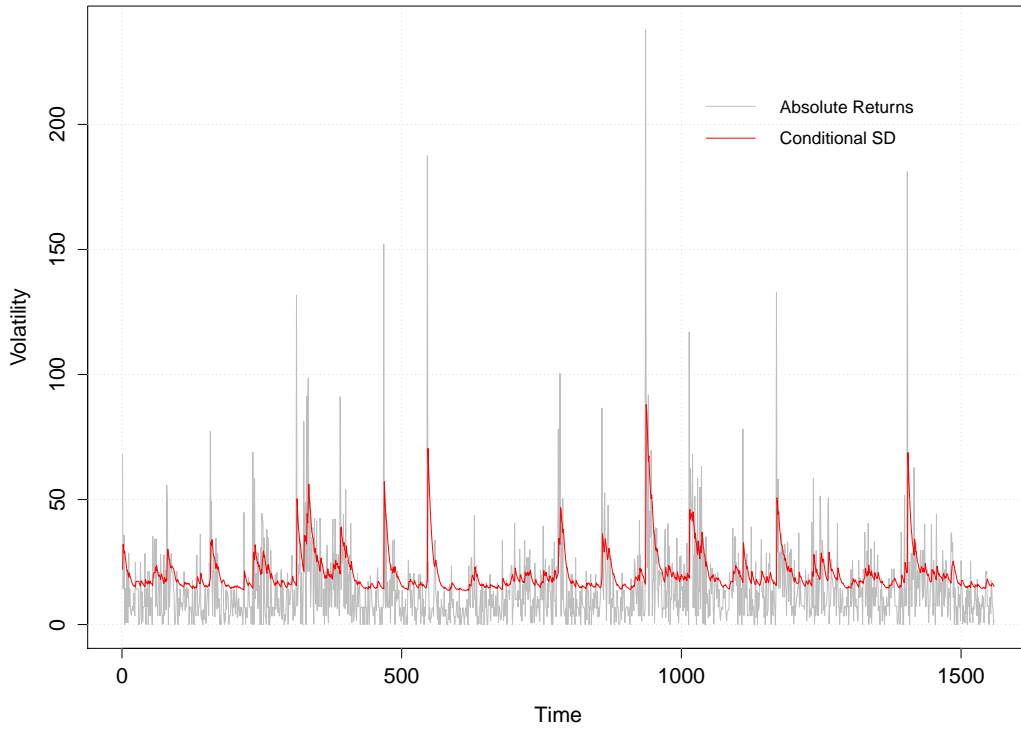
Number of trades series of WFC and ACI(2,1)



(b) Number of trades and estimated conditional intensity

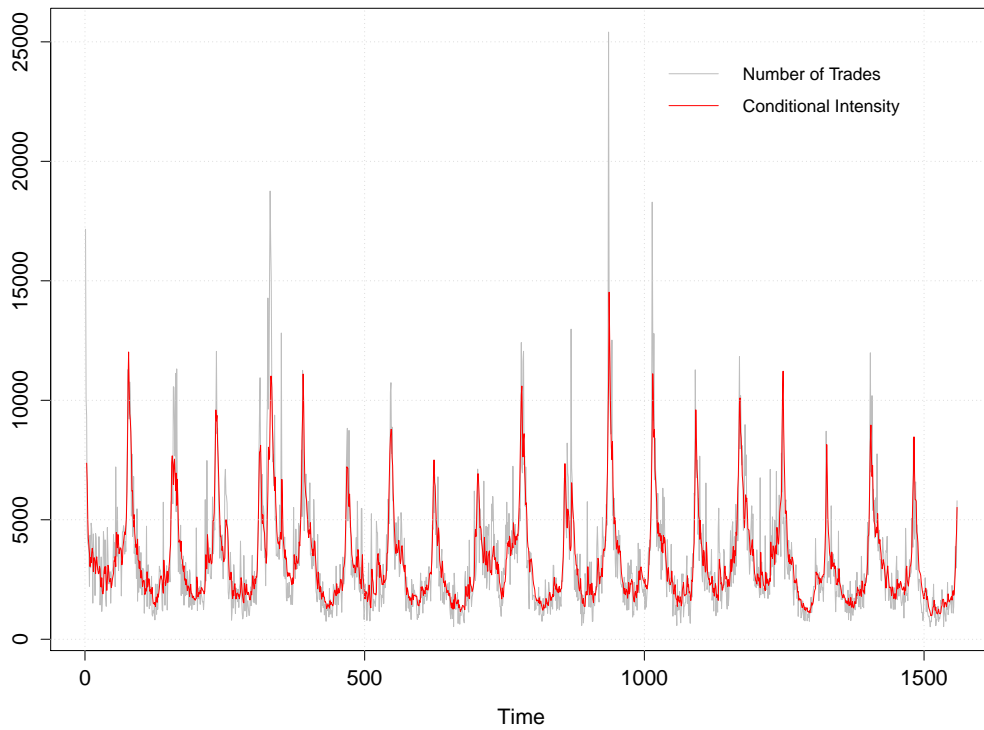
Figure 5: Time series plots of estimates in the first step

Absolute returns of BAC and GARCH(1,1)-GED fit



(a) Absolute return and estimated conditional standard deviation

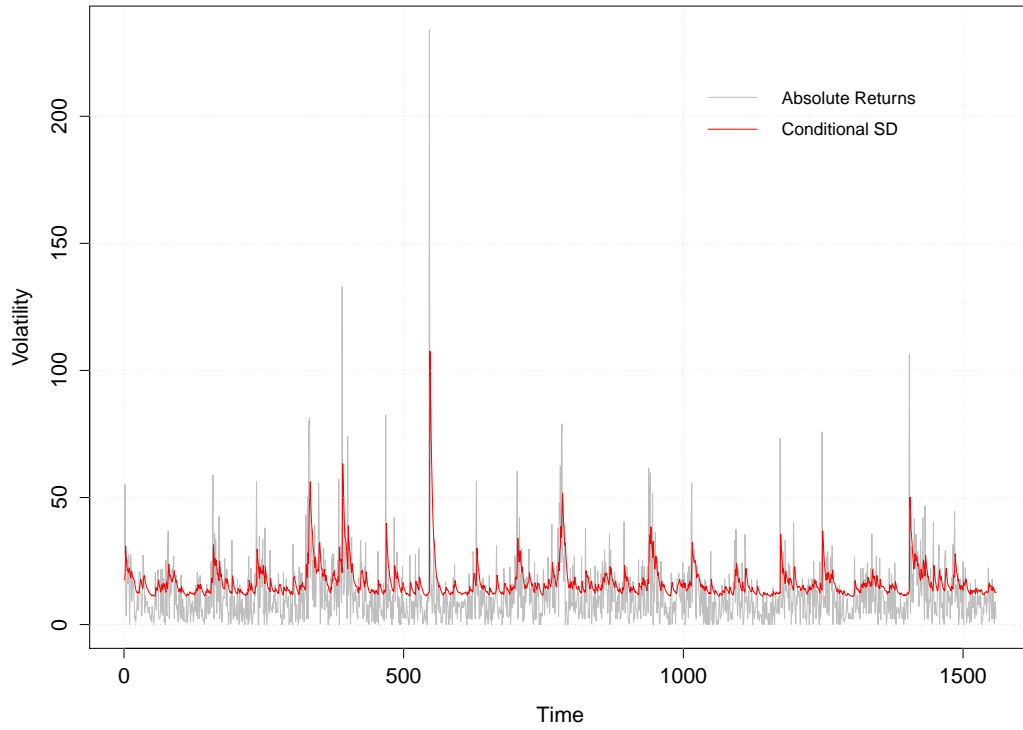
Number of trades series of BAC and ACI(3,1)



(b) Number of trades and estimated conditional intensity

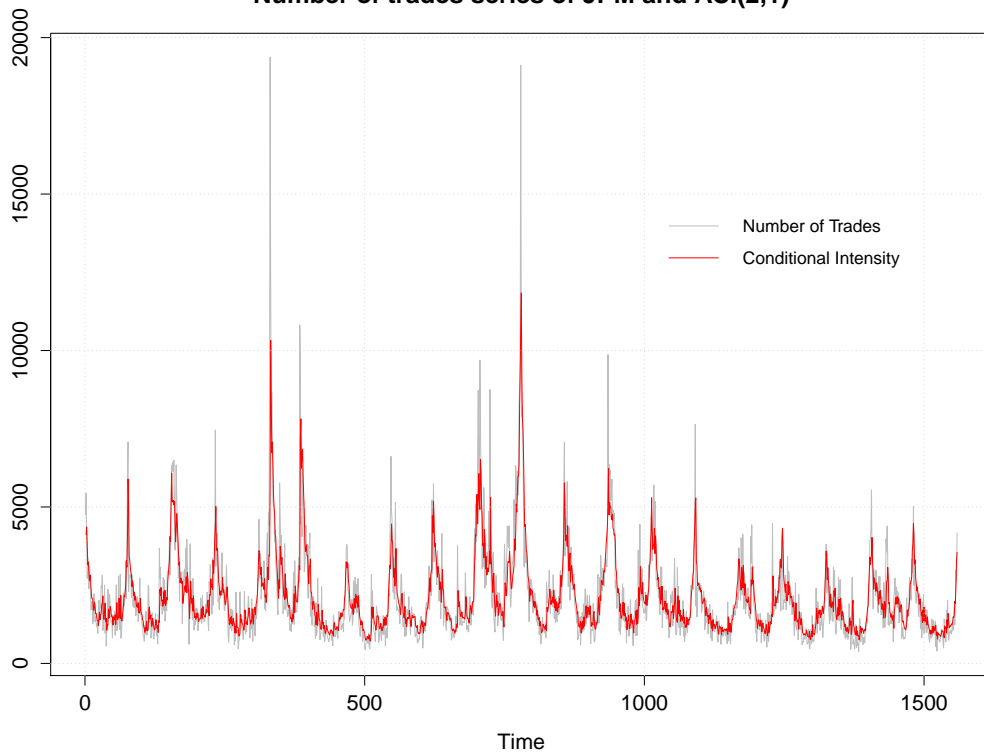
Figure 6: Time series plots of estimates in the first step

Absolute returns of JPM and GARCH(1,1)-GED fit



(a) Absolute return and estimated conditional standard deviation

Number of trades series of JPM and ACI(2,1)



(b) Number of trades and estimated conditional intensity

Figure 7: Time series plots of estimates in the first step

Nonparametric functions for expected extreme returns (WFC)

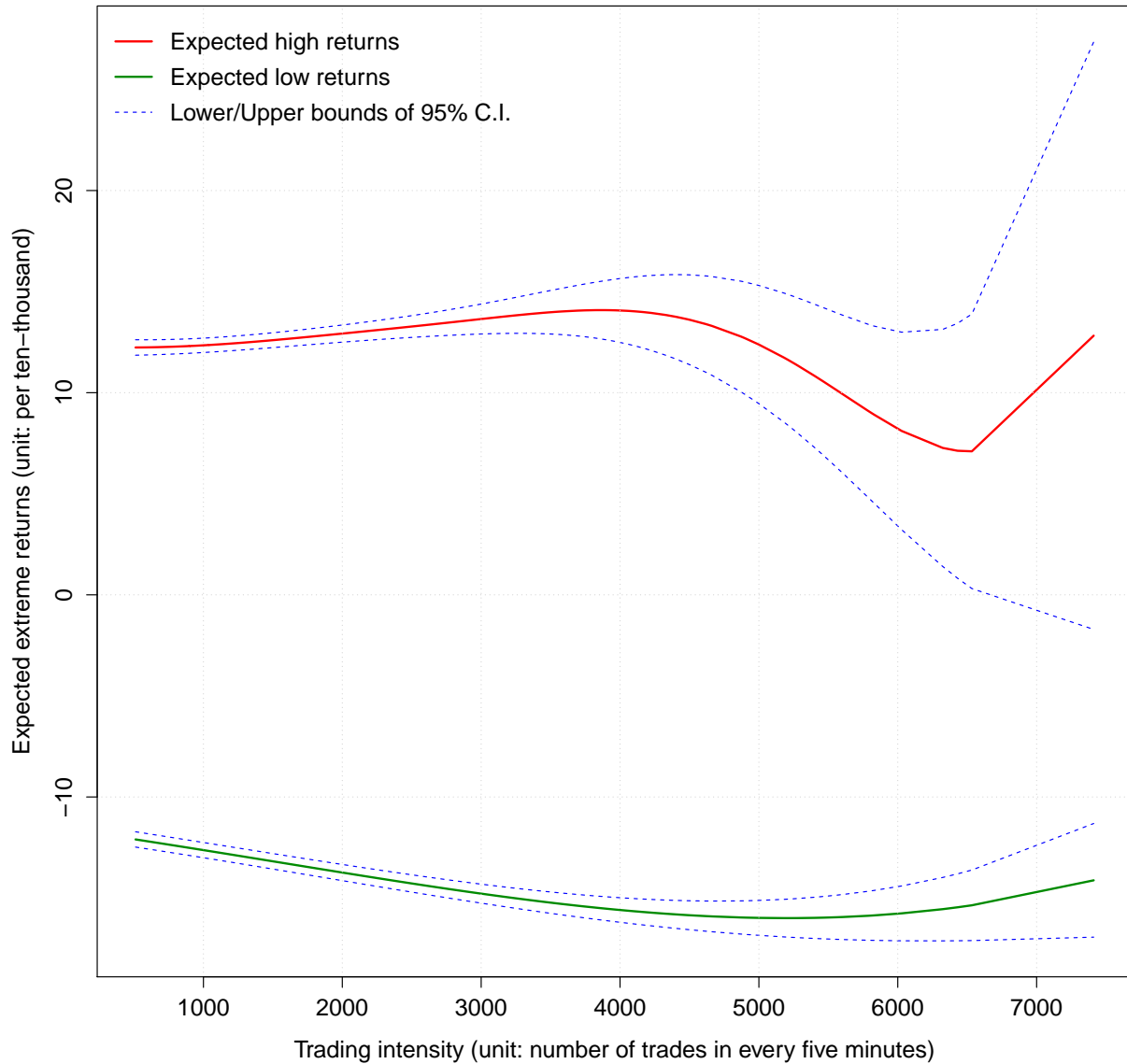


Figure 8: Nonparametric function of expected extreme returns versus trading intensity from Model 1 for WFC (Red solid line and green solid line are conditional extreme returns given trading intensity and fixing the other two regressors, center and range, at their unconditional means. Blue dotted lines are upper and lower 95% confidence intervals.)

Nonparametric functions for expected extreme returns (BAC)

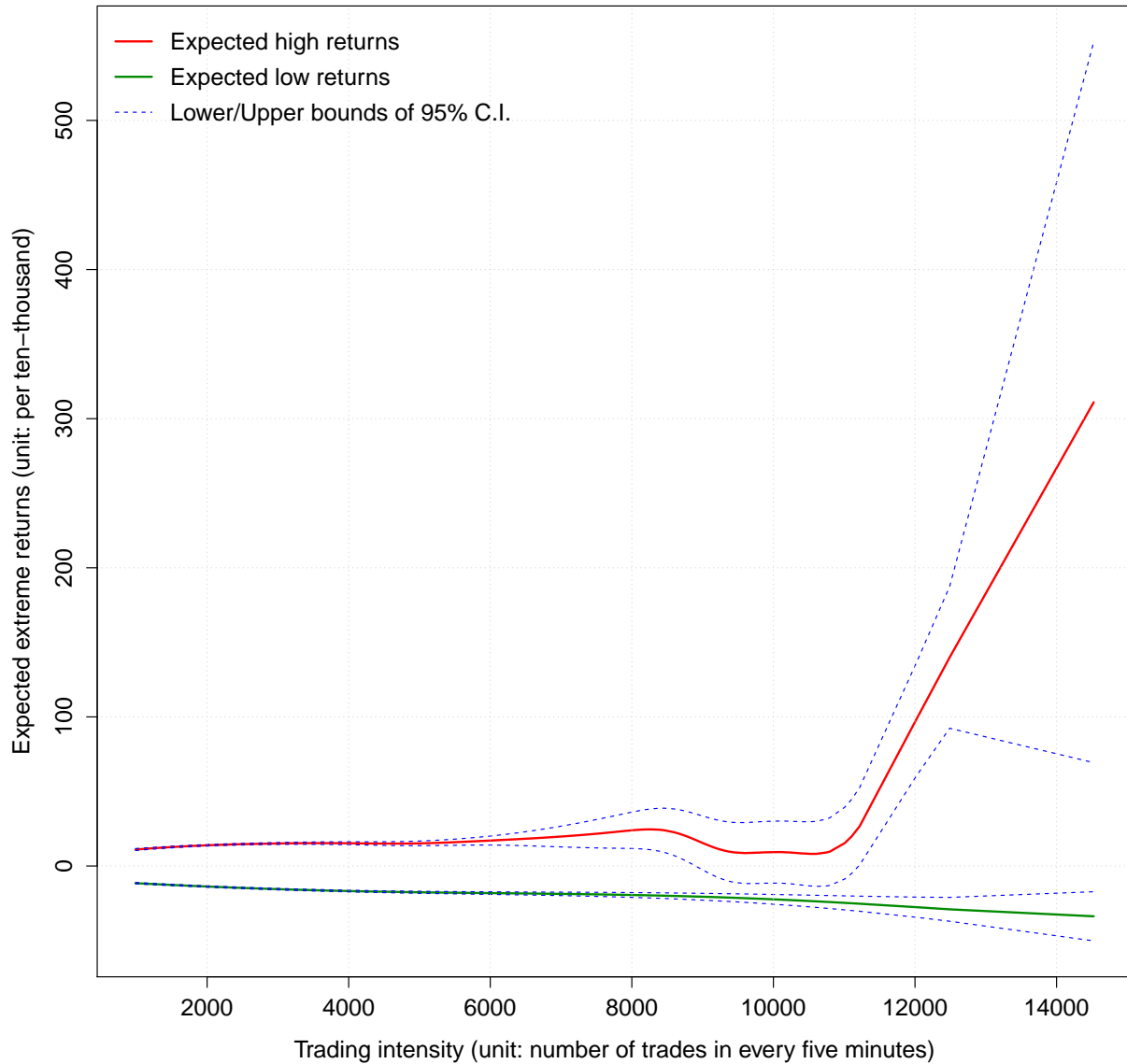


Figure 9: Nonparametric function of expected extreme returns versus trading intensity from Model 1 for BAC (Red solid line and green solid line are conditional extreme returns given trading intensity and fixing the other two regressors, center and range, at their unconditional means. Blue dotted lines are upper and lower 95% confidence intervals.)

Nonparametric functions for expected extreme returns (JPM)

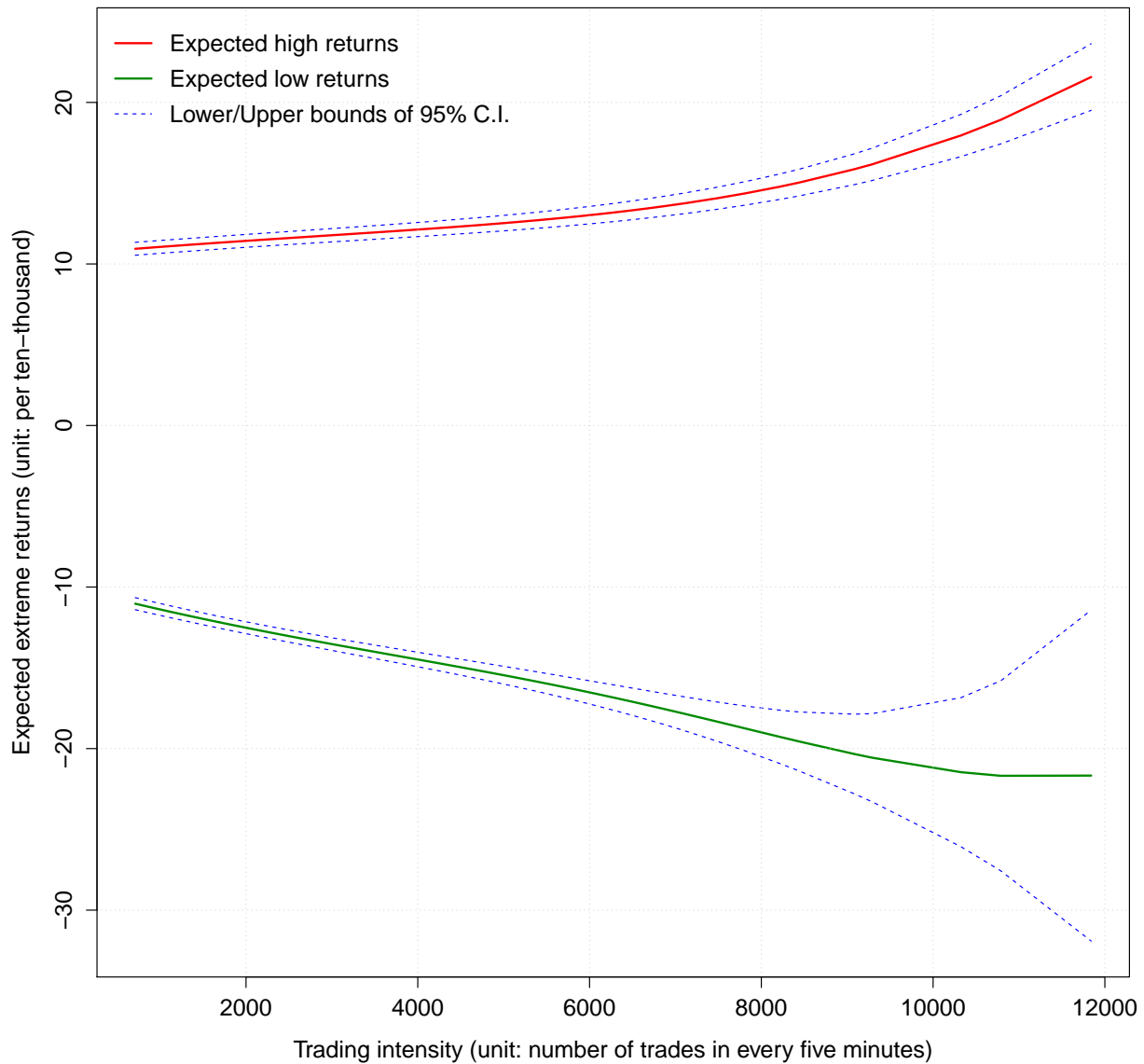


Figure 10: Nonparametric function of expected extreme returns versus trading intensity from Model 1 for JPM (Red solid line and green solid line are conditional extreme returns given trading intensity and fixing the other two regressors, center and range, at their unconditional means. Blue dotted lines are upper and lower 95% confidence intervals.)