Intraday Trades Profile Estimation: An Intensity Approach

Alessio Sancetta*

March 8, 2021

Abstract

The intraday trades profile is the expected intensity of a counting process where the counts measure the number of trades over an interval. It needs to capture the salient features of the trading activity, its spikes and periods of relative quietness. This calls for an estimator with a time varying resolution that allows us to identify jumps. The problem can be recast as a regression one, using a fused Lasso penalty. The framework allows us to identify jumps within possibly thousands different locations within a day when the number of trading days at disposal is in the order of hundreds. This can be done without imposing any conditions on the counting process except for certain regularity conditions on the expected intensity. The empirical results suggest that much of the trading activity in some liquid futures can be captured by a deterministic seasonal component in the trade arrival process.

Key Words: algorithmic trading; asymptotic distribution; consistency; counting process; fused Lasso estimator.

JEL Codes: C58, C52.

1 Introduction

This paper is concerned with estimation of intraday trades profiles. A trades profile shows how the daily number of trades are expected to be distributed within a day. Such profile, and its extension to include trade sizes, is an important input in high frequency execution algorithms such as volume weighted average algorithms. There, a large order is split into small orders and the frequency of trading is dictated by the expected trading activity. The use of a profile is often needed in order to plan how the algorithm will split the orders, have an ex ante estimate

*Acknowledgements: I thank Yuri Taranenko for useful conversations that motivated this paper, and in a related context, for suggesting an estimator similar to the one in Section 3.3. I also thank Luca Mucciante for technical discussions on this topic and for bringing some important references to my attention. I am also grateful to the associate Editor and the Referees whose comments led to considerable improvements in content and presentation. E-mail: <asancetta@gmail.com>, URL: <http://sites.google.com/site/wwwsancetta/>. Address for correspondence: Department of Economics, Royal Holloway University of London, Egham TW20 0EX, UK

of completion time, and establish the contribution of each period to the final weighted average price.

Given the irregular time series nature of trades arrival, it is natural to model trades profiles with a counting process. Then, the trades profile is just an another word for the expected intensity of the counting process. The difficulty in this problem is that the trades profile is not smooth. The profile needs to capture spikes in activity. Such spikes can be concentrated in as little as a minute of trading if not seconds. Scheduled events are typical examples. These include opening and close of trading for the product and related ones, option expiries, economic announcements, inventory reports for commodities, etc. Clearly, we may not know about all the scheduled events that affect the trades profile of a traded instrument. Moreover, we may not know for how long these events may affect the profile. The effect could last seconds, minutes or hours. Based on these remarks, from a mathematical point of view the profile is a non-smooth function with jump discontinuities at possibly unknown points.

The problem is similar to the one of estimating the seasonality of trade counts. A common way to estimate a seasonal component is to use a histogram using equally spaced bins. For the profile, we need to devise an adaptive way to choose the bin size at different locations. This would be simple if we had an oracle that presented us with all jump discontinuities. In general, the presence of jump discontinuities rules out smoothing estimators, as their strength is in the recovery of smooth functions (Stone, 1982). This paper uses an estimation technique that allows us to solve the problem as if we had an oracle, in large samples.

1.1 Paper Outline and Contribution

In Section 2, we discuss the model for trades arrival as given by a counting process with stochastic intensity. We suppose that the expectation of the stochastic intensity can be written as

$$h(t) = \sum_{j=1}^{J} h_j 1_{\mathcal{D}_j}(t).$$

Here, J is the number jump discontinuities, $\mathcal{D}_j$ is the set over which the intensity is constant with value h_j. All these quantities are unknown and need to be estimated from a sample of intraday trades over multiple days. We use an estimation procedure that identifies with probability going to one the sets $\mathcal{D}_j$ and computes an estimator $\hat{h}$ for h that is consistent under the uniform norm. This is achieved under weak conditions described in Section 2 and Theorem 1. In Theorem 2, we also derive an asymptotic normality result.

In Section 3, we include some extensions. These include estimation of volumes profiles and conditional estimation. The former requires to incorporate size information to trades arrival. Examples of conditional estimation include previous day number of trades and end of day trade count. The end of day trade count should be a predictable variable along the lines of

end of day volume and its prediction can be updated intradaily, as we collect trade information (Sancetta, 2019).

In the empirical part of the paper, Section 4, we consider estimation of the trades profile for six futures contracts traded on the CME. The aim of that section is to highlight features in the data that justify the suggested functional form of the expected intensity h. Interestingly, based on out of sample results, we find that a lot of the activity in trades arrival is captured by the trade profile. In Section 5, we rely on simulation results to gauge evidence on whether the estimator can be trusted when used in typical set ups as the one of the empirical section. Section 6 concludes. Proofs are deferred to the Appendix.

1.2 Relation to Existing Technical Results

In this paper, we are concerned with adaptive estimation of the expected intensity of a counting processes. We recast the problem into a panel data problem with martingale errors and construct a pooled estimator across days. Then, this becomes a signal recovery problem which has been studied extensively in the literature. In particular the asymptotic analysis relies on the arguments found in Rinaldo (2009). There, the context is different and the model is an independent Gaussian one. The Gaussian or sub-Gaussian assumption allows one to use a strong concentration inequality. We cannot do so here, as trades arrival are not independently distributed and they cannot be assumed to have Gaussian tails. Here, we only rely on existence of a second moment for the compensator of the counting process.

From a technical point of view, Alaya et al. (2015) study a problem close to the present one. There, the intensity/compensator is deterministic but subject to unknown break points, and the estimator is shown to be consistent in L_2. We cannot apply those results for two reasons. First, we aim at exact recovery of the interval of homogeneity of the expected intensity together with uniform convergence of the estimator. Second, we do not assume the intensity to be deterministic. Such an assumption would rule out data being generated from stochastic processes usually encountered in the finance and financial econometrics literature. Such models include Hawkes processes (Bauwens and Hautsch, 2009), autoregressive conditional duration (ACD) models (Engle and Russell, 1998) and models that depend on order book information (Sancetta, 2018). Finally, we also provide an asymptotic normality result.

2 The Model

Consider a sequence of counting processes $\left(N^{(i)}\right)_{i\geq 1}$, $N^{(i)} := \left(N^{(i)}(t)\right)_{t\in[0,1]}$, with intensity $\lambda^{(i)}$, where $h(t) = \mathbb{E}\lambda^{(i)}(t)$ is left continuous. The quantity h is the mean intensity and represents our quantity of interest. The index i represents a day and we suppose a sample of

n days. Then, the counting process admits the stochastic representation

$$N^{(i)}(t) = \int_0^t \lambda^{(i)}(s)\, ds + M^{(i)}(t), \tag{1}$$

where $\left(M^{(i)}(t)\right)_{t\in[0,1]}$ is a martingale with respect to the natural filtration of $N^{(i)}$ (Brémaud, 1980, Ch.II.T8). The first term on the right hand side (r.h.s.) is the compensator. We are interested in estimating $h(t) = \mathbb{E}\lambda^{(i)}(t)$, $t \in [0,1]$.

The goal of this paper is to provide a consistent estimator for h under the following conditions.

Regularity Conditions.

Condition 1 *(Counting Processes) The counting processes* $\left(N^{(i)}\right)_{i\geq 1}$ *have geometrically decaying beta mixing coefficients and stochastic intensity* $\lambda^{(i)}$.

Condition 2 *(Moments) There exist a finite absolute* $\bar{\lambda}_2 < \infty$, *such that* $\max_t \mathbb{E}\left|\lambda^{(i)}(t)\right|^2 \leq \bar{\lambda}_2$.

Condition 3 *(Model h) The expected intensity* $h(t) = \mathbb{E}\lambda^{(i)}(t)$ *satisfies*

$$h(t) = \sum_{j=1}^{J} h_j 1_{\mathcal{D}_j}(t) \tag{2}$$

for some unknown finite partition $\mathcal{D} = \{\mathcal{D}_j : j = 1, 2, ..., J\}$ *of* $[0,1]$ *where* $\mathcal{D}_j = (\delta_{j-1}, \delta_j]$, *for increasing rational numbers* δ_j *with* $\delta_0 = 0$ *and such that* $\delta_{\min} := \min_{j\leq J} |\mathcal{D}_j| > 0$.

For convenience, we refer to the above set of conditions as the Regularity Conditions.

Remarks on Dependence. Within each day, we do not impose any dependence condition on $N^{(i)}$. We only restrict dependence across days. Recall that random variables $(X_i)_{i\geq 1}$ have geometrically decaying beta mixing coefficients if $\beta(m) \leq be^{-cm}$ for some finite b and $c > 0$, where for every $m \geq 1$

$$\beta(m) := \sup_{j\geq 1} \mathbb{E} \sup_{B_j} |\Pr(B_j|A_j) - \Pr(B_j)|$$

for $A_j \in \mathcal{A}_j$ and $B_j \in \mathcal{B}_J$ where $\mathcal{A}_j$ and $\mathcal{B}_j$ are the sigma algebras generated by $\{X_i : i \leq j\}$ and $\{X_i : i \geq m + j\}$ respectively (see Rio, 2000, section 1.6, and eq. 1.69 for an equivalent definition). Hence, A_j and B_j are past and future events separated by m periods (days).

As an example, in our context we may think that the dependence is only via the total number of trades. This is an often used assumption by practitioners. Then, we would write

$\lambda^{(i)}(t) = \zeta_i \rho^{(i)}(t)$ where $\rho^{(i)} = \left(\rho^{(i)}(t)\right)_{t\in[0,1]}$ is i.i.d. across i, while $(\zeta_i)_{i\geq 1}$ are beta mixing real valued random variables. Many time series models have geometrically decaying beta mixing coefficients. For example, any finite order ARMA model with independent identically distributed (i.i.d.) innovations and law absolutely continuous with respect to (w.r.t.) the Lebesgue measure satisfies geometric mixing rates (Mokkadem, 1988). Similarly, GARCH models and more generally models that can be embedded in some stochastic recursive equations are also beta mixing with geometric mixing rate for innovations possessing a density w.r.t. the Lebesgue measure (Basrak et al., 2002: though for strong mixing, their result actually implies beta mixing). Many positive recurrent Markov chains also satisfy geometric absolute regularity (Mokkadem, 1990). Hence, the geometric mixing rate in the Regularity Conditions is satisfied by common time series models, though there are notable exceptions (Bradley, 1986, Example 6.2).

Summary of Main Result. The asymptotic results are derived letting $L, n \to \infty$ and possibly $J \to \infty$, where L^{-1} is an initial bin size defined in Section 2.2. Note that even if J is fixed, the challenge is that $\mathcal{D}$ is unknown and the bin size can be arbitrarily small. Under these weak conditions, the main theoretical result of the paper is that we can recover $\mathcal{D}$ with probability going to one and that the estimator $\hat{h}$ for h that we suggest for this problem is consistent under the uniform norm.

The assumptions are minimal. For example, we are able to recover $\mathcal{D}$ under no assumption on the rate at which L diverges to infinity. What is critical is that $\delta_{\min}$ is of larger order of magnitude than $\sqrt{n^{-1}\ln n}$ and this is formalised in (6) in Theorem 1.

2.1 Notation

Use r.h.s. and l.h.s. to mean right hand side and left hand side. The symbol $\lesssim$ means that the l.h.s. is bounded by a constant times the r.h.s., while $\gtrsim$ means the reverse; $\asymp$ means that the l.h.s. is bounded above and below by constants times the r.h.s. For a set $\mathcal{A}$ use $|\mathcal{A}|$ to denote its cardinality if $\mathcal{A}$ is a subset of the integers, or its Lebesgue measure if a subset of $\mathbb{R}$.

2.2 Estimation of the Mean Intensity

To set the scene for the estimation, let $\Delta_l := \Delta_l^{(L)} := ((l-1)/L, l/L]$. By assumption, $\min_{j\leq J} |\mathcal{D}_j| > 0$. Hence, we must have that the partition $\Delta = \{\Delta_l : l = 1, 2, ..., L\}$ is a refinement of $\mathcal{D}$ for L large enough. This is the case because $\lim_L |\Delta_l| = 0$ and the coefficients δ_j are rational numbers. Define $Y_l = \frac{1}{n}\sum_{i=1}^n \int_{\Delta_l} dN^{(i)}(t)$, $\theta_{0,l} = \int_{\Delta_l} h(s)\, ds$, $\eta_l = Y_l - \int_{\Delta_l} h(s)\, ds$. By definition of the intensity $\theta_{0,l} = \mathbb{E}Y_l$. Then, we write

$$Y = \theta_0 + \eta,$$

where Y, θ_0 and η are $L \times 1$ vectors, where their l^{th} entry is denoted with a subscript l. At all times, it is important to remember that Y and η depend on L and n, while θ_0 only on L. To keep the notation simple, we have suppressed this dependence.

We estimate the unknown vector θ_0. As $L \to \infty$, we shrink Δ_l. Given that h is left continuous, by Bochner's lemma, $L\int_{\Delta_l} h(t)\,dt$ is approximately equal to $h\left(\frac{l}{L}\right)$ for large L. Hence $\theta_{0,l} = L^{-1}h\left(\frac{l}{L}\right)$ if h has no jumps in Δ_l. An estimator for θ_0 is given by

$$\hat{\theta} = \arg\inf_{x\in\mathbb{R}^L} (Y-x)'(Y-x) + 2\tau\left|Dx\right|_1. \tag{3}$$

Here, D is an $(L-1)\times L$ matrix such that the (l,l) entry is -1, the $(l,l+1)$ entry is 1 and all the other entries are zero; $\tau \geq 0$ is a tuning parameter; $|\cdot|_1$ is the ℓ_1 norm; the prime symbol $'$ means transpose.

The estimator $\hat{\theta}$ from (3) can be used to derive an estimator for $\mathcal{D}$. Let $\mathcal{J}$ identify the jumps in θ_0, i.e. $\theta_{0,l} \neq \theta_{0,l-1}$ if $l \in \mathcal{J}$, and $\theta_{0,l} = \theta_{0,l-1}$ if $l \in \mathcal{J}^c$, where $\mathcal{J}^c$ is the complement of $\mathcal{J}$ in $\{1,2,...,L\}$. Note that $\mathcal{J}$ identifies the set of jumps of h only if Δ is a refinement of $\mathcal{D}$. From the estimator $\hat{\theta}$ we deduce the set of jumps $\hat{\mathcal{J}} \subset \{1,2,...,L\}$ in the fused estimator, i.e. $l \in \hat{\mathcal{J}}$ if $\hat{\theta}_l \neq \hat{\theta}_{l-1}$. We use $\hat{\mathcal{J}}$ to construct the partition $\hat{\mathcal{D}} := \left\{\hat{\mathcal{D}}_j : j = 1,2,...,\hat{J}\right\}$ which is an estimator for $\mathcal{D}$; here $\hat{J} = 1 + \left|\hat{\mathcal{J}}\right|$. Then, the estimator for $h(\cdot)$ is given by

$$\hat{h}(\cdot) = \sum_{j=1}^{\hat{J}} \hat{h}_j 1_{\hat{\mathcal{D}}_j}(\cdot), \tag{4}$$

where $\hat{h}_j = \frac{1}{n}\sum_{i=1}^n \left|\hat{\mathcal{D}}_j\right|^{-1} \int_{\hat{\mathcal{D}}_j} dN^{(i)}(t)$.

2.2.1 Numerical Computation of the Estimator

The estimator $\hat{\theta}$ is obtained using a dual formulation of the problem in (3) (Tibshirani and Taylor, 2011, eq. 13-14, among others). The problem is rewritten as

$$\hat{z} := \arg\min_{z\in\mathbb{R}^{L-1}:|z|_\infty \leq \tau} \left(Y - D'z\right)'\left(Y - D'z\right). \tag{5}$$

Then, $\hat{\theta} = Y - D'\hat{z}$. From $\hat{\theta}$ we identify the set $\hat{\mathcal{D}}_l$. While conceptually simple, this estimation method is not efficient, especially if we aim at finding solutions for different values of τ that lead to different values of the degrees of freedom of the estimator (see Section 2.4). In this respect, Algorithm 1 in Tibshirani and Taylor (2011) can be used to find the solution path for various values of τ. However, given that D is $(L-1)\times L$ with rank $L-1$, it is well known that the problem can be re-written as a standard Lasso problem. In this case existing methods to solve Lasso problems can also be used. We refer to Tibshirani and Taylor (2011)

for a discussion.

2.3 Consistency

By definition of Y, we place the trades in small equally spaced bins and recast the problem as a signal extraction problem. Here, we allow $L/n \to \infty$ and concern ourselves with the exact recovery of the discontinuities of the mean intensity h. Moreover, we consider convergence under the uniform norm. As already mentioned, the estimation problem is similar to Alaya et al. (2015). However, there are notable differences. In the current notation, those authors derive a sharp oracle inequality under the L_2 norm for $L/n \to 0$ using the total variation norm with data driven weights. There, h is the compensator for $dN^{(i)}$ $\forall i$. This means that $\left(N^{(i)}(t)\right)_{t\in[0,1]}$ is a doubly stochastic process. This rules out time series dependence of the durations. This assumption is not suitable for this problem, as it excludes commonly used high frequency models such ACD and Hawkes processes. We have the following consistency result.

Theorem 1 *Consider the estimator in (4). Suppose that the Regularity Conditions hold and that τ in (3) satisfies*

$$\sqrt{\frac{\ln n}{n}} = o(\tau)\,;\, \tau < \min_j |h_j - h_{j-1}| \frac{\delta_{\min}}{8}, \tag{6}$$

and $\sqrt{\frac{n}{\ln n}} \min_j |h_j - h_{j-1}| \delta_{\min} \to \infty$, $L \to \infty$.

Then, $\Pr(\mathcal{R}) \to 1$ *where*

$$\mathcal{R} = \left\{\hat{\mathcal{J}} = \mathcal{J}\right\} \cap \left\{\text{sign}\left(\hat{\theta}_l - \hat{\theta}_{l-1}\right) = \text{sign}\left(\theta_l - \theta_{l-1}\right), \forall l \in \mathcal{J}\right\}. \tag{7}$$

Moreover, $\max_{t\in[0,1]} \left|h(t) - \hat{h}(t)\right| \to 0$ *in probability if* $\frac{\delta_{\min}^{-2}+J}{n} \ln n \to 0$.

Throughout, for a real variable x, $\text{sign}(x)$ is the sign of x with $\text{sign}(0) = 0$. By the conditions of the theorem, (6) can only be satisfied if n is large enough. In particular, it is necessary that n is of larger order of magnitude than $\ln(n) \left(\min_j |h_j - h_{j-1}| \delta_{\min}\right)^{-2}$. While this might just be an artifact of the method of proof, it does show that we need the number of days n to be relatively large when $\delta_{\min}$ is small. Finally, if the data is m dependent, we can replace the $\ln n$ factor everywhere in the statement of Theorem 1 with a finite constant.

2.4 Choice of Penalty

The choice of penalty τ is not part of the estimation algorithm. We experimented with different criteria and found that simple versions of cross-validation (CV) led to the best results, reducing overfitting. In the simulations, given that computational speed was a concern, we randomly split the sample days into estimation and validation sample with the standard 70%/30% split

between days. We use the validation sample to estimate the penalty that minimizes the residual sum of the squares (RSS). This basic approach works well. In the empirical application, to minimise the dependency on the sample split, we repeat the split five times and average the RSS. We then choose the penalty that minimizes this averaged RSS.

Using the results from Theorem 2 in Section 2.5, we could suppose a Gaussian loglikelihood as approximation to the true one. Then, if computational speed is a concern, we also suggest to use a bias corrected AIC (Hurvich et al., 1998). This suggests choosing the penalty equal to

$$\hat{\tau} = \arg\min_{\tau \in \mathcal{T}} \left\{ \ln \left(\frac{\left(Y - \hat{\theta}^{(\tau)}\right)' \left(Y - \hat{\theta}^{(\tau)}\right)}{L} \right) + \frac{2\left(\hat{J}^{(\tau)} + 1\right)}{L - \hat{J}^{(\tau)} - 2} \right\}. \tag{8}$$

Here, the superscript τ makes explicit the dependence of the penalty τ in (3), and $\mathcal{T}$ is a set chosen by the user (e.g. using Algorithm 1 in Tibshirani and Taylor, 2011).

Instead of (8), the loglikelihood of the counting process with intensity equal to $\hat{h}$ can be used. However, this would be an approximation, unless we suppose that the intensities $\lambda^{(i)}$ are nonstochastic and hence equal to h. Moreover, this is more involved and slower when dealing with large samples. Nevertheless, we do use a loglikelihood approach for model evaluation in the empirical section (Section 4).

An alternative to the above procedures is Mallow's C_p, among others. Unfortunately, it requires to use an unbiased estimator of the variance, adding complexity to the procedure. We found that all the various procedures are equivalent to some extent, but cross-validation was the best. The biased corrected AIC procedure in (8) mitigated overfitting relative to AIC and Mallow's C_p, as expected. Additional details are in Section 5.

Tibshirani and Taylor (2011) derive an algorithm to find all the values of τ corresponding to a jump in the process: the fused Lasso path. This can be slow for large L. To speed up the search of τ, we use a different method. First, we find the smallest penalty that produces the most constrained model, i.e. the one where all coefficients are the same. To do so so, we use the unconstrained estimator of the dual (5). In this case, $\hat{z} = (DD')^{-1} DY$ implying that the smallest τ that still leads to a fully constrained estimator is $\tau = |\hat{z}|_\infty$, and denote this by $\tau_{\max}$. Given this value, we consider minimization of the cross-validated RSS and (8) w.r.t. τ in $\mathcal{T}_{\text{grid}} := \left\{\tau_{\max}10^{-4}, \tau_{\max}10^{-3}, \tau_{\max}10^{-2}, \tau_{\max}10^{-1}, \tau_{\max}\right\} \cap \left[\sqrt{n^{-1}\ln L}, \tau_{\max}\right]$, where the lower bound $\sqrt{n^{-1}\ln L}$ is chosen as a majorizer for the l.h.s. in (6). We then use the optimal value as starting value for τ, say τ_0, and use a derivatives free optimizer for τ_0 inside a suitable subset of $\mathcal{T}_{\text{grid}}$. The results presented in this paper used the Matlab function *patternsearch*. We explored the use of other procedures, but found this to be relatively fast and reliable relative to others. We also note that when carrying out calculations, we set the values of the estimated coefficients to be the same unless they differ by more than 2.5% in absolute value.

2.5 Asymptotic Distribution

Let $\text{sign}(\cdot)$ be the sign function and such that the sign of a vector is the vector of its signs.

Theorem 2 *Suppose that the Regularity Conditions hold and that* $\mathbb{E}\left|N^{(i)}(1)\right|^r < \infty$ *for some* $r > 2$. *Then*

$$\sqrt{n}\left(\hat{\theta} - \theta_0\right) + \sqrt{n}\tau D'\text{sign}\left(D\hat{\theta}\right) = \sqrt{n}\left(Y - \theta_0\right).$$

Let a *be an* $L \times 1$ *vector with entries bounded by one in absolute value. Then, as* $n \to \infty$, *for any* $L \geq 1$, *including* $L \to \infty$,

$$\sqrt{n}a'\left(Y - \theta_0\right) \to a'Z$$

in distribution where Z *is a mean zero Gaussian random vector whose covariance matrix* C *has* (k,l) *entry equal to* $\lim_n \frac{1}{n}\sum_{i,j=1}^n Cov\left(\int_{\Delta_k} dN^{(i)}, \int_{\Delta_l} dN^{(j)}\right)$, *and* $Var\left(a'Z\right) = a'Ca < \infty$.

The result is meant to hold for $L \to \infty$ at a rate that is independent of n. In this case, we mean that both $\sqrt{n}a'\left(Y - \theta_0\right)$ and $a'Z$ converge in distribution to the same Gaussian random variable. Depending on the choice of a the limiting Gaussian process can be degenerate. The simplest example is when $L \to \infty$ and a concentrates over a shrinking interval, e.g. $a_l = 1$ if $l = 1$ and zero otherwise.

For $L \to \infty$ faster than n, which is the case of interest in this paper, the covariance matrix cannot be estimated consistently. However, for inference, our interest lies in finding a consistent estimator of $Var\left(a'Z\right)$ for fixed a. In this case, we can use standard autoregressive consistent estimators based on a suitable kernel such as the Bartlett kernel. For definiteness,

$$\hat{\sigma}_a^2 := \sum_{|i-j| \leq m}\left(1 - \frac{|i-j|}{m+1}\right)\hat{\gamma}_a\left(|i-j|\right) \tag{9}$$

where m is an integer and

$$\hat{\gamma}_a\left(|i-j|\right) := \frac{1}{n}\sum_{i=1}^n\left(\sum_{l=1}^L a_l \int_{\Delta_l} dN^{(i)}\right)\left(\sum_{l=1}^L a_l \int_{\Delta_l} dN^{(j)}\right).$$

Lemma 1 *Suppose that the conditions of Theorem 1 hold and that the vector* a *is such that* $Var\left(a'Z\right) > 0$ *as* $L \to \infty$. *Then, for* $m = o\left(n^{1-\max\{2/r,1/2\}}\right)$, $\left|\hat{\sigma}_a^2 - a'Ca\right| \to 0$ *in probability.*

3 Extensions

We briefly discuss extensions to some problems of interest.

3.1 Volumes Profile

Volumes profiles are obtained if we associate trade sizes to each trade arrival. Let $V^{(i)}(t)$ be the volume cumulated until time t on day i. Suppose that $\mathbb{E}V^{(i)}(t) = \int_0^t v(s)\,ds$, where $v(s)$ is the traded volume intensity. We are interested in an estimator for $v(s)$. We could use the model

$$v(t) = \sum_{j=1}^{J} v_j h_j 1_{\mathcal{D}_j}(t).$$

Assuming that (2) holds true, we deduce that the coefficient v_j represent the expected trade size during the interval $\mathcal{D}_j$. These can be consistently estimated once $\hat{\mathcal{D}}$ is estimated.

3.2 Conditioning

Suppose that $\mathcal{G}$ is the sigma algebra generated by some random variable, and we are interested in $\mathbb{E}\left[N^{(i)}(t)\,|\mathcal{G}\right]$. We make the following modelling assumption:

$$\mathbb{E}\left[N^{(i)}(t)\,|\mathcal{G}\right] = \int_0^t \sum_{j=1}^{J} g_j 1_{\mathcal{D}_j}(s)\,ds.$$

Whether this is a strong assumption, it depends on the exact definition of $\mathcal{G}$. The simplest example is to condition on the day of the week. Because of news announcements, option expiries etc., the profile could be different on different days of the week. Next, we give two more examples.

Previous Day Trade Count. Suppose that the processes $N^{(i)}$ depend on the final value of the previous day process $N^{(i-1)}(1)$. We want to use $\mathcal{G} = \left\{N^{(i-1)}(1) \in G\right\}$ where G is a closed interval in the positive real line. Then, to estimate the coefficients g_j we use

$$\hat{g}_j = \frac{1}{n_G} \sum_{i=1}^{n} 1_{\left\{N^{(i-1)}(1)\in G\right\}} \left|\hat{\mathcal{D}}_j\right|^{-1} \int_{\hat{\mathcal{D}}_j} dN^{(i)}(t),$$

where $n_G = \sum_{i=1}^{n} 1_{\left\{N^{(i-1)}(1)\in G\right\}}$. If $n_G = 0$ $\hat{g}_j := 0$.

End of Day Trade Count. Suppose that $\mathcal{G} = \left\{N^{(i)}(1) \in G\right\}$. Then, $\hat{g}_j$ is as above with $1_{\left\{N^{(i-1)}(1)\in G\right\}}$ replaced by $1_{\left\{N^{(i)}(1)\in G\right\}}$. This case is of interest in some prediction problems if we have a prediction interval for $N^{(n+1)}(1)$ that can be updated during the trading day, say $G^{(n+1)} = [\nu(t) - \epsilon, \nu(t) + \epsilon]$, for some $\epsilon > 0$, where $\nu(t)$ defines a point prediction for $N^{(n+1)}(1)$ that is updated during the $n+1$ trading day. This allows us to split the prediction problem into a conditional profile estimation and a prediction of the daily trades count that

can be updated intraday (Sancetta, 2019, for the case of an intraday point prediction for end of day volume).

3.3 Profile Estimation for Illiquid Instruments

When instruments are illiquid, the partition Δ can be too fine and not much data could fall in Δ_l for some $l \in \{1, 2, ..., L\}$. To mitigate the problem, we can consider and adaptive partition $\tilde{\Delta} = \left\{\tilde{\Delta}_1, \tilde{\Delta}_2, ..., \tilde{\Delta}_I\right\}$ where I is an integer that depends on the sample. In particular, suppose that $\tilde{\Delta}_{p-1} = \left(\frac{d_{p-2}}{L}, \frac{d_{p-1}}{L}\right]$ is given, and let

$$d_p := \min\left\{x \in \{d_{p-1}+1, d_{p-1}+2, ..., L\} : \frac{1}{n}\sum_{i=1}^{n}\int_{d_{p-1}/L}^{x/L} dN^{(i)}(t) \geq \rho\bar{N}\right\}.$$

Here, $\bar{N} := \frac{1}{n}\sum_{i=1}^{n} N^{(i)}(1)$ and $\rho \in (0,1)$. We can choose $\rho = 0.01$ in which case at least 1% of the data falls in each bin. Hence, $\tilde{\Delta}$ depends on ρ, the sample and the initial partition Δ via L. This construction ensures that Δ is a refinement of $\tilde{\Delta}$. We define $Z_p = \left(L\left|\tilde{\Delta}_p\right|\right)^{-1}\frac{1}{n}\sum_{i=1}^{n}\int_{\tilde{\Delta}_p} dN^{(i)}(t)$ and set $Y_l = Z_{p(l)}$ where $p(l)$ is such that $\Delta_l \in \tilde{\Delta}_{p(l)}$. As usual Y_l is the l^{th} entry in Y used in (3). The quantity $L\left|\tilde{\Delta}_p\right|$ is the number of intervals of size L^{-1} contained in $\tilde{\Delta}_p$. This construction forces $Y_l = Y_{l+j}$ for $j \geq 1$ if not enough data can be found in a bin of size $1/L$. Work carried out by the author, but not reported here showed that for illiquid instruments this adjustment mitigates issues related to data sparsity. By setting $\rho = 0$ we recover the usual definition of Y.

4 The Intensity of Trades Arrival on Six CME Futures

We consider the front month of six futures contracts traded on the CME: EURUSD (6E), AUDUSD (6A), EURCHF (RF), EURGBP (RP), e-mini S&P500 (ES), Crude Oil (CL). The CME ticker for the Globex platform is reported in parenthesis. Globex is the electronic platform of CME, to be distinguished from the open outcry. Two of these six futures are not very liquid (RF, RP) and are included to gauge whether the estimator performs well in this case. The sample period is 2013.04.01-2013.10.01. The data is proprietary and was collected by a Chicago trading firm colocated at the Aurora data center in Chicago, with nanosecond precision timestamps. All messages sent by the CME, were recorded with no conflation.

The study aims to address the following four points: (1.) identify the salient features of the intraday profile based on the day of the week; (2.) produce a predictor of the profile for different days of the week; (3.) compare the out-of-sample predictions generated by the adaptive estimator to one-minute and thirty-minute bin estimators; (4.) verify if accounting for different days of the week is important.

4.1 Estimation and Performance Evaluation

The expected intensity h is estimated conditioning on different days of the week, as outlined in Section 3.2. We consider the estimator (4) of this paper with τ chosen by CV as described in Section 2.4. We choose the partition such that each Δ_l is one minute. Hence, we have 1440 bins in a day. On Fridays due to trading ending before midnight GMT, we have less bins. Figure 1 shows the one-minute binned profile for crude oil. When carrying out calculations, we also set the values of the estimated coefficients to be the same unless they differ by more than 2.5% in absolute value. To evaluate the performance we use the loglikelihood. The loglikelihood for the counting processes $\left(N^{(i)}\right)_{i\geq 1}$ is proportional to

$$\sum_{i=1}^{n}\sum_{j=1}^{N^{(i)}(1)}\ln\left(\lambda^{(i)}\left(T_j^{(i)}\right)\right)-\sum_{i=1}^{n}\int_0^1\lambda^{(i)}(t)\,dt,$$

where $T_j^{(i)}$ is the time of the j^{th} jump on the i^{th} day. However, given that we are not interested in the stochastic intensities $\lambda^{(i)}$, we replace these with our estimators for h. We denote by $\mathcal{L}_n(h)$ the loglikelihood using an intensity h in place of $\lambda^{(i)}$ in the above display, and consider the test statistic

$$S_n:=\frac{\mathcal{L}_n\left(h^{(1)}\right)-\mathcal{L}_n\left(h^{(2)}\right)}{\hat{\sigma}_{1,2}}. \tag{10}$$

Here, $h^{(1)}$ and $h^{(2)}$ are two competing models and

$$\hat{\sigma}_{1,2}^2=\sum_{i=1}^{n}\sum_{j=1}^{N^{(i)}(1)}\left|\ln\left(h^{(1)}\left(T_j^{(i)}\right)/h^{(2)}\left(T_j^{(i)}\right)\right)\right|^2.$$

Model parameters are estimated in the estimation sample, and the statistic S_n is computed in the subsequent test sample of n days. In particular, here we use 67% of the sample to compute the adaptive model, the one-minute, and the thirty-minute binned estimators conditioning on different days of the week. For comparison, we also estimate the adaptive model without distinguishing by day of the week, except for Fridays. We then use the remaining 33% of the sample to compute S_n. Under the null that the models perform equally well, assuming that the intensity is not stochastic, S_n is asymptotically standard normal (Sancetta, 2018, Proposition 1 for details).

4.2 Results

We compute summary statistics for the one-minute binned data in the form of the first four central moments and the autocorrelation function (acf). These show that the data are far from being generated from an homogeneous Poisson model and that time series dependence within each

day is strong. The persistency in the acf is often the result of heterogeneity/non-stationarity, as discussed by other authors in the related context of volatility estimation (Stărică and Granger, 2005). The theoretical setup of the paper covers such situations. Details are in Table 1.

We estimate the model for trade counts with penalty chosen using cross-validation. The resulting degrees of freedom/number of jumps tend to be unexpectedly high (see Table 2). With such large number of degrees of freedom, the estimated profiles capture the features of the trading activity. For example, the weekly Petroleum report from the U.S. Energy Information Administration is released on Wednesdays at 10:30am EST. The estimator for the crude oil profile on Wednesday captures the spike in trading activity around the release. This large spike at 10:30 EST is not present in the estimator for the other days. Figures 2 and 3 show the estimated trades profile for Wednesdays and Thursdays, for comparison. Due to daylight saving time, this corresponds in the plots to 14:30 GMT, which is 870 minutes after midnight.

Using the statistic (10) we find that there is clear gain in using the adaptive estimator as opposed either a one-minute or 30-minute binned estimator. Coarsely binning data at regular times is not the right way to obtain a less noisy estimator. In fact the results show that the performance is severely compromised. However, we also see that accounting for different days of the week may not always be crucial. In fact, we compare the adaptive estimator to the one where we do not condition on days of the week except for Fridays. This is because for our products, trading ends on Fridays at 4pm Central for the week. Table 3 reports the results. It is worth mentioning that considerations beyond the reported statistical fit may dictate the choice of model. For example, an algorithm that tries to exploit the liquidity around the Wednesday Petroleum report may require appropriate conditioning even if the overall fit is worse. More generally, if the cost of missing a spike in intensity is relatively high, one may need to resort to a loss function that accounts for such preferences in order to evaluate a model. Further discussion is beyond the scope of this paper.

The adaptive estimators appear to be relatively noisy. Hence, the analysis was also carried out imposing an upper bound on the maximum degrees of freedom. As opposed to Figure 2, Figure 4 shows the plot for such highly penalised estimator of crude oil. The plot is more regular. However, on an additional set of results, the highly penalised estimator fared considerably worse in terms of the test statistic (10) for all six futures. The reason is that a highly penalised estimator may miss important spikes and the cost of this in terms of loglikelihood is high.

5 Finite Sample Simulation Study

We simulate from a count process with deterministic intensity. We consider different scenarios in order to understand the strengths and weaknesses of the procedure. In particular we consider

the following nonstochastic intensity such that for every day i,

$$\lambda^{(i)}(t) = h(t) = \sum_{k=1}^{K} b_k 1_{\left(\frac{k-1}{K}, \frac{k}{K}\right]}(t) \tag{11}$$

where $K \in \{200, 2000\}$ and we vary the coefficients b_k. We define a partition $\{\mathcal{K}_0, \mathcal{K}_1\}$ of $\{1, 2, ..., K\}$ such that $b_k = K$ if $k \in \mathcal{K}_0$ and $b_k \in \{2K, 10K\}$ if $k \in \mathcal{K}_1$. Within this context, we discuss two configurations for $\mathcal{K}_0$ and $\mathcal{K}_1$.

Few discontinuities: J=3 in (2). We let $\mathcal{K}_1 = (c_1 K, c_2 K] \cap \{1, 2, ..., K\}$ with $(c_1, c_2) \in \{(0.45, 0.55), (0.49, 0.51)\}$. Hence the fraction of the sample where we have a higher intensity is equal to $d = c_2 - c_1$ and it is placed in the middle of the interval over which the intensity is defined. In this setup h is piecewise constant and larger for $t \in (c_1, c_2]$. This construction ensures that in each interval $\left(\frac{k-1}{K}, \frac{k}{K}\right]$ with $k \in \mathcal{K}_0$, we expect one event. The expected total number of events within each day is $O(K)$.

Many discontinuities: J=10,50 in (2). We let $\mathcal{K}_1 = \bigcup_{s \geq 1} (c_{1,s} K, c_{2,s} K] \cap \{1, 2, ..., K\}$, where $c_{1,s} = 2(s-1)d$, $c_{2,s} = (2s-1)d$, and $d \in \{0.1, 0.2\}$. Hence, we keep the fraction of the sample where intensity is constant equal to d and alternate between high and low intensity. According the previous remarks, the expected total number of events within each day is again $O(K)$, but larger than the previous setup.

Estimation details. For each simulation, $n \in \{50, 500\}$. We consider an estimator for h using the partition Δ with $L \in \{0.1K, 0.5K\}$. This means that within each bin $\Delta_l \in \Delta$, we expect a number b_k/L of events. The penalty is estimated using cross-validation, bias corrected AIC, and $\tau = \text{penScale} \times \tau_{\max}$ where $\text{penScale} \in \{1, 0.25, 0.1, 0\}$. When $\text{penScale} = 0$ we use $\tau = \sqrt{n^{-1} \ln L}$ (see Section 2.4).

This setup allows us to gauge the properties of the estimation procedure and penalty choice for numerous data generating processes.

Regarding the choice of τ, cross-validation gives the best results. Among various metrics we also look at false positives and negatives. A false positive happens when we estimate a jump when in fact h is constant. A false negative is when we do not detect a jump in h.

Cross validation produces a number of false positives, but not as many as the bias corrected AIC. The latter overfits and produces worst results. An illustrative subset of the results is reported in Table 4, where we look at the small sample n large L problem, which is one of the most difficult designs in our simulations. The full set of results is reported as supplementary material. While not reported there, we also note that standard AIC and Mallow's C_p performed worse.

We note that the price to pay for a false negative is large, particularly when the jump discontinuity is large as in the results from Table 4. In this case, it is always preferable to allow for excess false positive as long as the number of false negatives is minimised. To aid intuition, we also plotted the true intensity against the estimated ones. Visual inspection confirms that the procedure can successfully detect discontinuities (Figure 5 shows two examples using different sample sizes).

Irrespective of the method used to choose τ, we highlight a few main conclusions from the full set of simulations. As expected, the higher the true jump size in the intensity, the easier is to identify a jump discontinuity. Of course, the larger can be the error when missing a discontinuity. To estimate the trade profile, it is clear that we should use a large value of L. This means that instead of using L as initial smoothing parameter, L should be such that $L\mathbb{E}Y_l$ is a good approximation for $h\left(\frac{l}{L}\right)$. We would then rely on the estimator (4) to denoise Y. Table 5 substantiates these claims. We refer the reader to the supplementary material for the full set of results.

Finally, we also applied Theorem 2 to the simulated data. We considered $a_k = 1$ when $k \in \mathcal{K}_1$ and zero otherwise. We found that even when the cardinality of $\mathcal{K}_1$ is large, the Q-Q plot is relatively close to a normal distribution (Figure A.1.3.1 shows such result for the same simulation designs as in Figure 5).

6 Conclusion

This paper considered the problem of estimating the intraday trades profile of financial instruments. This is a quantity used as input in execution algorithms. For example, some implementation shortfall algorithms need to rely on such profile to determine ex ante how to break a large order into small ones and establish the overall required time to fully execute the order. The procedure is of interest when we need to identify significant recurrent events in an automated way for a large number of securities. In fact, the trades profile can be viewed as the intraday seasonal component in a counting process. The empirical results show that using traditional estimators such as thirty-minute averages of trades arrivals is suboptimal and misses the important features of the market.

The proposed procedure is simple to implement and only requires to solve a quadratic programming problem. The statistical framework is the one where we want to identify possible jumps in the true expected intensity, where the jump location is within possibly thousand different points within the day, and the number of days at disposal is possibly an order of magnitude lower. We impose no dependence conditions on the counting process within each day.

The theoretical results show that the estimator is consistent and simulation results show that in small samples the results are reasonably reliable. We also propose an extension to

less liquid instruments in order to mitigate noise in the estimation in thinly traded instruments. Our results also provide asymptotic normality of some function of the estimator. This asymptotic theory can be used to test if the profile may depend on the day of the week.

It is natural to think about estimation of these profiles as the first step in modelling the intensity of trade arrivals. The empirical results suggest that much of the activity that could be regarded as stochastic is deterministic. Hence, it is possible that the strong dependence in trade arrivals and the success of self exciting processes (Hawkes processes) might be due to unmodelled intraday seasonality. More work is needed in this direction and this will be the subject of future research.

References

[1] Alaya, M.Z., S. Gaïffas, A. Guilloux (2015) Learning the Intensity of Time Events With Change-Points. IEEE Transactions on Information Theory 61, 5148-5171.

[2] Aldous, D.J. (1976) A Characterisation of Hilbert Space Using the Central Limit Theorem. Journal of the London Mathematical Society 14, 376-380.

[1] Bauwens, L. and N. Hautsch (2009) Modelling Financial High Frequency Data Using Point Processes. In T.G. Andersen, R.A. Davis, J.-P. Kreiss and T. Mikosch (eds.), Handbook of Financial Time Series, 953-982. New York: Springer.

[3] Basrak, B., R.A. Davis and T. Mikosch (2002) Regular Variation of GARCH Processes, Stochastic Processes and their Applications 99, 95-115.

[4] Bradley, R.C. (1986) Basic Properties of Strong Mixing Conditions. In E. Eberlein and M.S. Taqqu (eds.), Dependence in Probability and Statistics, 165-192. Boston: Birkhauser.

[5] Brémaud, P. (1981) Point Processes and Queues: Martingale Dynamics. Berlin: Springer.

[6] Davidson, J. (1992) A Central Limit Theorem for Globally Nonstationary Near-Epoch Dependent Functions of Mixing Processes. Econometric Theory 8, 313-329.

[7] Davidson, J. (2020) A New Consistency Proof for HAC Variance Estimators. Economics Letters 186, 1-4.

[8] Engle, R.F. and J.R. Russell (1998) Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. Econometrica 66, 1127-1162.

[9] Hurvich, C.M., J.S. Simonoff and C.-L. Tsai (1998) Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion. Journal of the Royal Statistical Society B 60, 271-293.

[10] Mokkadem, A. (1988) Mixing Properties of ARMA Processes. Stochastic Processes and their Applications 29, 309-315.

[11] Mokkadem, A. (1990) Propriétés de Mélange des Processus autorégressifs Polynomiaux. Annales de l'I.H.P. Probabilités et Statistiques 26, 219-260.

[12] Rinaldo, A. (2009) Properties and Refinements of the Fused Lasso. Annals of Statistics 37, 2922-2952.

[13] Rio, E. (2000) Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants. Paris: Springer.

[14] Sancetta, A. (2018) Estimation for the Prediction of Point Processes with Many Covariates. Econometric Theory 34, 598-627.

[15] Sancetta, A. (2019) Intraday End-of-Day Volume Prediction. Journal of Financial Econometrics, https://doi.org/10.1093/jjfinec/nbz019.

[16] Stărică, C. and C. Granger (2005) The Review of Economics and Statistics 87 503-522.

[17] Stone, C.J. (1982) Optimal Global Rates of Convergence for Nonparametric Regression. Annals of Statistics 10, 1040-1053.

[18] Tibshirani, R.J. and J. Taylor (2011) The Solution Path of the Generalized Lasso. Annals of Statistics 39, 1335-1371.

[19] van der Vaart, A. and J.A. Wellner (2000) Weak Convergence and Empirical Processes. New York: Springer.

Appendix

A.1 Proofs

We introduce some additional notation. Define $\varepsilon_l = \frac{1}{n}\sum_{i=1}^n \int_{\Delta_l} dM^{(i)}(t)$, and $\theta_l = \frac{1}{n}\sum_{i=1}^n \int_{\Delta_l} \lambda^{(i)}(s)\, ds$. Note that $\theta_{0,l} = \mathbb{E}\theta_l$. Here, we shall just use $\mathbb{E}\theta_l$ when referring to $\theta_{0,l}$. Then, $\eta = (1-\mathbb{E})\,\theta + \varepsilon$, using vector notation. This decomposes the error into the difference between the stochastic intensity and the expected one, and a martingale difference. Once again, it is important to keep in mind the implicit dependence on n and L.

Define $\mathcal{B}_j := \{l : \Delta_l \in \mathcal{D}_j\}$, and $b_{\min} := \min_{j \le J} |\mathcal{B}_j|$ to be the smallest number of elements in Δ within each $\mathcal{D}_l$. Then, $b_{\min}$ depends on L and for $L \to \infty$, $b_{\min} = \lfloor L\delta_{\min} \rfloor$, where $\lfloor \cdot \rfloor$ is the integer part of its argument. This is because there are L elements in Δ and the size of the smallest $\mathcal{D}_j$ is $\delta_{\min}$.

To ease notation, we may use $b_j := |\mathcal{B}_j|$. For each $l \in \{1, 2, ..., L\}$, we let $\mathcal{B}_{s(l)}$ be $\mathcal{B}_j$ where j is such that $\Delta_l \in \mathcal{D}_j$, and $b_{s(l)} := \left|\mathcal{B}_{s(l)}\right|$. Hence, given $l \le L$, $\mathcal{B}_{s(l)}$ is the block of integers $\mathcal{B}_j$ such that $\Delta_l \in \mathcal{D}_j$. For arbitrary $l \le L$, this is only possible when $L \to \infty$ and surely larger than $\delta_{\min}^{-1}$.

The structure of the proof for the first part is the same as the proof of Theorem 2.3 in Rinaldo (2009). There, control of certain errors via maximal inequalities is shown to imply the event $\mathcal{R}$. Due to the different structure of the problem and different tail conditions, we need to apply different maximal inequalities and rely on the decomposition (1). To this end we need a few preliminary lemmas.

A.1.1 Preparatory Lemmas

We shall need the second moment of the process.

Lemma 2 *Under the Regularity Conditions,* $\mathbb{E}\left(\int dN^{(i)}(t)\right)^2 \le 2\int h(t)\, dt + 2\mathbb{E}\left(\int \lambda^{(i)}(t)\, dt\right)^2$.

Proof. Adding and subtracting $\int \lambda^{(i)}(t)\, dt$, and using a basic inequality,

$$\mathbb{E}\left(\int dN^{(i)}\right)^2 \le 2\mathbb{E}\left(\int dM^{(i)}\right)^2 + 2\mathbb{E}\left(\int \lambda^{(i)}(t)\, dt\right)^2.$$

The first expectation on the r.h.s. is equal to the expectation of the compensator which in turn equals $\int h(t)\, dt$, and we obtain the statement of the lemma. ■

We shall make use of the following basic inequality.

Lemma 3 *Under the Regularity Conditions,*

$$\mathbb{E}\left(\int_{\Delta_l} \lambda^{(i)}(s)\, ds\right)^2 \le L^{-2}\bar{\lambda}_2. \tag{A.1}$$

Proof. We prove this for arbitrary moments $r \geq 1$ assuming they exist. Write $\left(\int_{\Delta_l} \lambda^{(i)}(s)\, ds\right)^r$ as $|\Delta_l|^r \left(\int_{\Delta_l} \lambda^{(i)}(s) \frac{ds}{|\Delta_l|}\right)^r$, and use Jensen inequality to deduce that the latter quantity is bounded above by $L^{-(r-1)} \int_{\Delta_l} \left|\lambda^{(i)}(s)\right|^r ds$ using $|\Delta_l| = L^{-1}$. We then take expectation, and by Tonelli's theorem we can switch the order of integration and expectation. This implies that the l.h.s. of (A.1) is bounded above by $L^{-(r-1)} \int_{\Delta_l} \mathbb{E}\left|\lambda^{(i)}(s)\right|^r ds$. Using Condition 2 we deduce the statement of the lemma for $r = 2$. ■

We shall need the following Bernstein inequality for beta mixing random variables.

Lemma 4 *Let $(X_i)_{i\geq 1}$ be a sequence of random variables with values in a Polish space $\mathcal{X}$ and with beta mixing coefficient $\beta(m)$ for any $m \geq 1$. Let $\mathcal{F}$ be a finite set of functions on $\mathcal{X}$. Suppose that $\max_{i\leq n} \max_{f\in\mathcal{F}} Var(f(X_i)) < \infty$. Then, for any integer valued sequence $m \to \infty$ such that $n/m \to \infty$*

$$\Pr\left(\max_{f\in\mathcal{F}} \left|\frac{1}{n}\sum_{i=1}^{n} (1-\mathbb{E}) f(X_i)\right| > x\right) \leq 4\left(1+\frac{2m}{n}\right)\left(\frac{m}{nx^2}\right) \max_{i\leq n} \sum_{f\in\mathcal{F}} Var(f(X_i)) + \frac{2n\beta(m)}{m}.$$

Proof. The argument is well known, but outlined for clarity, as we make some slightly changes. For $j = 1, 2, 3, ...$, partition $\{1, 2, ..., n\}$ into blocks

$$\mathcal{H}_{1,j} := \{i \leq n : 1 + 2(j-1)m \leq i \leq m + (2j-1)m\},\ \mathcal{H}_{2,j} := \{i \leq n : 1 + (2j-1)m \leq i \leq 2jm\}.$$

Write $Z_{1,j}(f) := \frac{1}{m}\sum_{i\in\mathcal{H}_{1,j}} f(X_i)$, $Z_{2,j}(f) := \frac{1}{m}\sum_{i\in\mathcal{H}_{2,j}} f(X_i)$. To ease notation, assume that $\mathbb{E}f(X_i) = 0$ for all $f \in \mathcal{F}$ and $i \geq 1$. Let n_1 and n_2 be the number of non empty elements in $\{\mathcal{H}_{1,j} : j \geq 1\}$ and $\{\mathcal{H}_{2,j} : j \geq 1\}$, respectively. Note that $\{\mathcal{H}_{1,j} : j = 1, 2, ..., n_1\} \cup \{\mathcal{H}_{2,j} : j = 1, 2, ..., n_2\} = \{1, 2, ..., n\}$, and that $n_1, n_2 \in [\lfloor n/(2m)\rfloor, \lfloor n/(2m)\rfloor + 1]$. By a basic set inequality, $\Pr\left(\max_{f\in\mathcal{F}} \left|\frac{1}{n}\sum_{i=1}^{n} f(X_i)\right| > x\right)$ is bounded above by

$$\Pr\left(\max_{f\in\mathcal{F}} \left|\frac{m}{n}\sum_{j=1}^{n_1} Z_{1,j}(f)\right| > x/2\right) + \Pr\left(\max_{f\in\mathcal{F}} \left|\frac{m}{n}\sum_{j=1}^{n_2} Z_{2,j}(f)\right| > x/2\right). \tag{A.2}$$

To ease notation define $X_{\mathcal{H}_{1,j}} := (X_i)_{i\in\mathcal{H}_{1,j}}$. By Berbee's lemma (Rio, 2000, Lemma 5.1), there is a sequence $\left(X^*_{\mathcal{H}_{1,j}}\right)_{j\leq n_1}$ such that $X^*_{\mathcal{H}_{1,j}}$ and $X_{\mathcal{H}_{1,j}}$ have same law, $X^*_{\mathcal{H}_{1,j}}$ is independent of the sigma algebra generated by $\left\{X_{\mathcal{H}_{1,i}} : i \leq j-1\right\}$ and $\Pr\left(X_{\mathcal{H}_{1,j}} \neq X^*_{\mathcal{H}_{1,j}}\right) \leq \beta(m)$. Note that $X_{\mathcal{H}_{1,j}}$ and $X_{\mathcal{H}_{1,j-1}}$ are separated by m observations. Define $Z^*_{1,j}(f) := \frac{1}{m}\sum_{i\in\mathcal{H}_{1,j}} f(X^*_i)$. Then,

$$\Pr\left(\max_{f\in\mathcal{F}} \left|\sum_{j=1}^{n_1} Z_{1,j}(f)\right| > \frac{nx}{2m}\right) \leq \Pr\left(\max_{f\in\mathcal{F}} \left|\frac{m}{n}\sum_{j=1}^{n_1} Z^*_{1,j}(f)\right| > \frac{nx}{2m}\right) + (n_1 - 1)\beta(m)$$

setting $Z^*_{1,i}(f) = Z_{1,i}(f)$ and applying Berbee's lemma to the remaining $n_1 - 1$ blocks of variables. The same applies to $Z_{2,j}(f)$. In consequence, with an additional error $(n_1 + n_2 - 2)\beta(m)$, in (A.2) we can consider sums of independent random variables. We deduce that

$$\begin{aligned}\Pr\left(\max_{f\in\mathcal{F}}\left|\frac{1}{n}\sum_{i=1}^{n}(1-\mathbb{E})f(X_i)\right| > x\right) \quad \leq \quad & \Pr\left(\max_{f\in\mathcal{F}}\left|\sum_{j=1}^{n_1} Z^*_{1,j}(f)\right| > \frac{nx}{2m}\right) \\ & + \Pr\left(\max_{f\in\mathcal{F}}\left|\sum_{j=1}^{n_2} Z^*_{2,j}(f)\right| > \frac{nx}{2m}\right) \\ & + (n_1 + n_2 - 2)\beta(m). \qquad \text{(A.3)}\end{aligned}$$

Then, by the union bound, Chebyshev's inequality and independence, the first term on the r.h.s. is bounded above by

$$\left(\frac{2m}{nx}\right)^2 n_1 \sum_{f\in\mathcal{F}} \mathbb{E}\left|Z^*_{1,j}(f)\right|^2.$$

By convexity of the square we deduce that $\left|Z^*_{1,j}(f)\right|^2 \leq \frac{1}{m}\sum_{i\in\mathcal{H}_{1,j}}|f(X^*_i)|^2$ so that, swapping order of summation, the above display is bounded above by $\left(\frac{2m}{nx}\right)^2 n_1 \max_i \sum_{f\in\mathcal{F}} \mathbb{E}|f(X^*_i)|^2$. We repeat the same argument for the second term on the r.h.s. of (A.3). Then, the statement of the lemma follows once we recall that $n_1, n_2 \leq \lfloor n/(2m)\rfloor + 1$ so that $\max\{n_1, n_2\} \leq n(2m)^{-1} + 1$ and $(n_1 + n_2 - 2)\beta(m) \leq nm^{-1}\beta(m)$. ■

The following is an application of the above.

Lemma 5 *Under the Regularity Conditions, for any $x > 0$,*

$$\Pr\left(\max_{l\leq L}|\theta_l - \mathbb{E}\theta_l| \geq x\right) \leq 4\left(1 + \frac{2m}{n}\right)\left(\frac{m\bar{\lambda}_2}{nx^2L}\right) + \frac{2n\beta(m)}{m}.$$

Proof. Note that $\theta_l - \mathbb{E}\theta_l = \frac{1}{n}\sum_{i=1}^{n}(1-\mathbb{E})\int_{\Delta_l}\lambda^{(i)}(s)\,ds$. Moreover, from Lemma 3, $\mathbb{E}\left(\int_{\Delta_l}\lambda^{(i)}(s)\,ds\right)^2 \leq L^{-2}\bar{\lambda}_2$. Bounding the variance by the second moment and applying Lemma 4, we deduce the statement of the lemma. ■

We now control the error term.

Lemma 6 *Under the Regularity Conditions, for any $x > 0$,*

$$\Pr\left(\max_{l\leq L}|\varepsilon_l| \geq x\right) \leq 4\left(1 + \frac{2m}{n}\right)\left(\frac{m}{nx^2}\right)\int_0^1 h(s)\,ds + \frac{2n\beta(m)}{m}.$$

Proof. By definition, $\varepsilon_l = \frac{1}{n}\sum_{i=1}^{n}\int_{\Delta_l} dM^{(i)}(t)$. By standard isometry, $\mathbb{E}\left(\int_{\Delta_l} dM^{(i)}(t)\right)^2 = \mathbb{E}\int_{\Delta_l}\lambda^{(i)}(s)\,ds$ and the r.h.s. is equal to $\int_{\Delta_l} h(s)\,ds$. We apply Lemma 4, so that summing over $l \leq L$ gives the statement of the lemma. ■

We shall also need the following bound.

Lemma 7 *For each $l \leq L$, let $\eta_{\mathcal{B}(l)} = \frac{1}{b_{s(l)}} \sum_{k \in \mathcal{B}_{s(l)}} \eta_k$. Under the Regularity Conditions, for any $x > 0$, and large L,*

$$\Pr\left(\max_{l \in \mathcal{J}} \left|\eta_{\mathcal{B}(l)} - \eta_{\mathcal{B}(l-1)}\right| \geq x\right) \leq 16\left(1 + \frac{2m}{n}\right)\left(\frac{16m}{nx^2}\right)\left(\frac{\bar{\lambda}_2}{L} + \frac{\int_0^1 h(s)\, ds}{b_{\min}^2}\right) + \frac{4n\beta(m)}{m}.$$

Proof. Clearly,

$$\Pr\left(\max_{l \in \mathcal{J}} \left|\eta_{\mathcal{B}(l)} - \eta_{\mathcal{B}(l-1)}\right| \geq x\right) \leq 2\Pr\left(\max_{l \in \mathcal{J}} \left|\eta_{\mathcal{B}(l)}\right| \geq x/2\right)$$

and using the definition $\eta_k = (1 - \mathbb{E})\,\theta_k + \varepsilon_k$, the r.h.s. is bounded above by

$$2\Pr\left(\max_{l \in \mathcal{J}} \left|\frac{1}{b_{s(l)}} \sum_{k \in \mathcal{B}_{s(l)}} (1 - \mathbb{E})\,\theta_k\right| \geq \frac{x}{4}\right) + 2\Pr\left(\max_{l \in \mathcal{J}} \left|\frac{1}{b_{s(l)}} \sum_{k \in \mathcal{B}_{s(l)}} \varepsilon_k\right| \geq \frac{x}{4}\right) =: I + II.$$

The event in the first probability is contained in the event $\left\{\max_{l \leq L} |(1 - \mathbb{E})\,\theta_l| \geq \frac{x}{4}\right\}$, and we bound I using Lemma 5 with x replaced by $x/4$:

$$I \leq 8\left(1 + \frac{2m}{n}\right)\left(\frac{16m\bar{\lambda}_2}{nx^2 L}\right) + \frac{2n\beta(m)}{m}.$$

We now bound II. We note that

$$\frac{1}{b_{s(l)}} \sum_{k \in \mathcal{B}_{s(l)}} \varepsilon_k = \frac{1}{b_{s(l)}} \sum_{k \in \mathcal{B}_{s(l)}} \frac{1}{n} \sum_{i=1}^{n} \int_{\Delta_k} dM^{(i)}(t).$$

Then, we apply Lemma 4 with $f(X_i) = \frac{1}{b_{s(l)}} \int_{\Delta_{\mathcal{B}_{s(l)}}} dM^{(i)}(t)$, where $\Delta_{\mathcal{B}_{s(l)}} = \bigcup_{k \in \mathcal{B}_{s(l)}} \Delta_k$ and we used linearity of the sum and integral. The mean zero variable $f(X_i)$ has variance $b_{s(l)}^{-2} \int_{\Delta_{\mathcal{B}_{s(l)}}} h(s)\, ds$. We also note that $[0,1] = \bigcup_{l \in \mathcal{J}} \Delta_{\mathcal{B}_{s(l)}}$ because for each distinct k and l in $\mathcal{J}$, $\Delta_{\mathcal{B}_{s(k)}}$ and $\Delta_{\mathcal{B}_{s(l)}}$ have empty intersection by definition of $\mathcal{J}$ and in fact $\left\{\Delta_{\mathcal{B}_{s(l)}} : l \in \mathcal{J}\right\}$ is a partition of $[0,1]$. By these remarks, from Lemma 4, summing over $l \in \mathcal{J}$, and using the lower bound $b_{s(l)} \geq b_{\min}$, we deduce that

$$II \leq 8\left(1 + \frac{2m}{n}\right)\left(\frac{16m}{nb_{\min}^2 x^2}\right)\int_0^1 h(s)\, ds + \frac{2n\beta(m)}{m}.$$

Putting the upper bounds for I and II together we obtain the statement of the lemma. ■

Next, we relate the smallest absolute change in the adjacent elements in θ_0 to the smallest jump in h.

Lemma 8 *Under the Regularity Conditions,* $\alpha := \min_{l\in\mathcal{J}} |\theta_{0,l} - \theta_{0,l-1}| = \min_{j\leq J} |h_j - h_{j-1}| / L$ *for* $L \to \infty$.

Proof. Using the definition of $\theta_{0,l}$ and the functional form of h,

$$\theta_{0,l} - \theta_{0,l-1} = \sum_{j=1}^{J} h_j \left(\int_{\Delta_l} 1_{\mathcal{D}_j}(s)\, ds - \int_{\Delta_{l-1}} 1_{\mathcal{D}_j}(s)\, ds \right).$$

For $l \in \mathcal{J}$ we must have a jump so that $\min_{l\in\mathcal{J}} |\theta_{0,l} - \theta_{0,l-1}| = \min_j |h_j - h_{j-1}| / L$ for L large enough. ■

A.1.2 Proof of Theorem 1

We follow the proof of Theorem 2.3 in Rinaldo (2009) except for minor modifications due to the fact that the data arise from counting processes with a panel data type of structure. Hence, we only provide a sketch and refer to Rinaldo (2009) when needed.

Let $\mathcal{J}^c$ be the complement of $\mathcal{J}$ in $\{1, 2, ..., L\}$, so that $|\mathcal{J}^c| \leq L$ by definition. Define $d_l^\eta = \eta_l - \eta_{l-1}$, $\alpha = \min_{l\in\mathcal{J}} |\mathbb{E}\theta_l - \mathbb{E}\theta_{l-1}|$. From Rinaldo (2009), $\Pr(\mathcal{R}) \to 1$ if the following hold:

(1) $\Pr\left(\max_{l\in\mathcal{J}^c} \left|d_l^\eta\right| \geq 4\tau\right) \to 0$ (eq. A.4 in Rinaldo, 2009);

(2) $\tau < \alpha b_{\min}/4$ (eq. A.7 and the display immediately below that in Rinaldo, 2009);

(3) $\Pr\left(\max_{l\in\mathcal{J}} \left|\eta_{\mathcal{B}(l)} - \eta_{\mathcal{B}(l-1)}\right| \geq \alpha/2\right) \to 0$ (eq. A.6 in Rinaldo, 2009).

Here, $\eta_{\mathcal{B}(l)}$ is as defined in Lemma 7.

Proof of Point (1) Start with the simple bound

$$\begin{aligned} I := \Pr\left(\max_{l\in\mathcal{J}^c} \left|d_l^\eta\right| \geq 4\tau\right) &\leq 2\Pr\left(\max_{l\leq L} |\eta_l| \geq 2\tau\right) \\ &\leq 2\Pr\left(\max_{l\leq L} |\theta_l - \mathbb{E}\theta_l| \geq \tau\right) + 2\Pr\left(\max_{l\leq L} |\varepsilon_l| \geq \tau\right). \end{aligned}$$

We shall apply Lemmas 5 and 6. By the Regularity Conditions, in those lemmas we can choose m a constant multiple of $\ln n$ to ensure that $n\beta(m)$ go to zero. By these lemmas, we see that choosing $\tau^{-1} = o\left(\sqrt{\frac{n}{\ln n}}\right)$ both probabilities on the r.h.s. go to zero.

Proof of Point (2) Using Lemma 8, and multiplying both sides by $b_{\min}$, we have that $\alpha b_{\min} > \min_{j\leq J} |h_j - h_{j-1}|\, b_{\min}/L$ for L large enough. Then, $\alpha b_{\min} > \min_{j\leq J} |h_j - h_{j-1}|\, \delta_{\min}/2$, given that $b_{\min} \geq L\delta_{\min}/2$. By this last inequality and (6), we deduce that $\alpha b_{\min} > 4\tau$ so that Point 2 is satisfied.

Proof of Point (3) This directly follows from Lemma 7 and the inequality $b_{\min} \geq L\delta_{\min}/2$, as long as $\alpha^{-1} = o\left(L\delta_{\min}\sqrt{\frac{n}{m}}\right)$, which is the case by the conditions of the theorem if $\sqrt{\frac{n\delta_{\min}^2}{\ln n}} \min_{j\leq J} |h_j - h_{j-1}| \to \infty$. This concludes the proof that $\Pr(\mathcal{R}) \to 1$.

We now show that $\left|\hat{h}-h\right|_{\infty} \to 0$ in probability. It is sufficient show that $I := \Pr\left(\max_{j\leq J}\left|\hat{h}_j - h_j\right| > \epsilon\right)$ for any $\epsilon > 0$. On $\mathcal{R}$, $\mathcal{D} = \hat{\mathcal{D}}$, so that $\hat{h}_j - h_j = \frac{1}{n}\sum_{i=1}^{n}\left|\mathcal{D}_j\right|^{-1}(1-\mathbb{E})\int_{\mathcal{D}_j} dN^{(i)}(t)$. We apply Lemma 4 with $X_i(\cdot) = N^{(i)}(\cdot)$ and $f_j(X_i) = \left|\mathcal{D}_j\right|^{-1}\int_{\mathcal{D}_j} dN^{(i)}(t)$ for $j = 1, 2, ..., J$. With this notation, $Var\left(f_j(X_i)\right)$ is bounded above by $2\delta_{\min}^{-2}\int_{\mathcal{D}_j} h(t)\,dt + 2\bar{\lambda}_2$ using Lemma 2 and Lemma 3 together with $\left|\mathcal{D}_j\right|^{-1} \leq \delta_{\min}^{-1}$. Hence, by Lemma 4 we have that $I \leq 8\left(1+\frac{2m}{n}\right)\left(\frac{m}{n\epsilon^2}\right)\left(\delta_{\min}^{-2}\int_0^1 h(t)\,dt + J\bar{\lambda}_2\right) + 2nm^{-1}\beta(m)$ with m proportional to $\ln n$. This goes to zero as long as $\frac{\left(\delta_{\min}^{-2}+J\right)\ln n}{n} \to 0$ as stated in the theorem, so that the proof is completed.

A.1.3 Proof of Theorem 2

Set the derivative of the absolute value $|\cdot|$ at zero equal to zero. Hence, its derivative is the sign function $\text{sign}(\cdot)$, which we defined to be zero at zero. The first derivative of (3) evaluated at the optimum $x = \hat{\theta}$, satisfies

$$-a'\left(Y - \hat{\theta}\right) + \tau a' D' \text{sign}\left(D\hat{\theta}\right) = 0$$

for any a with entries bounded by one in absolute value. Recall that by $\text{sign}(\cdot)$ of a vector we mean the vector of its signs. Adding and subtracting $a'\theta_0$, and multiplying by $\sqrt{n}$, deduce that

$$\sqrt{n}a'\left(\hat{\theta} - \theta_0\right) + \tau\sqrt{n}a' D' \text{sign}\left(D\hat{\theta}\right) = -\sqrt{n}a'\left(Y - \theta_0\right). \tag{A.4}$$

To ease notation, let $X^{(i)}$ be the vector with l^{th} entry equal to $X_l^{(i)} = \int_{\Delta_l}\left(dN^{(i)}(t) - h(t)\,dt\right)$, $l = 1, 2, ..., L$. By definition of Y,

$$\sqrt{n}a'\left(Y - \theta_0\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} a' X^{(i)}.$$

Then, $\left(a'X^{(i)}\right)_{i\geq 1}$ is a sequence of real valued beta mixing random variables. Of course, there is a possible implicit dependence of L on n. Convergence of the r.h.s. to a normal random variable follows from Theorem 3.6 in Davidson (1992), as his dependence condition A is satisfied in our case. To see this, note that the strong mixing coefficients are bounded by the beta mixing coefficients (Rio, 2000), hence the mixingale assumption in Davidson is satisfied: see the discussion below his Definition 2.1. Moreover, $\max_i \mathbb{E}\left|a'X^{(i)}\right|^r < \infty$ for some $r > 2$. To see this, we use the fact that each entry in a is bounded by one in absolute value, so that it is sufficient to verify that $\mathbb{E}\left|N^{(i)}(1)\right|^r < \infty$, which is the case by assumption. We used the fact that $\bigcup_{l=1}^{L}\Delta_l = (0,1]$ and that $N^{(i)}(0) = 0$. Finally, we need to check that $a'Ca = \lim_n \frac{1}{n}\sum_{i,j=1}^{n} Cov\left(a'X^{(i)}, a'X^{(j)}\right) < \infty$. By a standard covariance inequality, $Cov\left(a'X^{(i)}, a'X^{(j)}\right) \leq \beta\left(|i-j|\right)^{\frac{r-2}{r}} \max_i \mathbb{E}\left(\left|a'X^{(i)}\right|^r\right)^{2/r}$ (Rio, 2000, eq. 1.12), which implies that $a'Ca < \infty$ by the geometric decay of the mixing coefficients. This concludes the proof of

the theorem.

A.1.3.1 Proof of Lemma 1

By the same argument as in the proof of Theorem 2, $\mathbb{E}\left|\sum_{l=1}^{L} a_l \int_{\Delta_l} dN^{(i)}\right|^r < \infty$. Then, the lemma is a re-statement of Theorem 2.1 in Davidson (2020). Note that the same remarks regarding the dependence condition used the proof of Theorem 2 apply here.

Figures

Figure 1: One-Minute Binned Estimator of the Intensity of Trades Arrivals for Crude Oil (CL) on Wednesdays. Time in the abscissa is in GMT. The sample period is 2013.04.01-2013.10.01.

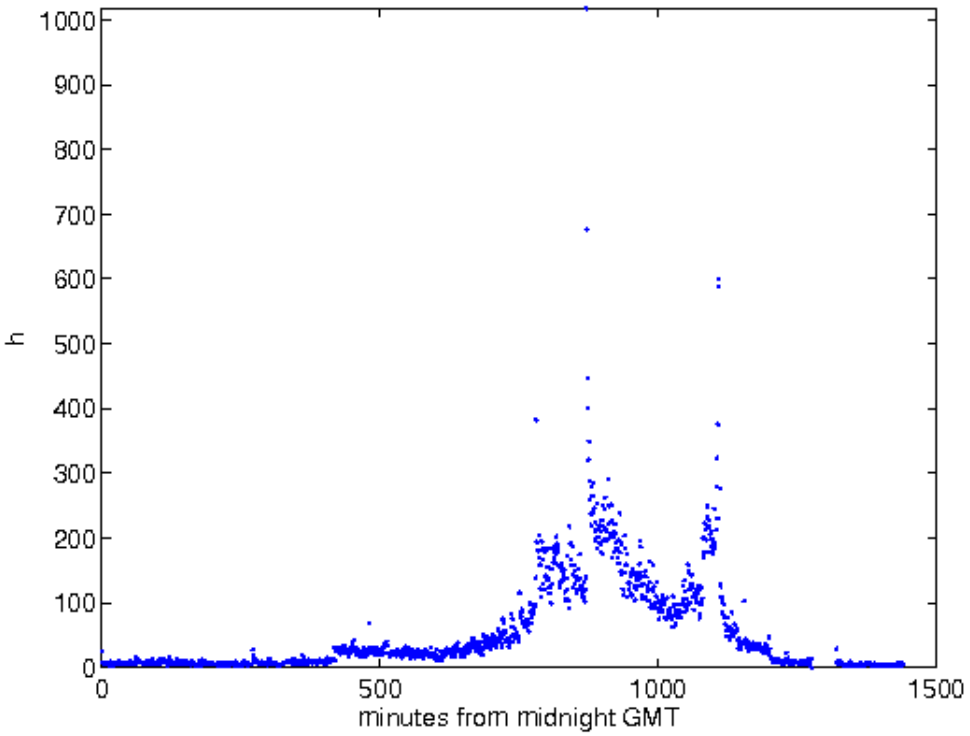


Figure 2: Adaptive Estimator of the Intensity of Trades Arrivals for Crude Oil (CL) on Wednesdays. Time in the abscissa is in GMT. The sample period is 2013.04.01-2013.10.01. The weekly Petroleum report from the U.S. Energy Information Administration is released on Wednesdays at 10:30am EST. Due to daylight saving time, 10:30am EST corresponds to 14:30 GMT, which is 870 minutes after midnight. At this time, a spike in intensity of trades arrival as high as 1000 contracts per minute can be seen.

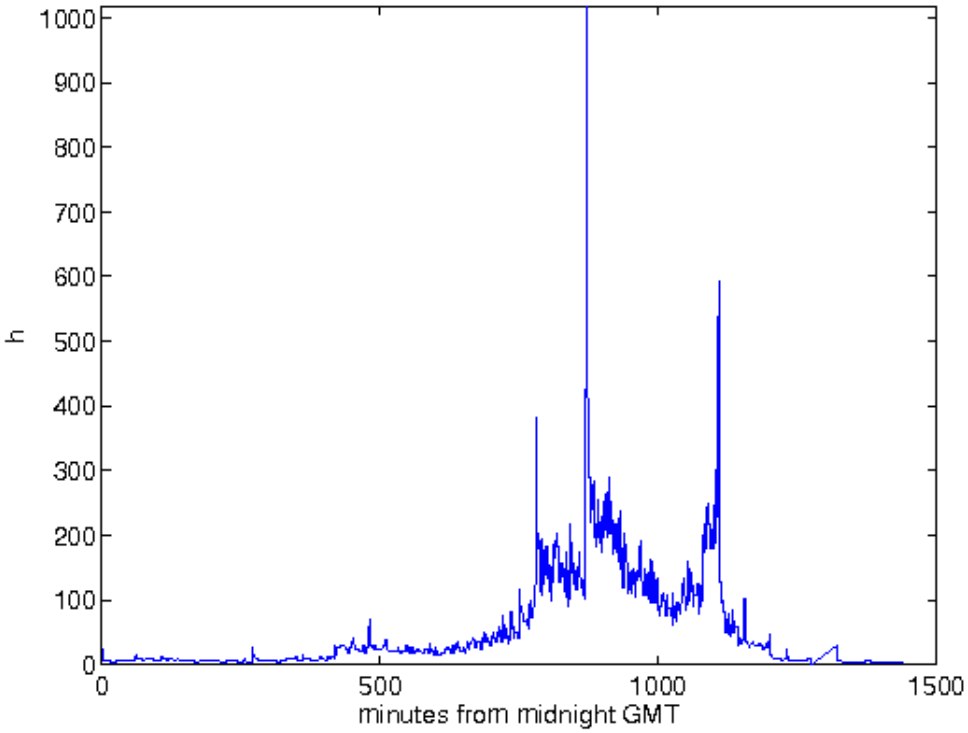


Figure 3: Adaptive Estimator of the Intensity of Trades Arrivals for Crude Oil (CL) on Thursdays. Time in the abscissa is in GMT. The sample period is 2013.04.01-2013.10.01.

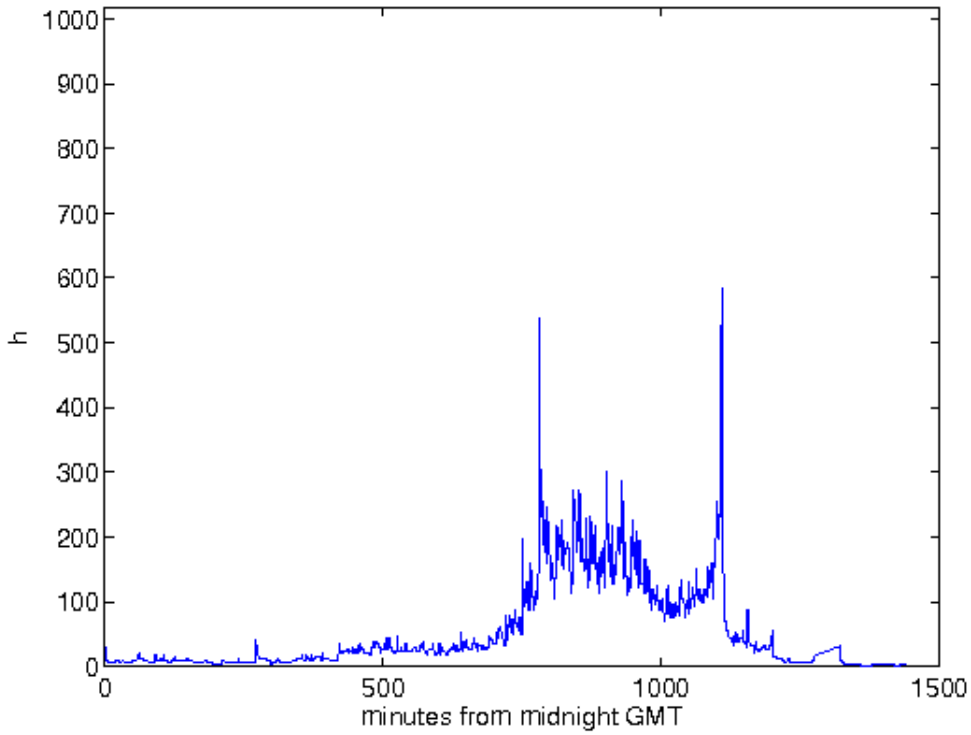


Figure 4: Highly Penalised Adaptive Estimator of the Intensity of Trades Arrivals for Crude Oil (CL) on Wednesdays. Details are as for Figure 2, but estimation was carried out increasing the penalty to reduce the maximum number of degrees of freedom.

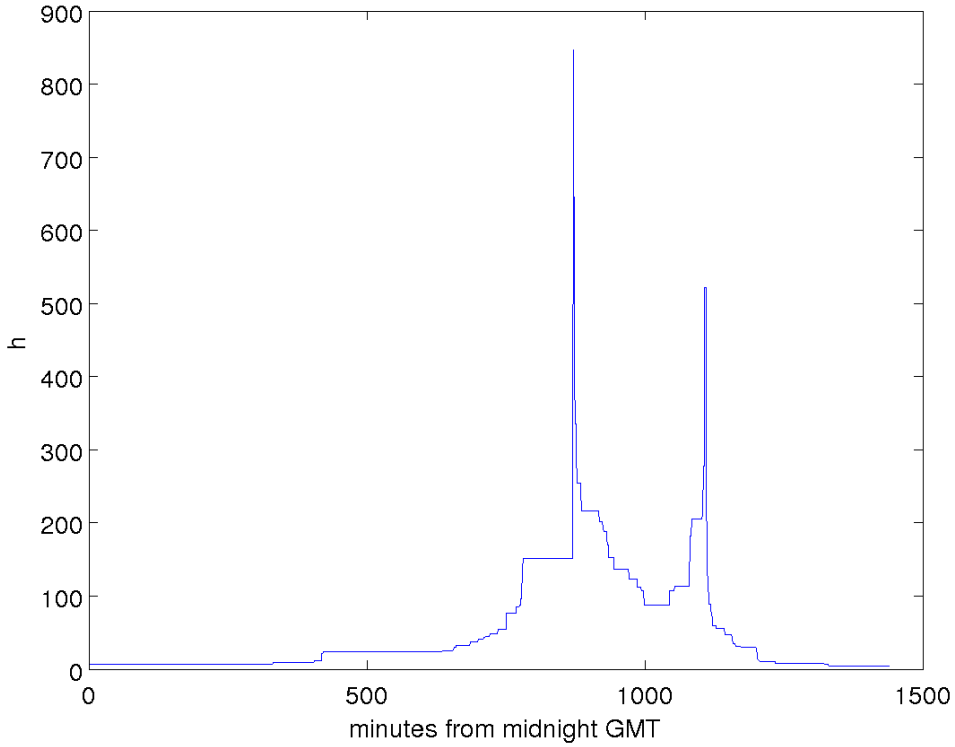


Figure 5: True and Estimated Intensities from Simulations. The true intensity is as in (11). The estimated intensity is based on the estimator in (3) using the penalty chosen by cross-validation with sample size $n = 50$ (top panels) and $n = 500$ (bottom panels). The left hand panels represent $K = 200$, $L = 20$, $J = 10$, $d = 0.1$ with $b_k = 400$ when $k \in \mathcal{K}_1$. The right hand side is for $K = 2000$, $L = 1000$, $J = 50$, $d = 0.02$, with $b_k = 4000$ when $k \in \mathcal{K}_1$. The estimated intensity with largest (max), smallest (min) median (med) error in L_1 norm are reported out of 250 simulations. For min and med the fit is indistinguishable from the true one in all plots.

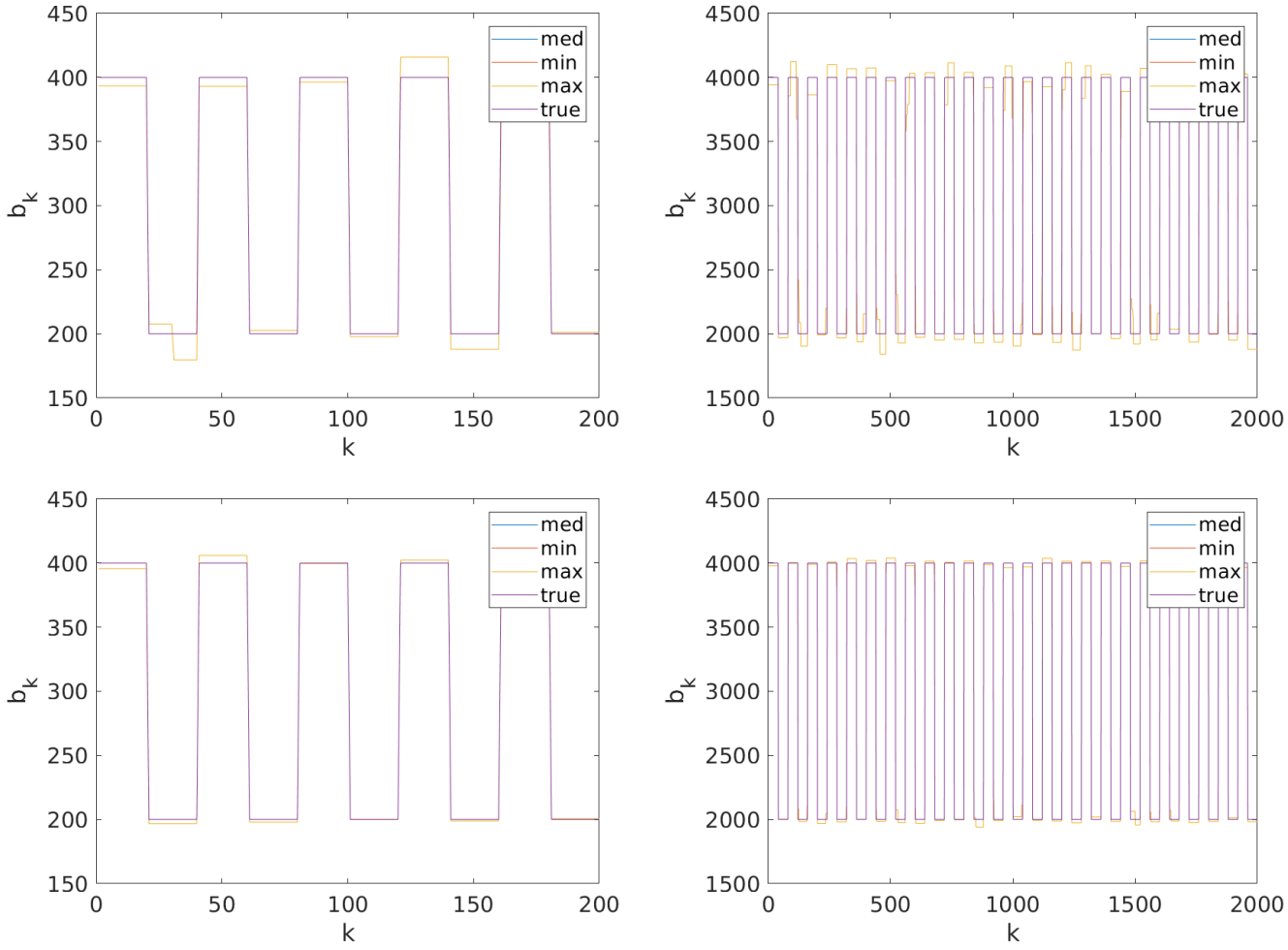
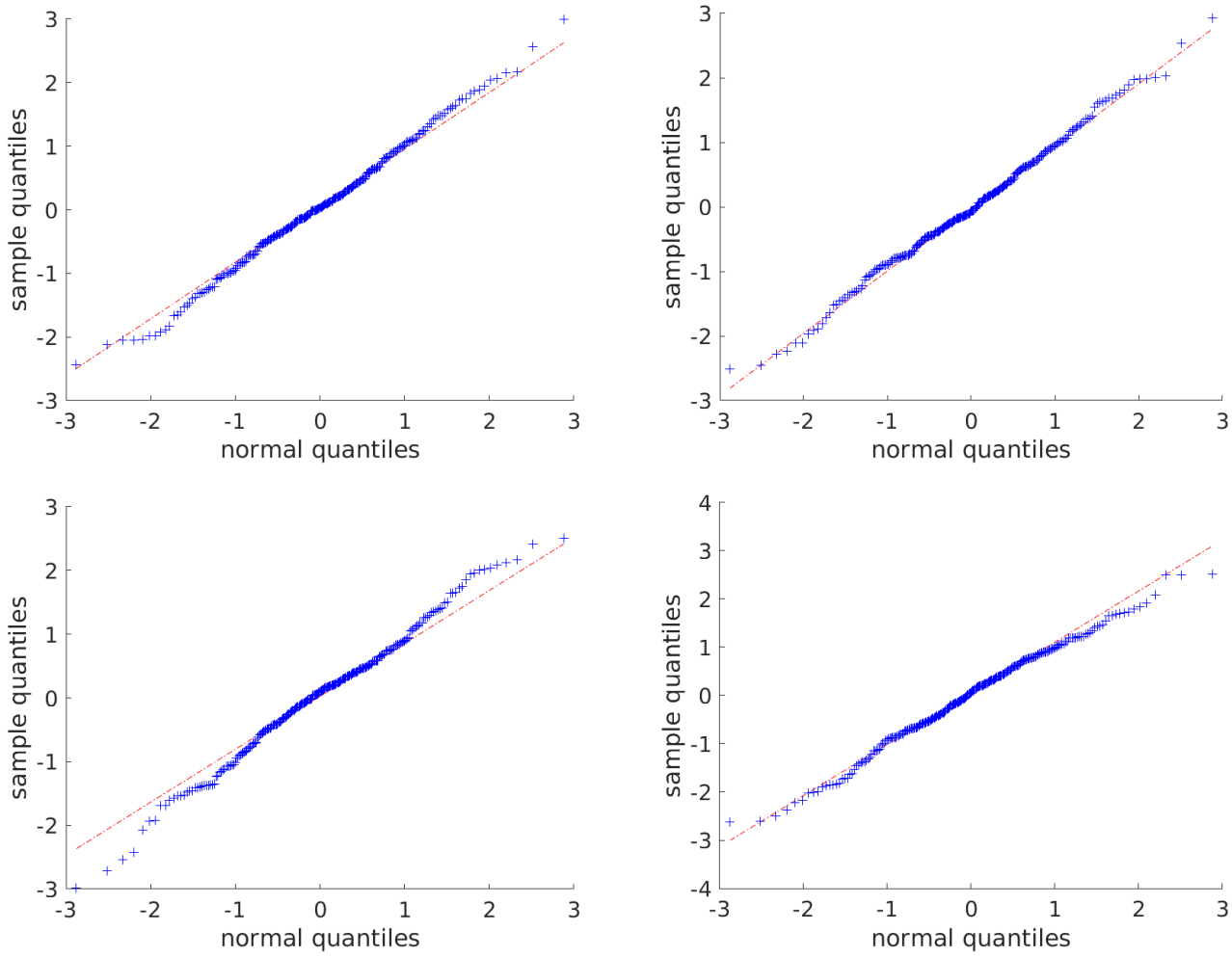


Figure 6: QQPlot for Statistics in Theorem 2. The plot is shown for the sample quantiles of $\sqrt{n}a'(Y-\theta_0)/\hat{\sigma}_a$ where $a_k = 1$ if $k \in \mathcal{K}_1$ and zero otherwise. The standard deviation $\hat{\sigma}_a$ in (9) is computed with $m = 0$. As by Theorem 2, the distribution is standard normal. The true intensity is as in (11). The estimated intensity is based on the estimator in (3) using the penalty chosen by cross-validation with sample size $n = 50$ (top panels) and $n = 500$ (bottom panels). The left hand panels represent $K = 200$, $L = 20$, $J = 10$, $d = 0.1$ with $b_k = 400$ when $k \in \mathcal{K}_1$. The right hand side is for $K = 2000$, $L = 1000$, $J = 50$, $d = 0.02$, with $b_k = 4000$ when $k \in \mathcal{K}_1$. The estimated intensity with largest (max), smallest (min) median (med) error in L_1 norm are reported out of 250 simulations.



Tables

Table 1: Summary Statistics. Summary statistics for one-minute binned number of trades are reported: mean, standard deviation (std), skewness (skew), kurtosis (kurt), the value of the sample autocorrelation function (acf) at lags 1, 10 and 20 (acf1, acf10, and acf20), the average number of trades within each day (avg.trd/day) and the total number of trades over the sample (tot.trd). The sample acf's are computed for each day and then averaged across all days. Standard errors for the acf's are not reported as all the numbers are significant at conventional levels.

	mean	std	skew	kurt	acf1	acf10	acf20	avg.trd/day	tot.trd
CL	49.72	84.89	4.29	35.55	0.76	0.55	0.49	68,971	5,862,563
6A	20.32	30.14	7.04	114.67	0.58	0.33	0.28	28,331	2,549,758
6E	40.11	71.16	8.89	167.48	0.62	0.39	0.33	55,913	5,032,155
RF	0.30	1.46	16.82	568.95	0.24	0.08	0.06	388	35,302
RP	0.39	1.65	13.86	381.19	0.26	0.09	0.07	510	46,437
ES	96.63	181.96	6.05	113.94	0.77	0.58	0.51	135,087	12,292,931

Table 2: Degrees of Freedom for Adaptive Estimated Model. The penalty is chosen using CV. Estimations are day of the week specific as discussed in Section 3.2.

	Mon	Tue	Wed	Thu	Fri
CL	653	588	540	582	531
6A	506	430	533	531	442
6E	570	542	628	562	534
RF	77	89	118	143	112
RP	89	119	158	162	117
ES	565	560	504	544	492

Table 3: Model Performance. The t-statistics for the standardised likelihood ratio statistic is reported. A large positive value favours the day of the week specific adaptive estimator. The statistic is asymptotically standard normal under the null of equality of model performance. A large value rejects a given model in favour of the day of the week specific adaptive intensity model. Binned30 and binned1 are equally spaced bin estimators with 30 and one-minute bin size. Adaptive_basic is the adaptive model that uses the same profile for Monday through Thursday and a separate one for Fridays.

	adaptive vs. binned30	adaptive vs. binned1	adaptive vs. adaptive_basic
CL	176.45	36.85	-42.15
6A	159.59	39.51	8.34
6E	155.32	33.34	34.75
RF	215.76	25.68	-1.32
RP	200.36	27.15	-5.52
ES	312.55	57.99	-5.80

Table 4: Simulations Sensitivity to Choice of τ. Results are reported for different designs with details in Section 5. In the details about the design, b1 denotes the value of b_k for $k \in \mathcal{K}_1$. The columns report the Monte Carlo approximation for the expected sup norm (sup), the expected L_1 norm, the expected number of false negative (fn), and the expected number of false positive (fp). An fn occurs when a true jump is missed, an fp when a jump is erroneously detect. The columns τ and df report the selected penalty and resulting degrees of freedom using CV and the modified AIC in (8), as described in Section 5.

	sup	L_1	fn	fp	τ	df
J=3 | n=50 |K=2000 | L=1000 | d=0.02 | b1=20000						
penScale =CV	486	0.007	0.8	5.9	2.3	8.1
penScale =AIC	327	0.018	0.7	27.6	0.4	29.9
penScale =1	17640	0.299	2.0	0.0	178.8	1.0
penScale =0.25	201	0.004	0.8	1.6	44.7	3.8
penScale =0.1	209	0.004	0.8	2.1	17.9	4.3
penScale =0	1042	0.029	0.8	58.1	0.4	60.3
J=50 | n=50 | K=2000 | L=1000 | d=0.02 | b1=20000						
penScale =CV	1411	0.010	2.0	79.7	1.3	80.7
penScale =AIC	2005	0.019	1.8	176.3	0.4	177.6
penScale =1	9012	0.818	2.0	0.0	187.3	1.0
penScale =0.25	323	0.007	2.0	49.0	46.8	50.0
penScale =0.1	358	0.007	2.0	49.0	18.7	50.0
penScale =0	2005	0.019	1.8	177.3	0.4	178.6

Table 5: Simulations for Penalty Chosen Using CV. Results are reported for different designs with details in Section 5. In the details about the design, b1 denotes the value of b_k for $k \in \mathcal{K}_1$. The columns report the Monte Carlo approximation for the expected sup norm (sup), the expected L_1 norm, the expected number of false negative (fn), and the expected number of false positive (fp). An fn occurs when a true jump is missed, an fp when a jump is erroneously detect. The columns τ and df report the selected penalty and resulting degrees of freedom using CV, as described in Section 5.

	sup	L_1	fn	fp	τ	df
J=10, 50 | n=50 | penScale =CV						
K=200 | L=20 | d=0.1 | b1=400	22	0.025	0.8	10.3	0.3	12.5
K=200 | L=20 | d=0.1 | b1=2000	44	0.011	0.9	9.4	0.5	11.5
K=200 | L=20 | d=0.02 | b1=400	137	0.321	0.9	9.4	5.6	11.5
K=200 | L=20 | d=0.02 | b1=2000	1112	0.785	0.8	9.5	0.5	11.7
K=200 | L=100 | d=0.1 | b1=400	50	0.035	1.9	18.8	0.5	19.9
K=200 | L=100 | d=0.1 | b1=2000	110	0.015	1.9	15.2	0.8	16.2
K=200 | L=100 | d=0.02 | b1=400	217	0.059	1.0	50.8	0.3	52.8
K=200 | L=100 | d=0.02 | b1=2000	1837	0.050	1.0	54.9	0.3	56.9
K=2000 | L=200 | d=0.1 | b1=4000	217	0.010	2.0	15.6	1.4	16.6
K=2000 | L=200 | d=0.1 | b1=20000	298	0.003	2.0	11.5	2.9	12.5
K=2000 | L=200 | d=0.02 | b1=4000	2062	0.022	1.7	74.5	0.4	75.8
K=2000 | L=200 | d=0.02 | b1=20000	18112	0.011	1.9	58.5	0.9	59.6
K=2000 | L=1000 | d=0.1 | b1=4000	466	0.011	2.0	22.9	1.6	23.9
K=2000 | L=1000 | d=0.1 | b1=20000	956	0.004	2.0	16.4	3.4	17.4
K=2000 | L=1000 | d=0.02 | b1=4000	646	0.027	2.0	112.9	0.6	113.9
K=2000 | L=1000 | d=0.02 | b1=20000	1411	0.010	2.0	79.7	1.3	80.7
J=10, 50 | n=500 | penScale =CV						
K=200 | L=20 | d=0.1 | b1=400	4	0.005	0.9	8.3	0.1	10.4
K=200 | L=20 | d=0.1 | b1=2000	11	0.003	0.9	8.2	0.2	10.2
K=200 | L=20 | d=0.02 | b1=400	124	0.320	0.9	8.4	0.1	10.5
K=200 | L=20 | d=0.02 | b1=2000	1087	0.785	1.0	8.0	0.2	10.0
K=200 | L=100 | d=0.1 | b1=400	15	0.007	2.0	12.5	0.1	13.6
K=200 | L=100 | d=0.1 | b1=2000	18	0.003	2.0	10.2	0.3	11.2
K=200 | L=100 | d=0.02 | b1=400	205	0.024	1.0	49.1	0.1	51.1
K=200 | L=100 | d=0.02 | b1=2000	1814	0.031	1.0	49.1	0.1	51.1
K=2000 | L=200 | d=0.1 | b1=4000	37	0.002	2.0	9.5	0.5	10.5
K=2000 | L=200 | d=0.1 | b1=20000	36	0.001	2.0	9.2	0.9	10.2
K=2000 | L=200 | d=0.02 | b1=4000	2014	0.006	2.0	52.7	0.1	53.7
K=2000 | L=200 | d=0.02 | b1=20000	18037	0.005	2.0	49.7	0.3	50.7
K=2000 | L=1000 | d=0.1 | b1=4000	140	0.002	2.0	12.3	0.5	13.3
K=2000 | L=1000 | d=0.1 | b1=20000	130	0.001	2.0	10.7	1.1	11.7
K=2000 | L=1000 | d=0.02 | b1=4000	199	0.006	2.0	67.7	0.2	68.7
K=2000 | L=1000 | d=0.02 | b1=20000	298	0.002	2.0	55.5	0.4	56.5

Supplement to "Intraday Trades Profile Estimation: An Intensity Approach" by Alessio Sancetta

Simulation Results

Results are reported for different designs with details in Section 5. In the details about the design, b1 denotes the value of b_k for $k \in \mathcal{K}_1$. The columns report the Monte Carlo approximation for the expected sup norm (sup), the expected L_1 norm, the expected number of false negative (fn), and the expected number of false positive (fp). An fn occurs when a true jump is missed, an fp when a jump is erroneously detect. The columns τ and df report the selected penalty and resulting degrees of freedom using different methods as described in Section 5

Table A: Simulation Results.

	sup	L_1	fn	fp	τ	df
J=3 | n=50 | penScale =AIC						
K=200 | L=20 | d=0.1 | b1=400	15	0.020	0.8	2.8	0.3	5.0
K=200 | L=20 | d=0.1 | b1=2000	23	0.015	0.7	3.0	0.3	5.2
K=200 | L=20 | d=0.02 | b1=400	163	0.050	0.8	3.2	0.3	5.4
K=200 | L=20 | d=0.02 | b1=2000	1443	0.259	0.7	2.7	0.3	5.0
K=200 | L=100 | d=0.1 | b1=400	192	0.035	1.6	7.7	0.4	9.2
K=200 | L=100 | d=0.1 | b1=2000	1791	0.049	1.6	9.8	0.3	11.2
K=200 | L=100 | d=0.02 | b1=400	32	0.035	0.8	7.5	0.3	9.7
K=200 | L=100 | d=0.02 | b1=2000	58	0.032	0.8	7.5	0.3	9.7
K=2000 | L=200 | d=0.1 | b1=4000	1968	0.020	1.5	28.2	0.4	29.8
K=2000 | L=200 | d=0.1 | b1=20000	17974	0.016	1.5	31.5	0.3	33.0
K=2000 | L=200 | d=0.02 | b1=4000	252	0.020	0.7	26.9	0.4	29.2
K=2000 | L=200 | d=0.02 | b1=20000	327	0.018	0.7	27.6	0.4	29.9
K=2000 | L=1000 | d=0.1 | b1=4000	516	0.029	0.7	50.8	0.4	53.0
K=2000 | L=1000 | d=0.1 | b1=20000	1585	0.023	0.7	65.7	0.4	68.0
K=2000 | L=1000 | d=0.02 | b1=4000	437	0.029	0.8	47.3	0.4	49.5
K=2000 | L=1000 | d=0.02 | b1=20000	1022	0.028	0.8	53.8	0.4	56.1
J=3 | n=50 | penScale =CV						
K=200 | L=20 | d=0.1 | b1=400	15	0.020	0.7	3.1	0.6	5.3
K=200 | L=20 | d=0.1 | b1=2000	22	0.014	0.8	2.8	1.1	5.0
K=200 | L=20 | d=0.02 | b1=400	165	0.050	0.9	2.9	8.6	5.0

Table A: Simulation Results.

	sup	L_1	fn	fp	τ	df
K=200 | L=20 | d=0.02 | b1=2000	1442	0.258	0.8	2.7	0.7	5.0
K=200 | L=100 | d=0.1 | b1=400	191	0.028	1.6	5.8	0.7	7.2
K=200 | L=100 | d=0.1 | b1=2000	1790	0.041	1.6	5.3	1.3	6.7
K=200 | L=100 | d=0.02 | b1=400	30	0.031	0.8	6.3	0.4	8.5
K=200 | L=100 | d=0.02 | b1=2000	52	0.026	0.8	5.9	0.6	8.2
K=2000 | L=200 | d=0.1 | b1=4000	1958	0.006	1.6	3.6	2.5	5.0
K=2000 | L=200 | d=0.1 | b1=20000	17967	0.007	1.6	3.7	4.5	5.1
K=2000 | L=200 | d=0.02 | b1=4000	127	0.006	0.8	3.3	1.4	5.6
K=2000 | L=200 | d=0.02 | b1=20000	192	0.006	0.8	3.9	2.1	6.1
K=2000 | L=1000 | d=0.1 | b1=4000	296	0.006	0.7	4.4	2.7	6.7
K=2000 | L=1000 | d=0.1 | b1=20000	524	0.004	0.7	3.9	5.0	6.2
K=2000 | L=1000 | d=0.02 | b1=4000	287	0.006	0.8	4.6	1.6	6.9
K=2000 | L=1000 | d=0.02 | b1=20000	486	0.007	0.8	5.9	2.3	8.1
J=3 | n=50 | penScale =1						
K=200 | L=20 | d=0.1 | b1=400	180	0.164	2.0	0.0	9.7	1.0
K=200 | L=20 | d=0.1 | b1=2000	1620	0.852	2.0	0.0	81.8	1.0
K=200 | L=20 | d=0.02 | b1=400	196	0.039	2.0	0.0	2.7	1.0
K=200 | L=20 | d=0.02 | b1=2000	1764	0.299	2.0	0.0	16.9	1.0
K=200 | L=100 | d=0.1 | b1=400	180	0.164	2.0	0.0	9.7	1.0
K=200 | L=100 | d=0.1 | b1=2000	1620	0.852	2.0	0.0	81.8	1.0
K=200 | L=100 | d=0.02 | b1=400	196	0.039	2.0	0.0	3.0	1.0
K=200 | L=100 | d=0.02 | b1=2000	1764	0.299	2.0	0.0	18.4	1.0
K=2000 | L=200 | d=0.1 | b1=4000	1801	0.163	2.0	0.0	92.2	1.0
K=2000 | L=200 | d=0.1 | b1=20000	16200	0.853	2.0	0.0	812.3	1.0
K=2000 | L=200 | d=0.02 | b1=4000	1961	0.038	2.0	0.0	22.1	1.0
K=2000 | L=200 | d=0.02 | b1=20000	17640	0.299	2.0	0.0	178.8	1.0
K=2000 | L=1000 | d=0.1 | b1=4000	1801	0.163	2.0	0.0	92.3	1.0
K=2000 | L=1000 | d=0.1 | b1=20000	16200	0.853	2.0	0.0	812.3	1.0
K=2000 | L=1000 | d=0.02 | b1=4000	1961	0.038	2.0	0.0	22.3	1.0
K=2000 | L=1000 | d=0.02 | b1=20000	17640	0.299	2.0	0.0	178.8	1.0
J=3 | n=50 | penScale =0.25						
K=200 | L=20 | d=0.1 | b1=400	9	0.013	0.9	1.2	2.4	3.3
K=200 | L=20 | d=0.1 | b1=2000	15	0.010	1.0	1.0	20.5	3.0
K=200 | L=20 | d=0.02 | b1=400	161	0.047	0.8	2.0	0.7	4.2
K=200 | L=20 | d=0.02 | b1=2000	1440	0.253	1.0	1.0	4.2	3.1
K=200 | L=100 | d=0.1 | b1=400	192	0.019	1.6	2.7	2.4	4.1

Table A: Simulation Results.

	sup	L_1	fn	fp	τ	df
K=200 | L=100 | d=0.1 | b1=2000	1800	0.033	2.0	2.0	20.5	3.0
K=200 | L=100 | d=0.02 | b1=400	24	0.020	0.8	2.9	0.7	5.1
K=200 | L=100 | d=0.02 | b1=2000	39	0.013	0.8	1.5	4.6	3.7
K=2000 | L=200 | d=0.1 | b1=4000	1974	0.005	1.7	2.1	23.1	3.4
K=2000 | L=200 | d=0.1 | b1=20000	18000	0.006	2.0	2.0	203.1	3.0
K=2000 | L=200 | d=0.02 | b1=4000	94	0.004	0.8	1.6	5.5	3.9
K=2000 | L=200 | d=0.02 | b1=20000	122	0.004	0.9	1.1	44.7	3.1
K=2000 | L=1000 | d=0.1 | b1=4000	202	0.004	0.7	2.0	23.1	4.2
K=2000 | L=1000 | d=0.1 | b1=20000	63	0.003	1.0	1.0	203.1	3.0
K=2000 | L=1000 | d=0.02 | b1=4000	213	0.004	0.8	2.6	5.6	4.8
K=2000 | L=1000 | d=0.02 | b1=20000	201	0.004	0.8	1.6	44.7	3.8
J=3 | n=50 | penScale =0.1						
K=200 | L=20 | d=0.1 | b1=400	10	0.015	0.8	1.6	1.0	3.8
K=200 | L=20 | d=0.1 | b1=2000	15	0.010	1.0	1.0	8.2	3.0
K=200 | L=20 | d=0.02 | b1=400	164	0.054	0.7	4.4	0.3	6.8
K=200 | L=20 | d=0.02 | b1=2000	1440	0.254	0.8	1.3	1.7	3.5
K=200 | L=100 | d=0.1 | b1=400	190	0.022	1.6	3.8	1.0	5.2
K=200 | L=100 | d=0.1 | b1=2000	1796	0.033	1.8	2.1	8.2	3.3
K=200 | L=100 | d=0.02 | b1=400	34	0.036	0.8	8.2	0.3	10.4
K=200 | L=100 | d=0.02 | b1=2000	39	0.015	0.8	2.1	1.8	4.3
K=2000 | L=200 | d=0.1 | b1=4000	1962	0.005	1.6	2.4	9.2	3.7
K=2000 | L=200 | d=0.1 | b1=20000	17999	0.006	2.0	2.0	81.2	3.0
K=2000 | L=200 | d=0.02 | b1=4000	98	0.005	0.8	1.8	2.2	4.0
K=2000 | L=200 | d=0.02 | b1=20000	130	0.004	0.8	1.3	17.9	3.5
K=2000 | L=1000 | d=0.1 | b1=4000	273	0.005	0.7	2.8	9.2	5.1
K=2000 | L=1000 | d=0.1 | b1=20000	141	0.003	0.9	1.2	81.2	3.4
K=2000 | L=1000 | d=0.02 | b1=4000	273	0.005	0.8	3.0	2.2	5.2
K=2000 | L=1000 | d=0.02 | b1=20000	209	0.004	0.8	2.1	17.9	4.3
J=3 | n=50 | penScale =0						
K=200 | L=20 | d=0.1 | b1=400	19	0.025	0.7	4.9	0.2	7.2
K=200 | L=20 | d=0.1 | b1=2000	27	0.018	0.7	4.9	0.2	7.3
K=200 | L=20 | d=0.02 | b1=400	164	0.055	0.7	4.9	0.2	7.3
K=200 | L=20 | d=0.02 | b1=2000	1444	0.264	0.7	4.8	0.2	7.1
K=200 | L=100 | d=0.1 | b1=400	192	0.041	1.6	10.8	0.3	12.2
K=200 | L=100 | d=0.1 | b1=2000	1791	0.051	1.6	11.4	0.3	12.8
K=200 | L=100 | d=0.02 | b1=400	35	0.038	0.8	8.9	0.3	11.2

Table A: Simulation Results.

	sup	L_1	fn	fp	τ	df
K=200 | L=100 | d=0.02 | b1=2000	61	0.035	0.8	9.2	0.3	11.5
K=2000 | L=200 | d=0.1 | b1=4000	1970	0.022	1.4	34.9	0.3	36.4
K=2000 | L=200 | d=0.1 | b1=20000	17975	0.016	1.5	33.9	0.3	35.4
K=2000 | L=200 | d=0.02 | b1=4000	263	0.022	0.7	33.4	0.3	35.7
K=2000 | L=200 | d=0.02 | b1=20000	335	0.020	0.7	33.4	0.3	35.7
K=2000 | L=1000 | d=0.1 | b1=4000	547	0.031	0.7	59.3	0.4	61.6
K=2000 | L=1000 | d=0.1 | b1=20000	1590	0.023	0.7	67.1	0.4	69.3
K=2000 | L=1000 | d=0.02 | b1=4000	470	0.032	0.8	57.2	0.4	59.4
K=2000 | L=1000 | d=0.02 | b1=20000	1042	0.029	0.8	58.1	0.4	60.3
J=10, 50 | n=50 | penScale =AIC						
K=200 | L=20 | d=0.1 | b1=400	21	0.034	1.0	8.1	0.9	10.1
K=200 | L=20 | d=0.1 | b1=2000	41	0.010	1.0	8.7	0.4	10.8
K=200 | L=20 | d=0.02 | b1=400	106	0.333	2.0	0.4	1.7	1.4
K=200 | L=20 | d=0.02 | b1=2000	1108	0.786	0.9	8.7	0.5	10.8
K=200 | L=100 | d=0.1 | b1=400	50	0.037	1.8	20.0	0.3	21.2
K=200 | L=100 | d=0.1 | b1=2000	144	0.021	1.8	23.7	0.3	25.0
K=200 | L=100 | d=0.02 | b1=400	129	0.271	1.8	10.9	1.8	12.2
K=200 | L=100 | d=0.02 | b1=2000	1836	0.049	1.0	53.5	0.3	55.5
K=2000 | L=200 | d=0.1 | b1=4000	326	0.019	1.7	40.7	0.3	42.0
K=2000 | L=200 | d=0.1 | b1=20000	735	0.008	1.8	38.9	0.3	40.1
K=2000 | L=200 | d=0.02 | b1=4000	2062	0.023	1.7	79.2	0.4	80.6
K=2000 | L=200 | d=0.02 | b1=20000	18136	0.013	1.6	76.4	0.3	77.7
K=2000 | L=1000 | d=0.1 | b1=4000	683	0.028	1.9	76.4	0.4	77.5
K=2000 | L=1000 | d=0.1 | b1=20000	2008	0.017	1.8	119.1	0.4	120.3
K=2000 | L=1000 | d=0.02 | b1=4000	744	0.033	1.9	145.7	0.4	146.8
K=2000 | L=1000 | d=0.02 | b1=20000	2005	0.019	1.8	176.3	0.4	177.6
J=10, 50 | n=50 | penScale =CV						
K=200 | L=20 | d=0.1 | b1=400	22	0.025	0.8	10.3	0.3	12.5
K=200 | L=20 | d=0.1 | b1=2000	44	0.011	0.9	9.4	0.5	11.5
K=200 | L=20 | d=0.02 | b1=400	137	0.321	0.9	9.4	5.6	11.5
K=200 | L=20 | d=0.02 | b1=2000	1112	0.785	0.8	9.5	0.5	11.7
K=200 | L=100 | d=0.1 | b1=400	50	0.035	1.9	18.8	0.5	19.9
K=200 | L=100 | d=0.1 | b1=2000	110	0.015	1.9	15.2	0.8	16.2
K=200 | L=100 | d=0.02 | b1=400	217	0.059	1.0	50.8	0.3	52.8
K=200 | L=100 | d=0.02 | b1=2000	1837	0.050	1.0	54.9	0.3	56.9
K=2000 | L=200 | d=0.1 | b1=4000	217	0.010	2.0	15.6	1.4	16.6

Table A: Simulation Results.

	sup	L_1	fn	fp	τ	df
K=2000 | L=200 | d=0.1 | b1=20000	298	0.003	2.0	11.5	2.9	12.5
K=2000 | L=200 | d=0.02 | b1=4000	2062	0.022	1.7	74.5	0.4	75.8
K=2000 | L=200 | d=0.02 | b1=20000	18112	0.011	1.9	58.5	0.9	59.6
K=2000 | L=1000 | d=0.1 | b1=4000	466	0.011	2.0	22.9	1.6	23.9
K=2000 | L=1000 | d=0.1 | b1=20000	956	0.004	2.0	16.4	3.4	17.4
K=2000 | L=1000 | d=0.02 | b1=4000	646	0.027	2.0	112.9	0.6	113.9
K=2000 | L=1000 | d=0.02 | b1=20000	1411	0.010	2.0	79.7	1.3	80.7
J=10, 50 | n=50 | penScale =1						
K=200 | L=20 | d=0.1 | b1=400	102	0.333	2.0	0.0	10.9	1.0
K=200 | L=20 | d=0.1 | b1=2000	903	0.818	2.0	0.0	91.6	1.0
K=200 | L=20 | d=0.02 | b1=400	102	0.333	2.0	0.0	2.9	1.0
K=200 | L=20 | d=0.02 | b1=2000	903	0.818	2.0	0.0	19.7	1.0
K=200 | L=100 | d=0.1 | b1=400	102	0.333	2.0	0.0	10.9	1.0
K=200 | L=100 | d=0.1 | b1=2000	903	0.818	2.0	0.0	91.6	1.0
K=200 | L=100 | d=0.02 | b1=400	102	0.333	2.0	0.0	3.2	1.0
K=200 | L=100 | d=0.02 | b1=2000	903	0.818	2.0	0.0	20.3	1.0
K=2000 | L=200 | d=0.1 | b1=4000	1006	0.333	2.0	0.0	102.7	1.0
K=2000 | L=200 | d=0.1 | b1=20000	9012	0.818	2.0	0.0	905.1	1.0
K=2000 | L=200 | d=0.02 | b1=4000	1006	0.333	2.0	0.0	23.9	1.0
K=2000 | L=200 | d=0.02 | b1=20000	9012	0.818	2.0	0.0	187.3	1.0
K=2000 | L=1000 | d=0.1 | b1=4000	1006	0.333	2.0	0.0	102.7	1.0
K=2000 | L=1000 | d=0.1 | b1=20000	9012	0.818	2.0	0.0	905.1	1.0
K=2000 | L=1000 | d=0.02 | b1=4000	1006	0.333	2.0	0.0	23.9	1.0
K=2000 | L=1000 | d=0.02 | b1=20000	9012	0.818	2.0	0.0	187.3	1.0
J=10, 50 | n=50 | penScale =0.25						
K=200 | L=20 | d=0.1 | b1=400	14	0.019	1.0	8.0	2.7	10.0
K=200 | L=20 | d=0.1 | b1=2000	31	0.009	1.0	8.0	22.9	10.0
K=200 | L=20 | d=0.02 | b1=400	133	0.323	1.2	6.5	0.7	8.3
K=200 | L=20 | d=0.02 | b1=2000	1103	0.785	1.0	8.0	4.9	10.0
K=200 | L=100 | d=0.1 | b1=400	22	0.020	2.0	9.2	2.7	10.2
K=200 | L=100 | d=0.1 | b1=2000	31	0.009	2.0	9.0	22.9	10.0
K=200 | L=100 | d=0.02 | b1=400	208	0.112	1.2	40.4	0.8	42.2
K=200 | L=100 | d=0.02 | b1=2000	1829	0.045	1.0	48.0	5.1	50.0
K=2000 | L=200 | d=0.1 | b1=4000	49	0.006	2.0	9.0	25.7	10.0
K=2000 | L=200 | d=0.1 | b1=20000	100	0.003	2.0	9.0	226.3	10.0
K=2000 | L=200 | d=0.02 | b1=4000	2042	0.015	2.0	49.0	6.0	50.0

Table A: Simulation Results.

	sup	L_1	fn	fp	τ	df
K=2000 | L=200 | d=0.02 | b1=20000	18106	0.009	2.0	49.0	46.8	50.0
K=2000 | L=1000 | d=0.1 | b1=4000	219	0.007	2.0	9.4	25.7	10.4
K=2000 | L=1000 | d=0.1 | b1=20000	100	0.003	2.0	9.0	226.3	10.0
K=2000 | L=1000 | d=0.02 | b1=4000	397	0.014	2.0	49.6	6.0	50.6
K=2000 | L=1000 | d=0.02 | b1=20000	323	0.007	2.0	49.0	46.8	50.0
J=10, 50 | n=50 | penScale =0.1						
K=200 | L=20 | d=0.1 | b1=400	15	0.020	1.0	8.1	1.1	10.1
K=200 | L=20 | d=0.1 | b1=2000	31	0.009	1.0	8.0	9.2	10.0
K=200 | L=20 | d=0.02 | b1=400	139	0.320	0.8	10.2	0.3	12.4
K=200 | L=20 | d=0.02 | b1=2000	1103	0.785	1.0	8.0	2.0	10.0
K=200 | L=100 | d=0.1 | b1=400	45	0.025	2.0	12.0	1.1	13.0
K=200 | L=100 | d=0.1 | b1=2000	32	0.009	2.0	9.0	9.2	10.0
K=200 | L=100 | d=0.02 | b1=400	217	0.058	1.0	50.2	0.3	52.2
K=200 | L=100 | d=0.02 | b1=2000	1829	0.045	1.0	48.0	2.0	50.0
K=2000 | L=200 | d=0.1 | b1=4000	110	0.007	2.0	9.3	10.3	10.3
K=2000 | L=200 | d=0.1 | b1=20000	100	0.003	2.0	9.0	90.5	10.0
K=2000 | L=200 | d=0.02 | b1=4000	2042	0.015	2.0	49.3	2.4	50.3
K=2000 | L=200 | d=0.02 | b1=20000	18106	0.009	2.0	49.0	18.7	50.0
K=2000 | L=1000 | d=0.1 | b1=4000	457	0.007	2.0	12.5	10.3	13.5
K=2000 | L=1000 | d=0.1 | b1=20000	132	0.003	2.0	9.0	90.5	10.0
K=2000 | L=1000 | d=0.02 | b1=4000	637	0.017	2.0	61.7	2.4	62.7
K=2000 | L=1000 | d=0.02 | b1=20000	358	0.007	2.0	49.0	18.7	50.0
J=10, 50 | n=50 | penScale =0						
K=200 | L=20 | d=0.1 | b1=400	23	0.026	0.7	10.8	0.2	13.1
K=200 | L=20 | d=0.1 | b1=2000	47	0.012	0.7	10.2	0.2	12.5
K=200 | L=20 | d=0.02 | b1=400	139	0.320	0.8	10.6	0.2	12.9
K=200 | L=20 | d=0.02 | b1=2000	1114	0.785	0.7	10.3	0.2	12.6
K=200 | L=100 | d=0.1 | b1=400	54	0.039	1.8	22.3	0.3	23.5
K=200 | L=100 | d=0.1 | b1=2000	146	0.021	1.7	25.2	0.3	26.5
K=200 | L=100 | d=0.02 | b1=400	217	0.059	1.0	50.8	0.3	52.8
K=200 | L=100 | d=0.02 | b1=2000	1837	0.050	1.0	54.9	0.3	56.9
K=2000 | L=200 | d=0.1 | b1=4000	331	0.020	1.6	44.7	0.3	46.1
K=2000 | L=200 | d=0.1 | b1=20000	736	0.008	1.8	39.5	0.3	40.7
K=2000 | L=200 | d=0.02 | b1=4000	2064	0.024	1.6	83.5	0.3	84.9
K=2000 | L=200 | d=0.02 | b1=20000	18137	0.013	1.6	77.3	0.3	78.7
K=2000 | L=1000 | d=0.1 | b1=4000	729	0.029	1.8	83.9	0.4	85.1

Table A: Simulation Results.

	sup	L_1	fn	fp	τ	df
K=2000 | L=1000 | d=0.1 | b1=20000	2010	0.017	1.8	120.3	0.4	121.5
K=2000 | L=1000 | d=0.02 | b1=4000	763	0.033	1.9	149.9	0.4	151.1
K=2000 | L=1000 | d=0.02 | b1=20000	2005	0.019	1.8	177.3	0.4	178.6
J=3 | n=500 | penScale =AIC						
K=200 | L=20 | d=0.1 | b1=400	3	0.003	0.9	1.3	0.1	3.4
K=200 | L=20 | d=0.1 | b1=2000	4	0.003	0.9	1.5	0.1	3.6
K=200 | L=20 | d=0.02 | b1=400	160	0.034	0.9	1.3	0.1	3.4
K=200 | L=20 | d=0.02 | b1=2000	1440	0.246	0.9	1.3	0.1	3.4
K=200 | L=100 | d=0.1 | b1=400	198	0.009	1.7	3.5	0.1	4.8
K=200 | L=100 | d=0.1 | b1=2000	1798	0.027	1.7	3.4	0.1	4.7
K=200 | L=100 | d=0.02 | b1=400	8	0.005	0.9	2.3	0.1	4.4
K=200 | L=100 | d=0.02 | b1=2000	11	0.004	0.9	2.4	0.1	4.5
K=2000 | L=200 | d=0.1 | b1=4000	1997	0.002	1.9	4.5	0.1	5.6
K=2000 | L=200 | d=0.1 | b1=20000	17994	0.004	1.9	4.4	0.1	5.6
K=2000 | L=200 | d=0.02 | b1=4000	49	0.002	0.9	3.7	0.1	5.7
K=2000 | L=200 | d=0.02 | b1=20000	50	0.002	0.9	3.6	0.1	5.7
K=2000 | L=1000 | d=0.1 | b1=4000	109	0.003	0.9	7.7	0.1	9.9
K=2000 | L=1000 | d=0.1 | b1=20000	165	0.002	0.9	7.3	0.1	9.5
K=2000 | L=1000 | d=0.02 | b1=4000	103	0.003	0.9	7.5	0.1	9.6
K=2000 | L=1000 | d=0.02 | b1=20000	128	0.002	0.9	7.3	0.1	9.5
J=3 | n=500 | penScale =CV						
K=200 | L=20 | d=0.1 | b1=400	3	0.003	1.0	1.3	0.2	3.3
K=200 | L=20 | d=0.1 | b1=2000	4	0.003	1.0	1.2	0.3	3.3
K=200 | L=20 | d=0.02 | b1=400	160	0.034	1.0	1.2	0.2	3.3
K=200 | L=20 | d=0.02 | b1=2000	1440	0.246	0.9	1.2	0.2	3.3
K=200 | L=100 | d=0.1 | b1=400	198	0.008	1.7	2.6	0.2	3.9
K=200 | L=100 | d=0.1 | b1=2000	1798	0.027	1.7	2.5	0.4	3.7
K=200 | L=100 | d=0.02 | b1=400	7	0.004	0.9	1.9	0.1	4.0
K=200 | L=100 | d=0.02 | b1=2000	11	0.004	0.9	1.9	0.2	4.0
K=2000 | L=200 | d=0.1 | b1=4000	1998	0.002	1.9	2.1	0.7	3.1
K=2000 | L=200 | d=0.1 | b1=20000	17996	0.003	1.9	2.1	1.4	3.1
K=2000 | L=200 | d=0.02 | b1=4000	22	0.001	1.0	1.1	0.4	3.1
K=2000 | L=200 | d=0.02 | b1=20000	29	0.001	1.0	1.1	0.6	3.1
K=2000 | L=1000 | d=0.1 | b1=4000	60	0.001	0.9	1.5	0.8	3.7
K=2000 | L=1000 | d=0.1 | b1=20000	54	0.001	0.9	1.3	1.5	3.5
K=2000 | L=1000 | d=0.02 | b1=4000	59	0.001	0.9	1.5	0.5	3.6

Table A: Simulation Results.

	sup	L_1	fn	fp	τ	df
K=2000 | L=1000 | d=0.02 | b1=20000	67	0.001	0.9	1.5	0.7	3.6
J=3 | n=500 | penScale =1						
K=200 | L=20 | d=0.1 | b1=400	180	0.164	2.0	0.0	9.2	1.0
K=200 | L=20 | d=0.1 | b1=2000	1620	0.853	2.0	0.0	81.3	1.0
K=200 | L=20 | d=0.02 | b1=400	196	0.038	2.0	0.0	2.0	1.0
K=200 | L=20 | d=0.02 | b1=2000	1764	0.299	2.0	0.0	16.4	1.0
K=200 | L=100 | d=0.1 | b1=400	180	0.164	2.0	0.0	9.2	1.0
K=200 | L=100 | d=0.1 | b1=2000	1620	0.853	2.0	0.0	81.3	1.0
K=200 | L=100 | d=0.02 | b1=400	196	0.038	2.0	0.0	2.1	1.0
K=200 | L=100 | d=0.02 | b1=2000	1764	0.299	2.0	0.0	17.8	1.0
K=2000 | L=200 | d=0.1 | b1=4000	1800	0.164	2.0	0.0	90.7	1.0
K=2000 | L=200 | d=0.1 | b1=20000	16200	0.853	2.0	0.0	810.7	1.0
K=2000 | L=200 | d=0.02 | b1=4000	1960	0.038	2.0	0.0	20.4	1.0
K=2000 | L=200 | d=0.02 | b1=20000	17640	0.299	2.0	0.0	177.3	1.0
K=2000 | L=1000 | d=0.1 | b1=4000	1800	0.164	2.0	0.0	90.7	1.0
K=2000 | L=1000 | d=0.1 | b1=20000	16200	0.853	2.0	0.0	810.7	1.0
K=2000 | L=1000 | d=0.02 | b1=4000	1960	0.038	2.0	0.0	20.4	1.0
K=2000 | L=1000 | d=0.02 | b1=20000	17640	0.299	2.0	0.0	177.3	1.0
J=3 | n=500 | penScale =0.25						
K=200 | L=20 | d=0.1 | b1=400	2	0.003	1.0	1.0	2.3	3.0
K=200 | L=20 | d=0.1 | b1=2000	3	0.003	1.0	1.0	20.3	3.0
K=200 | L=20 | d=0.02 | b1=400	160	0.034	1.0	1.0	0.5	3.0
K=200 | L=20 | d=0.02 | b1=2000	1440	0.246	1.0	1.0	4.1	3.0
K=200 | L=100 | d=0.1 | b1=400	199	0.007	1.9	2.0	2.3	3.1
K=200 | L=100 | d=0.1 | b1=2000	1800	0.026	2.0	2.0	20.3	3.0
K=200 | L=100 | d=0.02 | b1=400	6	0.003	0.9	1.3	0.5	3.4
K=200 | L=100 | d=0.02 | b1=2000	9	0.003	1.0	1.0	4.5	3.0
K=2000 | L=200 | d=0.1 | b1=4000	2000	0.002	2.0	2.0	22.7	3.0
K=2000 | L=200 | d=0.1 | b1=20000	18000	0.003	2.0	2.0	202.7	3.0
K=2000 | L=200 | d=0.02 | b1=4000	20	0.001	1.0	1.0	5.1	3.1
K=2000 | L=200 | d=0.02 | b1=20000	26	0.001	1.0	1.0	44.3	3.0
K=2000 | L=1000 | d=0.1 | b1=4000	29	0.001	0.9	1.1	22.7	3.2
K=2000 | L=1000 | d=0.1 | b1=20000	17	0.001	1.0	1.0	202.7	3.0
K=2000 | L=1000 | d=0.02 | b1=4000	42	0.001	0.9	1.3	5.1	3.4
K=2000 | L=1000 | d=0.02 | b1=20000	29	0.001	1.0	1.0	44.3	3.0
J=3 | n=500 | penScale =0.1						

Table A: Simulation Results.

	sup	L_1	fn	fp	τ	df
K=200 | L=20 | d=0.1 | b1=400	2	0.003	1.0	1.0	0.9	3.0
K=200 | L=20 | d=0.1 | b1=2000	3	0.003	1.0	1.0	8.1	3.0
K=200 | L=20 | d=0.02 | b1=400	160	0.034	1.0	1.1	0.2	3.1
K=200 | L=20 | d=0.02 | b1=2000	1440	0.246	1.0	1.1	1.6	3.1
K=200 | L=100 | d=0.1 | b1=400	198	0.007	1.7	2.1	0.9	3.4
K=200 | L=100 | d=0.1 | b1=2000	1800	0.026	2.0	2.0	8.1	3.0
K=200 | L=100 | d=0.02 | b1=400	6	0.003	0.9	1.3	0.2	3.4
K=200 | L=100 | d=0.02 | b1=2000	10	0.003	0.9	1.1	1.8	3.2
K=2000 | L=200 | d=0.1 | b1=4000	1999	0.002	2.0	2.0	9.1	3.0
K=2000 | L=200 | d=0.1 | b1=20000	18000	0.003	2.0	2.0	81.1	3.0
K=2000 | L=200 | d=0.02 | b1=4000	20	0.001	1.0	1.0	2.0	3.1
K=2000 | L=200 | d=0.02 | b1=20000	26	0.001	1.0	1.0	17.7	3.0
K=2000 | L=1000 | d=0.1 | b1=4000	38	0.001	0.9	1.2	9.1	3.3
K=2000 | L=1000 | d=0.1 | b1=20000	17	0.001	1.0	1.0	81.1	3.0
K=2000 | L=1000 | d=0.02 | b1=4000	42	0.001	0.9	1.3	2.0	3.4
K=2000 | L=1000 | d=0.02 | b1=20000	44	0.001	0.9	1.2	17.7	3.2
J=3 | n=500 | penScale =0						
K=200 | L=20 | d=0.1 | b1=400	3	0.003	0.9	1.5	0.1	3.5
K=200 | L=20 | d=0.1 | b1=2000	4	0.003	0.9	1.6	0.1	3.7
K=200 | L=20 | d=0.02 | b1=400	160	0.034	0.9	1.5	0.1	3.5
K=200 | L=20 | d=0.02 | b1=2000	1440	0.247	0.9	1.5	0.1	3.6
K=200 | L=100 | d=0.1 | b1=400	198	0.009	1.7	3.8	0.1	5.1
K=200 | L=100 | d=0.1 | b1=2000	1798	0.028	1.7	3.6	0.1	4.9
K=200 | L=100 | d=0.02 | b1=400	8	0.005	0.9	2.5	0.1	4.6
K=200 | L=100 | d=0.02 | b1=2000	12	0.004	0.9	2.6	0.1	4.7
K=2000 | L=200 | d=0.1 | b1=4000	1997	0.002	1.9	4.6	0.1	5.7
K=2000 | L=200 | d=0.1 | b1=20000	17994	0.004	1.9	4.5	0.1	5.7
K=2000 | L=200 | d=0.02 | b1=4000	50	0.002	0.9	3.8	0.1	5.9
K=2000 | L=200 | d=0.02 | b1=20000	50	0.002	0.9	3.7	0.1	5.8
K=2000 | L=1000 | d=0.1 | b1=4000	110	0.003	0.9	7.9	0.1	10.1
K=2000 | L=1000 | d=0.1 | b1=20000	166	0.002	0.9	7.4	0.1	9.6
K=2000 | L=1000 | d=0.02 | b1=4000	103	0.003	0.9	7.7	0.1	9.8
K=2000 | L=1000 | d=0.02 | b1=20000	128	0.002	0.9	7.4	0.1	9.6
J=10, 50 | n=500 | penScale =AIC						
K=200 | L=20 | d=0.1 | b1=400	4	0.005	0.9	8.2	0.1	10.3
K=200 | L=20 | d=0.1 | b1=2000	11	0.003	0.9	8.3	0.1	10.4

Table A: Simulation Results.

	sup	L_1	fn	fp	τ	df
K=200 | L=20 | d=0.02 | b1=400	124	0.320	1.0	8.3	0.1	10.3
K=200 | L=20 | d=0.02 | b1=2000	1087	0.785	1.0	8.0	0.1	10.0
K=200 | L=100 | d=0.1 | b1=400	15	0.008	2.0	13.4	0.1	14.4
K=200 | L=100 | d=0.1 | b1=2000	23	0.004	2.0	12.2	0.1	13.2
K=200 | L=100 | d=0.02 | b1=400	205	0.024	1.0	48.8	0.1	50.8
K=200 | L=100 | d=0.02 | b1=2000	1814	0.031	1.0	48.9	0.1	50.9
K=2000 | L=200 | d=0.1 | b1=4000	65	0.003	2.0	12.1	0.1	13.1
K=2000 | L=200 | d=0.1 | b1=20000	55	0.001	2.0	11.0	0.1	12.1
K=2000 | L=200 | d=0.02 | b1=4000	2015	0.006	2.0	54.6	0.1	55.7
K=2000 | L=200 | d=0.02 | b1=20000	18037	0.005	2.0	52.7	0.1	53.7
K=2000 | L=1000 | d=0.1 | b1=4000	164	0.003	2.0	19.6	0.1	20.6
K=2000 | L=1000 | d=0.1 | b1=20000	340	0.001	2.0	17.9	0.1	18.9
K=2000 | L=1000 | d=0.02 | b1=4000	206	0.006	2.0	76.1	0.1	77.1
K=2000 | L=1000 | d=0.02 | b1=20000	485	0.003	2.0	69.2	0.1	70.2
J=10, 50 | n=500 | penScale =CV						
K=200 | L=20 | d=0.1 | b1=400	4	0.005	0.9	8.3	0.1	10.4
K=200 | L=20 | d=0.1 | b1=2000	11	0.003	0.9	8.2	0.2	10.2
K=200 | L=20 | d=0.02 | b1=400	124	0.320	0.9	8.4	0.1	10.5
K=200 | L=20 | d=0.02 | b1=2000	1087	0.785	1.0	8.0	0.2	10.0
K=200 | L=100 | d=0.1 | b1=400	15	0.007	2.0	12.5	0.1	13.6
K=200 | L=100 | d=0.1 | b1=2000	18	0.003	2.0	10.2	0.3	11.2
K=200 | L=100 | d=0.02 | b1=400	205	0.024	1.0	49.1	0.1	51.1
K=200 | L=100 | d=0.02 | b1=2000	1814	0.031	1.0	49.1	0.1	51.1
K=2000 | L=200 | d=0.1 | b1=4000	37	0.002	2.0	9.5	0.5	10.5
K=2000 | L=200 | d=0.1 | b1=20000	36	0.001	2.0	9.2	0.9	10.2
K=2000 | L=200 | d=0.02 | b1=4000	2014	0.006	2.0	52.7	0.1	53.7
K=2000 | L=200 | d=0.02 | b1=20000	18037	0.005	2.0	49.7	0.3	50.7
K=2000 | L=1000 | d=0.1 | b1=4000	140	0.002	2.0	12.3	0.5	13.3
K=2000 | L=1000 | d=0.1 | b1=20000	130	0.001	2.0	10.7	1.1	11.7
K=2000 | L=1000 | d=0.02 | b1=4000	199	0.006	2.0	67.7	0.2	68.7
K=2000 | L=1000 | d=0.02 | b1=20000	298	0.002	2.0	55.5	0.4	56.5
J=10, 50 | n=500 | penScale =1						
K=200 | L=20 | d=0.1 | b1=400	100	0.333	2.0	0.0	10.2	1.0
K=200 | L=20 | d=0.1 | b1=2000	901	0.818	2.0	0.0	90.5	1.0
K=200 | L=20 | d=0.02 | b1=400	100	0.333	2.0	0.0	2.2	1.0
K=200 | L=20 | d=0.02 | b1=2000	901	0.818	2.0	0.0	18.5	1.0

Table A: Simulation Results.

	sup	L_1	fn	fp	τ	df
K=200 | L=100 | d=0.1 | b1=400	100	0.333	2.0	0.0	10.2	1.0
K=200 | L=100 | d=0.1 | b1=2000	901	0.818	2.0	0.0	90.5	1.0
K=200 | L=100 | d=0.02 | b1=400	100	0.333	2.0	0.0	2.3	1.0
K=200 | L=100 | d=0.02 | b1=2000	901	0.818	2.0	0.0	18.7	1.0
K=2000 | L=200 | d=0.1 | b1=4000	1002	0.333	2.0	0.0	100.8	1.0
K=2000 | L=200 | d=0.1 | b1=20000	9003	0.818	2.0	0.0	901.6	1.0
K=2000 | L=200 | d=0.02 | b1=4000	1002	0.333	2.0	0.0	21.2	1.0
K=2000 | L=200 | d=0.02 | b1=20000	9003	0.818	2.0	0.0	182.3	1.0
K=2000 | L=1000 | d=0.1 | b1=4000	1002	0.333	2.0	0.0	100.8	1.0
K=2000 | L=1000 | d=0.1 | b1=20000	9003	0.818	2.0	0.0	901.6	1.0
K=2000 | L=1000 | d=0.02 | b1=4000	1002	0.333	2.0	0.0	21.2	1.0
K=2000 | L=1000 | d=0.02 | b1=20000	9003	0.818	2.0	0.0	182.3	1.0
J=10, 50 | n=500 | penScale =0.25						
K=200 | L=20 | d=0.1 | b1=400	3	0.005	1.0	8.0	2.6	10.0
K=200 | L=20 | d=0.1 | b1=2000	11	0.003	1.0	8.0	22.6	10.0
K=200 | L=20 | d=0.02 | b1=400	123	0.320	1.0	8.0	0.6	10.0
K=200 | L=20 | d=0.02 | b1=2000	1087	0.785	1.0	8.0	4.6	10.0
K=200 | L=100 | d=0.1 | b1=400	3	0.005	2.0	9.0	2.6	10.0
K=200 | L=100 | d=0.1 | b1=2000	11	0.003	2.0	9.0	22.6	10.0
K=200 | L=100 | d=0.02 | b1=400	205	0.023	1.0	48.0	0.6	50.0
K=200 | L=100 | d=0.02 | b1=2000	1813	0.031	1.0	48.0	4.7	50.0
K=2000 | L=200 | d=0.1 | b1=4000	16	0.002	2.0	9.0	25.2	10.0
K=2000 | L=200 | d=0.1 | b1=20000	29	0.001	2.0	9.0	225.4	10.0
K=2000 | L=200 | d=0.02 | b1=4000	2013	0.005	2.0	49.0	5.3	50.0
K=2000 | L=200 | d=0.02 | b1=20000	18037	0.005	2.0	49.0	45.6	50.0
K=2000 | L=1000 | d=0.1 | b1=4000	16	0.002	2.0	9.0	25.2	10.0
K=2000 | L=1000 | d=0.1 | b1=20000	29	0.001	2.0	9.0	225.4	10.0
K=2000 | L=1000 | d=0.02 | b1=4000	50	0.004	2.0	49.0	5.3	50.0
K=2000 | L=1000 | d=0.02 | b1=20000	107	0.002	2.0	49.0	45.6	50.0
J=10, 50 | n=500 | penScale =0.1						
K=200 | L=20 | d=0.1 | b1=400	3	0.005	1.0	8.0	1.0	10.0
K=200 | L=20 | d=0.1 | b1=2000	11	0.003	1.0	8.0	9.1	10.0
K=200 | L=20 | d=0.02 | b1=400	123	0.320	1.0	8.0	0.2	10.0
K=200 | L=20 | d=0.02 | b1=2000	1087	0.785	1.0	8.0	1.9	10.0
K=200 | L=100 | d=0.1 | b1=400	4	0.005	2.0	9.0	1.0	10.0
K=200 | L=100 | d=0.1 | b1=2000	11	0.003	2.0	9.0	9.1	10.0

Table A: Simulation Results.

	sup	L_1	fn	fp	τ	df
K=200 | L=100 | d=0.02 | b1=400	205	0.023	1.0	48.0	0.2	50.0
K=200 | L=100 | d=0.02 | b1=2000	1813	0.031	1.0	48.0	1.9	50.0
K=2000 | L=200 | d=0.1 | b1=4000	16	0.002	2.0	9.0	10.1	10.0
K=2000 | L=200 | d=0.1 | b1=20000	29	0.001	2.0	9.0	90.2	10.0
K=2000 | L=200 | d=0.02 | b1=4000	2013	0.005	2.0	49.0	2.1	50.0
K=2000 | L=200 | d=0.02 | b1=20000	18037	0.005	2.0	49.0	18.2	50.0
K=2000 | L=1000 | d=0.1 | b1=4000	23	0.002	2.0	9.0	10.1	10.0
K=2000 | L=1000 | d=0.1 | b1=20000	29	0.001	2.0	9.0	90.2	10.0
K=2000 | L=1000 | d=0.02 | b1=4000	56	0.004	2.0	49.0	2.1	50.0
K=2000 | L=1000 | d=0.02 | b1=20000	107	0.002	2.0	49.0	18.2	50.0
J=10, 50 | n=500 | penScale =0						
K=200 | L=20 | d=0.1 | b1=400	5	0.006	0.9	8.5	0.1	10.5
K=200 | L=20 | d=0.1 | b1=2000	11	0.003	0.9	8.4	0.1	10.5
K=200 | L=20 | d=0.02 | b1=400	124	0.320	0.9	8.5	0.1	10.6
K=200 | L=20 | d=0.02 | b1=2000	1087	0.785	1.0	8.0	0.1	10.0
K=200 | L=100 | d=0.1 | b1=400	15	0.008	2.0	13.6	0.1	14.7
K=200 | L=100 | d=0.1 | b1=2000	23	0.004	2.0	12.3	0.1	13.3
K=200 | L=100 | d=0.02 | b1=400	205	0.024	1.0	49.1	0.1	51.1
K=200 | L=100 | d=0.02 | b1=2000	1814	0.031	1.0	49.1	0.1	51.1
K=2000 | L=200 | d=0.1 | b1=4000	65	0.003	2.0	12.2	0.1	13.2
K=2000 | L=200 | d=0.1 | b1=20000	55	0.001	2.0	11.1	0.1	12.1
K=2000 | L=200 | d=0.02 | b1=4000	2015	0.006	2.0	54.7	0.1	55.8
K=2000 | L=200 | d=0.02 | b1=20000	18037	0.005	2.0	52.7	0.1	53.8
K=2000 | L=1000 | d=0.1 | b1=4000	165	0.003	2.0	19.8	0.1	20.8
K=2000 | L=1000 | d=0.1 | b1=20000	340	0.001	2.0	17.9	0.1	18.9
K=2000 | L=1000 | d=0.02 | b1=4000	207	0.006	2.0	76.2	0.1	77.2
K=2000 | L=1000 | d=0.02 | b1=20000	485	0.003	2.0	69.3	0.1	70.3