UNIVERSITY OF
CAMBRIDGE

Cambridge Working Papers in Economics

Modeling dynamic diurnal patterns in high frequency financial data

Ryoko Ito

CWPE 1315

Modeling dynamic diurnal patterns in high frequency financial data

Ryoko Ito*
Faculty of Economics, University of Cambridge

April 12, 2013

Abstract

A spline-DCS model is developed to forecast the conditional distribution of high-frequency financial data with periodic behavior. The dynamic cubic spline of Harvey and Koopman (1993) is applied to allow diurnal patterns to evolve stochastically over time. An empirical application illustrates the practicality and impressive predictive performance of the model.

KEYWORDS: outlier; robustness; score; calendar effect; spline; trade volume.

JEL Classification: C22

1 Introduction

Intra-day periodicity caused by periodic trading patterns is a stylized feature of high-frequency financial data. Standard methods for modeling diurnal patterns include the use of the Fourier representation or a deterministic spline and computing sample moments for each intra-day bin.[1] In the existing literature, the shape of diurnal patterns is generally assumed to be a deterministic function of time and remains the same for every trading day. It is also a standard approach to estimate the intra-day periodic component first and *diurnally adjust* data before estimating

*Email: ri239@cam.ac.uk. I would like to thank my supervisor Andrew Harvey for his helpful guidance and comments on this research throughout the year. I would also like to thank Philipp Andres, Stephen Thiele, Michele Caivano, Oliver Linton, Donald Robertson, and the participants of the Score Workshop in 2013 at Tinbergen Institute, especially Andre Lucas, for providing thoughtful comments on a preliminary version of this paper.

[1]See, for instance, Andersen and Bollerslev (1998), Engle and Russell (1998), Zhang et al. (2001), Campbell and Diebold (2005), Engle and Rangel (2008), Brownlees et al. (2011), and Engle and Sokalska (2012).

other stochastic components by a two-step procedure, but such a procedure can render the asymptotic properties of statistical tests invalid. This paper proposes to use the dynamic cubic spline of Harvey and Koopman (1993) to allow for the possibility that diurnal patterns may evolve stochastically over time. The dynamic cubic spline is a parsimonious way of capturing evolving diurnal patterns, and it can be estimated simultaneously with all other components of the model by the method of maximum likelihood.

Harvey and Koopman (1993) originally developed the dynamic cubic spline to estimate the intra-weekly patterns of hourly electricity demand. It was employed later by Harvey et al. (1997) to model a changing seasonal component of weekly money supply in the U.K., and also by Bowsher and Meeks (2008) to forecast zero-coupon yield curves. Bowsher and Meeks (2008) interpret the spline as a special type of "dynamic factor model", where the knots of the spline are the factors and the factor loadings are treated as given and specified according to the requirement that the model has to be a cubic spline. In order to capture dynamic diurnal patterns in high-frequency financial data, we integrate the dynamic cubic spline with a model for forecasting conditional density and volatility (or scale) dynamics. The model we use for this purpose is the dynamic conditional score (DCS) model formally defined and studied by Harvey (2013) and Andres and Harvey (2012). The DCS model is also termed the generalized autoregressive score (GAS) model and studied independently by Creal et al. (2011, 2013).

The basic DCS model for scale is defined as follows. Given a sequence of T observations $(y_t)_{t=0}^T$, suppose we denote its underlying random data generating process by $(Y_t)_{t=0}^T$. We assume that this process is defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $(\mathcal{F}_t)_{t=0}^T$, and that $\mathcal{F}_0$ is trivial so that Y_0 is almost surely constant. The scale DCS model assumes that there exist a location factor $c \in \mathbb{R}$ and a standard cumulative distribution function (cdf) denoted by F such that

$$(Y_t - c)/\alpha_{t|t-1} | \mathcal{F}_{t-1} \sim \texttt{iid}\, F \qquad t = 1, \dots T \tag{1}$$

for a sequence of scale factors $\left(\alpha_{t|t-1}\right)_{t=1}^T$, where $\alpha_{t|t-1} > 0$ for all $t = 1, \dots, T$. The notation $\cdot_{t|t-1}$ reflects the conditionality of the variable at time t on $\mathcal{F}_{t-1}$. F is a short-hand notation for the standard cdf $F(\cdot; \boldsymbol{\theta})$ re-centered around the origin with a constant vector $\boldsymbol{\theta}$ of distribution parameters that does not include c and $\alpha_{t|t-1}$. The probability density/mass function (pdf) of F is denoted by f. The first-order DCS filter for scale is

$$\alpha_{t|t-1} = \exp(\lambda_{t|t-1}), \qquad \lambda_{t|t-1} = \delta + \phi \lambda_{t-1|t-2} + \kappa u_{t-1}, \qquad t = 1, \dots T, \tag{2}$$

where u_t is the conditional *score*

$$u_t = \frac{\partial \log \left[e^{-\lambda_{t|t-1}} f(y_t e^{-\lambda_{t|t-1}}; \boldsymbol{\theta}) \right]}{\partial \lambda_{t|t-1}}, \qquad t = 1, \ldots, T.$$

We also have $\delta \in \mathbb{R}$, $\kappa > 0$, and $|\phi| < 1$ if $(\lambda_{t|t-1})_{t=1}^{T}$ is stationary. $\omega = \delta/(1-\phi)$ is the unconditional mean of $\lambda_{t|t-1}$ when it is stationary. The use of the exponential link with the canonical link parameter $\lambda_{t|t-1} \in \mathbb{R}$ ensures that the scale parameter $\alpha_{t|t-1}$ remains positive for all t without any parameter restrictions.[2]

Andres and Harvey (2012) study the case where y_t is non-negative for all t, and consider a number of non-negative distributions for F including generalized gamma (GG), generalized beta of the second kind (GB2), and log-normal.[3] The Beta-t-EGARCH model of Harvey and Chakravarty (2008) is also a special case of DCS in which the Student's t-distribution is assigned to F. Beta-t-EGARCH can be used to model volatility of asset returns.

The advantages of DCS are threefold. First, the parameters of the model can be estimated easily by the method of maximum likelihood. The consistency and asymptotic normality of the maximum likelihood estimators (MLEs) are established for many distributions, including a number of heavy- or long-tailed distributions, by Harvey (2013). Second, the analytic expressions of the multi-step optimal predictors[4] as well as their mean square errors are available whenever the corresponding moments of F and the moment generating function (mgf) of u_t exist. Third, the model is robust to extreme observations because of the use of the conditional score in the filter. See Harvey (2013) for more details.

This paper contributes to the literature by illustrating the practicality and forecasting performance of our proposed spline-DCS model. Aside from diurnal patterns, our spline-DCS model also captures other stylized features of high-frequency financial data such as a frequency mass of zero-valued observations, the *day-of-the-week effect*, and the *overnight effect*.

The day-of-the-week effect is a type of calendar effects, and refers to the dependency of the dynamics of a given variable on the day of the week. (See Taylor (2005) and Hautsch (2012) for discussions.) Generally, this effect is modeled in the existing literature using day-of-the-week dummies, see, for instance, Andersen

[2]The choice of link function should always be consistent with the parametric choice of F for the asymptotic properties of DCS to go through. See Harvey (2013).

[3]If F is characterized by the GG distribution, u_t is a linear function of a gamma distributed random variable. If F is characterized by the GB2 distribution, u_t is a linear function of a beta distributed random variable. In both cases, u_t is `iid` given $\mathcal{F}_{t-1}$ for all t. See Appendix A.

[4]This optimality is in terms of minimum mean square error (MMSE).

and Bollerslev (1998) and Lo and Wang (2010).[5] In contrast, using a special case of the dynamic cubic spline termed *weekly spline*, our model captures the day-of-the-week effect by allowing for a change in the level *as well as* the overall shape of diurnal patterns on different weekdays without using any dummy variables.

We capture the overnight effect by relaxing the continuity condition of the spline between the end and the beginning of any two consecutive trading days. This means that the spline on any trading day can begin from a level different from where it ended on the previous trading day. This method allows for the possibility that it may take time for the overnight effect to diminish completely during the day, and eliminates the need to identify which morning observations are due to overnight information. Our way of adjusting for the overnight effect can be viewed as an alternative to the existing methods in the literature including differentiating day and overnight jumps, the use of dummy variables, and treating events that occur in a specified period immediately after market opening as censored. See, for instance, Gerhard and Hautsch (2007), and Boes et al. (2007). It seems that our model handles extreme morning observations very well. This is partly due to the use of the exponential link function: highly volatile and dynamic movements in the scale parameter can be induced by small movements in the link parameter to which the DCS filter is applied.

We fit the spline-DCS model to intra-day observations of *trade volume* of IBM stock traded on the New York Stock Exchange (NYSE). Trade volume is a measure of intensity of trading activity. There are a variety of volume measures used in the literature including the number of shares traded, dollar volume, number of transactions, turnover (shares traded divided by shares outstanding), and dollar turnover. As the choice of volume measure is not very important given the stated purpose of this paper, we arbitrarily choose the number of shares traded. Although this study deals with non-negative time series, the spline-DCS model can be also applied to variables with support over the entire real line, in which case our model should be viewed as an extension of Beta-t-EGARCH and can be used to contribute to studies of high-frequency asset returns including the seminal paper by Andersen and Bollerslev (1998).

Although we use the DCS model as a vehicle for illustrating the usefulness of the dynamic cubic spline for high-frequency financial analysis, the dynamic cubic spline can be easily put in the GARCH or ACD framework. Due to the empirical

[5]Andersen and Bollerslev (1998) use day-of-the-week dummies in the analysis of high-frequency returns of the Deutsche Mark-Dollar exchange rate (DM-Dollar). Lo and Wang (2010) use day-of-the-week dummies in the analysis of the weighted daily turnover and returns of indices of NYSE and AMEX ordinary common shares.

Series	Smpl. Size	Mean	Max.	Min.	S.D.	Skewness	99.9% quantile	Freq. of zero
IBM30s	19,500	10,539	1,652,100	0	26,071	29	293,654	0.47%
IBM1m	9,750	21,297	1,652,100	0	39,114	18	604,295	0.06%

Table 1 – Table of sample statistics of the IBM trade volume. Sampling period: Monday 28 February - Friday 31 March 2000. (25 trading days with 6.5 trading hours per trading day.)

feature of our data, the estimated cubic spline for our application turns out to be U-shaped on any trading day. This U-shape is obtained without imposing any restrictions on the shape of the spline. As the spline can reflect any periodic shape including multi-modal and concave ones, it can be used to estimate diurnal patterns of other financial variables including asset returns and duration.

The structure of this paper is as follows. Section 2 gives an initial investigation of the trade volume data and motivates the construction of our model. Section 3 outlines our modeling assumptions and formally constructs the spline-DCS model. Sections 3.6 and 4 discuss methods for estimation and model selection. The in-sample and out-of-sample estimation results are reported in Sections 4 and 5.

2 Data characteristics

Before proceeding to detailed modeling and forecasting results, it is useful to get an overall feel for the trade volume data. This section provides an initial investigation of our data to motivate our formal model in Section 3.

We analyze the *trade volume* (in the number of shares) of IBM stock traded on NYSE during the market opening hours (9.30am-4pm in the New York local time) between Monday 28 February and Friday 31 March 2000, which includes 25 trading days and no public holidays. Our raw data set is in tick-format and consists of the record of every trade in the order of occurrence. The tick-data is irregularly spaced and often has multiple transactions in one second. In order to explore the effects of marginal changes in the aggregation interval on our inference, we aggregate the tick-data by 30 seconds and 1 minute, respectively, to generate two aggregated series. If the aggregation interval is 30 seconds, there are 780 observations per trading day. For convenience, we refer to the aggregated series as IBM30s if the aggregation interval is 30 seconds and IBM1m if 1 minute.

In the top panel of Figure 1, we observe several recurrent spikes in trade volume near the moment of market opening or closure. These extreme observations make the upper tail of empirical distribution very long. (See Table 1.) The smoothed IBM30s series in the right column of Figure 1 reflects that there is a diurnal U-

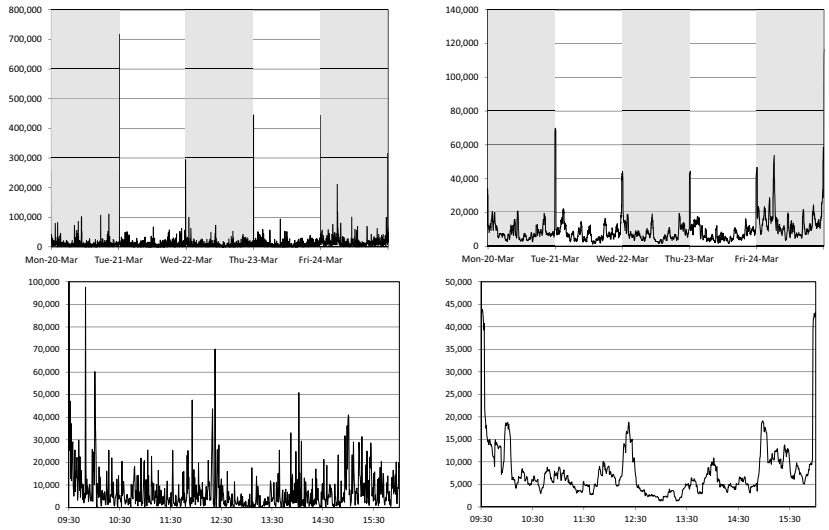

Figure 1 – IBM30s (left column) and the same series smoothed by the simple moving average of nearest 20 observations (right column). Time on the x-axis. Top panel: Monday 20 - Friday 24 March 2000. Each day covers market opening hours between 9.30am-4pm (in the New York local time). Bottom panel: Wednesday 22 March 2000, covering 9.30am-4pm (in the New York local time).

shaped pattern in trading activity on every trading day. Extreme movements in trading activity in the first hour of trading day can be caused by news transmitted over night. Trading activity slows down towards lunch time of around 1pm as overnight information is processed, but picks up again in the afternoon as traders re-balance their positions before market closure. One may naturally suspect that the shape of diurnal pattern may be slowly changing over time. Other measures of trading activity based on durations (such as trade durations, midquote change durations, and volume durations) exhibit similar diurnal patterns, but the shape is inverted so that the duration is long during lunch. See Hautsch (2012, p.41).

The left and middle columns of Figure 2 show that our series are heavily right-skewed and have extremely long upper-tail. The length of upper-tail can be also seen in the distance between the maximum and the 99.9% sample quantile in Table 1.[6] The right column of Figure 2 shows the highly persistent nature of our series. IBM30s exhibits statistically significant autocorrelation that is very slow to decay. The autocorrelation of IBM1m is faster to decay than that of IBM30s. Moreover, there is a non-negligible number of zero-valued observations. (See Table 1.)

The discussion so far suggests that our model needs a periodic component to capture the diurnal U-shaped patterns. This component should allow for the possibility that the shape of diurnal patterns may change over time. Moreover,

[6]Note that the sample skewness statistics in Table 1 must be interpreted with care as the theoretical skewness may not exist.

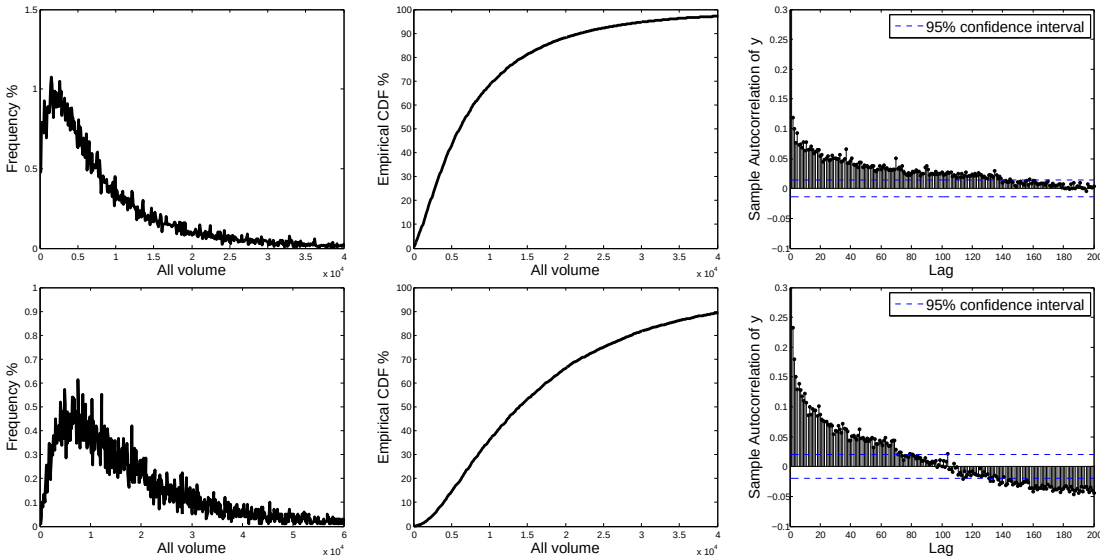

Figure 2 – Frequency distribution (left column), empirical cdf (middle column), and sample autocorrelation (right column). IBM30s (top row), and IBM1m (bottom row). The axes of cdf and autocorrelation charts are fixed so that they are easily comparable. The 200th lag corresponds approximately to 1.5 hours prior for IBM30s and 3 hours prior for IBM1m.

non-periodic factors may coexist with the periodic component. One such factor is a highly persistent low-frequency component. The empirical autocorrelation structure suggests that a highly persistent behavior similar to long memory may be inherent in data. This can be captured by a combination of autoregressive components. The presence of non-negligible number of zero-valued observations can be explained by a binary component governing whether the next observation is zero or non-zero. Given the empirical distribution of data, we find that fixing the location parameter (c) at zero works for our application. The next section gives the formal definition of each of these components.

3 Spline-DCS model

3.1 Redefining time indices

Henceforth, we use the notation $\tau = 1, \ldots, I$ to denote the location of intra-day bins and $t = 1, \ldots, T$ to denote trading days in our sample. We use the subscript $\cdot_{t,\tau}$ to denote the τth intra-day bin on the tth trading day. We assume that $\mathcal{F}_{1,0}$ is trivial. We let $\tau = 1$ to be the location of the first aggregated observation for each trading day. The total sample size is $I \times T$. In order to avoid superfluous subscripts, we merely write $\cdot_{t,\tau}$ instead of $\cdot_{t,\tau|t,\tau-1}$ even when the variable is conditional on

$\mathcal{F}_{t,\tau-1}$. Finally, we introduce the following set notations

$$\begin{aligned}\Psi_{T,\,I} &= \{(t,\tau) \in \{1,2,\ldots,T\} \times \{1,2,\ldots,I\}\},\\ \Psi_{T,I>0} &= \{(t,\tau) \in \{1,2,\ldots,T\} \times \{1,2,\ldots,I\} : y_{t,\tau} > 0\}.\end{aligned}$$

3.2 A probability mass at the origin

The presence of a non-negligible number of zero-valued observations cannot be explained by conventional continuous distributions as the probability of observing a particular value is zero by definition. Although one can eliminate zero-valued observations by widening the aggregation interval, a better approach is to define a distribution such that its strictly positive support is captured by a conventional continuous distribution and the origin has a discrete probability mass.

Formally, we define the cdf $F : \mathbb{R}_{\geq 0} \to [0,1]$ with a constant parameter vector $\boldsymbol{\theta}$ of a standard random variable $X \sim F$ as

$$\mathbb{P}_F(X = 0) = p, \quad \mathbb{P}_F(X > 0) = 1 - p, \quad \mathbb{P}_F(X \leq x | X > 0) = F^*(x; \boldsymbol{\theta}^*) \tag{3}$$

for some $p \in [0,1]$ and any $x > 0$, where $F^* : \mathbb{R}_{>0} \to [0,1]$ is the cdf of a conventional standard continuous random variable with the time-invariant parameter vector $\boldsymbol{\theta}^*$.[7] Thus we have $\boldsymbol{\theta} = (p, \boldsymbol{\theta}^{*\top})^\top$. The properties of this type of distributions are studied formally in Hautsch et al. (2010).[8] For non-negative series, F^* can be a number of distributions including Weibull, Gamma, Burr, and log-normal, many of which are special cases of the GG and GB2 distributions. See Figure 3 and, more formally, Kleiber and Kotz (2003) for the relationship between these distributions.

Under this definition of F in (3), $u_{t,\tau}$ becomes the conditional score of F^* and it is defined only for $y_{t,\tau} > 0$. Thus, we set $u_{t,\tau} = \inf_{s \in \Omega} u_{t,\tau}(s)$ whenever $y_{t,\tau} = 0$. If F^* is the GB2 distribution, this means that we set $u_{t,\tau} = -\nu\xi$ whenever $y_{t,\tau} = 0$. It is natural to suspect that p may change over time because the probability of zero-volume must be lower during morning trading hours than during quiet lunch hours. Rydberg and Shephard (2003) and Hautsch et al. (2010) independently study decomposition models for estimating the conditional dynamics of p via the logit link. An interesting extension of our model is a hybrid spline-DCS model

[7] The unconditional nth moment of X is well-defined as long as it is well-defined for F^* because

$$\mathbb{E}_F[X^n] = p\,\mathbb{E}_F[X^n|X=0] + (1-p)\mathbb{E}_F[X^n|X>0] = (1-p)\mathbb{E}_{F^*}[X^n].$$

[8] This decomposition technique is a standard one in econometrics and essentially the same as the classical censored regression models and the decomposition models of McCulloch and Tsay (2001) and Rydberg and Shephard (2003).

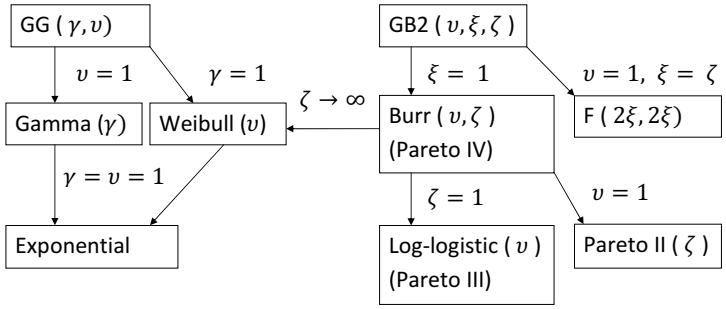

Figure 3 – Nested diagram of some of the useful non-negative distributions. Scale factor is assumed to be one in all cases.

for the conditional dynamics of p as well as the scale parameter. However, in this paper, we assume p to be constant for simplicity and leave this extension for future research. This restriction on p is inconsequential in our application as the number of zero-valued observations is relatively small compared to the total number of observations.

Denoting the standardized observations as $\varepsilon_{t,\tau} \equiv y_{t,\tau}/\alpha_{t,\tau}$ for all $(t,\tau) \in \Psi_{T,I}$, where $\alpha_{t,\tau} > 0$ is the time-varying scale parameter, the joint likelihood function based on F for the set of observations $(y_{t,\tau})_{(t,\tau)\in\Psi_{T,I}}$ is

$$L\left((y_{t,\tau})_{(t,\tau)\in\Psi_{T,I}};\boldsymbol{\theta}\right) = \prod_{(t,\tau)\in\Psi_{T,I>0}} (1-p)\exp(-\lambda_{t,\tau}) f^*(\varepsilon_{t,\tau};\boldsymbol{\theta}^*) \prod_{(t,\tau)\in\Psi_{T,I}\cap\Psi^c_{T,I>0}} p,$$

where f^* is the pdf of F^*. The log-likelihood is

$$\log L = A\log(1-p) + (T\times I - A)\log p + \sum_{(t,\tau)\in\Psi_{T,I>0}} \log\left(\exp(-\lambda_{t,\tau}) f^*(\varepsilon_{t,\tau};\boldsymbol{\theta}^*)\right), \quad (4)$$

where $A = |\Psi_{T,I>0}|$. It is easy to check that the MLE of p is $\widehat{p} = (T\times I - A)/(T\times I)$.

3.3 Unobserved components of scale

We assume that the diurnal U-shaped patterns and autocorrelation in our data are entirely due to the periodicity and autocorrelation in $(\alpha_{t,\tau})_{(t,\tau)\in\Psi_{T,I}}$. The standardized series $\varepsilon_{t,\tau}$ given $\mathcal{F}_{t,\tau-1}$ is assumed to be `iid` and free of periodic behavior for all $(t,\tau) \in \Psi_{T,I}$. As we will estimate $\alpha_{t,\tau}$ via the link parameter $\lambda_{t,\tau}$, this means that $\lambda_{t,\tau}$ exhibits periodicity and autocorrelation.[9] Given these assumptions, we specify $\lambda_{t,\tau}$ as a sum of (i) the periodic component $s_{t,\tau}$, which explains the diurnal U-shaped patterns that may or may not evolve over time; (ii) the random-walk

[9]We note that, by the stationarity/nonstationarity property of $\lambda_{t,\tau}$, we do not infer the true stationarity property of the underlying process as the spirit of our model is merely to obtain a good local approximation of the true data generating process. Our analysis is local in the sense that it is limited to working with discrete-time observations of what is actually a continuous-time process collected over a short sampling period.

component $\mu_{t,\tau}$ capturing the slowly changing low-frequency movements that is non-periodic; and (iii) the stationary high-frequency component $\eta_{t,\tau}$, which is a mixture of several autoregressive components and captures the highly persistent nature of data. Formally, our spline-DCS model is:

$$y_{t,\tau} = \varepsilon_{t,\tau} \exp(\lambda_{t,\tau}), \qquad \lambda_{t,\tau} = \omega + \mu_{t,\tau} + \eta_{t,\tau} + s_{t,\tau}, \quad \varepsilon_{t,\tau} | \mathcal{F}_{t,\tau-1} \sim \texttt{iid}\, F$$

for $\omega \in \mathbb{R}$, $\tau = 1, \ldots, I$ and $t = 1, \ldots, T$. $\mu_{t,\tau}$ is defined as

$$\mu_{t,\tau} = \mu_{t,\tau-1} + \kappa_\mu u_{t,\tau-1}, \qquad \forall\, (t,\tau) \in \Psi_{T,\,I},$$

where $u_{t,\tau}$ is the score of F^* and $u_{t,\tau} = \inf_{s\in\Omega} u_{t,\tau}(s)$ whenever $y_{t,\tau} = 0$. The estimation results in Section 4 suggest that this random-walk component does a good job in capturing the low-frequency dynamics of our data.[10] $\eta_{t,\tau}$ is the stationary component defined as:

$$\begin{aligned}
\eta_{t,\tau} &= \sum_{j=1}^{J} \eta_{t,\tau}^{(j)}, \qquad \forall\, (t,\tau) \in \Psi_{T,\,I} \\
\eta_{t,\tau}^{(j)} &= \phi_1^{(j)} \eta_{t,\tau-1}^{(j)} + \phi_2^{(j)} \eta_{t,\tau-2}^{(j)} \cdots + \phi_{m^{(j)}}^{(j)} \eta_{t,\tau-m^{(j)}}^{(j)} + \kappa_\eta^{(j)} u_{t,\tau-1}, \qquad j = 1, \ldots, J
\end{aligned}$$

for some $J \in \mathbb{N}_{>0}$. We assume that $m^{(j)} \in \mathbb{N}_{>0}$ and $\eta_{t,\tau}^{(j)}$ is stationary for all $j = 1, \ldots, J$. The stationarity requirements on the autoregressive coefficients are the same as in an ARMA model. By making $\eta_{t,\tau}$ a mixture of components with autoregressive structures, we are allowing for a highly persistent behavior similar to long memory. Although the specification of $\eta_{t,\tau}$ above is in a general form of J autoregressive components, $J = 2$ works well for our application. For the identifiability of each component, we assume that $\mu_{t,\tau}$ is less sensitive to changes in $u_{t,\tau-1}$ than $\eta_{t,\tau}^{(1)}$, which is in turn less sensitive than $\eta_{t,\tau}^{(2)}$. Thus we assume that $\kappa_\mu < \kappa_\eta^{(1)} < \kappa_\eta^{(2)}$. Moreover, the scale of trade volume should increase in the wake of positive news. Thus we have $\kappa_\mu > 0$.

3.4 Dynamic cubic spline $s_{t,\tau}$

The periodic component $s_{t,\tau}$ captures the diurnal U-shaped patterns. We define this component based on the cubic spline model of Harvey and Koopman (1993). In order to formally define $s_{t,\tau}$, we first give the definition of a special case, termed a *static* spline, in which the pattern of periodicity does not change over time. Some of the technical details are omitted in the following sections, but we give the complete mathematical construction in Appendix B.

[10] In other applications, the low-frequency component can be an integrated random walk. See Harvey (2013, p.92).

3.4.1 Static daily spline

The cubic spline is termed a *daily* spline if the periodicity is complete over one trading day. The static daily spline assumes that the shape of diurnal patterns is the same for every trading day. The daily spline is a continuous piecewise function of time and connected at $k+1$ knots for some $k \in \mathbb{N}_{>0}$ such that $k < I$. The coordinates of the knots along the time axis are denoted by $\tau_0 < \cdots < \tau_k$, where $\tau_0 = 1$, $\tau_k = I$, and $\tau_j \in \{2, \ldots, I-1\}$ for $j = 1, \ldots, k-1$. The set of knots is also called *mesh*. The y-coordinates (height) of the knots are denoted by $\boldsymbol{\gamma} = (\gamma_0, \ldots, \gamma_k)^\top$. The static daily spline ($s_{t,\tau} = s_\tau$) is defined as

$$s_\tau = \sum_{j=1}^{k} \mathbb{1}_{\{\tau \in [\tau_{j-1}, \tau_j]\}} \, \boldsymbol{z}_j(\tau) \cdot \boldsymbol{\gamma}, \qquad \tau = 1, \ldots, I, \tag{5}$$

where $\boldsymbol{z}_j : [\tau_{j-1}, \tau_j]^{k+1} \to \mathbb{R}^{k+1}$ for $j = 1, \ldots k$ is a $(k+1)$-dimensional vector of deterministic functions that conveys all information about the *polynomial order*, *continuity*, and *zero-sum conditions* of the spline. See Appendix B for the definition of these conditions and the derivation of $\boldsymbol{z}_j(\tau)$. The zero-sum condition ensures that the parameters in $\boldsymbol{\gamma}$ are identified. To impose the zero-sum condition, we also need to set $\gamma_k = -\sum_{i=0}^{k-1} w_{*i}\gamma_i / w_{*k}$, where $\boldsymbol{w}_* = (w_{*0}, \ldots, w_{*k})^\top$ is defined in Appendix B.

3.4.2 Location of daily knots and overnight effect

The locations of internal knots $\tau_1, \ldots, \tau_{k-1}$ and the size of k depend on the empirical shape of diurnal patterns, the number of intra-day observations, and estimation efficiency. Increasing k can sometimes improve the fit of the model, but using too many knots deteriorates estimation efficiency. In the subsequent analysis, we find that positioning internal knots at 11am, 12.30pm, and 2.30pm works well for a daily spline. The shape of the spline up to 12.30pm captures the busy trading hours in the morning, between 12.30pm and 2.30pm captures the quiet lunch hours, and after 2.30pm captures any acceleration in trading activities before closure. There is little to no improvement in the goodness of fit when the number of knots per day increases from this specification. The periodicity condition of the original spline by Harvey and Koopman (1993) sets (τ_k, γ_k) to be the same as (τ_0, γ_0) as their hourly electricity demand data is collected 24 hours on each sampling day. We capture the overnight effect by relaxing this condition in order to allow for a discrepancy in the spline between the end and the beginning of any

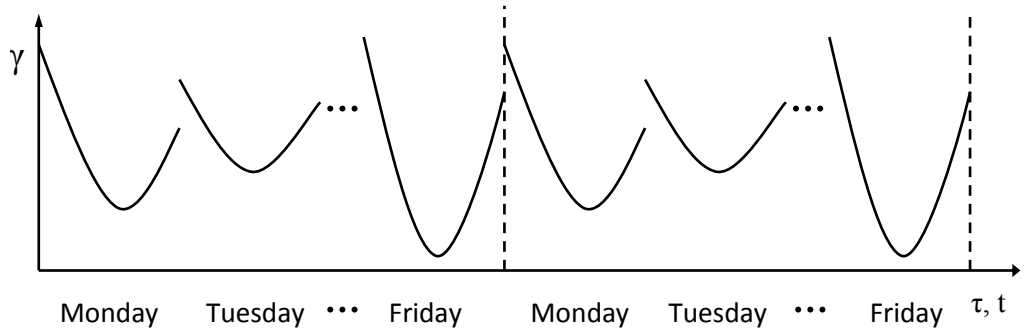

Figure 4 – Picture illustration of the static weekly spline. Both the level and the shape depend on the day of the week.

two consecutive trading days.[11]

3.4.3 Dynamic daily spline

The static daily spline (5) becomes dynamic by letting $\boldsymbol{\gamma}$ be time-varying as

$$s_{t,\tau} = \sum_{j=1}^{k} \mathbb{1}_{\{\tau \in [\tau_{j-1}, \tau_j]\}} \, \boldsymbol{z}_j(\tau) \cdot \boldsymbol{\gamma}_{t,\tau}, \qquad \boldsymbol{\gamma}_{t,\tau} = \boldsymbol{\gamma}_{t,\tau-1} + \boldsymbol{\kappa}^* \cdot u_{t,\tau-1} \tag{6}$$

for $\tau = 1, \ldots, I$ and $t = 1, \ldots, T$, where $\boldsymbol{\kappa}^* = (\kappa_0^*, \ldots, \kappa_k^*)^\top$. Harvey and Koopman (1993) use a set of contemporaneous Gaussian disturbances to drive the dynamics of $\boldsymbol{\gamma}_{t,\tau}$ instead of the lagged score.[12] In terms of parameter identification, our dynamic spline in (6) still satisfies the zero-sum condition over one complete period due to the construction of $\boldsymbol{z}_j(\tau)$, but we also need to impose the parameter restrictions

$$\gamma_{k;1,0} = -\frac{1}{w_{*k}} \sum_{i=0}^{k-1} w_{*i} \gamma_{i;1,0} \qquad\qquad \kappa_k^* = -\frac{1}{w_{*k}} \sum_{i=0}^{k-1} w_{*i} \kappa_i^*. \tag{7}$$

where $\gamma_{i;1,0}$ denotes the ith element of $\boldsymbol{\gamma}_{1,0}$. See Appendix B.

3.4.4 Static weekly spline

The static spline becomes a static *weekly spline* if we set the periodicity of the spline to be complete over one trading week instead of one day. The static weekly spline assumes that the shape of diurnal patterns may be different for each weekday, but the overall shape for the whole week is fixed and remains the same for every trading week. For this spline, we redefine $\tau_0, \tau_1, \ldots, \tau_k$ as follows. We let $\widetilde{\tau}_0 < \widetilde{\tau}_1 < \cdots < \widetilde{\tau}_{k'}$ denote the coordinates along the time-axis of the *intra-day*

[11] This makes the definition of our $\boldsymbol{z}_j(\tau)$ to be different from the one in Harvey and Koopman (1993). See Appendix B.

[12] This makes the identification restrictions on our spline parameters to be different from the ones specified by Harvey and Koopman (1993).

mesh, where $k' < I$, $\widetilde{\tau}_0 = 1$, $\widetilde{\tau}_{k'} = I$, and $\widetilde{\tau}_j \in \{1, \ldots, I-1\}$ for $j = 1, \ldots, k'-1$. Then the coordinates $\tau_0, \tau_1, \ldots, \tau_k$ along the time-axis of the total mesh for the whole week is defined as $\tau_{i(k'+1)+j} = \widetilde{\tau}_j$ for $i = 0, \ldots, 4$ and $j = 0, \ldots, k'$. This means that $(\tau_j)_{j=0}^k$ is no longer an increasing sequence. The total number of knots for one whole week is $k+1 = 5(k'+1)$. The height of the knots are $\gamma_0, \gamma_1, \ldots, \gamma_{k'}$ for Monday, $\gamma_{k'+1}, \gamma_{k'+2}, \ldots, \gamma_{2(k'+1)}$ for Tuesday, and so on. As before, there is no periodicity condition between τ_k and τ_0 so that we can allow for the effect of weekend news on trading patterns. Moreover, we capture the overnight effect of weeknights (between Monday and Friday) by relaxing the continuity and polynomial order restrictions between $\widetilde{\tau}_{k'}$ and $\widetilde{\tau}_0$ of any two successive weekdays. Thus, the procedure for computing $\boldsymbol{z}_j(\tau)$ is different from the daily spline. See Appendix B.

The weekly spline can be used to capture the day-of-the-week effect. It allows for the level as well as the overall shape of the diurnal U-shaped patterns to depend on the day of the week by varying the height of the knots for different weekdays. That is, the weekly spline allows for $(\gamma_0, \gamma_1, \ldots, \gamma_{k'})^\top \neq (\gamma_{k'+1}, \gamma_{k'+2}, \ldots, \gamma_{2(k'+1)})^\top \neq \cdots \neq (\gamma_{4k'+4}, \gamma_{4k'+5}, \ldots, \gamma_{5(k'+1)})^\top$. See Figure 4. We can test for the day-of-the-week effect by a likelihood ratio test under the null hypothesis:

$$H_0 : (\gamma_0, \gamma_1, \ldots, \gamma_{k'})^\top = (\gamma_{k'+1}, \gamma_{k'+2}, \ldots, \gamma_{2(k'+1)})^\top = \cdots = (\gamma_{4k'+4}, \gamma_{4k'+5}, \ldots, \gamma_{5(k'+1)})^\top.$$

The alternative hypothesis replaces $=$ by $\neq$. The likelihood ratio statistic under the null asymptotically has the chi-square distribution with $4(k'+1)$ degrees of freedom. The weekly spline may be preferred over day-of-the-week dummies as a method for capturing the day-of-the-week effect if there are reasons to believe that the day-of-the-week effect not only shifts the level, but also changes the overall shape of the diurnal U-shaped patterns.

One disadvantage of the weekly spline is that the number of total knots for one week increases quickly (fivefold) with the number of daily knots. If we use 5 knots per trading day ($k' = 4$) as we specified for the daily spline, the total number of knots for the week is 25 ($k = 24$). Estimating such a high number of knots can be computationally inefficient, especially when the sampling frequency is high.

3.4.5 Restricted weekly spline

Instead of letting the coordinates of the mesh be free for all weekdays, we can restrict them to be the same on selected days. We term this special case the *restricted weekly spline*. We restrict the diurnal pattern to be the same on mid-weekdays (Tuesday-Thursday) and let the pattern be different on Monday and Friday. This restricted weekly spline captures a special type of the day-of-the-

week effect called the weekend effect, which, in the context of this paper, refers to the tendency of trade volume before and after the weekend to display distinct patterns compared to mid-weekdays. This effect is due to the amount of news disseminated over weekend, which may be more than any of the weeknights.

The evidence for the day-of-the-week effect in financial data is generally mixed and largely depends on the estimation method and the variable being estimated. For instance, Andersen and Bollerslev (1998) found that the day-of-the-week effect is insignificant in the DM-Dollar returns once the calendar effect (e.g. daylight saving and public holidays) and the effects of major macroeconomic announcements are taken into account. However, they found indications of a weak (but clear) seasonality on Monday mornings and Friday afternoons. Lo and Wang (2010) found that some financial returns exhibit a strong day-of-the-week effect. Their volume data as measured by turnover is roughly constant on all days except on Mondays and Fridays when turnover is slightly lower than mid-weekdays. The weekend effect is well-documented in many other studies (mainly in the context of asset returns) and appears to be more pronounced than the more general day-of-the-week effect. Thus, given the computational cost of the full-blown weekly spline, estimating the restricted weekly spline may be adequate and sufficient if one's ultimate goal is to establish a good forecasting model. Formally, we define (the static version of) the restricted weekly spline as:

$$s_\tau = \sum_{j=1}^{k} \mathbb{1}_{\{\tau \in [\tau_{j-1}, \tau_j]\}} \, \boldsymbol{z}_j(\tau) \cdot \mathbf{S}\boldsymbol{\gamma} \tag{8}$$

where $\boldsymbol{\gamma} = (\widetilde{\boldsymbol{\gamma}}_1^\top, \widetilde{\boldsymbol{\gamma}}_2^\top, \widetilde{\boldsymbol{\gamma}}_3^\top)^\top$ and $\widetilde{\boldsymbol{\gamma}}_i$ for $i = 1, 2, 3$ are $(k'+1)$-dimensional mesh vectors for Monday, mid-weekdays (Tuesday-Thursday), and Friday, respectively, and $\boldsymbol{z}_j(\tau)$ is of the weekly spline. $\mathbf{S}$ is the following $5(k'+1) \times 3(k'+1)$ matrix of zeros and ones:

$$\mathbf{S} = \begin{bmatrix} \mathbf{I}_{(k'+1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{(k'+1)} & \mathbf{0} \\ \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{I}_{(k'+1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{(k'+1)} \end{bmatrix}.$$

where $\mathbf{I}_{(k'+1)}$ is the identity matrix of size $(k'+1)$. We can rewrite (8) as

$$s_\tau = \sum_{j=1}^{k} \mathbb{1}_{\{\tau \in [\tau_{j-1}, \tau_j]\}} \, \widetilde{\boldsymbol{z}}_j(\tau) \cdot \boldsymbol{\gamma} \tag{9}$$

where $\widetilde{\boldsymbol{z}}_j(\tau) = \mathbf{S}^\top \boldsymbol{z}_j(\tau)$. Finally, we can let $\boldsymbol{\gamma}$ be time-varying according to the dynamics in (6) with appropriate adjustments to the zero-sum conditions. Then

we obtain the dynamic restricted weekly spline. In this case, we use the notations $\boldsymbol{\kappa}^* = (\widetilde{\boldsymbol{\kappa}}_1^{*\top}, \widetilde{\boldsymbol{\kappa}}_2^{*\top}, \widetilde{\boldsymbol{\kappa}}_3^{*\top})^\top$ and $\boldsymbol{\gamma}_{t,\tau} = (\widetilde{\boldsymbol{\gamma}}_{1;t,\tau}^\top, \widetilde{\boldsymbol{\gamma}}_{2;t,\tau}^\top, \widetilde{\boldsymbol{\gamma}}_{3;t,\tau}^\top)^\top$ for $(t,\tau) \in \Psi_{T,I}$, where $\dim(\widetilde{\boldsymbol{\kappa}}_j^*) = \dim(\widetilde{\boldsymbol{\gamma}}_{j;t,\tau}) = (k'+1)$ for $j = 1,2,3$. If we place 5 knots per trading day (i.e. $k' = 4$) as we specified for the daily spline, we have 15 unrestricted knots for one week.

3.5 Summary of the model specification

To summarize, our spline-DCS model is

$$y_{t,\tau} = \varepsilon_{t,\tau} \exp(\lambda_{t,\tau}), \quad \lambda_{t,\tau} = \omega + \mu_{t,\tau} + \eta_{t,\tau} + s_{t,\tau}, \quad \varepsilon_{t,\tau} | \mathcal{F}_{t,\tau-1} \sim \texttt{iid} F,$$

where the components are

$$\begin{aligned}
\mu_{t,\tau} &= \mu_{t,\tau-1} + \kappa_\mu u_{t,\tau-1} \\
\eta_{t,\tau} &= \eta_{t,\tau}^{(1)} + \eta_{t,\tau}^{(2)} \\
\eta_{t,\tau}^{(j)} &= \phi_1^{(j)} \eta_{t,\tau-1}^{(j)} + \phi_2^{(j)} \eta_{t,\tau-2}^{(j)} \cdots + \phi_{m^{(j)}}^{(j)} \eta_{t,\tau-m^{(j)}}^{(j)} + \kappa_\eta^{(j)} u_{t,\tau-1}, \quad j = 1,2 \\
s_{t,\tau} &= \sum_{j=1}^{k} \mathbb{1}_{\{\tau \in [\tau_{j-1}, \tau_j]\}} \boldsymbol{z}_j(\tau) \cdot \boldsymbol{\gamma}_{t,\tau} \\
\boldsymbol{\gamma}_{t,\tau} &= \boldsymbol{\gamma}_{t,\tau-1} + \boldsymbol{\kappa}^* \cdot u_{t,\tau-1}
\end{aligned}$$

for $(t,\tau) \in \Psi_{T,I}$, where F denotes the distribution defined in (3) and has the parameter $p = \mathbb{P}(y_{t,\tau} = 0)$ for any $(t,\tau) \in \Psi_{T,I}$. $u_{t,\tau}$ is the score of conditional distribution F^*, and we set $u_{t,\tau} = \inf_{s \in \Omega} u_{t,\tau}(s)$ whenever $y_{t,\tau} = 0$. $\boldsymbol{z}_j(\tau)$ is to be replaced by $\widetilde{\boldsymbol{z}}_j(\tau)$ for the restricted weekly spline.

3.5.1 Public holidays

Public holidays can be treated in the same way as overnight periods or weekends where data is missing due to market closure. Whenever $t \in \mathcal{H} \equiv \{t \in \mathbb{N}_{>0} : t \text{ is a public holiday}\}$, we set

$$\mu_{t,\tau} = \mu_{t,\tau-1}, \quad \eta_{t,\tau}^{(1)} = \eta_{t,\tau-1}^{(1)}, \quad \eta_{t,\tau}^{(2)} = \eta_{t,\tau-1}^{(2)}, \quad \boldsymbol{\gamma}_{t,\tau} = \boldsymbol{\gamma}_{t,\tau-1}$$

for all $\tau = 1, \ldots, I$ so that $\lambda_{t,\tau}$ is unchanged. The joint log-likelihood is defined only for days $t \in \mathcal{H}^c$.

3.6 Maximum likelihood estimation

All parameters of the model can be estimated by the method of maximum likelihood using the log-likelihood function in (4). We do not formally verify that the asymptotic results of the MLEs shown in Harvey (2013) also extend to our particular specification in which the scale parameter $\alpha_{t,\tau}$ (or $\lambda_{t,\tau}$) is nonstationary. This is left for future research. We choose the initial values of the components of

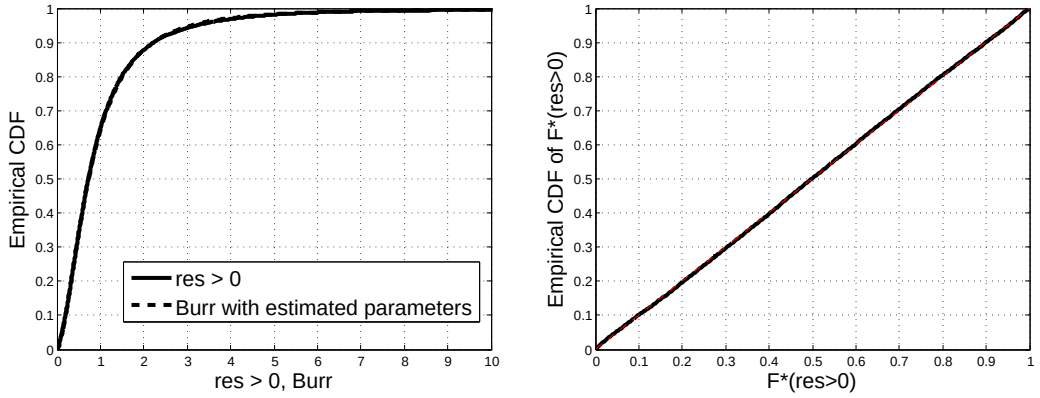

Figure 5 – Empirical cdf of $\widehat{\varepsilon}_{t,\tau} > 0$ against cdf of Burr$(\widehat{\nu}, \widehat{\zeta})$ (left). Empirical cdf of the PIT of $\widehat{\varepsilon}_{t,\tau} > 0$ when $F^*(\cdot; \widehat{\boldsymbol{\theta}}^*)$ is Burr$(\widehat{\nu}, \widehat{\zeta})$ (right). IBM30s over 28 February - 31 March 2000. Using Model 2 specified in Table 5.

$\lambda_{1,0}$ as follows. We set $\eta_{1,0} = 0$ as we have $\mathbb{E}[\eta^{(1)}_{t,\tau}] = \mathbb{E}[\eta^{(2)}_{t,\tau}] = 0$. As $\mu_{t,\tau}$ is assumed to be a random-walk, we have $\mathbb{E}[\mu_{t,\tau}] = \mu_{1,0}$. So we impose the constraint $\mu_{1,0} = 0$ for ω to be identified. For $s_{1,0}$ or $\boldsymbol{\gamma}_{1,0}$, we treat the elements of $\boldsymbol{\gamma}_{1,0}$ as unknown constant parameters to be estimated simultaneously with all other parameters of the model. We impose the zero-sum constraint $\boldsymbol{w}_* \cdot \boldsymbol{\gamma}_{1,0} = 0$ as specified in (7).

4 Estimation results

This section reports the estimation results of fitting our model to the series we investigated in Section 2. In order to test the sensitivity of our inference to changes in the aggregation interval, we report the estimation results for both IBM30s and IBM1m. We denote the estimated quantities by $\widehat{\cdot}$.

4.1 Choice of F^*

Given the shape of empirical distribution of data, distributions in the GG, GB2, and Pareto classes are among our candidate F^* for both IBM1m and IBM30s. Note that the GB2 distribution is closely related to the Pareto class of distributions. See Appendix A and Figure 3 for the formal definitions of these distributions and the relationship between them. Taking a general-to-specific approach, we sequentially estimate GB2, Burr, and then log-logistic from the GB2 class of distributions. Within the GG class of distributions, we sequentially estimate GG, and then Gamma and Weibull.[13]

We find that the GB2 distribution is difficult to estimate as it has three param-

[13]We find this general-to-specific approach easy to implement as GB2 and GG nest many useful distributions. One is typically required to write estimation programs for GB2 and GG first, and then simply impose parameter restrictions to fit many distributions nested within GB2 and GG.

eters in $\boldsymbol{\theta}^*$. Burr fits both IBM30s and IBM1m very well. As the GG distribution is a limiting distribution of GB2 for when ζ is large, GG can be considered as a special class of GB2. However, the goodness of fit of the GG class of distributions are found to be inferior to the GB2 class of distributions. The inferior fit of GG is consistent with the estimated size of the GB2 parameter ζ, which is far from being large for both IBM1m and IBM30s. (See Table 2.)

Figure 5 illustrates the impressive fit of Burr for IBM30s. The quality of fit for IBM1m is equally as impressive. The empirical cdf of non-zero $\widehat{\varepsilon}_{t,\tau}$ overlaps the cdf of Burr($\widehat{\nu}$,$\widehat{\zeta}$) so that these two lines are visually indistinguishable. The empirical cdf of the probability integral transform (PIT) of non-zero $\widehat{\varepsilon}_{t,\tau}$ when $F^*(\cdot;\widehat{\boldsymbol{\theta}}^*)$ is Burr($\widehat{\nu},\widehat{\zeta}$) lies along the diagonal, indicating that the PIT values are remarkably close to being standard uniformly distributed (denoted by $U[0,1]$). The computing time taken for a maximum likelihood estimation procedure to converge when F^* is Burr is generally very short (a few seconds) and the results are robust to the choice of initial parameter values.

As an experiment, we have also fitted many other distributions including inverse-Gamma and log-Cauchy. However, none achieved the closeness of fit of the Burr distribution.

4.2 Comparing with log-normality

The log-normal distribution is widely used to fit non-negative time series whenever the logarithm of observations roughly resembles normality. In such cases, the degree of efficiency and bias in the estimated parameters depends on how far the distribution of the log-variable is from normality. See, for example, Alizadeh et al. (2002) for a discussion and an empirical example.

In our case, the fit of log-normal turns out to be inferior to that of Burr for both IBM1m and IBM30s.[14] This is due to the departure of the logarithm of non-zero data (hereafter denoted by log(IBM1m) and log(IBM30s)) from normality particularly around the tail regions. This is illustrated in Figure 6. While the frequency distribution of log(IBM1m) roughly resembles normality, log(IBM30s) is clearly far from normal. The QQ-plots show that the normal distribution appears to put too much weight on the lower-tail and too little weight on the upper-tail.

Burr fits better than log-normal presumably because the shape of log-normal is determined by only one parameter (σ) while Burr has two shape parameters (ν and ζ). To compare the shape of Burr and log-normal, Figure 7 shows the pdf of

[14]The Bowman-Shenton (Jarque-Bera) test comfortably rejects normality of the logarithm of the estimation residuals when F^* is log-normal for both IBM1m and IBM30s.

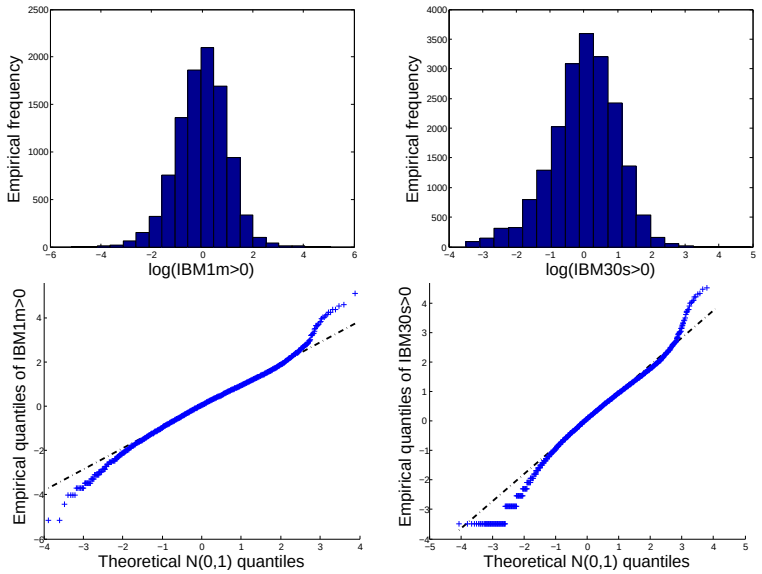

Figure 6 – The frequency distribution (top panel) and the QQ-plots (bottom panel): log(IBM1m) (left column) and log(IBM30s) (right column). The logarithm series are re-centered around mean and standardized by one standard deviation.

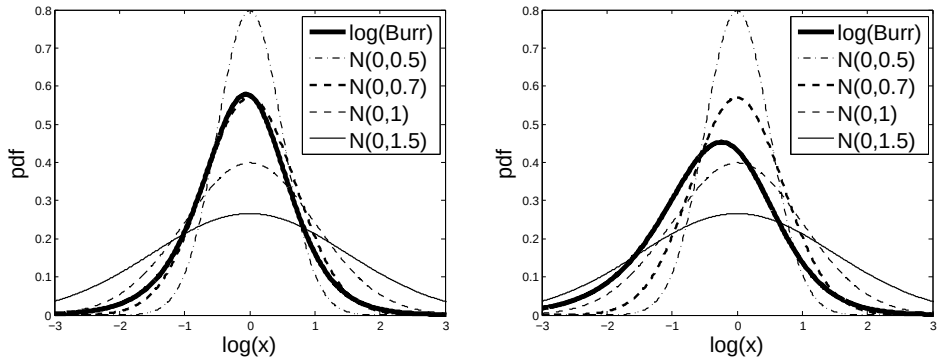

Figure 7 – The pdf of $\log(X)$: $X \sim \text{Burr}(\widehat{\nu}, \widehat{\zeta})$ against $X \sim \text{log-normal}(\sigma)$ for different values of σ. $\widehat{\nu}$ and $\widehat{\zeta}$ are obtained by fitting Model 2 of Table 5 to IBM1m (left) and IBM30s (right).

$\log(X)$ when $X \sim \text{Burr}(\widehat{\nu}, \widehat{\zeta})$ against the case when X is log-normally distributed with different values of σ. For both IBM1m and IBM30s, the asymmetric shape of the pdf of $\log(X)$ for $X \sim \text{Burr}(\widehat{\nu}, \widehat{\zeta})$ contrasts with the symmetric shape of the normal distribution.

4.3 Estimated coefficients and diagnostics

We find that empirically a two component specification (i.e. $J = 2$) for $\eta_{t,\tau}$ is the most effective for capturing the autocorrelation structure of data. The autoregressive structure of order two for $\eta^{(1)}_{t,\tau}$ and one for $\eta^{(2)}_{t,\tau}$ was found to be appropriate for both IBM1m and IBM30s. That is, our preferred specification for

	IBM30s	IBM1m		IBM30s	IBM1m
κ_μ	0.006 (0.001)	0.007 (0.002)	$\gamma_{0;1,0}$	1.273 (0.106)	1.195 (0.098)
$\phi_1^{(1)}$	0.557 (0.136)	0.377 (0.093)	$\gamma_{1;1,0}$	0.078 (0.058)	0.075 (0.057)
$\phi_2^{(1)}$	0.410 (0.135)	0.567 (0.096)	$\gamma_{2;1,0}$	-0.469 (0.070)	-0.450 (0.069)
$\kappa_\eta^{(1)}$	0.049 (0.007)	0.045 (0.008)	$\gamma_{3;1,0}$	-0.227 (0.047)	-0.244 (0.047)
$\phi_1^{(2)}$	0.688 (0.041)	0.621 (0.057)	ω	9.146 (0.174)	9.752 (0.155)
$\kappa_\eta^{(2)}$	0.092 (0.008)	0.069 (0.008)	ν	1.631 (0.016)	2.230 (0.033)
κ_0^*	0.003 (0.002)	0.003 (0.002)	ζ	1.486 (0.045)	1.142 (0.044)
κ_1^*	0.001 (0.001)	0.000 (0.001)	p	0.0047 (0.0005)	0.0006 (0.0003)
κ_2^*	-0.002 (0.001)	-0.002 (0.001)			
κ_3^*	0.000 (0.001)	0.000 (0.001)			

Table 2 – Estimated coefficients when $\eta_{t,\tau}$ has two components and $s_{t,\tau}$ is the daily spline ('Model 2' in Table 5). F^* is Burr. Standard errors in parenthesis are computed using numeric derivatives of the likelihood function with respect to the parameters.

Series	Model specification	Type of spline	*df*	Loglike $\zeta = 1$	Loglike $\zeta \neq 1$	Likelihood ratio stat	χ^2_{df}, p-value
IBM30s	Model 2	Daily	1	-195,584	-195,487	195.6	0.000
	Model 4	Weekly	1	-195,572	-195,475	195.4	0.000
IBM1m	Model 2	Daily	1	-104,119	-104,113	11.8	0.001
	Model 4	Weekly	1	-104,103	-104,097	11.8	0.001

Table 3 – Likelihood ratio statistics to test the null $H_0 : \zeta = 1$ (log-logistic) against the alternative $H_1 : \zeta \neq 1$ (Burr). Model specifications are in Table 5.

$\eta_{t,\tau}$ is $\eta_{t,\tau} = \eta_{t,\tau}^{(1)} + \eta_{t,\tau}^{(2)}$ with

$$\eta_{t,\tau}^{(1)} = \phi_1^{(1)}\eta_{t,\tau-1}^{(1)} + \phi_2^{(1)}\eta_{t,\tau-2}^{(1)} + \kappa_\eta^{(1)}u_{t,\tau-1}, \qquad \eta_{t,\tau}^{(2)} = \phi_1^{(2)}\eta_{t,\tau-1}^{(2)} + \kappa_\eta^{(2)}u_{t,\tau-1} \tag{10}$$

for $(t,\tau) \in \Psi_{T,\,I}$.

Table 2 shows the estimated coefficients of our model when $\eta_{t,\tau}$ has two components and $s_{t,\tau}$ is the daily spline ('Model 2' in Table 5). The coefficient estimates for all other specifications we considered are reported in Appendix C. For both IBM1m and IBM30s, we have $\widehat{\kappa}_\eta^{(2)} > \widehat{\kappa}_\eta^{(1)} > \widehat{\kappa}_\mu > 0$ as expected, which means that $\eta_{t,\tau}^{(2)}$ is more sensitive to changes in $u_{t,\tau-1}$ than $\eta_{t,\tau}^{(1)}$, and that $\eta_{t,\tau}^{(1)}$ is more sensitive to $u_{t,\tau-1}$ than $\mu_{t,\tau}$. We also have $0 < \widehat{\phi}_1^{(2)} < 1$ so $\eta_{t,\tau}^{(2)}$ is stationary. The stationarity of $\eta_{t,\tau}^{(1)}$ requires[15]

$$\phi_1^{(1)} + \phi_2^{(1)} < 1, \quad -\phi_1^{(1)} + \phi_2^{(1)} < 1, \quad \text{and} \quad \phi_2^{(1)} > -1. \tag{11}$$

It is easy to check that these conditions are satisfied by $\widehat{\phi}_1^{(1)}$ and $\widehat{\phi}_2^{(1)}$ so that $\eta_{t,\tau}^{(1)}$ is also stationary. The estimate of the probability mass at zero is $\widehat{p} = 0.0006$

[15] See, for example, Harvey (1993, p.19).

	Model spec.	No. of coef.	Type of spline	AIC	SIC
IBM30s	——	14	Static ($\underline{\gamma}^{\dagger}$)	20.02608	20.032
	Model 2	18	Dynamic ($\underline{\gamma}^{\dagger}_{t,\tau}$)	20.02608	20.033
IBM1m	——	14	Static ($\underline{\gamma}^{\dagger}$)	21.30550	21.316
	Model 2	18	Dynamic ($\underline{\gamma}^{\dagger}_{t,\tau}$)	21.30553	21.319

Table 4 – Static versus dynamic splines. The criterions for static spline are computed by setting $\boldsymbol{\kappa}^* = \mathbf{0}$ in Model 2. Model 2 is as specified in Table 5. F^* is Burr.

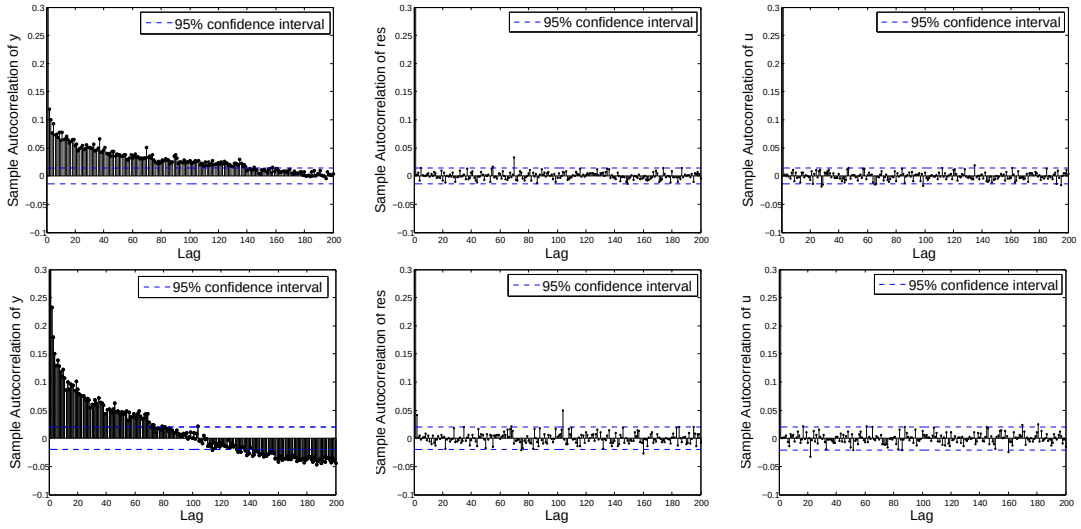

Figure 8 – Sample autocorrelation of trade volume (left), of $\widehat{\varepsilon}_{t,\tau}$ (middle), and of $\widehat{u}_{t,\tau}$ (right). IBM30s (top row) and IBM1m (bottom row). Using Model 2 as specified in Table 5. The 95% confidence interval is computed at ± 2 standard errors.

for IBM1m and $\widehat{p} = 0.0047$ for IBM30s. These estimates are consistent with the sample statistics in Table 1. The estimate of the Burr parameter ($\widehat{\zeta}$) is not far from one. If $\zeta = 1$, Burr becomes log-logistic. A likelihood test rejects the null hypothesis that $\zeta = 1$. See Table 3.[16] We have $\widehat{\nu}\widehat{\zeta} \approx 2.5$ for IBM1m and $\widehat{\nu}\widehat{\zeta} \approx 2.4$ for IBM30s, implying that only the first and second moments exist under the assumption that F^* is Burr. (See Appendix A.1.) Thus the theoretical skewness does not exist for both IBM1m and IBM30s.

The estimates of some of the elements in $\boldsymbol{\kappa}^*$ are very close to zero and appear statistically insignificant. We can assess whether a dynamic spline is preferred over a static one by comparing AIC and SIC.[17] They are tabulated in Table 4. We find that SIC is slightly larger for the dynamic spline, while the results are mixed for AIC. Hence the static spline is preferred over the dynamic spline according to SIC, whereas AIC is inconclusive. SIC penalizes the inclusion of $\boldsymbol{\kappa}^*$ more severely

[16] This can be also tested using other likelihood based tests such as a Wald test.
[17] AIC and SIC stand for the Akaike and Schwarz information criteria, respectively.

	IBM30s				IBM1m			
Model #	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
Spline	Daily	Daily	Weekly	Weekly	Daily	Daily	Weekly	Weekly
$\eta_{t,\tau}$[a]	AR(1)	(10)	AR(1)	(10)	AR(1)	(10)	AR(1)	(10)
Dyn.coef.	7	10	17	20	7	10	17	20
AIC	20.0289	20.0261	20.0289	20.0269	21.3094	21.3055	21.3087	21.3065
SIC	20.0350	20.0333	20.0430	20.0422	21.3204	21.3188	21.3345	21.3344
$\widehat{\varepsilon}_{t,\tau}$: Q(100)[b]	—	—	—	—	93.3	87.0	94.5	90.5
Q(150)	106.7	108.1	105.2	103.2	152.8	149.8	158.2	154.9
Q(200)	143.5	143.0	137.9	134.8	195.2	194.3	199.4	197.9
$\widehat{u}_{t,\tau}$: Q(100)	—	—	—	—	124.0*	88.2	122.2**	91.1
Q(150)	191.9**	141.9	190.9**	142.4	161.9	124.0	158.3	126.9
Q(200)	236.2*	182.7	233.5**	183.2	204.5	167.3	201.5	169.9

[a] $\eta_{t,\tau}$ has either the one-component AR(1) structure or the two-component specification in (10).
[b] $Q(K)$ is the Box-Ljung statistic based on the first K autocorrelation computed under the null hypothesis of no autocorrelation up to order K. $Q(K)$ is compared against the chi-square distribution with $K-q$ degrees of freedom (denoted by χ^2_{K-q}), where q is the number of dynamic parameters of the model (e.g. the basic DCS in (2) has two dynamic parameters ϕ and κ). We report the results for $K \approx \sqrt{T \times I}$ and some deviations above it following Harvey (1993, p.76). * and ** denote the significance at the 5% and 1% levels, respectively.

Table 5 – Goodness-of-fit statistics. F^* is Burr and all specifications use the dynamic cubic spline with 5 knots per day. "Weekly" refers to the restricted weekly spline.

	Model specification	*df*	Loglike Restricted	Loglike Unrestricted	Likelihood ratio stat	χ^2_{df}, p-value
IBM30s	Model 2 & 4	18	-195,487	-195,475	24.2	0.149
IBM1m	Model 2 & 4	18	-104,113	-104,097	31.0	0.029

Table 6 – Likelihood ratio test for the weekend effect. Using the dynamic spline. Sample period: 28 February - 31 March 2000. Model specifications are in Table 5.

than AIC.

Figure 8 shows that the slowly decaying autocorrelation in IBM30s and IBM1m is captured very well by the model as the estimation residuals $\widehat{\varepsilon}_{t,\tau}$ and the score $\widehat{u}_{t,\tau}$ show no obvious signs of serial correlation. The absence of autocorrelation is also verified more formally by the Box-Ljung statistics in Table 5. Note that $\widehat{u}_{t,\tau}$ exhibits stronger serial correlation than $\widehat{\varepsilon}_{t,\tau}$. This is because the score down-weighs (and thus it is robust to) the effects of extreme observations. We find that both AIC and SIC fall and the autocorrelation in $\widehat{u}_{t,\tau}$ is removed whenever we change $\eta_{t,\tau}$ from the one-component AR(1) structure to the two component specification in (10).

4.3.1 Test for the weekend effect

In Table 5, we observe that SIC always increases (while the results are mixed for AIC) when we change the specification from daily to restricted weekly spline. This is largely due to the inclusion of ten extra parameters in the restricted weekly spline, which is penalized severely by SIC. We can test whether the restricted weekly spline is preferred over the daily spline by a likelihood ratio test. The null hypothesis[18] of the test is

$$H_0 : \widetilde{\boldsymbol{\kappa}}_1^* = \widetilde{\boldsymbol{\kappa}}_2^* = \widetilde{\boldsymbol{\kappa}}_3^* \quad \text{and} \quad \widetilde{\boldsymbol{\gamma}}_{1;1,0} = \widetilde{\boldsymbol{\gamma}}_{2;1,0} = \widetilde{\boldsymbol{\gamma}}_{3;1,0}.$$

The alternative hypothesis replaces $=$ by $\neq$. As $\dim(\widetilde{\boldsymbol{\gamma}}_{j;1,0}) = \dim(\widetilde{\boldsymbol{\kappa}}_j^*) = 5$ for $j = 1, 2, 3$, the likelihood ratio statistics are compared against the chi-square distribution with eighteen degrees of freedom.[19] If the null hypothesis is rejected, there is statistical evidence for the weekend effect over our sampling period. The test results reported in Table 6 depend on the aggregation interval. The null cannot be rejected at the 5% level for IBM30s, but it is rejected at the 5% level for IBM1m.

4.4 Estimated components

Figure 9 shows $\widehat{\varepsilon}_{t,\tau}$ and the estimated components of $\widehat{\lambda}_{t,\tau}$ of Model 2 for IBM30s. (The results for IBM1m are very similar to Figure 9 and so they are omitted.) All series are displayed over the entire sampling period between 28 February and 31 March 2000. While the time series plot of IBM30s clearly exhibits periodic patterns, $\widehat{\varepsilon}_{t,\tau}$ appears free of periodicity. The diurnal U-shaped patterns are captured by $\widehat{\lambda}_{t,\tau}$ through the spline $\widehat{s}_{t,\tau}$. $\widehat{\mu}_{t,\tau}$ and $\widehat{\eta}_{t,\tau}$ appear to satisfy their structural assumptions as $\widehat{\mu}_{t,\tau}$ resembles a random-walk while $\widehat{\eta}_{t,\tau}^{(1)}$ and $\widehat{\eta}_{t,\tau}^{(2)}$ resemble stationary AR(2) and AR(1) processes, respectively.

Figure 10 shows the estimated daily spline, $\widehat{s}_{t,\tau}$, for IBM30s. We find that $\widehat{s}_{t,\tau}$ successfully captures the tendency of trade volume to be high in the morning, fall during the quiet lunch hours of around 1pm, and pick-up again in the afternoon. As it is a dynamic spline, the shape of the diurnal U-shaped pattern varies over time. The spline takes a step-increase between the end and the beginning of any two consecutive trading days, reflecting the overnight effect.

[18] We can also perform the test using the static spline, in which case the null and alternative hypotheses of the test become

$$H_0 : \widetilde{\gamma}_1 = \widetilde{\gamma}_2 = \widetilde{\gamma}_3 \text{ (daily spline)}, \quad H_1 : \widetilde{\gamma}_1 \neq \widetilde{\gamma}_2 \neq \widetilde{\gamma}_3 \text{ (restricted weekly spline)}.$$

[19] Not twenty degrees of freedom due to the zero-sum conditions.

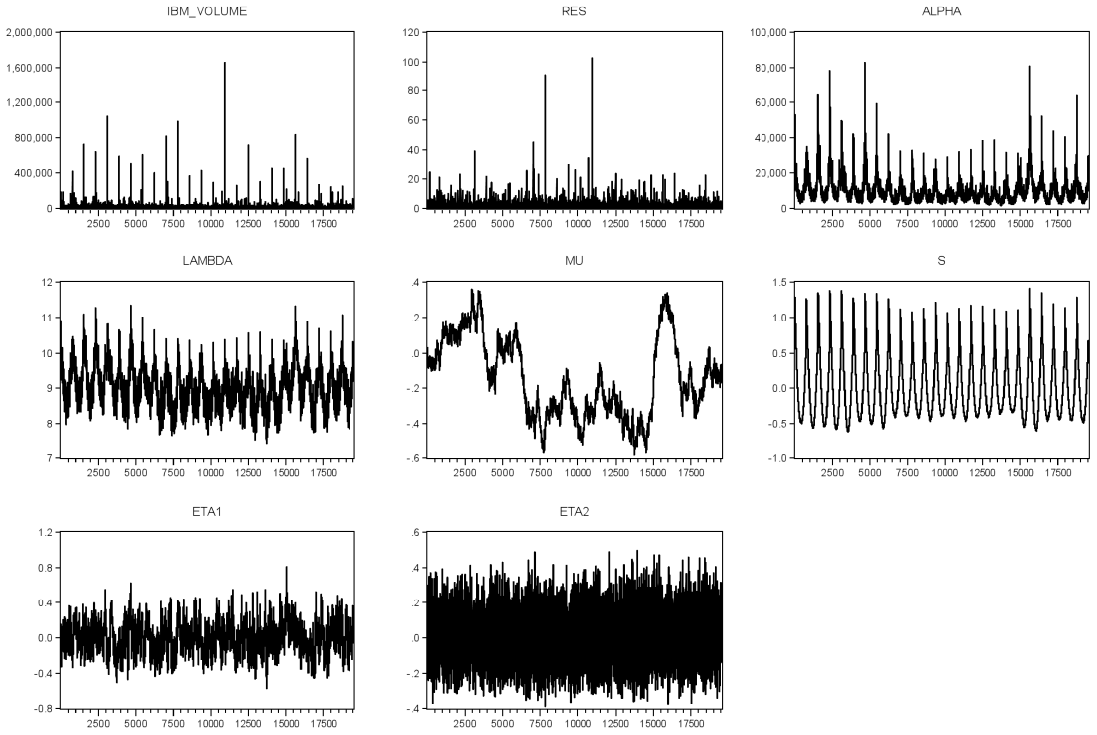

Figure 9 – IBM30s (top left), $\widehat{\varepsilon}_{t,\tau}$ (top middle), $\widehat{\alpha}_{t,\tau}$ (top right), $\widehat{\lambda}_{t,\tau}$ (second-row left), $\widehat{\mu}_{t,\tau}$ (second-row middle), $\widehat{s}_{t,\tau}$ (second-row right), $\widehat{\eta}^{(1)}_{t,\tau}$ (bottom left), and $\widehat{\eta}^{(2)}_{t,\tau}$ (bottom middle) of Model 2 fitted to IBM30s over 28 February - 31 March 2000. Time along the x-axes.

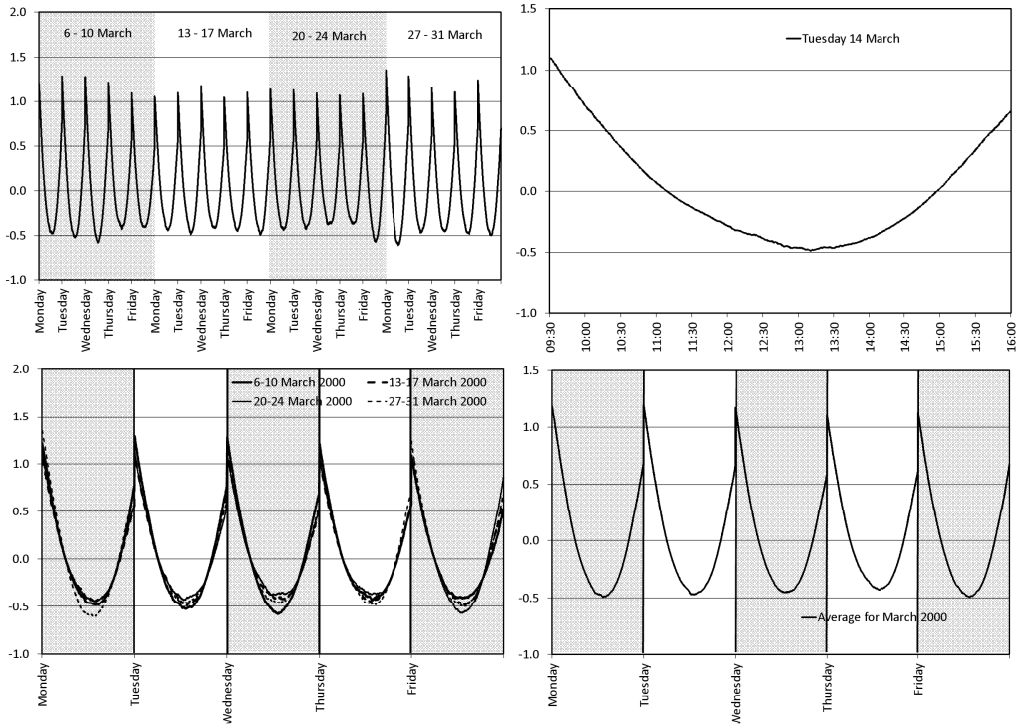

Figure 10 – $\widehat{s}_{t,\tau}$ of Model 2 for IBM30s. Over 6 - 31 March 2000 (top left). $\widehat{s}_{t,\tau}$ of a typical day, Tuesday 14 March, from market open to close (top right). $\widehat{s}_{t,\tau}$ of different weeks over 6 - 31 March 2000 superimposed (bottom left). The weekly average of $\widehat{s}_{t,\tau}$ over 6 - 31 March 2000 (bottom right). Time along the x-axes.

5 Out-of-sample performance

5.1 Model stability: one-step ahead forecasts

We use the predictive cdf to assess the stability of the estimated distribution and parameters as well as the ability of our model to produce good one-step ahead forecasts over a given out-of-sample prediction period. The procedure is as follows. Henceforth, we use the following notations

$$\begin{aligned}\Psi_h &= \{(t,\tau) \in \{T+1,\ldots,T+h\} \times \{1,\ldots,I\}\} \\ \Psi_{h,>0} &= \{(t,\tau) \in \{T+1,\ldots,T+h\} \times \{1,\ldots,I\} : y_{t,\tau} > 0\}\end{aligned}$$

Given a set of in-sample observations up to trading day T, we compute the predictive cdf for the next h trading-days at each positive observation as

$$F^*(\widehat{\varepsilon}_{t,\tau}; \widehat{\boldsymbol{\theta}}^*), \quad \widehat{\varepsilon}_{t,\tau} = y_{t,\tau}/\widehat{\alpha}_{t,\tau} \qquad \forall (t,\tau) \in \Psi_{h,>0} \tag{12}$$

where all parameters are the maximum likelihood estimates computed using the observations up to day T. We take the estimation results from Section 4 to compute (12). The predictive cdf in (12) simply gives the PIT values of future observations.[20] Figure 11 shows the empirical cdf of the PIT values for IBM30s over different forecast horizons up to $h = 20$ days ahead. (The results for IBM1m are very similar to Figure 11.) As we have 390 observations per day for IBM1m and 780 observations per day for IBM30s, $h = 20$ corresponds to 7,800 steps ahead for IBM1m and 156,000 steps ahead for IBM30s. We find that the distribution of the PIT values is roughly $U[0,1]$ for much of the 20 days of forecast horizon for both IBM1m and IBM30s, although there is non-negligible deterioration in the quality of fit. For IBM1m, the empirical cdf of the PIT values is particularly close to uniformity when $h = 5$ and $h = 20$. For IBM30s, it is the closest when $h = 5$.

In summary, our estimated distributions and parameter values appear to be fairly stable and able to provide reasonable one-step-ahead forecasts of the conditional distribution of our series for up to 20 days of prediction horizon.

[20] We assess the closeness of the PIT values to uniformity only qualitatively by inspecting the empirical cdf of (12). One can formally test this using the predictive likelihood methods discussed in Mitchell and Wallis (2011). In our application, the predictive (log-)likelihood of a single observation is

$$\log l_{t,\tau} = \mathbb{1}_{\{y_{t,\tau}>0\}} \log(1-\widehat{p}\,) + \big(1 - \mathbb{1}_{\{y_{t,\tau}>0\}}\big) \log \widehat{p} + \mathbb{1}_{\{y_{t,\tau}>0\}} \log\Big(\widehat{\alpha}_{t,\tau}^{-1} f^*(\widehat{\varepsilon}_{t,\tau}; \widehat{\boldsymbol{\theta}}^*)\Big)$$

for all $(t,\tau) \in \Psi_h$.

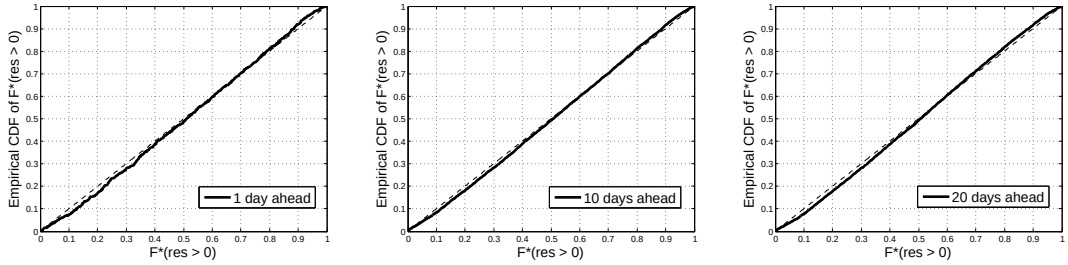

Figure 11 – Empirical cdf of the PIT values (predictive cdf) for one-step ahead forecasts. Forecast horizons: 1 day ahead (left), 10 days ahead (middle), 20 days ahead (right). IBM30s for forecast horizons between 3 - 23 April 2000. Using Model 2 as specified in Table 5.

5.2 Multi-step ahead forecasts

We now examine multi-step forecasts, which are of greater interest than one-step forecasts as they give us an ultimate assessment of our model's predictive ability. We produce multi-step ahead density forecasts over a long forecast horizon using the estimation results obtained in Section 4. The procedure is as follows. We make optimal forecasts of $\alpha_{t,\tau} = \exp(\lambda_{t,\tau})$ for $(t,\tau) \in \Psi_h$ conditional on $\mathcal{F}_{T,I}$. This optimality is in terms of MMSE so that the prediction is $\mathbb{E}\left[\exp(\lambda_{t,\tau})|\mathcal{F}_{T,I}\right] \equiv \widetilde{\alpha}_{t,\tau}$ for $(t,\tau) \in \Psi_h$.[21] We then standardize the actual future observations $y_{t,\tau}$ by $\widetilde{\alpha}_{t,\tau}$ for all $(t,\tau) \in \Psi_h$. The standardized future observations $\widetilde{\varepsilon}_{t,\tau} \equiv y_{t,\tau}/\widetilde{\alpha}_{t,\tau}$ should be conditionally distributed as Burr($\widehat{\nu}$, $\widehat{\zeta}$) at least approximately. Then $F^*(\widetilde{\varepsilon}_{t,\tau}; \widehat{\boldsymbol{\theta}}^*)$ for $(t,\tau) \in \Psi_{h,>0}$ gives the PIT values of future observations, and they should be distributed approximately as $U[0,1]$ and be free of autocorrelation if the predictions are good. We make predictions for up to $h = 8$ future trading days, which corresponds to 3,120 steps ahead for IBM1m and 6,240 steps ahead for IBM30s.

We find that the distribution of the PIT values is very close to $U[0,1]$ for the first day of forecast horizon for both IBM1m and IBM30s. See Figure 12. (We omit the results for IBM1m again as they are very similar to IBM30s.) This means that the density prediction produced by our model is very good for the first 390 steps for IBM1m and 780 steps for IBM30s. A Box-Ljung test indicates that the PIT values for the first quarter of the first forecast day, equivalent to around 200 steps (100 steps) ahead for IBM30s (IBM1m), are not serially correlated. See Figure 13. By the end of the first forecast day, the PIT becomes autocorrelated, although the degree of autocorrelation remains very small. These results are again similar for IBM1m. If we consider this predictive performance in terms of the number of

[21] Harvey (2013) outlines the forecasting methodology for the basic scale DCS model (1)-(2). The difference between notations $\widehat{\cdot}_{t,\tau}$ and $\widetilde{\cdot}_{t,\tau}$ is that the computation of the former does not require the use of conditional moment conditions.

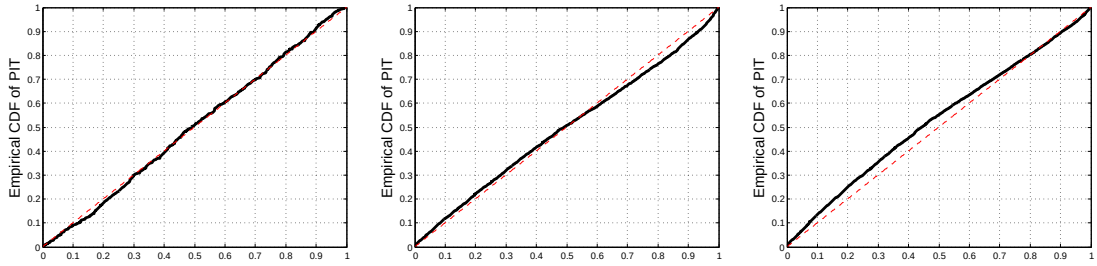

Figure 12 – Empirical cdf of the PIT of multi-step forecasts. Forecast horizons: 1 day ahead (left), 5 days ahead (middle), 8 days ahead (right). IBM30s for forecast horizons between 3 - 23 April 2000. Using Model 2 as specified in Table 5.

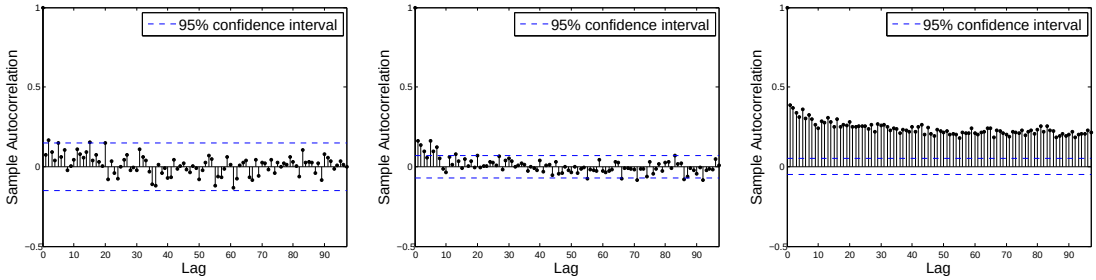

Figure 13 – Autocorrelation of the PIT of multi-step forecasts. Forecast horizons: a quarter of a day ahead (left), one day ahead (middle), two days ahead (right). IBM30s for forecast horizons between 3 - 23 April 2000. Using Model 2 as specified in Table 5.

steps, these results are impressive as they imply that the PIT values are close to being `iid` $U[0,1]$ for several hundred steps of the forecast horizon.

Beyond the first forecast day, the quality of density forecasts and the degree of autocorrelation deteriorate with the length of forecast horizon. Any deterioration in the goodness of fit over a given out-of-sample period can be attributed to changes in fundamental factors such as the nature of periodicity, autocorrelation, and the shape of the underlying distribution. In order to reflect these changes, the model should be re-estimated frequently as more data becomes available. When updating the model, one may want to keep the length of the sampling period roughly fixed and roll it forward instead of arbitrarily increasing the sample size because the estimation efficiency deteriorates relatively quickly as the number of sampling days T increases in a high-frequency environment with large I.

6 Concluding remarks

This paper developed a spline-DCS model for high-frequency observations of a financial variable with diurnal patterns that may evolve over time. We also showed that DCS becomes a powerful tool for financial analysis and forecasting, provided that we thoroughly understand the main characteristics of data and the parametric assumptions are formulated to reflect them.

Our estimation results are robust to the choice of aggregation interval, but the aggregation interval affects whether we can detect the presence of the-day-of-the-week effect. The Burr distribution, which is a special case of the GB2 distribution, achieves an impressive fit to the data. In particular, the PIT values of the estimation residuals are remarkably close to being `iid` standard uniformly distributed. In contrast, the log-normal distribution is found to give an inferior fit due to the departure of the log-transformed data from normality especially around the tail regions. Although the evidence for changing diurnal patterns appeared to be mixed or weak for our data, diurnal patterns may evolve more significantly for other data sets collected over longer sampling periods. The out-of-sample forecast results show that the in-sample estimation results are stable, and that our model is able to provide good one-step and multi-step ahead forecasts for several hundred steps of prediction horizon.

The object of our empirical analysis is trade volume as measured by the number of shared traded, and, as such, this study also contributes to the literature dedicated to the analysis of market activity and intensity. We studied the movements of volume in complete isolation from price, which is ultimately not satisfactory if one is interested in studying the interaction of price and quantity dynamics. Thus, the next natural step is to construct multivariate intra-day DCS that models price and volume simultaneously. One can also extend our framework to study correlation of scale with other variables, or construct a model for panel-data using a composite likelihood.

References

Alizadeh, S., Brandt, M. W. and Diebold, F. X. (2002). Range-based estimation of stochastic volatility models, *Journal of Finance* **57**(3): 1047–1091.

Andersen, T. G. and Bollerslev, T. (1998). Deutsche mark-dollar volatility: intraday activity patterns, macroeconomic annoucements, and longer run dependencies, *Journal of Finance* **53**(1): 219–265.

Andres, P. and Harvey, A. C. (2012). The dynamic location/scale model: with applications to intra-day financial data, *Cambridge Working Papers in Economics CWPE1240*, University of Cambridge.

Boes, M.-J., Drost, F. C. and Werker, B. J. M. (2007). The impact of overnight periods on option pricing, *Journal of Financial and Quantitative Analysis* **42**(2): 517–533.

Bowsher, C. G. and Meeks, R. (2008). The dynamics of economic functions: Modelling and forecasting the yield curve, *Journal of the American Statistical*

Association **103**(484): 1419–1437.

Brownlees, C. T., Cipollini, F. and Gallo, G. M. (2011). Intra-daily volume modelling and prediction for algorithmic trading, *Journal of Financial Economics* **9**(3): 489–518.

Campbell, S. D. and Diebold, F. X. (2005). Weather forecasting for weather derivatives, *Journal of American Statistical Association* **100**(469): 6–16.

Creal, D. D., Koopman, S. J. and Lucas, A. (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations, *Journal of Business and Economic Statistics* **29**(4): 552–563.

Creal, D. D., Koopman, S. J. and Lucas, A. (2013). Generalized autoregressive score models with applications, *Journal of Applied Econometrics* **28**: (forthcoming).

Engle, R. F. and Rangel, J. G. (2008). The spline-GARCH model for low-frequency volatility and its global macroeconomic causes, *Review of Financial Studies* **21**(3): 1187–1222.

Engle, R. F. and Russell, J. R. (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data, *Econometrica* **66**(5): 1127–1162.

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

Gerhard, F. and Hautsch, N. (2007). A dynamic semiparametric proportional hazard model, *Studies in Nonlinear Dynamics and Econometrics* **11**(2).

Harvey, A. C. (1993). *Time Series Models*, 2nd edn, Harvester, Wheatsheaf.

Harvey, A. C. (2013). *Dynamic models for volatility and heavy tails: with applications to financial and economic time series*, Econometric Society monograph, Cambridge University Press (forthcoming).

Harvey, A. C. and Chakravarty, T. (2008). Beta-t-(E)GARCH, *Cambridge Working Papers in Economics CWPE0840*, University of Cambridge.

Harvey, A. C. and Koopman, S. J. (1993). Forecasting hourly electricity demand using time-varying splines, *Journal of the American Statistical Association* **88**(424): 1228–1236.

Harvey, A. C., Koopman, S. J. and Riani, M. (1997). The modeling and seasonal adjustment of weekly observations, *Journal of Business & Economic Statistics* **15**(3): 354–68.

Hautsch, N. (2012). *Econometrics of financial high-frequency data*, Springer, Berlin.

Hautsch, N., Malec, P. and Schienle, M. (2010). Capturing the zero: a new class of

zero-augmented distributions and multiplicative error processes, *SFB 649 Discussion Papers SFB649DP2010-055*, Sonderforschungsbereich 649, Humboldt University, Berlin.

Kleiber, C. and Kotz, S. (2003). *Statistical Size Distributions in Economics and Actuarial Sciences*, Wiley, New York.

Lo, A. W. and Wang, J. (2010). Stock market trading volume, *in* Y. At-Sahalia and L. Hansen (eds), *Handbook of financial econometrics*, Vol. 2, North-Holland, New York.

McCulloch, R. E. and Tsay, R. S. (2001). Nonlinearity in high-frequency financial data and hierarchical models, *Studies in Nonlinear Dynamics and Econometrics* **5**(1): 1–17.

Mitchell, J. and Wallis, K. F. (2011). Evaluating density forecasts: forecast combinations, model mixtures, calibration and sharpness, *Journal of Applied Econometrics* **26**(6): 1023–1040.

Poirier, D. (1976). *The econometrics of structural change: with special emphasis on spline functions*, North-Holland.

Rydberg, T. H. and Shephard, N. (2003). Dynamics of trade-by-trade price movements: decomposition and models, *Journal of Financial Econometrics* **1**(1): 2–25.

Taylor, S. (2005). *Asset price dynamics, volatility, and prediction*, Princeton University Press, Princeton, N.J.

Zhang, M. Y., Russell, J. R. and Tsay, R. S. (2001). A nonlinear autoregressive conditional duration model with applications to financial transaction data, *Journal of Econometrics* **104**(1): 179–207.

A List of distributions and their properties

A.1 Generalized beta distribution

The (standard) generalized beta distribution of the second kind (GB2) has the pdf:

$$f(x;\nu,\xi,\zeta) = \frac{\nu x^{\nu\xi-1}(x^\nu+1)^{-\xi-\zeta}}{B(\xi,\zeta)}, \qquad x>0, \text{ and } \nu,\xi,\zeta>0$$

where $B(\cdot,\cdot)$ denotes the Beta function.[22] If a non-standardized random variable Y follows the GB2 distribution, its pdf $f_Y : \mathbb{R}_{>0} \to \mathbb{R}$ with the scale parameter $\alpha > 0$ is $f_Y(y;\alpha,\nu,\xi,\zeta) = f(y/\alpha;\nu,\xi,\zeta)/\alpha$ for $y > 0$. For a set of `iid` observations $y_1,\ldots,y_T$ where each follows the non-standardized GB2 distribution, the log-likelihood function of a single observation y_t can be written using the exponential link function $\alpha = \exp(\lambda)$ with the link parameter $\lambda \in \mathbb{R}$ as:

$$\log f_Y(y_t) = \log(\nu) - \nu\xi\lambda + (\nu\xi - 1)\log(y_t) - \log B(\xi,\zeta) - (\xi+\zeta)\log[(y_t e^{-\lambda})^\nu + 1].$$

The *score* u_t of the non-standardized GB2 computed at y_t is

$$u_t \equiv \frac{\partial \log f_Y(y_t)}{\partial\lambda} = \frac{\nu(\xi+\zeta)(y_t e^{-\lambda})^\nu}{(y_t e^{-\lambda})^\nu + 1} - \nu\xi = \nu(\xi+\zeta)b_t(\xi,\zeta) - \nu\xi$$

where we used the notation $b_t(\xi,\zeta) \equiv (y_t e^{-\lambda})^\nu/((y_t e^{-\lambda})^\nu + 1)$. By the property of the GB2 distribution, we know that $b_t(\xi,\zeta)$ follows the beta distribution with parameters ξ and ζ.[23] It is easy to check that $\mathbb{E}[u_t] = 0$. $b_t(\xi,\zeta)$ is bounded between 0 and 1, which means that we have $-\nu\xi \le u_t \le \nu\zeta$.

A.2 Generalized gamma distribution

The (standard) generalized gamma (GG) distribution has the pdf:

$$f(x;\gamma,\nu) = \frac{\nu}{\Gamma(\gamma)} x^{\nu\gamma-1}\exp\left(-x^\nu\right), \qquad 0 < x, \text{ and } \gamma,\nu > 0,$$

where $\Gamma(\cdot)$ is the gamma function.[24] If a non-standardized random variable Y follows the GG distribution, its pdf $f_Y : \mathbb{R}_{>0} \to \mathbb{R}$ with the scale parameter $\alpha > 0$ is $f_Y(y;\alpha,\gamma,\nu) = f(y/\alpha;\gamma,\nu)/\alpha$ for $y > 0$. For a set of `iid` observations $y_1,\ldots,y_T$ where each follows the non-standardized GG distribution, the log-likelihood function of a single observation y_t can be written using the exponential link function $\alpha = \exp(\lambda)$ as:

$$\log f_Y(y_t) = \log(\nu) - \lambda + (\nu\gamma - 1)\log(y_t e^{-\lambda}) - (y_t e^{-\lambda})^\nu - \log\Gamma(\gamma).$$

[22] GB2 becomes the Burr distribution when $\xi = 1$ and the log-logistic distribution when $\xi = \zeta = 1$. The Burr distribution is also called the Pareto Type IV distribution (Pareto IV). Log-logistic is also called Pareto III. Burr becomes Pareto II when $\nu = 1$. Burr becomes Weibull (defined in Appendix A.2) when $\zeta \to \infty$. GB2 with $\nu = 1$ and $\xi = \zeta$ is a special case of the F distribution with the degrees of freedom $\nu_1 = \nu_2 = 2\xi$.

[23] The beta distribution characterized by the mgf is

$$M_b(z;\xi,\zeta) \equiv \mathbb{E}[e^{bz}] = 1 + \sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\left(\frac{\xi+r}{\xi+\zeta+r}\right)\frac{z^k}{k!}\right).$$

[24] The GG distribution becomes the gamma distribution when $\nu = 1$, the Weibull distribution when $\gamma = 1$, and the exponential distribution when $\nu = \gamma = 1$.

The score u_t of the non-standardized GG computed at y_t is

$$u_t \equiv \frac{\partial \log f_Y(y_t)}{\partial \lambda} = \nu(y_t e^{-\lambda})^\nu - \nu\gamma = \nu g_t(\nu) - \nu\gamma,$$

where we used the notation $g_t(\nu) \equiv (y_t e^{-\lambda})^\nu$. By the property of the GG distribution, we know that $g_t(\nu)$ follows the (standard) gamma distribution with parameter γ, which is characterized by the mgf $\mathbb{E}[e^{gz}] = (1-z)^{-\gamma}$ for $z < 1$. We also have $\mathbb{E}[u_t] = 0$ and $u_t > -\nu\gamma$ by the property of the gamma distribution.

B Dynamic cubic spline

In this section, we formally explain the mathematical construction of the dynamic cubic spline from Section 3.4. The time indices are as defined in Section 3.3.

B.1 Daily spline

The specification of the knots are the same as Section 3.4.1. We denote the distance between knots along the time-axis by $h_j = \tau_j - \tau_{j-1}$ for $j = 1, \dots k$. Our cubic spline function $g : [\tau_0, \tau_k] \to \mathbb{R}$ is a piecewise function of the form

$$g(\tau) = \sum_{j=1}^{k} g_j(\tau) \mathbb{1}_{\{\tau \in [\tau_{j-1}, \tau_j]\}}, \qquad \forall\, \tau \in [\tau_0, \tau_k],$$

where each function $g_j : [\tau_{j-1}, \tau_j] \to \mathbb{R}$ is a polynomial of order up to three for all $j = 1, \dots, k$. We can set g to be continuous at each knot (τ_j, γ_j); that is, $g_j(\tau_j) = \gamma_j$ and $g_j(\tau_{j-1}) = \gamma_{j-1}$ for all $j = 1, \dots, k$. This means we have

$$g_j(\tau_{j-1}) = g_{j-1}(\tau_{j-1}) \qquad \text{and} \qquad g_j'(\tau_{j-1}) = g_{j-1}'(\tau_{j-1}). \tag{13}$$

for $j = 2, \dots, k$. (13) is the *continuity condition* of g. The polynomial order of each g_j means that $g_j''(\cdot)$ is a linear function on $[\tau_{j-1}, \tau_j]$ for $j = 1, \dots, k$. This implies that

$$g_j''(\tau) = a_{j-1} + \frac{\tau - \tau_{j-1}}{h_j}(a_j - a_{j-1}) = \frac{(\tau_j - \tau)}{h_j} a_{j-1} + \frac{(\tau - \tau_{j-1})}{h_j} a_j, \tag{14}$$

for $\tau \in [\tau_{j-1}, \tau_j]$ and $j = 1, \dots, k$, where $a_0 = g_1''(\tau_0)$ and $a_j = g_j''(\tau_j)$ for $j = 1, \dots, k$. (14) is the *polynomial order condition* of g.[25]

We integrate (14) with respect to τ to find the expressions for g_j' and g_j. That is, we evaluate $g_j'(\tau) = \int g_j''(\tau) d\tau$ and $g_j(\tau) = \int \int g_j''(\tau) d\tau$ for each $j = 1, \dots, k$,

[25] In addition to the continuity and polynomial order conditions, Harvey and Koopman (1993) include the periodicity condition; that is, g_1 and g_k satisfy $\gamma_0 = \gamma_k$, $g_1'(\tau_0) = g_k'(\tau_k)$, and $g_1''(\tau_0) = g_k''(\tau_k)$ so that $a_0 = a_k$. This is because their hourly electricity demand data is collected 24 hours on every sampling day.

where we recover the integration constant using (13). Then we obtain

$$g_j'(\tau) = -\left[\frac{1}{2}\frac{(\tau_j-\tau)^2}{h_j} - \frac{h_j}{6}\right] a_{j-1} + \left[\frac{1}{2}\frac{(\tau-\tau_{j-1})^2}{h_j} - \frac{h_j}{6}\right] a_j, \quad (15)$$

$$g_j(\tau) = \boldsymbol{r}_j(\tau)\cdot\boldsymbol{\gamma} + \boldsymbol{s}_j(\tau)\cdot\boldsymbol{a} \quad (16)$$

for $\tau \in [\tau_{j-1}, \tau_j]$ and $j = 1, \ldots, k$, where $\boldsymbol{a} = (a_0, a_1, \ldots, a_k)^\top$, and $\boldsymbol{r}_j(\tau)$ and $\boldsymbol{s}_j(\tau)$ are the following k-dimensional vectors

$$\boldsymbol{r}_j(\tau) = \left(0, \ldots, 0, \frac{(\tau_j-\tau)}{h_j}, \frac{(\tau-\tau_{j-1})}{h_j}, 0, \ldots, 0\right)^\top,$$

$$\boldsymbol{s}_j(\tau) = \left(0, \ldots, 0, (\tau_j-\tau)\frac{(\tau_j-\tau)^2 - h_j^2}{6h_j}, (\tau-\tau_{j-1})\frac{(\tau-\tau_{j-1})^2 - h_j^2}{6h_j}, 0, \ldots, 0\right)^\top.$$

The non-zero elements of $\boldsymbol{r}_j(\tau)$ and $\boldsymbol{s}_j(\tau)$ are at the jth and $(j+1)$th entries. The conditions for g_j' in (13) and (15) give

$$\frac{h_j}{h_j+h_{j+1}}a_{j-1} + 2a_j + \frac{h_{j+1}}{h_j+h_{j+1}}a_{j+1} = \frac{6\gamma_{j-1}}{h_j(h_j+h_{j+1})} - \frac{6\gamma_j}{h_j h_{j+1}} + \frac{6\gamma_{j+1}}{h_{j+1}(h_j+h_{j+1})}$$

for $j = 1, \ldots, k-1$. From these, we obtained a system of $k-1$ equations with $k+1$ unknowns $a_0, \ldots, a_k$. Following Poirier (1976) we set $a_0 = a_k = 0$ (the *natural condition* for a spline). We can write this system of equations in a matrix form as $\mathbf{P}\boldsymbol{a} = \mathbf{Q}\boldsymbol{\gamma}$, where $\mathbf{P}$ and $\mathbf{Q}$ are the following square matrices of size $(k+1)$:

$$\mathbf{P} = \begin{bmatrix} 2 & 0 & 0 & 0 & \ldots & 0 & 0 \\ \frac{h_1}{h_1+h_2} & 2 & \frac{h_2}{h_1+h_2} & 0 & \ldots & 0 & 0 \\ 0 & \frac{h_2}{h_2+h_3} & 2 & \frac{h_3}{h_2+h_3} & \cdots & 0 & 0 \\ 0 & 0 & \frac{h_3}{h_3+h_4} & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 2 & \frac{h_k}{h_{k-1}+h_k} \\ 0 & 0 & 0 & 0 & \ldots & 0 & 2 \end{bmatrix},$$

$$\mathbf{Q} = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & 0 \\ \frac{6}{h_1(h_1+h_2)} & -\frac{6}{h_1 h_2} & \frac{6}{h_2(h_1+h_2)} & \ldots & 0 & 0 \\ 0 & \frac{6}{h_2(h_2+h_3)} & -\frac{6}{h_2 h_3} & \ldots & 0 & 0 \\ 0 & 0 & \frac{6}{h_3(h_3+h_4)} & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -\frac{6}{h_{k-1}h_k} & \frac{6}{h_k(h_{k-1}+h_k)} \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{bmatrix}.$$

The first and the last rows of $\mathbf{P}$ and $\mathbf{Q}$ ensure that $a_0 = a_k = 0$. For a non-singular $\mathbf{P}$, we have $\boldsymbol{a} = \mathbf{P}^{-1}\mathbf{Q}\boldsymbol{\gamma}$. Then (16) can be written as $g_j(\tau) = \boldsymbol{w}_j(\tau)\cdot\boldsymbol{\gamma}$ for $\tau \in [\tau_{j-1}, \tau_j]$, where $\boldsymbol{w}_j(\tau)^\top = \boldsymbol{r}_j(\tau)^\top + \boldsymbol{s}_j(\tau)^\top\mathbf{P}^{-1}\mathbf{Q}$. Finally, we obtain the

following expression for the daily cubic spline

$$s_\tau = g(\tau) = \sum_{j=1}^{k} \mathbb{1}_{\{\tau\in[\tau_{j-1},\tau_j]\}} \boldsymbol{w}_j(\tau)\cdot\boldsymbol{\gamma}, \qquad \forall\,\tau\in[\tau_0,\tau_k]. \tag{17}$$

The elements of $\boldsymbol{\gamma}$ are the parameters of the model to be estimated. For the parameters to be identified, we impose the following zero-sum constraint on the elements of $\boldsymbol{\gamma}$

$$\sum_{\tau\in[\tau_0,\tau_k]} s_\tau = \sum_{\tau\in[\tau_0,\tau_k]} \sum_{j=1}^{k} \mathbb{1}_{\{\tau\in[\tau_{j-1},\tau_j]\}} \boldsymbol{w}_j(\tau)\cdot\boldsymbol{\gamma} = \boldsymbol{w}_*\cdot\boldsymbol{\gamma} = 0,$$

where

$$\boldsymbol{w}_* = (w_{*0}, w_{*1}, \ldots, w_{*k})^\top = \sum_{\tau\in[\tau_0,\tau_k]} \sum_{j=1}^{k} \mathbb{1}_{\{\tau\in[\tau_{j-1},\tau_j]\}} \boldsymbol{w}_j(\tau).$$

We can impose this condition by setting $\gamma_k = -\sum_{i=0}^{k-1} w_{*i}\gamma_i/w_{*k}$. Then (17) becomes

$$s_\tau = \sum_{j=1}^{k} \mathbb{1}_{\{\tau\in[\tau_{j-1},\tau_j]\}} \sum_{i=0}^{k-1} \left(w_{ji}(\tau) - \frac{w_{jk}(\tau)w_{*i}}{w_{*k}} \right) \gamma_i = \sum_{j=1}^{k} \mathbb{1}_{\{\tau\in[\tau_{j-1},\tau_j]\}} \boldsymbol{z}_j(\tau)\cdot\boldsymbol{\gamma} \tag{18}$$

for $\tau\in[\tau_0,\tau_k]$. $w_{ji}(\tau)$ denotes the ith element of $\boldsymbol{w}_j(\tau)$, and the ith element of $\boldsymbol{z}_j(\tau)$ is

$$z_{ji}(\tau) = \begin{cases} w_{ji}(\tau) - w_{jk}(\tau)w_{*i}/w_{*k} & i\neq k \\ 0 & i = k \end{cases}$$

for $\tau\in[\tau_{j-1},\tau_j]$ and each $i=0,\ldots,k$ and $j=1,\ldots,k$. When estimating the model, it is convenient to compute $\boldsymbol{w}_*$ using the equation $\boldsymbol{w}_*^\top = \boldsymbol{r}_*^\top + \boldsymbol{s}_*^\top\mathbf{P}^{-1}\mathbf{Q}$, where $\boldsymbol{r}_*$ and $\boldsymbol{s}_*$ are k-dimensional vectors computed using the rules of arithmetic and polynomial series as

$$\begin{aligned} \boldsymbol{r}_* &= \left(\frac{\tau_1-\tau_0+1}{2}, \frac{\tau_2-\tau_0}{2}, \ldots, \frac{\tau_{k-1}-\tau_{k-3}}{2}, \frac{\tau_k-\tau_{k-1}+1}{2} \right)^\top, \\ \boldsymbol{s}_* &= \left(\frac{h_1-h_1^3}{24}, \frac{\tau_2-\tau_0-h_2^3-h_1^3}{24}, \ldots, \frac{\tau_{k-1}-\tau_{k-3}-h_{k-1}^3-h_{k-2}^3}{24}, \frac{h_k-h_k^3}{24} \right)^\top. \end{aligned}$$

Note that these formulae for computing $\boldsymbol{w}_*$, $\boldsymbol{r}_*$, and $\boldsymbol{s}_*$ are different from those of Harvey and Koopman (1993) due to the removal of the periodicity condition. The static spline in (18) becomes dynamic by letting $\boldsymbol{\gamma}$ be time-varying according to

$$s_{t,\tau} = \sum_{j=1}^{k} \mathbb{1}_{\{\tau\in[\tau_{j-1},\tau_j]\}}\, \boldsymbol{z}_j(\tau)\cdot\boldsymbol{\gamma}_{t,\tau}, \qquad \boldsymbol{\gamma}_{t,\tau} = \boldsymbol{\gamma}_{t,\tau-1} + \boldsymbol{\kappa}^*\cdot u_{t,\tau-1} \tag{19}$$

for $\tau=1,\ldots,I$ and $t=1,\ldots,T$, where $\boldsymbol{\kappa}^* = (\kappa_0^*,\ldots,\kappa_k^*)^\top$ is a vector of parame-

ters. The dynamic spline (19) needs to sum to zero over one complete period for the parameters to be identified. That is, $s_{t,\tau}$ must satisfy $\sum_{\tau=1}^{I} s_{t,\tau} = \boldsymbol{w}_* \cdot \boldsymbol{\gamma}_{t,\tau} = 0$ for $t = 1, \ldots, T$. The construction of $\boldsymbol{z}_j(\tau)$ ensures that this constraint holds, but we also need to set $\boldsymbol{w}_* \cdot \boldsymbol{\gamma}_{1,0} = 0$ and $\boldsymbol{w}_* \cdot \boldsymbol{\kappa}^* = 0$. We can impose these conditions on $\boldsymbol{\gamma}_{1,0}$ and $\boldsymbol{\kappa}^*$ by setting

$$\gamma_{k;1,0} = -\frac{1}{w_{*k}} \sum_{i=0}^{k-1} w_{*i} \gamma_{i;1,0} \qquad \text{and} \qquad \kappa_k^* = -\frac{1}{w_{*k}} \sum_{i=0}^{k-1} w_{*i} \kappa_i^*,$$

where $\gamma_{i;1,0}$ denotes the ith element of $\boldsymbol{\gamma}_{1,0}$.

B.2 Weekly spline

The specification of the knots for the weekly spline are as we defined in Section 3.4.4. As before, we allow for $\gamma_k \neq \gamma_0$ in order to capture the weekend effect. We capture the overnight effect of weeknights by relaxing the continuity and polynomial order conditions (13)-(14) between $\widetilde{\tau}_{k'}$ and $\widetilde{\tau}_0$ of any two successive weekdays. This redefines $\mathbf{P}$ and $\mathbf{Q}$ matrices as follows. For the $\mathbf{P}$ matrix, we replace the off-diagonal entries in the $i(k'+1)$th and $(i(k'+1)+1)$th rows by zeros for each $i = 1, \ldots (k+1)/(k'+1)$. For the $\mathbf{Q}$ matrix, we replace *all* entries in the $i(k'+1)$th and $(i(k'+1)+1)$th rows by zeros for each $i = 1, \ldots (k+1)/(k'+1)$.

C Estimated coefficients of other model specifications in Table 5

	IBM30s						IBM1m					
	Model 1		Model 3		Model 4		Model 1		Model 3		Model 4	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
κ_μ	0.019	0.002	0.016	0.002	0.008	0.002	0.017	0.002	0.014	0.002	0.010	0.002
$\phi_1^{(1)}$	0.795	0.016	0.797	0.016	0.556	0.143	0.721	0.025	0.710	0.026	0.346	0.107
$\phi_2^{(1)}$	—	—	—	—	0.397	0.143	—	—	—	—	0.553	0.120
$\kappa_\eta^{(1)}$	0.116	0.005	0.117	0.005	0.048	0.009	0.097	0.005	0.097	0.005	0.043	0.012
$\phi_1^{(2)}$	—	—	—	—	0.676	0.048	—	—	—	—	0.607	0.065
$\kappa_\eta^{(2)}$	—	—	—	—	0.089	0.009	—	—	—	—	0.065	0.012
κ_0^*	0.006	0.004	0.017	0.012	0.008	0.007	0.004	0.004	0.013	0.011	0.010	0.009
κ_1^*	0.003	0.002	0.000	0.007	0.002	0.005	0.002	0.002	0.002	0.007	0.003	0.007
κ_2^*	-0.005	0.002	-0.012	0.003	-0.005	0.003	-0.005	0.002	-0.010	0.004	-0.007	0.004
κ_3^*	-0.002	0.002	-0.005	0.004	-0.003	0.003	-0.001	0.002	-0.007	0.004	-0.005	0.004
κ_4^*	—	—	0.017	0.011	0.006	0.006	—	—	0.018	0.010	0.010	0.008
κ_5^*	—	—	0.000	0.004	0.001	0.003	—	—	0.002	0.004	0.002	0.003
κ_6^*	—	—	0.000	0.003	0.000	0.002	—	—	0.000	0.003	-0.001	0.002
κ_7^*	—	—	-0.003	0.002	-0.002	0.002	—	—	-0.004	0.002	-0.003	0.002
κ_8^*	—	—	-0.003	0.002	0.000	0.002	—	—	-0.001	0.002	0.000	0.002
κ_9^*	—	—	0.005	0.004	0.003	0.003	—	—	0.004	0.004	0.003	0.003
κ_{10}^*	—	—	0.011	0.009	0.006	0.005	—	—	0.012	0.009	0.008	0.006
κ_{11}^*	—	—	0.013	0.008	0.002	0.004	—	—	0.002	0.006	-0.001	0.004
κ_{12}^*	—	—	-0.001	0.005	0.000	0.004	—	—	0.001	0.005	0.002	0.005
κ_{13}^*	—	—	0.006	0.005	0.003	0.004	—	—	0.007	0.006	0.005	0.005
$\gamma_{0;1,0}$	1.219	0.091	1.542	0.221	1.563	0.192	1.118	0.075	1.486	0.179	1.517	0.171
$\gamma_{1;1,0}$	0.119	0.062	0.052	0.142	0.008	0.129	0.112	0.052	0.007	0.130	-0.007	0.134
$\gamma_{2;1,0}$	-0.483	0.069	-0.419	0.161	-0.444	0.128	-0.442	0.061	-0.396	0.135	-0.401	0.127
$\gamma_{3;1,0}$	-0.231	0.042	-0.137	0.111	-0.143	0.109	-0.246	0.036	-0.115	0.111	-0.131	0.109
$\gamma_{4;1,0}$	—	—	0.547	0.307	0.589	0.198	—	—	0.665	0.298	0.699	0.228
$\gamma_{5;1,0}$	—	—	1.106	0.077	1.120	0.086	—	—	1.034	0.083	1.041	0.086
$\gamma_{6;1,0}$	—	—	0.037	0.068	0.016	0.057	—	—	0.041	0.058	0.029	0.055
$\gamma_{7;1,0}$	—	—	-0.479	0.056	-0.488	0.061	—	—	-0.446	0.052	-0.450	0.060
$\gamma_{8;1,0}$	—	—	-0.178	0.053	-0.197	0.050	—	—	-0.217	0.047	-0.224	0.049
$\gamma_{9;1,0}$	—	—	0.655	0.117	0.607	0.113	—	—	0.658	0.101	0.645	0.098
$\gamma_{10;1,0}$	—	—	1.030	0.215	1.033	0.193	—	—	0.779	0.213	0.849	0.192
$\gamma_{11;1,0}$	—	—	0.273	0.234	0.324	0.119	—	—	0.346	0.106	0.334	0.100
$\gamma_{12;1,0}$	—	—	-0.328	0.115	-0.233	0.099	—	—	-0.252	0.106	-0.224	0.109
$\gamma_{13;1,0}$	—	—	-0.414	0.151	-0.350	0.128	—	—	-0.423	0.145	-0.394	0.136
ω	9.136	0.225	8.976	0.203	9.008	0.166	9.735	0.193	9.578	0.173	9.588	0.154
ν	1.627	0.016	1.628	0.016	1.632	0.016	2.224	0.033	2.231	0.033	2.234	0.033
ζ	1.492	0.045	1.493	0.045	1.487	0.045	1.143	0.044	1.141	0.044	1.141	0.044
p	0.0047	0.0005	0.0047	0.0005	0.0047	0.0005	0.0006	0.0003	0.0006	0.0003	0.0006	0.0003

Table 7 – Estimated coefficients of other model specifications in Table 5. Standard errors are computed using numeric derivatives of the likelihood function with respect to the parameters. See Table 5 for the model specifications.