Interval-valued Time Series Models: Estimation based on Order Statistics Exploring the Agriculture Marketing Service Data

Wei Lin[†]

[†]International School of Economics and Management

Capital University of Economics and Business

Beijing, China 100070

Gloria González-Rivera[*]

[*]Department of Economics

University of California, Riverside

Riverside, CA 92521

Revised on May 1, 2015

Abstract

The current regression models for interval-valued data ignore the extreme nature of the lower and upper bounds of intervals. We propose a new estimation approach that considers the bounds of the interval as realizations of the max/min order statistics coming from a sample of n_t random draws from the conditional density of an underlying stochastic process $\{Y_t\}$. This approach is important for data sets for which the relevant information is only available in interval format, e.g., low/high prices. We are interested in the characterization of the latent process as well as in the modeling of the bounds themselves. We estimate a dynamic model for the conditional mean and conditional variance of the latent process, which is assumed to be normally distributed, and for the conditional intensity of the discrete process $\{n_t\}$, which follows a negative binomial density function. Under these assumptions, together with the densities of order statistics, we obtain maximum likelihood estimates of the parameters of the model, which are needed to estimate the expected value of the bounds of the interval. We implement this approach with the time series of livestock prices, of which only low/high prices are recorded making the price process itself a latent process. We find that the proposed model provides an excellent fit of the intervals of low/high returns with an average coverage rate of 83%. We also offer a comparison with current models for interval-valued data.

Key Words: Interval-valued Data, Order Statistics, Intensity, Maximum Likelihood Estimation.

JEL Classification: C01, C13, C32.

1 Introduction

Since the work on symbolic data by Billard and Diday (2003, 2006), a variety of regression models have been proposed to fit interval-valued data, see the survey article by Arroyo, González-Rivera and Maté (2011) for an extensive review. A first approach proposed by Billard and Diday was to regress the centers of the intervals of the dependent variable on the centers of the intervals of the regressors. Subsequent approaches considered two separate regressions, one for the lower bound and another for the upper bound of intervals (Brito, 2007), or one regression for the center and another for the range of the interval (Lima Neto and de Carvalho, 2010). None of these approaches guarantees that the fitted values from the regressions will satisfy the natural order of an interval, i.e., by $\widehat{y}_l \leq \widehat{y}_u$, for all observations in the sample. A solution came from Lima Neto and de Carvalho (2010) who modified the previous regression models by imposing non-negative constraints on the regression coefficients of the model for the range. González-Rivera and Lin (2013) argued that these *ad hoc* constraints limit the usefulness of the model and proposed a constrained regression model that generalizes the previous regression models for lower/upper bounds or center/radius of intervals, and naturally guarantees that the proper order of the fitted intervals is satisfied.

A common thread to these approaches is that they consider the lower and upper bound as distinct stochastic processes. In this paper we propose an alternative approach and argue that there is only one stochastic process, say $\{Y_t\}$, that generates the upper and lower bounds of the interval. When we analyze interval-valued data, we only observe the bounds and these are extreme realizations of a latent random variable. This is our conceptual setup. At a fixed time t, we consider a random variable Y_t with a given conditional density function from which we draw randomly n_t realizations. The lower and upper bounds of the interval, i.e. (y_{lt} and y_{ut}) are the realized minimum and maximum values coming from the set of realizations associated with the n_t draws. As such, our interest moves towards the analysis of these two order statistics and their probability density functions. As an example, consider a time series of daily prices. In a given day t, from opening to closing time, there are n_t transactions, each one generating a market price. If we consider the daily number of trades as the n_t random draws, their corresponding intra daily prices are the realizations of the random variable daily price Y_t, and the highest/lowest prices are the realizations of the max/min order statistics of Y_t. Observe that we are not interested in the dynamics of the intra daily prices; only the lowest/highest prices carry information on the daily market activity. To start the modeling exercise, we require a set of assumptions regarding the density of the underlying stochastic process and the density of the number of draws. We will assume that the first process is continuous and it follows a conditional normal density function, and that the second process is naturally discrete and it follows a negative binomial density. Under these assumptions, we will obtain the expected values of the lower and upper bounds of the interval.

Furthermore, this modeling approach will also provide information on the latent process because we will be able to model its conditional mean and conditional variance. This is an advantage in those instances in which there are no records of opening or closing prices, typically the object of analysis, or when those prices are not very representative of the state of the market. In this paper, we model such a time series: agricultural and livestock prices provided by the US Department of Agriculture. We look into beef sales prices; the daily information provided is low price, high price, weighted average price, number of trades and total pounds traded. We could model the weighted average price but this is not very informative for potential sellers and buyers. Instead, we construct the daily interval-valued time series of low/high beef prices, which we manually dig from several archives provided by the US Department of Agriculture, and implement our approach to discover the characteristics of the latent price as well as the expected values of the low and high prices.

The paper is organized as follows. In Section 2, we discuss the key ideas of our modeling approach and its implementation under a set of assumptions. In Section 3, we use Monte Carlo simulation to investigate the properties of the proposed maximum likelihood estimator. In Section 4, we model the dynamics of the daily beef sales and prices. In Section 5, we compare our proposed model with some existing approaches on modeling interval-valued time series using both simulated and real data. Finally, in Section 6, we conclude by summarizing our findings.

2 General Framework

We assume that there is an underlying stochastic process for the interval-valued time series, and in a given time t, e.g. day, month, etc. the high/low values of intervals are the realized highest and lowest order statistics based on the random draws from the conditional densities of the underlying stochastic process. Formally,

Assumption 1. *(DGP) Let $\{Y_t : t = 1, \cdots, T\}$ be the underlying stochastic process. The latent random variable Y_t at time t has a conditional probability density function $f(y_t|\mathfrak{F}_t)$. At each time t, from the conditional density of Y_t we draw n_t observations. The number of draws has a discrete density function $h(n_t|\mathfrak{F}_t)$. Let y_{lt} and y_{ut} be the smallest and largest value of the* *random sample* $\mathcal{S}_t \equiv \{y_{it} : i = 1, 2, \cdots, n_t\}$*:*

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

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

Then, $\{(y_{lt}, y_{ut}, n_t) : t = 1, \cdots, T\}$ *forms the observed interval time series and number of random draws, and* $\mathfrak{F}_t \equiv \{(y_{ls}, y_{us}, n_s) : s = 1, \cdots, t-1\}$ *is the information set available at time* t*.*

At time t, the low and high observations (y_{lt} and y_{ut}) are the lowest and highest ranked order statistics of the random sample $\mathcal{S}_t$ formed by the n_t draws or trades. The joint conditional probability density of $(y_{lt},\ y_{ut})$ given n_t and information set $\mathfrak{F}_t$ is,

$$\begin{aligned} g(y_{lt}, y_{ut}|n_t, \mathfrak{F}_t) &= n_t(n_t-1)\left[F(y_{ut}|\mathfrak{F}_t) - F(y_{lt}|\mathfrak{F}_t)\right]^{n_t-2} \\ &\quad \times f(y_{lt}|\mathfrak{F}_t) f(y_{ut}|\mathfrak{F}_t) \end{aligned}$$

where $F(\cdot|\mathfrak{F}_t)$ is the cumulative distribution function corresponding to the conditional density $f(\cdot|\mathfrak{F}_t)$. Then, the joint probability density of (y_{lt}, y_{ut}, n_t) conditional on information set $\mathfrak{F}_t$ is,

$$p(y_{lt}, y_{ut}, n_t|\mathfrak{F}_t) = g(y_{lt}, y_{ut}|n_t, \mathfrak{F}_t) h(n_t|\mathfrak{F}_t).$$

We still need to specify the conditional densities $f(y_t|\mathfrak{F}_t)$ and $h(n_t|\mathfrak{F}_t)$ and their dependence on the information set. Therefore, we have Assumptions 2 and 3.

Assumption 2. *(Distributions) The conditional densities of the underlying stochastic process* $\{Y_t\}$ *and of the number of random draws* n_t *are normal and negative binomial respectively, i.e.,*

$$\begin{aligned} f(y_t|\mathfrak{F}_t) &\equiv \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left\{-\frac{(y_t-\mu_t)^2}{2\sigma_t^2}\right\}, \\ h(n_t|\mathfrak{F}_t) &\equiv \frac{\Gamma(n_t+d-2)}{(n_t-2)!\Gamma(d)} \left(\frac{d}{\lambda_t+d}\right)^d \left(\frac{\lambda_t}{\lambda_t+d}\right)^{n_t-2}, \end{aligned}$$

where μ_t *and* σ_t^2 *are the conditional mean and conditional variance of* Y_t*; and* λ_t *and* d *are the intensity function and dispersion parameter of the discrete process* n_t*.*

In the absence of any information about the underlying stochastic process, it seems sensible to assume normality as the most innocuous density. We could also rely on results provided by extreme value theory, which claim that maximal and minimal order statistics (properly normalized and centered) generated by distributions such as normal, exponential, Weibull, gamma, log-normal etc., weakly converge to the same Gumbel distribution as the number of random draws increases. Therefore, asymptotically the distribution of the order statistics from underlying normal is the same as that from many other types of underlying distributions up to some affine transformations. In this sense, the normality assumption on the underlying stochastic process that generates interval-valued time series is not too restrictive, at least *asymptotically*, i.e. as n_t gets large.

We assume negative binomial distribution for n_t as a robust alternative to the Poisson distribution because the additional dispersion parameter d will capture potential over dispersion in the data. When d goes to infinity, the negative binomial converges to Poisson.

Assumption 3. *(Dependence) The conditional mean, variance and intensity of the underlying random processes y_t and the discrete process n_t are parametric functions of the information set, i.e.,*

$$\begin{aligned}
\mu_t(\alpha) &= f_\mu(\mathbf{w}_t;\, \alpha), && (2.1)\\
\log \sigma_t^2(\beta) &= f_\sigma(\mathbf{w}_t;\, \beta), && (2.2)\\
\lambda_t(\gamma) &= f_\lambda(\mathbf{w}_t;\, \gamma). && (2.3)
\end{aligned}$$

The functions $f_\mu(\cdot)$, $f_\sigma(\cdot)$, and $f_\lambda(\cdot)$ represent the dependence on the information set, and the parameters α, β, and γ will be estimated. The random vector $\mathbf{w}_t$ is a subset of information set $\mathfrak{F}_t$, i.e., $\mathbf{w}_t \subset \mathfrak{F}_t$.

The information set $\mathbf{w}$ will consist of past low/high intervals and past number of draws (trades). With this information, it makes sense to model the conditional mean $\mu_t(\alpha)$ mainly as a function of the past centers of the intervals, i.e. $(y_l + y_h)/2$ and the conditional variance $\sigma_t^2(\beta)$ as a function of the past ranges of the intervals, i.e $(y_h - y_l)$. The conditional intensity λ_t will be a function of the past number of draws. It is also possible that there will be interactions among the three functions, for instance, the number of trades (volume) may influence volatility. Eventually, the final specifications will be driven by the characteristics of the data.

Let $\theta_1 \equiv (\alpha, \beta)$ and $\theta_2 \equiv (\gamma, d)$. Given Assumptions 2 and 3, the joint density of (y_{lt}, y_{ut}, n_t) can be explicitly written as,

$$\begin{aligned}
p(y_{lt}, y_{ut}, n_t|\mathbf{w}_t;\theta) &= g(y_{lt}, y_{ut}|n_t, \mathbf{w}_t;\theta_1)h(n_t|\mathbf{w}_t;\theta_2)\\
&= n_t(n_t-1)\left[\Phi\left(\frac{y_{ut}-\mu_t(\alpha)}{\sigma_t(\beta)}\right) - \Phi\left(\frac{y_{lt}-\mu_t(\alpha)}{\sigma_t(\beta)}\right)\right]^{n_t-2}\\
&\quad \times \frac{1}{\sigma_t(\beta)}\phi\left(\frac{y_{lt}-\mu_t(\alpha)}{\sigma_t(\beta)}\right)\frac{1}{\sigma_t(\beta)}\phi\left(\frac{y_{ut}-\mu_t(\alpha)}{\sigma_t(\beta)}\right)
\end{aligned}$$

$$\times \frac{\Gamma(n_t + d - 2)}{(n_t - 2)!\Gamma(d)} \left(\frac{d}{\lambda_t + d}\right)^d \left(\frac{\lambda_t}{\lambda_t + d}\right)^{n_t - 2}, \tag{2.4}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal probability density and cumulative distribution functions.

The estimation of the model proceeds by maximum likelihood (ML). For a sample $\{(y_{lt}, y_{ut}, n_t) : t = 1, 2, \cdots, T\}$, the log-likelihood function is

$$L(\theta|\mathbf{y}_u, \mathbf{y}_l, \mathbf{n}) = L_1(\theta_1|\mathbf{y}_u, \mathbf{y}_l, \mathbf{n}) + L_2(\theta_2|\mathbf{n})$$

where

$$\begin{aligned} L_1(\theta_1) &\equiv \sum_{t=1}^{T} \log g(y_{lt}, y_{ut}|n_t, \mathbf{w}_t; \theta_1) \\ &= \sum_{t=1}^{T} \log n_t + \sum_{t=1}^{T} \log(n_t - 1) - 2\sum_{t=1}^{T} \log \sigma_t(\beta) \\ &\quad + \sum_{t=1}^{T} (n_t - 2) \log \left[\Phi\left(\frac{y_{ut} - \mu_t(\alpha)}{\sigma_t(\beta)}\right) - \Phi\left(\frac{y_{lt} - \mu_t(\alpha)}{\sigma_t(\beta)}\right)\right] \\ &\quad + \sum_{t=1}^{T} \log \phi\left(\frac{y_{ut} - \mu_t(\alpha)}{\sigma_t(\beta)}\right) + \sum_{t=1}^{T} \log \phi\left(\frac{y_{lt} - \mu_t(\alpha)}{\sigma_t(\beta)}\right) \end{aligned} \tag{2.5}$$

and

$$\begin{aligned} L_2(\theta_2) &\equiv \sum_{t=1}^{T} \log h(n_t|\mathbf{w}_t; \theta_2) \\ &= \sum_{t=1}^{T} \log \Gamma(n_t + d - 2) - \sum_{t=1}^{T} \log \Gamma(d) - \sum_{t=1}^{T} \log(n_t - 2)! \\ &\quad + d\sum_{t=1}^{T} \log\left(\frac{d}{\lambda_t + d}\right) + \sum_{t=1}^{T} (n_t - 2) \log\left(\frac{\lambda_t}{\lambda_t + d}\right). \end{aligned} \tag{2.6}$$

and the ML estimator $\widehat{\theta}_{\text{ML}}$ is the solution that maximizes the log-likelihood function, i.e., $\widehat{\theta}_{\text{ML}} = \arg\max_{\Theta} L(\theta|\mathbf{y}_u, \mathbf{y}_l, \mathbf{n})$. The estimation procedure may be simplified when the parameters in their respective conditional densities are exogenous from each other, and the ML estimator $\widehat{\theta}_{\text{ML}}$ that maximizes the joint log-likelihood function $L(\theta|\mathbf{y}_u, \mathbf{y}_l, \mathbf{n})$ is equivalent to the ML estimators $\widehat{\theta}_1$ and $\widehat{\theta}_2$ that maximize $L_1(\theta_1|\mathbf{y}_u, \mathbf{y}_l, \mathbf{n})$ and $L_2(\theta_2|\mathbf{n})$ separately. When the two set of parameters are exogenous for each other, the estimation and properties of $\widehat{\theta}_2$ are standard and have been extensively studied within the standard Negative Binomial Generalized Linear

Model (GLM) with link function $\lambda_t(\gamma)$.[1]

Under general settings, the asymptotic properties of the ML estimator $\widehat{\theta}_1$ have been studied for cross-sectional data in Newey and McFadden (1994) and for dependent processes in Wooldridge (1994). Under regularity conditions stated in Theorems 5.1 and 5.2 in Wooldridge (1994), the ML estimator is weakly consistent and asymptotic normal, i.e. $\sqrt{T}(\widehat{\theta}_1 - \theta_1^*) \xrightarrow{d} N(\mathbf{0}, \mathbf{V}^{-1})$, where $\mathbf{V} \equiv -\lim_{T\to\infty} T^{-1} E(\nabla^2_{\theta_1} L_1(\theta_1)|_{\theta_1=\theta_1^*})$. We believe that a quasi-ML estimator is not feasible for the following reason. According to Gourieroux, Monfort and Trognon (1984) and White (1994), for correctly specified conditional mean and variance models and under some regularity conditions, the consistency of QML estimators is guaranteed when the density function belongs to the quadratic exponential family regardless of whether or not the true underlying density lies within that family. Though we assume that the underlying process $\{y_{it}\}_{i=1}^{n_t}$ is normal and, as such, it belongs to the quadratic exponential family, we build the log-likelihood function for the only observed data we have, which is the maximum and minimum values of a collection of draws. The function ((2.5)) contains the following term

$$\log\left[\Phi\left(\frac{y_{ut}-\mu_t(\alpha)}{\sigma_t(\beta)}\right) - \Phi\left(\frac{y_{lt}-\mu_t(\alpha)}{\sigma_t(\beta)}\right)\right],$$

which cannot be written as a quadratic function of $(y_{lt},\, y_{ut})$. Therefore, the joint density of the ordinal statistics does not belong to the quadratic exponential family.

Given the model specification and the density of the order statistics, and by calling the law of iterated expectations, we obtain the conditional means of the lower and upper bounds of interval, μ_{lt} and μ_{ut} respectively, as follows

$$\begin{aligned}
\mu_{lt} &\equiv E(y_{lt}|\mathfrak{F}_t) = E[E(y_{lt}|\mathfrak{F}_t, n_t)|\mathfrak{F}_t] \\
&= E\left[n_t \int_{-\infty}^{+\infty} s\left(1-\Phi\left(\frac{s-\mu_t(\alpha)}{\sigma_t(\beta)}\right)\right)^{n_t-1} \frac{1}{\sigma_t(\beta)}\phi\left(\frac{s-\mu_t(\alpha)}{\sigma_t(\beta)}\right) ds \middle| \mathfrak{F}_t\right] \\
&= \sum_{n_t=2}^{\infty}\left[n_t \int_{-\infty}^{+\infty} s\left(1-\Phi\left(\frac{s-\mu_t(\alpha)}{\sigma_t(\beta)}\right)\right)^{n_t-1} \frac{1}{\sigma_t(\beta)}\phi\left(\frac{s-\mu_t(\alpha)}{\sigma_t(\beta)}\right) ds \cdot h(n_t|\mathfrak{F}_t)\right]
\end{aligned}$$

and

$$\begin{aligned}
\mu_{ut} &\equiv E(y_{ut}|\mathfrak{F}_t) \\
&= \sum_{n_t=2}^{\infty}\left[n_t \int_{-\infty}^{+\infty} s\left(\Phi\left(\frac{s-\mu_t(\alpha)}{\sigma_t(\beta)}\right)\right)^{n_t-1} \frac{1}{\sigma_t(\beta)}\phi\left(\frac{s-\mu_t(\alpha)}{\sigma_t(\beta)}\right) ds \cdot h(n_t|\mathfrak{F}_t)\right] \quad (2.7)
\end{aligned}$$

[1] In R, there is a package `glm.nb` that provides estimates of $\widehat{\theta}_2$.

Similarly, the conditional variances of lower and upper bounds of the interval are as follows,

$$\begin{aligned}\sigma_{lt}^2 &\equiv E(y_{lt}^2|\mathfrak{F}_t) - E(y_{lt}|\mathfrak{F}_t)^2 \\ &= \sum_{n_t=2}^{\infty}\left[n_t\int_{-\infty}^{+\infty} s^2\left(1-\Phi\left(\frac{s-\mu_t(\alpha)}{\sigma_t(\beta)}\right)\right)^{n_t-1}\frac{1}{\sigma_t(\beta)}\phi\left(\frac{s-\mu_t(\alpha)}{\sigma_t(\beta)}\right)ds\cdot h(n_t|\mathfrak{F}_t)\right]-\mu_{lt}^2\end{aligned}$$

and

$$\begin{aligned}\sigma_{ut}^2 &\equiv E(y_{ut}^2|\mathfrak{F}_t) - E(y_{ut}|\mathfrak{F}_t)^2 \\ &= \sum_{n_t=2}^{\infty}\left[n_t\int_{-\infty}^{+\infty} s^2\left(\Phi\left(\frac{s-\mu_t(\alpha)}{\sigma_t(\beta)}\right)\right)^{n_t-1}\frac{1}{\sigma_t(\beta)}\phi\left(\frac{s-\mu_t(\alpha)}{\sigma_t(\beta)}\right)ds\cdot h(n_t|\mathfrak{F}_t)\right]-\mu_{ut}^2\end{aligned} \tag{2.8}$$

After estimation, we plug the ML estimates $\hat{\theta}$ into expressions (2.7) – (2.8) and obtain the estimates of conditional means and conditional variances of the lower and upper bounds, denoted as $\hat{y}_{lt}$, $\hat{y}_{ut}$, $\widehat{\sigma}_{lt}^2$, and $\widehat{\sigma}_{ut}^2$, which in turn permits the construction of confidence bands for the bounds. [2]

3 Simulation

We perform Monte Carlo simulations to assess the finite sample performance of the proposed maximum likelihood estimators. The data generating processes (DGP) satisfy Assumptions 1-3 and the conditional moments are specified as follows

$$\mu_t = \alpha_0 + \sum_i \alpha_{li} y_{l,t-i} + \sum_i \alpha_{hi} y_{h,t-i} + \sum_i \alpha_{ni}\log n_{t-i}, \tag{3.1}$$

$$\log\sigma_t^2 = \beta_0 + \sum_i \beta_{ri}\log(y_{h,t-i}-y_{l,t-i})^2 + \sum_i \beta_{ni}\log n_{t-i}, \tag{3.2}$$

$$\log\lambda_t = \gamma_0 + \sum_i \gamma_{ri}\log(y_{h,t-i}-y_{l,t-i})^2 + \sum_i \gamma_{ni}\log n_{t-i}. \tag{3.3}$$

In Table 1 we summarize all the DGPs specifications. We consider three specifications for each conditional moment, mean, variance, and intensity. DGP1 has no dependence in y_t or in n_t. DGP2 has higher persistence than DGP3 but with less number of lagged regressors.

[2] For the infinite integral, we use the built-in R function `integrate` for the one-dimensional infinite integrals.The function performs a globally adaptive interval subdivision in connection with extrapolation by Wynn's Epsilon algorithm and using the Gauss-Kronrod quadrature as a basic step. The infinite sums are approximated by finite summations. In the summation, we let n_t take value from $\max(2, \lfloor\lambda_t - 5\nu_t\rfloor)$ to $\lceil\lambda_t + 5\nu_t\rceil$ where λ_t and $\nu_t^2 = \lambda_t + \lambda_t^2/d$ are the conditional intensity (mean) and variance of random variable n_t respectively. According to Chebyshev's inequality, the interval $[\max(2, \lfloor\lambda_t - 5\nu_t\rfloor), \lceil\lambda_t + 5\nu_t\rceil]$ provides a coverage of *at least* 96% of the outcomes of n_t.

The simulated data, $\{(y_{lt},\, y_{ut}, n_t)\}_{t=1}^{T}$, for each DGP are obtained by the following procedure:

1. Choose some arbitrary initial past values to $\{y_{l,0},\, y_{l,-1}, \cdots,\, y_{l,-L}\}$, $\{y_{u,0},\, y_{u,-1}, \cdots,\, y_{u,-L}\}$, and $\{n_0,\, n_{-1}, \cdots,\, n_{-L}\}$, such that L is larger than the maximal lag in the autoregressive specifications of the conditional mean, variance, and intensity.
2. At time $t = 1$, using the DGP specifications and past values of the variables, generate the conditional mean μ_1, conditional variance σ_1^2, and conditional intensity λ_1.
3. Generate a random number n_1' from the negative binomial distribution $NB(\lambda_1,\, d)$, and use $n_1 = n_1' + 2$ as the simulated number of random draws.
4. Generate n_1 random numbers from the normal distribution $N(\mu_1,\, \sigma_1^2)$. The simulated interval is obtained as

$$(y_{l,1},\, y_{h,1}) \equiv \left(\min_{1\le j\le n_1} \{y_{j1}\},\ \max_{1\le j\le n_1} \{y_{j1}\} \right),$$

5. Repeat steps 2 to 4 to obtain a sample size of $T + 100$ times, and discard the first 100 simulated observations as the "burn-in" period.

For each DGP, we consider both small ($T = 200$) and large ($T = 2,000$) sample sizes.[3] Finally, we replicate each DGP 5,000 times.,

[TABLE 1] [TABLE 2]

In Table 2, we report the average of the absolute bias of the ML estimates and their average mean squared error (AMSE). For all DGPs, the average absolute bias is very close to zero even for small samples, and it goes to zero when the sample size increases from 200 to 2,000. As expected, the AMSE decreases with sample size. We observe that for DGP2 and DGP3, the estimates of the constant terms i.e., α_0, β_0, and γ_0, are much less accurate (large AMSEs) than those of the slope coefficients, in particular for small samples.

We also test the normality of the ML estimates (with 5000 replications) by implementing the Anderson-Darling (AD) test and the popular Jarque-Bera test (JB). The AD test asseses the distance between the empirical distribution function $F_n(x)$ and the hypothesized distribution

[3] When generating the sample, we produce additional 100 observations and use only the last 200 or 2,000 observations as the effective sample for estimation.

function $F_0(x)$ under the null hypothesis, i.e.,

$$AD = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F_0(x))^2}{[F_0(x)(1 - F_0(x))]} dx.$$

The asymptotic distribution of the AD statistic is nonstandard and its critical values or p-values are available in Pearson and Hartley (1972, Table 54). Stephens (1974) found that the AD test is a powerful statistic to detect most departures from normality.

[TABLE 3]

In Table 3 we report both the AD and JB statistics with their p-values. Those statistics with p-values less than 1% are written in bold. For DGP1 both statistics fail to reject normality for all estimates in small and large samples except for the dispersion estimate in small samples. For DGP2 and DGP3, there is rejection of normality for a few estimates when the sample size is small but, as the sample becomes large, both statistics fail to reject normality. Overall, AD and JB reach the same decision. These results are in agreement with the asymptotic properties of the ML estimators discussed above.

4 Modeling Interval-valued Beef Prices

The Agriculture Marketing Service (AMS) within the United States Department of Agriculture (USDA) provides current, unbiased price and sales information to assist in the orderly marketing and distribution of farm commodities. Its daily market news reports include information on prices, volume, quality, condition, and other market data on farm products in specific markets and marketing areas. Reports cover both domestic and international markets.

4.1 Description of the Data

The specific data set that we analyze is the national daily boxed beef cuts negotiated sales prices. The historical data are archived and it is downloaded from the following website:[4] `http://goo.gl/76WYQ`. The name of the daily beef report is "Boxed Beef Cutout & Cuts-Negotiated Sales PM CSV", coded as "LM_XB403". In the daily report, sales prices of different parts (choice cuts) of beef are provided, and we select item "109E". The available information includes number of trades, total pounds, low price, high price, and weighted average price, where prices reflect U.S. dollars per 100 pounds.

[4] We shorten the original url by Google url shortener.

There is not special reason to work with this particular commodity. Our interest is to show that there are time series for which the relevant information is not given by a single number, in this case, a daily price. The weighted price is not very informative for potential sellers and buyers as it includes sales volume and it is not representative of any given transaction. If we are interested in the forecasting of beef prices, we are bound to work with the low and high prices, and thus the importance of modeling the interval. To this end, we apply the estimation methodology proposed in the previous sections. There are other areas within economics and other sciences where the relevant and only format of the data is the interval format. For instance, electricity prices, fair market price within real estate markets, bid/ask prices in the bond markets, low/high temperature records in a given location, blood pressure measurements in health records, etc. all these are examples of data sets where a point-valued format does not exist because in most cases it would be meaningless.

The AMS archives are very data rich and, as we did with beef prices, sales data on any other livestock or farm products can be retrieved, though it requires a non-trivial manual effort. We construct a daily time series that ranges form January 4th, 2010 to September 30th 2013 for a total of 950 observations. In Figures 1a and 1b, we plot the time series of daily low/high prices and number of trades, respectively.

[FIGURE 1a] [FIGURE 1b]

The prices are nonstationary, they have an upward trend, thus we need to work with returns. To preserve the interval format, we calculate daily returns with respect to the previous day weighted average price, that is,

$$
\begin{aligned}
r_{ht} &= \frac{P_{high,t} - P_{avg,t-1}}{P_{avg,t-1}} \times 100\% \\
r_{lt} &= \frac{P_{low,t} - P_{avg,t-1}}{P_{avg,t-1}} \times 100\%.
\end{aligned}
$$

In Figure 1c, we plot the low and high returns. In Table 4, we report the descriptive statistics for the three series: low percentage change r_{lt}, high percentage change r_{ht} and number of trades n_t. The median low return is -3.41% and the median high return is 5.83%. Both series have similar standard deviation of about 3.5%, and both are skewed and leptokurtic. The median number of daily trades is 30. We observe that over dispersion is present, i.e. the variance 166.92 is much larger than the mean 31.54, so that the assumed negative binomial distribution for number of trades is a plausible assumption.

[TABLE 4]

In Figures 2 and 3, we plot the autocorrelation and partial autocorrelation functions for each of the three times series. Both returns series, r_{lt} and r_{ht}, exhibit moderate but significant autocorrelation, in contrast with the zero autocorrelation customarily found in financial returns. The dependence observed in the series $\log n_t$ is predominantly weekly seasonality (5-day week).

[FIGURE 2] [FIGURE 3]

4.2 Model Selection and Estimation

Taking into account the information provided in the previous section, we propose a specification search going from a rather general to a more parsimonious model. The general model involves an autoregressive representation for the conditional mean, conditional variance, and conditional intensity. The starting most general specification is as follows,

$$\begin{aligned}
\mu_t &= \alpha_0 + \sum_{\xi=1}^{p_1} \alpha_{l\xi} r_{l,t-\xi} + \sum_{\xi=1}^{p_1} \alpha_{h\xi} r_{h,t-\xi} \\
&\quad + s_{1\mu} \log n_{t-5} + s_{2\mu} \log n_{t-10} + s_{3\mu} \log n_{t-15} \\
&\quad + \left(1 + s_{1\mu} L^5 + s_{2\mu} L^{10} + s_{3\mu} L^{15}\right) \sum_{\xi=1}^{p_2} \alpha_{n\xi} \log n_{t-\xi} \qquad (4.1) \\
\log \sigma_t^2 &= \beta_0 + \sum_{\xi=1}^{q_1} \beta_{r\xi} \log(r_{h,t-\xi} - r_{l,t-\xi})^2 \\
&\quad + s_{1\sigma} \log n_{t-5} + s_{2\sigma} \log n_{t-10} + s_{3\sigma} \log n_{t-15} \\
&\quad + \left(1 + s_{1\sigma} L^5 + s_{2\sigma} L^{10} + s_{3\sigma} L^{15}\right) \sum_{\xi=1}^{q_2} \beta_{n\xi} \log n_{t-\xi} \qquad (4.2) \\
\log \lambda_t &= \gamma_0 + \sum_{\xi=1}^{k_1} \gamma_{r\xi} \log(r_{h,t-\xi} - r_{l,t-\xi})^2 \\
&\quad + s_{1\lambda} \log n_{t-5} + s_{2\lambda} \log n_{t-10} + s_{3\lambda} \log n_{t-15} \\
&\quad + \left(1 + s_{1\lambda} L^5 + s_{2\lambda} L^{10} + s_{3\lambda} L^{15}\right) \sum_{\xi=1}^{k_2} \gamma_{n\xi} \log n_{t-\xi}, \qquad (4.3)
\end{aligned}$$

According to the information contained in the autocorrelation functions, the number of trades exhibits marked weekly seasonality of order 2 or 3 [5]. As a starting point, we choose order 3 and we include the non-seasonal and seasonal component (in a multiplicative fashion) of the number of trades as regressors in the conditional mean and variance. In the conditional mean, we also include the past centers of the interval in an unrestricted format, i.e. $r_{l,t-\xi}$ and $r_{h,t-\xi}$ separately. The conditional variance and conditional intensity are assumed to be functions

[5] Since we consider a 5-day week, autoregressive seasonality of order 1, 2, and 3 means a 5-period, 10-period and 15-period lag, respectively i.e., L^5, L^{10}, L^{15}.

of the past ranges of the intervals. Note that the range is also an estimator of volatility, so that, a priori, we claim that past volatility may affect the intensity of trades, and vice versa, i.e. information of past number of trades may affect volatility. The specification of the conditional variance and conditional intensity are in logs to avoid imposing positivity restrictions in the parameters of such equations. Our task is to find the order of the several polynomial lags in the three equations.

We estimate jointly by maximum likelihood the mean and variance equations, (4.1) and (4.2). The order of the polynomial lags are restricted to the following large set

$$p_1, p_2, q_1, q_2 \in \mathcal{A} \equiv \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\},$$

The preferred model is selected by minimizing the Bayesian Information Criterion (BIC). The optimal orders are $p_1 = 7$, $p_2 = 9$, $q_1 = 5$, and $q_2 = 1$. Similarly, we estimate by MLE the intensity equation (4.3) with lag orders (k_1, k_2) to be chosen from the following set

$$k_1, k_2 \in \mathcal{B} \equiv \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

By minimizing the BIC, the optimal orders are $k_1 = 0$ and $k_2 = 4$, which seems to indicate that there is no volatility effect on the intensity of trading. We provide the estimation results of the three equations in Tables 5, 6, and 7.

[TABLE 5] [TABLE 6] [TABLE 7]

We implement a stationary block bootstrap procedure to obtain the standard errors of the ML estimates, which are reported in the third column of each table. We also report the 95% bootstrapped confidence intervals in the fourth and fifth columns of each table, and test whether the ML estimates are statistically significant. The ML estimate that are statistically significant at the 5% significance level are written in bold.

For the conditional mean equation, we observe that neither the seasonal nor the non-seasonal components of trading activity have any effect in the mean. The relevant dynamics are short and mostly confined to the first lag of the lower and upper bounds of past intervals. Given this result, we will entertain additional more parsimonious specifications with $p_1 = 5$ (SPEC2), and $p_1 = 1$ (SPEC3). For the conditional variance, the dynamics of the past ranges are important and to a lesser extent, past trading activity, which is negatively correlated with volatility. For the conditional intensity equation, only the seasonal component is most relevant.

4.3 Performance Evaluation

We compare three specifications: SPEC1 chosen by BIC ($p_1 = 7$, $p_2 = 9$, $q_1 = 5$, $q_2 = 1$, $k_1 = 0$, and $k_2 = 4$) and two more parsimonious models, SPEC2 ($p_1 = 5$, $p_2 = 5$, $q_1 = 5$, $q_2 = 1$, $k_1 = 0$, and $k_2 = 4$), and SPEC3 ($p_1 = 1$, $p_2 = 1$, $q_1 = 5$, $q_2 = 1$, $k_1 = 0$, and $k_2 = 4$).

First, we check the autocorrelation of the residuals of the three models. Since the true conditional mean and variance of the latent stochastic process r_t are unobservable, we use center values $c_t = (r_{lt} + r_{ht})/2$ and squared ranges $(r_{h,t} - r_{l,t})^2$ as proxies for realized conditional mean $\hat{\mu}_t$ and variance $\hat{\sigma}_t^2$. Then, we construct the following two pseudo-Pearson residuals

$$\hat{\varepsilon}_{1t} = \frac{c_t - \widehat{\mu}_t}{\widehat{\sigma}_t}, \qquad \hat{\varepsilon}_{2t} = \frac{(r_{h,t} - r_{l,t})^2}{\widehat{\sigma}_t^2}$$

and check whether these “standardized residuals” are uncorrelated. In Table 8, we report the p-values of Ljung-Box tests for the pseudo-Pearson residuals.

[TABLE 8]

For SPEC1 and SPEC2, the residuals $\hat{\varepsilon}_{1t}$ do not show any autocorrelation but for SPEC3, we reject the null hypothesis of no autocorrelation at any significance level. We conclude that the conditional mean return has significant dynamics at least of order 5. This is very different from financial returns which are basically a white noise. For SPEC1 and SPEC2, the residuals $\hat{\varepsilon}_{2t}$ show a bit of autocorrelation at the lower lags 1 and 2 but the magnitude of the autocorrelation coefficients is very small, much less than 0.1. On the contrary, SPEC3 generates autocorrelated residuals $\hat{\varepsilon}_{2t}$ at any lags. In summary, SPEC1 and SPEC2 have very similar performance in capturing the dynamics of center values and squared ranges and they are preferred to SPEC3.

Secondly, we evaluate the performance of the three models by computing several measures of fit for interval-valued data. 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) Root Mean Squared Error (RMSE) for upper and lower bounds separately.
$RMSE_l = \sqrt{\sum_{t=1}^{T} (\hat{y}_{lt} - y_{lt})^2 / T}$ and $RMSE_u = \sqrt{\sum_{t=1}^{T} (\hat{y}_{ut} - y_{ut})^2 / T}$;

(ii) Coverage Rate (CR) and Efficiency Rate (ER) of the estimated intervals (Rodrigues and Salish, 2014).
$CR = \frac{1}{T} \sum_{t=1}^{T} w(y_t \cap \hat{y}_t)/w(y_t)$, $ER = \frac{1}{T} \sum_{t=1}^{T} w(y_t \cap \hat{y}_t)/w(\hat{y}_t)$, where $y_t \cap \hat{y}_t$ is the intersection of actual and fitted intervals, and $w(\cdot)$ is the width of the interval. The coverage rate (CR) is the average proportion of the actual interval covered by the fitted

interval, and the efficiency rate (ER) is the average proportion of the fitted interval covered by the actual interval. Both rates are between zero and one and a large rate means a better fit. Given an actual interval, a wide fitted interval implies a large coverage rate but a low efficiency rate, on the contrary, a tight fitted interval implies a low coverage rate but a high efficiency rate. Therefore, we take into account the potential trade-off between the two rates by calculating an average of the two, i.e., $(CR + ER)/2$.

(iii) Multivariate Loss Functions (MLF) for the vector of lower and upper bounds (Komunjer and Owyang, 2011).
We implement the following multivariate loss function $L_p(\tau, \mathbf{e}) \equiv (\parallel \mathbf{e} \parallel_p + \tau'\mathbf{e}) \parallel \mathbf{e} \parallel_p^{p-1}$ where $\| \cdot \|_p$ is the l_p-norm, τ is two-dimensional parameter vector bounded by the unit ball $\mathcal{B}_q$ in $\mathbb{R}^2$ with l_q-norm (where p and q satisfy $1/p+1/q = 1$), 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 τ parameter vectors within the unit balls $\mathcal{B}_\infty$ and $\mathcal{B}_2$ respectively, $MLF_1 = \int_{\tau \in \mathcal{B}_\infty} (|e_l| + |e_u| + \tau_1 e_l + \tau_2 e_u) d\tau$, $MLF_2 = \int_{\tau \in \mathcal{B}_2} \left[e_l^2 + e_u^2 + (\tau_1 e_l + \tau_2 e_u)(e_l^2 + e_u^2)^{1/2} \right] d\tau$.

(iv) 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 $MDE^q(\{\hat{y}_t\}, \{y_t\}) = [\sum_{t=1}^{T} D^q(\hat{y}_t, y_t)/T]^{1/q}$. We consider $q = 1$ and $q = 2$, with a distance measure such as $D_(\hat{y}_t, y_t) = \frac{1}{\sqrt{2}}[(\hat{y}_{lt} - y_{lt})^2 + (\hat{y}_{ut} - y_{ut})^2]^{1/2}$.

The evaluation results for the three specifications considered are reported in Table 9.

[TABLE 9]

Across the four measures, SPEC1 offers the best fit though it is only marginally better than SPEC2. Both specifications dominate SPEC3. In summary, SPEC2 ($p_1 = 5$, $p_2 = 5$, $q_1 = 5$, $q_2 = 1$, $k_1 = 0$, and $k_2 = 4$) provides a more parsimonious specification than SPEC1 without sacrificing a good data fit. The following figures offer a visual aspect of the fitting of the data. Based on the estimation results, we calculate the estimated conditional mean $\widehat{\mu}_t$, conditional variance $\widehat{\sigma}_t^2$ and conditional intensity $\widehat{\lambda}_t$. In addition, we calculate the estimated expected low and high returns ($\widehat{r}_{lt}$ and $\widehat{r}_{ht}$) and their corresponding estimated conditional variances ($\widehat{\sigma}_{lt}^2$ and $\widehat{\sigma}_{ht}^2$) according to expressions (2.7)- (2.8). The time series of all these estimates and of the actual values are plotted in Figure 4.

[FIGURE 4]

In Figure 4a, we observe that $\widehat{\mu}_t$ lies very much around the center of the actual intervals. The estimated expected low and high returns also follow very closely the profiles of the actual low and high returns. According to their RMSEs, we find a better fit for the upper bounds than for the lower bounds across the three specifications considered. The average coverage rate is about 83% and the average efficiency rate about 76%, which is a very good fitting. In Figure 4b, we plot the estimated variance, which shows that there is substantial heteroskedasticity in the series. The three most prominent bursts of volatility corresponds to those instances where the ranges of the intervals are the widest. In Figure 4c, we plot the actual number of trades n_t and the estimated intensity $\widehat{\lambda}_t$. Although these two series are not directly comparable, we see that $\widehat{\lambda}_t$ as a measure of the expected number of trades follows very closely the actual number.

5 Comparison with Existing Approaches

We compare our proposed approach with most current models for interval-valued time series. Specifically, we consider the following six competing models:

- CR: Center-Radius model (Lima Neto and De Carvalho, 2008)
- CCR: Constrained Center-Radius model (Lima Neto and De Carvalho, 2010)
- VAR: Vector AutoRegression model (García-Ascanio and Maté, 2009; Arroyo, González-Rivera, and Maté, 2010)
- IAR-TS: Interval AutoRegressive-Two Step model (González-Rivera and Lin, 2013)
- IAR-MTS: Interval AutoRegressive-Modified Two Step model (González-Rivera and Lin, 2013)
- STAR: Space-Time AutoRegressive model (Teles and Brito, 2013).

CR and CCR are autoregressive specifications for the center and radius of the intervals; CR is unrestricted and CCR restricts the parameters of the radius equation to be non-negative. VAR is an unrestricted autoregressive specification for the lower and upper bounds. IAR-TS and IAR-MTS is a restricted autoregressive specification where the bivariate density function of the upper and lower bound is a truncated function to preserve the natural order of the interval. STAR is a bivariate linear system of equations for the lower and upper bounds that allows contemporaneous correlation between both bounds; it resembles a structural vector autoregression with constraints.

Observe that the competing approaches only model the conditional mean of the upper and lower bounds of the interval. Our proposed model offers a richer structure as we aim to discover the underlying stochastic process from which the extreme bounds are observed. We analyze the interplay of the conditional mean, conditional variance, and conditional intensity. In this sense, the competing approaches are not strictly comparable with our approach. Nevertheless, we simulate interval-valued data and we use the real data set of section 4 to estimate the competing models We can only compare the fitted values of the lower and upper bounds across specifications. Obviously, we cannot compare variance and intensity estimates.

5.1 Comparison Using Simulated Data

We generate time series data, $\{(y_{lt}, y_{ut}, n_t)\}_{t=1}^{T}$, following the procedure described in section 3. The sample size is $T = 200$ and the number of replications is 1,000. For each replication, we calculate the four measures of fit (RMSE, CR & ER, MLF, MDE) described in section 4.3. Once we finish the 1,000 replications, we calculate the average of each fit measure across replications and call them A-RMSE, A-CR & A-ER, A-MLF, and A-MDE. In Table 10, we present the comparison across models.

[TABLE 10]

Not very surprisingly, the proposed model based on order-statistics, which is correctly specified, outperforms the other six competing models across all evaluation criteria. It offers similar interval coverage, A-ER and A-CR, to the competing models but it substantially overperforms with respect to the other three measures of fit. Overall, CR or CCR are the worst performers. VAR, IAR-TS, and IAR-MTS deliver almost identical results because there are not binding constraints and, in this case, the IAR collapses to an unrestricted VAR. The STAR model performs worst than VAR-type models and marginally better than CR-type models.

5.2 Comparison Using Real Data

We fit the competing models to the interval-valued time series of beef prices that we have modeled in section 4. Our preferred model is *Spec 1* described in section 4.2. For the specification of the competing models, the BIC selects 4 lags. In Table 11, we report the results. For a given evaluation criteria, we write in bold the value corresponding to the best performer.

[TABLE 11]

Overall, there is not a best performer across all the evaluation criteria. As in the case of simulated data, CR and CCR models are the worst performers. The restrictions on the model specification (centers are only regressed on past centers and radius only on past radius) seem to be too stringent. We note that both specifications deliver the same fit across evaluation criteria. This indicates that the non-negative restriction on the parameters of the radius equation are not binding.

The rest of the specifications are very similar in performance. VAR, IAR-TS and IAR-MTS models deliver similar results across all the loss functions. The IAR models impose that the density of the error term follows a truncated bivariate normal distribution, which guarantees that the lower and upper bounds do not cross over. However, if the no-crossing-over restriction is not binding, the IAR models collapse to VARs. Across criteria, IAR-TS offers the best performance in many instances. The STAR model seems to provide better coverage than the rest. The performance of our preferred model (Spec 1) is very similar to the STAR and IAR models. Note that Spec 1 is a highly non-linear specification as it consists of the joint modeling of the conditional mean, variance, and intensity (see equation 2.7). The estimation of a much larger number of parameters may disadvantage this model compared to the competing specifications. Nevertheless, it delivers similar performance to that of more parsimonious specifications. Most importantly, it provides extra information (variance and intensity) about the characteristics of the underlying process that generates the interval bounds.

6 Conclusion

By focusing on the lower and upper bound of an interval as two different stochastic processes, the current literature on model estimation has ignored the extreme nature of such bounds. Our main contribution is a modeling approach that exploits such extreme property. At the core, we have argued that there is only one stochastic process $\{Y_t\}$ from which the lower and upper bounds of the intervals (y_{lt} and y_{ut}) are the realized extreme observations (minima and maxima) coming from the n_t random draws from the conditional density of the process. A key point is to understand that the researcher is interested in the characterization of this latent stochastic process as much as in the modeling of the bounds themselves. This question is important because there are time series data sets for which the relevant information is only available in interval format e.g. low/high prices. Yet in these instances, it is of interest to know the expected price or any other conditional moment of the price process, e.g. variance. skewness, etc. We have introduced a data set of daily beef prices and sales for which the opening or closing prices are not reported because they are not very informative to potential sellers and buyers, and consequently we are restricted to work with the interval low/high prices.

The modeling approach is based on the theory of order statistics. For implementation purposes, we need to assume a conditional density for the latent stochastic process. We have assumed normality as a first approximation but this assumption can be refined according to the researcher's needs. One advantage of the data set that we have analyzed is that contains information on the daily number of trades, so that we are able to model the conditional intensity of trading in addition to the modeling of the conditional mean and variance. The standard distributional assumption for counts is a Poisson density but we have assumed a more robust alternative, the Negative Binomial, as it takes into account potential over dispersion of the data. Given these distributions, we have estimated the model by maximum likelihood. Monte Carlo simulations indicate that the estimators are well-behaved and, in large samples, they do not show any apparent biases and they seem to be normally distributed.

The modeling of beef prices shows interesting features. The conditional mean of the corresponding returns exhibits relevant dynamics in contrast to financial returns. The return process is heteroskedastic and the dynamics of the conditional variance are driven by the range of past intervals and the past number of trades. The conditional intensity function is mainly driven by the seasonal component of the number of trades. We have also estimated the expected low and high returns to construct the fitted intervals. When these are compared with the actual intervals, we find that the model provides a very good fitting with an average coverage rate of 83% and an average efficiency rate of about 76%.

Finally, we have compared the performance of the proposed approach with other competing models. Our approach offers much more information about the underlying process that generates the interval bounds and, in this sense, it is not strictly comparable to current models. Nevertheless, by only comparing performance based on fitted values of the bounds, our model is superior to or, in some instances, as good as the IAR-type models.

References

[1] Anderson, T.W. and Darling, D.A. (1954), "A Test of Goodness-of-Fit," *Journal of the American Statistical Association*. Vol. 49, pp. 765–769.

[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] Billard, L., and Diday, E. (2003), "From the statistics of data to the statistics of knowledge: symbolic data analysis," *Journal of the American Statistical Association*. Vol. 98, pp. 470–487.

[4] Billard, L., and Diday E. (2006), *Symbolic Data Analysis: Conceptual Statistics and Data Mining*, 1st ed., Wiley and Sons, Chichester.

[5] Brito, P. (2007), "Modelling and analysing interval data," *Proceedings of the 30th Annual Conference of GfKl*. Springer, Berlin, pp. 197–208.

[6] García-Ascanio, C. and Maté, C. (2009). "Electric power demand forecasting using interval time series: A comparison between VAR and iMLP," *Energy Policy,* 38: 715–725.

[7] 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.

[8] Gourieroux, C., A. Monfort, and A. Trognon (1984). "Pseudo Maximum Likelihood Methods: Theory," *Econometrica*,Vol. 52, No. 3, pp. 681–700.

[9] Jarque, Carlos M.; Bera, Anil K. (1980), "Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals," *Economics Letters*. Vol. 6 (3), pp. 255–259.

[10] Komunjer, I., and Owyang, M. (2012), "Multivariate Forecast Evaluation and Rationality Testing," *Review of Economics and Statistics*. Vol. 94, No. 4, pp. 1066–1080.

[11] Lima Neto, E., and de Carvalho, F. (2008), "Centre and Range Method for Fitting a Linear Regression Model to Symbolic Interval Data," *Computational Statistics and Data Analysis*. Vol. 52, pp. 1500–1515.

[12] Lima Neto, E., and de Carvalho, F. (2010), "Constrained linear regression models for symbolic interval-valued variables," *Computational Statistics and Data Analysis*. Vol. 54, pp. 333–347.

[13] Newey, W., and MacFadden, D. (1994), "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics, Vol. IV, R. Engle and D. McFadden, (eds.)*. North Holland, Amsterdam, pp. 2111–2245.

[14] Rodrigues, P. M. M. and Salish, N. (2015), "Modeling and forecasting interval time series with threshold models," *Advances in Data Analysis and Classification*. Vol. 9, No. 1, pp. 41–57.

[15] Stephens, M. A. (1974), "EDF Statistics for Goodness of Fit and Some Comparisons" *Journal of the American Statistical Association,* Vol 69, pp. 730–737.

[16] Teles, P. and Brito, P. (2013). "Modeling interval time series with spacetime processes," *Communications in Statistics: Theory and Methods*. (in press) DOI: 10.1080/03610926.2013.782200

[17] White, H. (1994). *Estimation, Inference and Specification Analysis.* Cambridge University Press, Cambridge.

[18] Wooldridge, Jeffrey M., (1986). "Estimation and inference for dependent processes," in *Handbook of Econometrics, R. F. Engle & D. McFadden (ed.)* Edition 1, Vol 4, Chapter 45, pp. 2639–2738 Elsevier.

Table 1: Specification of Data Generating Processes

Parameters	DGP1	DGP2	DGP3
α_0	1	1	1
α_{l1}	–	0.6	0.2
α_{h1}	–	0.6	0.2
α_{l2}	–	-0.3	-0.2
α_{h2}	–	-0.3	-0.2
α_{l3}	–	–	0.1
α_{h3}	–	–	0.1
α_{l4}	–	–	-0.1
α_{h4}	–	–	-0.1
α_{l5}	–	–	0.1
α_{h5}	–	–	0.1
α_{n1}	–	0.6	0.2
α_{n2}	–	-0.3	-0.2
α_{n3}	–	–	0.1
α_{n4}	–	–	-0.1
α_{n5}	–	–	0.1
β_0	1	1	1
β_{r1}	–	0.6	0.2
β_{r2}	–	-0.3	-0.2
β_{r3}	–	–	0.1
β_{r4}	–	–	-0.1
β_{r5}	–	–	0.1
β_{n1}	–	0.6	0.2
β_{n2}	–	-0.3	-0.2
β_{n3}	–	–	0.1
β_{n4}	–	–	-0.1
β_{n5}	–	–	0.1
γ_0	5	1	1
γ_{r1}	–	0.6	0.2
γ_{r2}	–	-0.3	-0.2
γ_{r3}	–	–	0.1
γ_{r4}	–	–	-0.1
γ_{r5}	–	–	0.1
γ_{n1}	–	0.6	0.2
γ_{n2}	–	-0.3	-0.2
γ_{n3}	–	–	0.1
γ_{n4}	–	–	-0.1
γ_{n5}	–	–	0.1
$1/d$	0.1	0.1	0.1
Sample Size	200/2,000	200/2,000	200/2,000
Replications	5,000	5,000	5,000

Table 2: Simulation Results

Parameters	DGP1		DGP1		DGP2		DGP2		DGP3		DGP3	
	$\lvert\overline{bias}\rvert$ $(\times 10^{-4})$	$amse$ $(\times 10^{-4})$	$\lvert\overline{bias}\rvert$ $(\times 10^{-4})$	$amse$ $(\times 10^{-4})$	$\lvert\overline{bias}\rvert$	$amse$	$\lvert\overline{bias}\rvert$	$amse$	$\lvert\overline{bias}\rvert$	$amse$	$\lvert\overline{bias}\rvert$	$amse$
α_0	6.2033	9.8389	0.2491	0.9565	0.0856	10.4447	0.0146	0.9191	0.0463	0.5128	0.0020	0.0438
α_{l1}	–	–	–	–	0.0054	0.0013	0.0005	0.0001	0.0054	0.0029	0.0005	0.0002
α_{h1}	–	–	–	–	0.0066	0.0013	0.0009	0.0001	0.0047	0.0028	0.0003	0.0002
α_{l2}	–	–	–	–	0.0030	0.0011	0.0003	0.0001	0.0010	0.0029	0.0002	0.0003
α_{h2}	–	–	–	–	0.0031	0.0011	0.0004	0.0001	0.0002	0.0028	0.0002	0.0003
α_{l3}	–	–	–	–	–	–	–	–	0.0062	0.0031	0.0006	0.0003
α_{h3}	–	–	–	–	–	–	–	–	0.0046	0.0032	0.0007	0.0003
α_{l4}	–	–	–	–	–	–	–	–	0.0019	0.0028	0.0001	0.0003
α_{h4}	–	–	–	–	–	–	–	–	0.0003	0.0027	0.0000	0.0003
α_{l5}	–	–	–	–	–	–	–	–	0.0077	0.0029	0.0006	0.0003
α_{h5}	–	–	–	–	–	–	–	–	0.0069	0.0027	0.0011	0.0002
α_{n1}	–	–	–	–	0.0083	0.4796	0.0029	0.0437	0.0028	0.0432	0.0001	0.0038
α_{n2}	–	–	–	–	0.0245	0.6436	0.0022	0.0590	0.0042	0.0441	0.0005	0.0039
α_{n3}	–	–	–	–	–	–	–	–	0.0040	0.0454	0.0007	0.0041
α_{n4}	–	–	–	–	–	–	–	–	0.0008	0.0441	0.0004	0.0040
α_{n5}	–	–	–	–	–	–	–	–	0.0031	0.0435	0.0009	0.0040
β_0	4.2093	2.3476	1.3849	0.2298	0.0482	0.0715	0.0080	0.0068	0.0275	0.1841	0.0072	0.0157
β_{r1}	–	–	–	–	0.0090	0.0032	0.0014	0.0003	0.0018	0.0031	0.0001	0.0003
β_{r2}	–	–	–	–	0.0044	0.0021	0.0006	0.0002	0.0062	0.0027	0.0006	0.0002
β_{r3}	–	–	–	–	–	–	–	–	0.0020	0.0030	0.0004	0.0003
β_{r4}	–	–	–	–	–	–	–	–	0.0044	0.0029	0.0008	0.0003
β_{r5}	–	–	–	–	–	–	–	–	0.0007	0.0029	0.0001	0.0003
β_{n1}	–	–	–	–	0.0042	0.0026	0.0008	0.0003	0.0080	0.0164	0.0022	0.0016
β_{n2}	–	–	–	–	0.0088	0.0036	0.0010	0.0003	0.0003	0.0166	0.0000	0.0015
β_{n3}	–	–	–	–	–	–	–	–	0.0003	0.0170	0.0011	0.0015
β_{n4}	–	–	–	–	–	–	–	–	0.0053	0.0160	0.0007	0.0014
β_{n5}	–	–	–	–	–	–	–	–	0.0008	0.0162	0.0003	0.0015
γ_0	2.5165	5.2969	1.5468	0.5265	0.0057	0.1536	0.0016	0.0141	0.0722	0.1567	0.0067	0.0143
γ_{r1}	–	–	–	–	0.0055	0.0071	0.0011	0.0007	0.0000	0.0027	0.0002	0.0003
γ_{r2}	–	–	–	–	0.0081	0.0046	0.0010	0.0005	0.0025	0.0022	0.0003	0.0002
γ_{r3}	–	–	–	–	–	–	–	–	0.0008	0.0025	0.0000	0.0002
γ_{r4}	–	–	–	–	–	–	–	–	0.0021	0.0024	0.0004	0.0002
γ_{r5}	–	–	–	–	–	–	–	–	0.0000	0.0025	0.0000	0.0002
γ_{n1}	–	–	–	–	0.0122	0.0059	0.0018	0.0006	0.0128	0.0138	0.0009	0.0013
γ_{n2}	–	–	–	–	0.0098	0.0076	0.0015	0.0008	0.0080	0.0135	0.0018	0.0013
γ_{n3}	–	–	–	–	–	–	–	–	0.0081	0.0139	0.0008	0.0013
γ_{n4}	–	–	–	–	–	–	–	–	0.0108	0.0128	0.0009	0.0013
γ_{n5}	–	–	–	–	–	–	–	–	0.0121	0.0135	0.0014	0.0013
$1/d$	4.1933	1.0871	0.6572	0.1130	0.0026	0.0001	0.0003	0.0000	0.0186	0.0013	0.0019	0.0001
Sample Size	200		2,000		200		2,000		200		2,000	
Replications	5,000		5,000		5,000		5,000		5,000		5,000	

Table 3: Anderson-Darling(AD) and Jarque-Bera(JB) Normality Tests†

	DGP1(S)		DGP1(L)		DGP2(S)		DGP2(L)		DGP3(S)		DGP3(L)	
	AD	JB	AD	JB	AD	JB	AD	JB	AD	JB	AD	JB
α_0	0.18	0.33	0.34	0.60	1.05	**10.9**	0.45	1.92	1.11	**18.1**	0.31	1.23
	(0.92)	(0.85)	(0.50)	(0.74)	(0.01)	**(0.00)**	(0.27)	(0.38)	(0.01)	**(0.00)**	(0.55)	(0.54)
α_{l1}	–	–	–	–	**1.56**	8.03	0.62	3.24	0.38	0.54	0.31	0.39
	–	–	–	–	**(0.00)**	(0.02)	(0.11)	(0.20)	(0.40)	(0.76)	(0.56)	(0.82)
α_{h1}	–	–	–	–	**1.26**	**18.7**	0.17	0.17	0.19	1.85	0.35	0.65
	–	–	–	–	**(0.00)**	**(0.00)**	(0.93)	(0.92)	(0.90)	(0.40)	(0.47)	(0.72)
α_{l2}	–	–	–	–	0.63	3.15	**1.17**	6.15	0.80	2.57	0.28	1.66
	–	–	–	–	(0.10)	(0.21)	**(0.00)**	(0.05)	(0.04)	(0.28)	(0.64)	(0.44)
α_{h2}	–	–	–	–	**1.79**	**21.8**	0.51	4.87	0.23	0.55	0.28	0.76
	–	–	–	–	**(0.00)**	**(0.00)**	(0.20)	(0.09)	(0.80)	(0.76)	(0.66)	(0.68)
α_{l3}	–	–	–	–	–	–	–	–	0.44	3.37	0.31	2.64
	–	–	–	–	–	–	–	–	(0.28)	(0.19)	(0.55)	(0.27)
α_{h3}	–	–	–	–	–	–	–	–	0.84	7.97	0.32	2.17
	–	–	–	–	–	–	–	–	(0.03)	(0.02)	(0.52)	(0.34)
α_{l4}	–	–	–	–	–	–	–	–	0.43	1.82	0.52	2.85
	–	–	–	–	–	–	–	–	(0.31)	(0.40)	(0.18)	(0.24)
α_{h4}	–	–	–	–	–	–	–	–	0.20	1.28	0.44	1.25
	–	–	–	–	–	–	–	–	(0.89)	(0.53)	(0.29)	(0.54)
α_{l5}	–	–	–	–	–	–	–	–	0.17	0.08	0.81	3.14
	–	–	–	–	–	–	–	–	(0.93)	(0.96)	(0.04)	(0.21)
α_{h5}	–	–	–	–	–	–	–	–	0.63	**11.37**	0.19	0.63
	–	–	–	–	–	–	–	–	(0.10)	**(0.00)**	(0.90)	(0.73)
α_{n1}	–	–	–	–	0.84	1.77	0.62	3.93	0.18	0.88	0.40	4.20
	–	–	–	–	(0.03)	(0.41)	(0.11)	(0.14)	(0.92)	(0.64)	(0.36)	(0.12)
α_{n2}	–	–	–	–	0.40	3.24	0.34	0.34	0.24	6.04	0.23	1.58
	–	–	–	–	(0.36)	(0.20)	(0.49)	(0.84)	(0.77)	(0.05)	(0.81)	(0.45)
α_{n3}	–	–	–	–	–	–	–	–	0.47	1.13	0.42	0.72
	–	–	–	–	–	–	–	–	(0.24)	(0.57)	(0.33)	(0.70)
α_{n4}	–	–	–	–	–	–	–	–	0.25	0.50	0.43	1.10
	–	–	–	–	–	–	–	–	(0.76)	(0.78)	(0.31)	(0.58)
α_{n5}	–	–	–	–	–	–	–	–	0.92	4.20	0.39	0.44
	–	–	–	–	–	–	–	–	(0.02)	(0.12)	(0.38)	(0.80)
β_0	0.40	0.95	0.27	0.44	0.94	**12.5**	0.27	1.19	1.09	**20.64**	0.76	1.61
	(0.35)	(0.62)	(0.68)	(0.80)	(0.02)	**(0.00)**	(0.69)	(0.55)	(0.01)	**(0.00)**	(0.05)	(0.45)
β_{r1}	–	–	–	–	**1.54**	**15.8**	0.70	4.77	0.40	4.87	0.22	1.31
	–	–	–	–	**(0.00)**	**(0.00)**	(0.07)	(0.09)	(0.36)	(0.09)	(0.83)	(0.52)
β_{r2}	–	–	–	–	0.59	2.93	0.88	0.99	0.84	**25.65**	0.56	4.69
	–	–	–	–	(0.12)	(0.23)	(0.02)	(0.61)	(0.03)	**(0.00)**	(0.15)	(0.10)
β_{r3}	–	–	–	–	–	–	–	–	**1.25**	**19.26**	0.31	0.64
	–	–	–	–	–	–	–	–	**(0.00)**	**(0.00)**	(0.55)	(0.72)
β_{r4}	–	–	–	–	–	–	–	–	0.30	1.95	0.64	0.16
	–	–	–	–	–	–	–	–	(0.58)	(0.38)	(0.10)	(0.92)
β_{r5}	–	–	–	–	–	–	–	–	0.25	1.26	0.49	4.38
	–	–	–	–	–	–	–	–	(0.76)	(0.53)	(0.22)	(0.11)
β_{n1}	–	–	–	–	0.41	1.54	0.60	1.24	0.74	3.90	1.07	8.68
	–	–	–	–	(0.34)	(0.46)	(0.12)	(0.54)	(0.05)	(0.14)	(0.01)	(0.01)
β_{n2}	–	–	–	–	0.19	1.52	0.41	0.97	0.30	0.48	0.73	6.39
	–	–	–	–	(0.89)	(0.47)	(0.34)	(0.62)	(0.58)	(0.79)	(0.06)	(0.04)
β_{n3}	–	–	–	–	–	–	–	–	0.55	2.38	0.60	0.04
	–	–	–	–	–	–	–	–	(0.16)	(0.30)	(0.12)	(0.98)
β_{n4}	–	–	–	–	–	–	–	–	0.59	3.66	0.36	1.31
	–	–	–	–	–	–	–	–	(0.12)	(0.16)	(0.45)	(0.52)
β_{n5}	–	–	–	–	–	–	–	–	1.15	**16.77**	0.27	2.69
	–	–	–	–	–	–	–	–	(0.01)	**(0.00)**	(0.68)	(0.26)
γ_0	0.43	2.34	0.44	3.20	0.45	7.70	0.20	0.33	**5.65**	**88.98**	0.77	9.59
	(0.32)	(0.31)	(0.30)	(0.20)	(0.28)	(0.02)	(0.88)	(0.85)	**(0.00)**	**(0.00)**	(0.04)	(0.01)
γ_{r1}	–	–	–	–	0.28	6.13	0.36	0.49	0.71	8.04	0.26	0.70
	–	–	–	–	(0.65)	(0.05)	(0.45)	(0.78)	(0.06)	(0.02)	(0.71)	(0.70)
γ_{r2}	–	–	–	–	0.44	1.00	0.18	0.73	0.70	**15.76**	0.31	0.62
	–	–	–	–	(0.30)	(0.61)	(0.92)	(0.69)	(0.07)	**(0.00)**	(0.56)	(0.73)
γ_{r3}	–	–	–	–	–	–	–	–	0.60	7.54	0.31	1.83
	–	–	–	–	–	–	–	–	(0.12)	(0.02)	(0.55)	(0.40)
γ_{r4}	–	–	–	–	–	–	–	–	0.53	2.77	0.36	1.80
	–	–	–	–	–	–	–	–	(0.17)	(0.25)	(0.45)	(0.41)
γ_{r5}	–	–	–	–	–	–	–	–	0.48	0.97	0.65	9.43
	–	–	–	–	–	–	–	–	(0.23)	(0.62)	(0.09)	(0.01)
γ_{n1}	–	–	–	–	0.32	0.46	0.35	1.00	0.53	0.69	0.23	0.30
	–	–	–	–	(0.53)	(0.79)	(0.48)	(0.61)	(0.18)	(0.71)	(0.80)	(0.86)
γ_{n2}	–	–	–	–	0.38	1.08	0.81	2.27	0.30	4.79	0.29	0.53
	–	–	–	–	(0.41)	(0.58)	(0.04)	(0.32)	(0.58)	(0.09)	(0.61)	(0.77)
γ_{n3}	–	–	–	–	–	–	–	–	0.28	2.07	0.12	0.55
	–	–	–	–	–	–	–	–	(0.63)	(0.35)	(0.99)	(0.76)
γ_{n4}	–	–	–	–	–	–	–	–	0.58	0.75	0.59	3.26
	–	–	–	–	–	–	–	–	(0.13)	(0.69)	(0.12)	(0.20)
γ_{n5}	–	–	–	–	–	–	–	–	0.74	5.40	0.54	0.13
	–	–	–	–	–	–	–	–	(0.05)	(0.07)	(0.17)	(0.94)
$1/d$	**1.85**	**32.3**	0.18	2.37	**1.22**	**20.8**	0.32	2.84	**4.67**	**50.46**	0.80	5.22
	(0.00)	**(0.00)**	(0.91)	(0.31)	**(0.00)**	**(0.00)**	(0.53)	(0.24)	**(0.00)**	**(0.00)**	(0.04)	(0.07)

† The p-values are in parentheses. S and L stand for small and large sample size respectively.

Table 4: Descriptive Statistics

Statistics	low % change (r_{lt})	high % change (r_{ht})	# of trades (n_t)
Minimum	-29.000	-2.984	6
1st Quartile	-5.565	3.999	22
Median	-3.415	5.832	30
3rd Quartile	-1.604	8.211	40
Maximum	24.480	44.750	130
Mean	-3.8380	6.3350	31.540
Variance	12.7687	11.7219	166.92
Correlation	0.3750		——
Skewness	-0.7585	2.0796	0.8657
Kurtosis	11.9538	19.5875	5.9716

Table 5: Estimation Results of Conditional Mean Equation

Conditional Mean Equation				
			95% C.I.	
	estimate	s.e.†	lower	upper
α_0	-1.57	1.82	-5.14	2.31
α_{l1}	-0.35	0.04	**-0.41**	**-0.26**
α_{l2}	0.07	0.03	**0.00**	**0.14**
α_{l3}	0.06	0.03	-0.01	0.11
α_{l4}	0.11	0.03	**0.04**	**0.15**
α_{l5}	0.03	0.03	-0.04	0.07
α_{l6}	0.07	0.03	-0.01	0.11
α_{l7}	0.03	0.04	-0.05	0.10
α_{h1}	0.22	0.04	**0.13**	**0.28**
α_{h2}	0.04	0.04	-0.05	0.09
α_{h3}	0.02	0.03	-0.04	0.07
α_{h4}	-0.01	0.03	-0.08	0.05
α_{h5}	0.07	0.03	**0.00**	**0.11**
α_{h6}	-0.02	0.03	-0.07	0.04
α_{h7}	-0.03	0.03	-0.08	0.03
α_{n1}	0.18	0.21	-0.29	0.52
α_{n2}	0.12	0.17	-0.29	0.40
α_{n3}	0.19	0.16	-0.24	0.40
α_{n4}	-0.14	0.17	-0.37	0.32
α_{n5}	0.31	0.41	-0.67	0.86
α_{n6}	-0.44	0.23	-0.78	0.11
α_{n7}	0.47	0.22	-0.15	0.74
α_{n8}	-0.08	0.20	-0.43	0.38
α_{n9}	0.46	0.22	-0.11	0.75
$s_{1\mu}$	-0.30	0.38	-0.79	0.64
$s_{2\mu}$	0.25	0.23	-0.22	0.69
$s_{3\mu}$	-0.29	0.20	-0.52	0.29

†Standard errors are obtained by stationary block bootstrapping.

Table 6: Estimation Results of Conditional Variance Equation

Conditional Variance Equation				
			95% C.I.	
	estimate	s.e.†	lower	upper
β_0	-0.43	0.39	-1.15	0.34
β_{r1}	0.42	0.08	**0.28**	**0.56**
β_{r2}	0.08	0.04	-0.01	0.15
β_{r3}	0.08	0.05	**0.00**	**0.18**
β_{r4}	0.09	0.03	**0.01**	**0.14**
β_{r5}	0.12	0.05	-0.01	0.20
β_{n1}	-0.10	0.06	-0.2	0.02
$s_{1\sigma}$	-0.21	0.09	**-0.37**	**-0.04**
$s_{2\sigma}$	0.02	0.07	-0.11	0.15
$s_{3\sigma}$	-0.11	0.07	-0.20	0.07

†Standard errors are obtained by stationary block bootstrapping.

Table 7: Estimation Results of Conditional Intensity Equation

Conditional Intensity Equation				
			95% C.I.	
	estimate	s.e.†	lower	upper
γ_0	1.01	0.25	**1.05**	**1.97**
γ_{n1}	-0.04	0.03	-0.07	0.02
γ_{n2}	-0.04	0.02	-0.08	0.00
γ_{n3}	0.02	0.02	-0.03	0.05
γ_{n4}	0.08	0.02	**0.02**	**0.12**
$s_{1\lambda}$	0.37	0.03	**0.30**	**0.42**
$s_{2\lambda}$	0.18	0.04	**0.05**	**0.20**
$s_{3\lambda}$	0.11	0.03	-0.02	0.11
$1/d$	0.10	0.01	**0.10**	**0.14**

†Standard errors are obtained by stationary block bootstrapping.

Table 8: Ljung-Box Tests for pseudo-Pearson Residuals $\hat{\varepsilon}_{1t}$ and $\hat{\varepsilon}_{2t}$

	$\hat{\varepsilon}_{1t}$ (p-values)				$\hat{\varepsilon}_{2t}$ (p-values)		
lags	*SPEC1*	*SPEC2*	*SPEC3*	lags	*SPEC1*	*SPEC2*	*SPEC3*
1	0.60	0.62	0.12	1	0.01	0.01	0.00
2	0.82	0.74	0.00	2	0.02	0.01	0.00
3	0.88	0.88	0.00	3	0.04	0.03	0.01
4	0.95	0.95	0.00	4	0.08	0.06	0.02
5	0.98	0.96	0.00	5	0.14	0.10	0.04
6	0.99	0.76	0.00	6	0.21	0.15	0.07
7	0.96	0.75	0.00	7	0.23	0.17	0.08
8	0.98	0.75	0.00	8	0.11	0.08	0.04
9	0.81	0.50	0.00	9	0.12	0.08	0.04
10	0.87	0.60	0.00	10	0.14	0.10	0.05
11	0.61	0.45	0.00	11	0.12	0.08	0.04
12	0.67	0.52	0.00	12	0.15	0.10	0.05
13	0.73	0.58	0.00	13	0.12	0.08	0.04
14	0.69	0.53	0.00	14	0.11	0.07	0.04
15	0.75	0.60	0.00	15	0.08	0.05	0.03
16	0.78	0.61	0.00	16	0.09	0.06	0.03
17	0.71	0.49	0.00	17	0.11	0.08	0.04
18	0.76	0.54	0.00	18	0.14	0.10	0.06
19	0.61	0.41	0.00	19	0.18	0.13	0.08
20	0.66	0.47	0.00	20	0.22	0.16	0.10

Table 9: Measures of Goodness of Fit

	RMSE		CR & ER			MLF		MDE	
	Lower	Upper	CR	ER	$\frac{\text{CR+ER}}{2}$	$p=1$	$p=2$	$q=1$	$q=2$
Spec 1	**3.3226**	**2.8191**	0.8388	0.7648	**0.8018**	**4.5586**	**18.9874**	**2.5054**	**3.0812**
Spec 2	3.3385	2.8397	0.8385	0.7624	0.8005	4.5903	19.2095	2.5204	3.0992
Spec 3	3.4431	2.8886	0.8356	0.7568	0.7962	4.7041	20.1987	2.5808	3.1779

Table 10: Comparison with Existing Approaches using Simulated Data*

	A-RMSE		A-CR & A-ER			A-MLF		A-MDE	
	Lower	Upper	A-CR	A-ER	$\frac{\text{A-CR+A-ER}}{2}$	$p=1$	$p=2$	$q=1$	$q=2$
CR	6.4701	6.7084	0.9310	0.9172	0.9241	10.2730	87.2878	5.6477	6.5952
CCR	7.1343	7.3709	0.9263	0.9090	0.9176	11.3340	105.8595	6.1847	7.2583
VAR	6.4305	6.6703	0.9313	0.9176	0.9244	10.2163	86.2639	5.6167	6.5564
IAR-TS	6.4183	6.6581	0.9314	0.9182	0.9248	10.1871	85.9428	5.6009	6.5443
IAR-MTS	6.4306	6.6704	0.9313	0.9176	0.9244	10.2165	86.2669	5.6169	6.5566
STAR	6.9327	6.9922	0.9159	0.9128	0.9143	10.8039	98.0272	5.4865	6.9634
Proposed model	**5.1801**	**5.2324**	0.9420	0.9348	**0.9384**	**8.0387**	**54.5267**	**4.4742**	**5.2124**

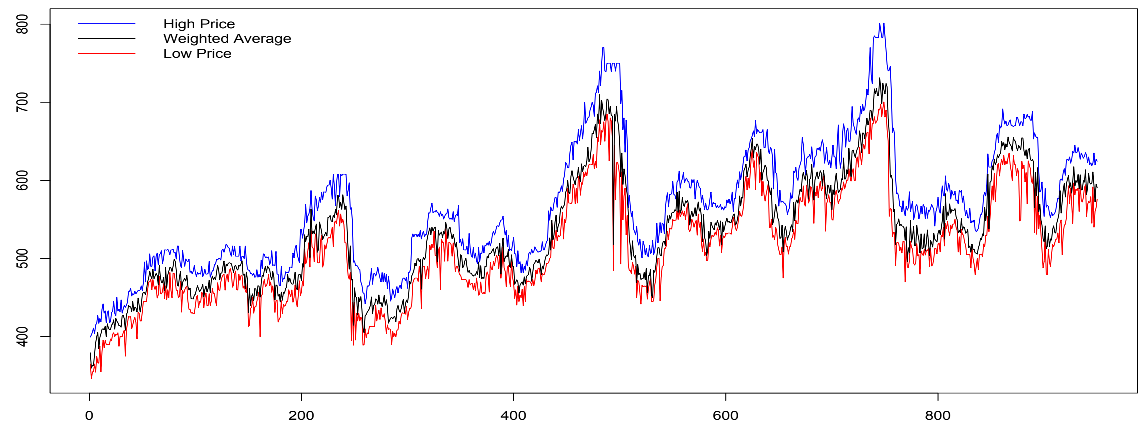
* Simulated data are generated by DGP2 with 200 observations. Simulation is replicated 1000 times.

Table 11: Comparison with Existing Approaches using Real Data[†]

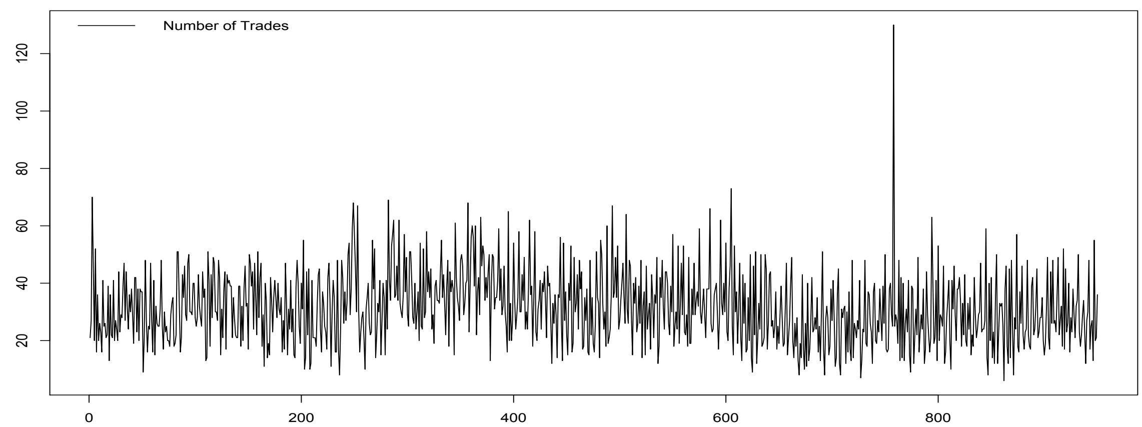
	RMSE		CR & ER			MLF		MDE	
	Lower	Upper	CR	ER	$\frac{\text{CR+ER}}{2}$	$p=1$	$p=2$	$q=1$	$q=2$
CR	3.5201	2.9698	0.8024	0.7678	0.7851	4.7513	21.2110	2.5920	3.2566
CCR	3.5201	2.9698	0.8024	0.7678	0.7851	4.7513	21.2110	2.5920	3.2566
VAR	3.3440	2.7688	0.8136	0.7793	0.7964	4.5226	18.8481	2.4790	3.0699
IAR-TS	3.3332	**2.7687**	0.8145	0.7796	0.7970	**4.5062**	**18.7754**	2.4697	**3.0639**
IAR-MTS	3.3426	2.7689	0.8137	0.7793	0.7965	4.5214	18.8399	2.4782	3.0692
STAR	**3.1551**	2.9975	0.8439	0.8172	**0.8306**	4.5372	18.9400	**2.3186**	3.0773
Preferred model*	3.3226	2.8191	0.8388	0.7648	0.8018	4.5586	18.9874	2.5054	3.0812

[†] The real data is the interval-valued time series of beef prices.

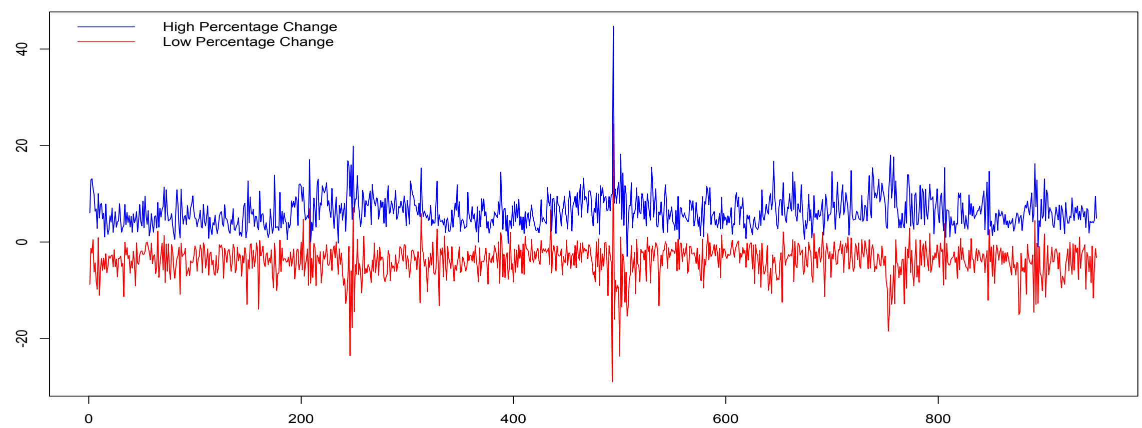
* Our preferred model is *Spec 1* in section 4.2



(a) Daily Low and High Prices, and Weighted Average Price



(b) Daily Number of Trades



(c) Daily Low and High Returns

Figure 1: Daily Prices and Returns and Number of Trades

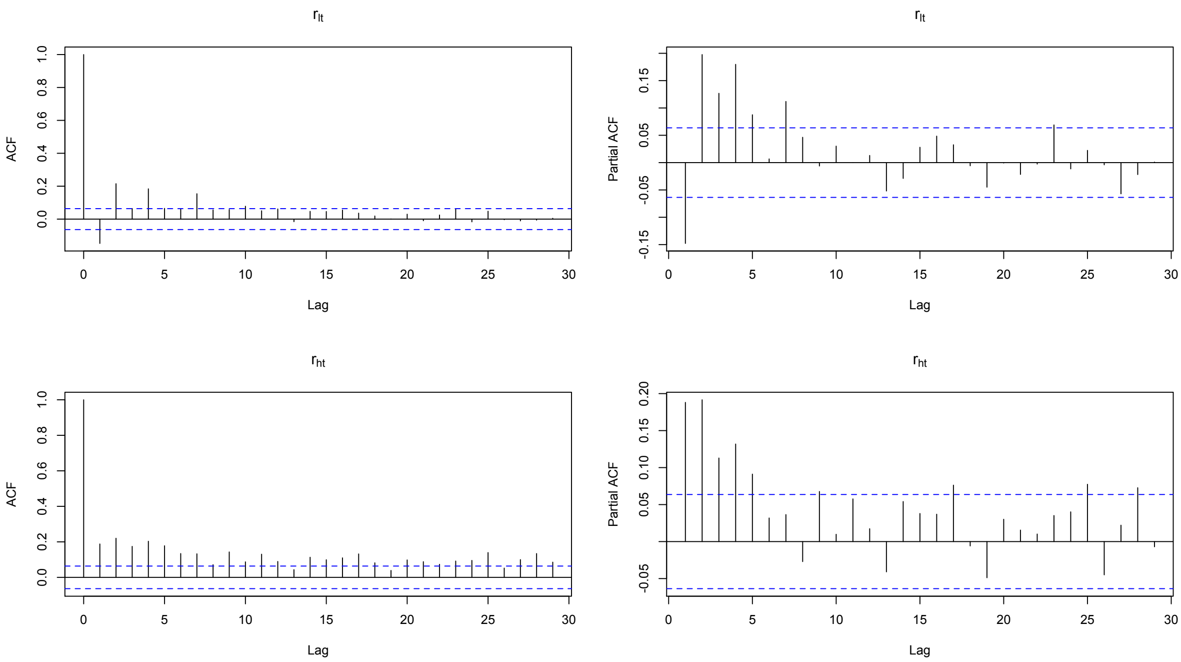


Figure 2: ACF and PACF of Low/High Daily Returns

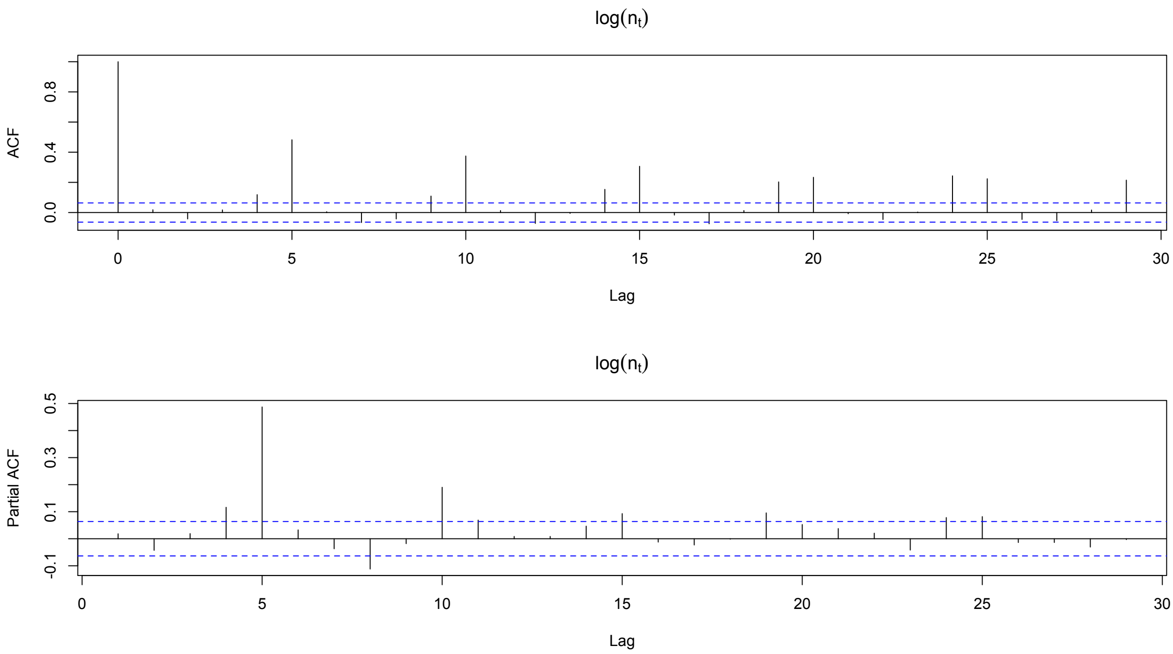
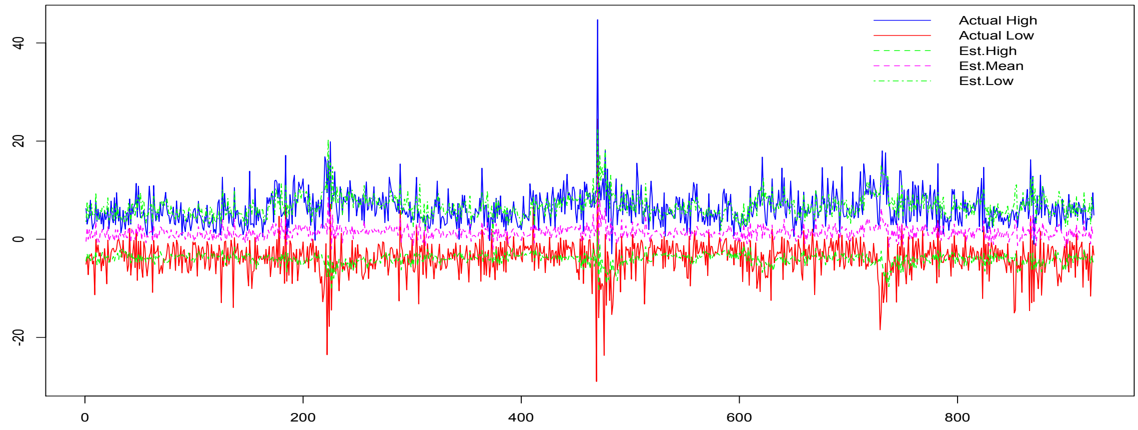
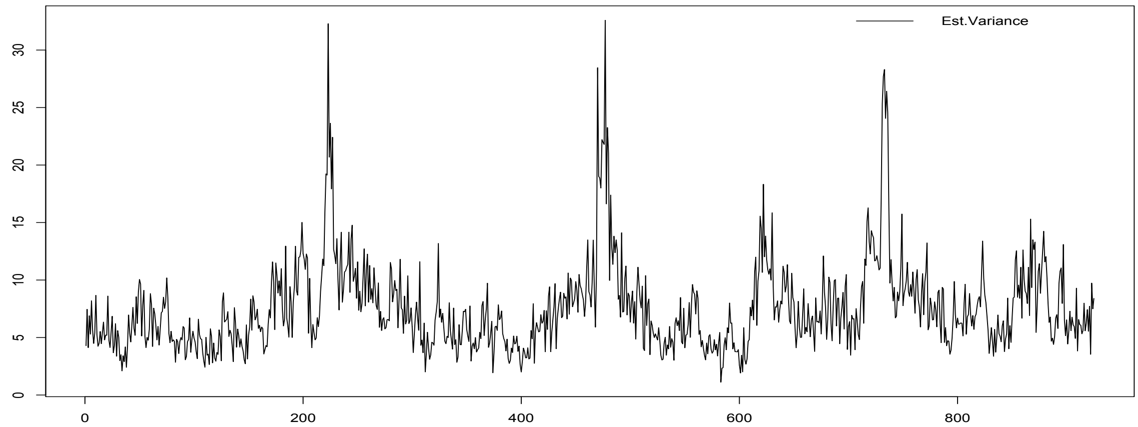


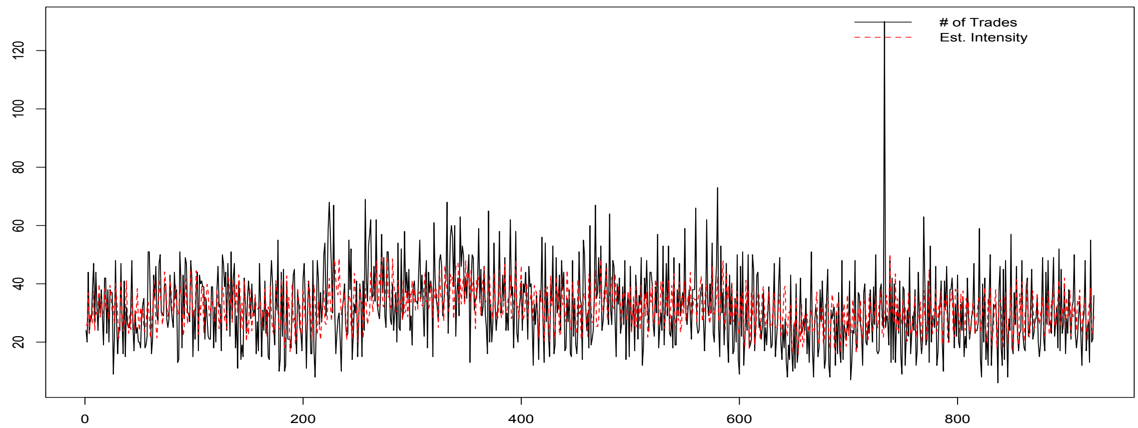
Figure 3: ACF and PACF of Logarithm of Number of Trades



(a) Estimated Daily Returns: $\widehat{\mu}_t$, $\widehat{r}_{lt}$, $\widehat{r}_{ht}$



(b) Estimated Variance: $\widehat{\sigma}_t^2$



(c) Estimated Intensity: $\widehat{\lambda}_t$

Figure 4: Estimated Conditional Mean/Variance/Intensity