Journal of Econometrics 160 (2011) 204–219

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Subsampling realised kernels[☆]

Ole E. Barndorff-Nielsen[a,b], Peter Reinhard Hansen[c], Asger Lunde[d,b,*], Neil Shephard[e,f]

[a] The T.N. Thiele Centre for Mathematics in Natural Science, Department of Mathematical Sciences, Aarhus University, Ny Munkegade, DK-8000 Aarhus C, Denmark
[b] CREATES, Aarhus University, Denmark
[c] Department of Economics, Stanford University, Landau Economics Building, 579 Serra Mall, Stanford, CA 94305-6072, USA
[d] School of Economics and Management, Aarhus University, Building 1322 Bartholins Allé 10, DK-8000 Aarhus C, Denmark
[e] Oxford-Man Institute, University of Oxford, Eagle House, Walton Well Road, Oxford OX2 6EE, UK
[f] Department of Economics, University of Oxford, UK

ARTICLE INFO

Article history:
Available online 6 March 2010

Keywords:
Long run variance estimator
Market frictions
Quadratic variation
Realised kernel
Realised variance
Subsampling

ABSTRACT

In a recent paper we have introduced the class of realised kernel estimators of the increments of quadratic variation in the presence of noise. We showed that this estimator is consistent and derived its limit distribution under various assumptions on the kernel weights. In this paper we extend our analysis, looking at the class of subsampled realised kernels and we derive the limit theory for this class of estimators. We find that subsampling is highly advantageous for estimators based on discontinuous kernels, such as the truncated kernel. For *kinked kernels*, such as the Bartlett kernel, we show that subsampling is impotent, in the sense that subsampling has no effect on the asymptotic distribution. Perhaps surprisingly, for the efficient *smooth kernels*, such as the Parzen kernel, we show that subsampling is harmful as it increases the asymptotic variance. We also study the performance of subsampled realised kernels in simulations and in empirical work.

1. Introduction

High frequency financial data allow us to estimate the increments to quadratic variation, the usual ex post measure of the variation of asset prices (e.g. Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002)). Common estimators, such as the realised variance, can be sensitive to market frictions when applied to returns recorded over shorter time intervals such as 1 min, or even more ambitiously, 1 s (e.g. Zhou (1996), Fang (1996) and Andersen et al. (2000)). In response, two non-parametric generalisations have been proposed: *subsampling* and *realised kernels* by Zhang et al. (2005) and Barndorff-Nielsen et al. (2008), respectively. Here we partially unify these approaches by studying the properties of *subsampled realised kernels*.

Our interest is the estimation of the increment to quadratic variation over some arbitrary fixed time period written as $[0, t]$, which could represent a day say, using estimators of the realised kernel type. For a continuous time log-price process X and time gap $\delta > 0$, the flat-top[1] *realised kernels* of Barndorff-Nielsen et al. (2008) take on the following form:

$$K(X_\delta) = \gamma_0(X_\delta) + \sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\{\gamma_h(X_\delta) + \gamma_{-h}(X_\delta)\}, \quad H \geq 1.$$

Here $k(x)$, $x \in [0, 1]$, is a weight function with $k(0) = 1$, $k(1) = 0$, while

$$\gamma_h(X_\delta) = \sum_{j=1}^{n_\delta} x_j x_{j-h},$$

$$x_j = X_{\delta j} - X_{\delta(j-1)}, \; h = -H, \ldots, -1, 0, 1, \ldots, H,$$

with $n_\delta = \lfloor t/\delta \rfloor$. Think of δ as being small and so x_j represents the j-th high frequency return, while $\gamma_0(X_\delta)$ is the realised variance of X. The above authors gave a relatively exhaustive treatment of $K(X_\delta)$ when X is a Brownian semimartingale plus noise.

It is important to distinguish three classes of kernel functions $k(x)$: *smooth*, *kinked*, and *discontinuous*. Examples are the Parzen,

☆ The second and fourth author are also affiliated with CREATES, a research centre funded by the Danish National Research Foundation. The Ox language of Doornik (2006) was used to perform the calculations reported here. We are grateful to Joseph Romano, Per Mykland and Nour Meddahi (the editors), and two anonymous referees for their suggestions that improved this manuscript. All errors remain our responsibility.

* Corresponding author at: School of Economics and Management, Aarhus University, Building 1322 Bartholins Allé 10, DK-8000 Aarhus C, Denmark.
E-mail addresses: oebn@imf.au.dk (O.E. Barndorff-Nielsen), peter.hansen@stanford.edu (P.R. Hansen), alunde@econ.au.dk (A. Lunde), neil.shephard@economics.ox.ac.uk (N. Shephard).

[1] It is called a flat-top estimator as it imposes that the weight at lag 1 is 1. The motivation for this is discussed extensively in Barndorff-Nielsen et al. (2008).

0304-4076/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2010.03.031

the Bartlett and the truncated kernel, respectively. Barndorff-Nielsen et al. (2008) have shown that the smooth class, which satisfy $k'(0) = k'(1) = 0$, lead to realised kernels that converge at the efficient rate $n_\delta^{1/4}$, whereas the kinked kernels, which do not satisfy $k'(0) = k'(1) = 0$, lead to realised kernels that converge at rate $n_\delta^{1/6}$. The discontinuous kernels lead to inconsistent estimators as we show in Section 3.4.

Realised kernels use returns computed starting at $t = 0$. There may be efficiency gains by jittering the initial value S times, illustrated in Fig. 1, producing S sets of high frequency returns x_j^s, $s = 1, 2, \ldots, S$. Zhang et al. (2005) made this point for realised variances. We can then average the resulting S realised kernel estimators:

$$K(X_\delta; S) = \frac{1}{S}\sum_{s=1}^{S} K^s(X_\delta),$$

where

$$K^s(X_\delta) = \gamma_0^s(X_\delta) + \sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)\left\{\gamma_h^s(X_\delta) + \gamma_{-h}^s(X_\delta)\right\},$$

$$\gamma_h^s(X_\delta) = \sum_{j=1}^{n_\delta} x_j^s x_{j-h}^s, \quad x_j^s = X_{\delta\left(j+\frac{(s-1)}{S}\right)} - X_{\delta\left(j+\frac{(s-1)}{S}-1\right)}.$$

We call $K(X_\delta; S)$ the *subsampled realised kernel*—noting that this form of subsampling is different from the conventional form of subsampling, as we discuss below.

Here we show that subsampling is very useful for the class of discontinuous kernels, because subsampling makes these estimators consistent and converge in distribution at rate $n^{1/6}$, where $n = Sn_\delta$ is the *effective sample size*. Zhou (1996) used a simple discontinuous kernel and gave a brief discussion of subsampling of that kernel. We will see that his estimator can be made consistent by allowing $S \to \infty$ as $n \to \infty$, a result which is implicit in his paper, but one that he did not explicitly draw out. For the class of kinked kernels, we show that subsampling is impotent, in the sense that the asymptotic distribution is the same whether subsampling is used or not. Finally, we show that subsampling is harmful when applied to smooth kernels. In fact, if the number of subsamples, S, increases with the sample size, n, the best rate of convergence is reduced to less than the efficient one, $n^{1/4}$.

The intuition for these results follows from Lemma A.1 in the Appendix. It shows that

$$\gamma_h(X_\delta; S) = \frac{1}{S}\sum_{s=1}^{S}\gamma_h^s(X_\delta) \simeq \sum_{s=-S+1}^{S-1} k_B\left(\frac{s}{S}\right)\gamma_{Sh+s}(X_{\delta/S}),$$
$$\text{where } k_B(x) = 1 - |x|,$$

where the approximation is due to subtle end effects. The implication is that

$$K(X_\delta; S) \simeq \sum_{s=-S+1}^{S-1} k_B\left(\frac{s}{S}\right)\gamma_s(X_{\delta/S}) + \sum_{h=1}^{H} k\left(\frac{h-1}{H}\right)$$
$$\times \sum_{s=-S}^{S} k_B\left(\frac{s}{S}\right)\left\{\gamma_{Sh+s}(X_{\delta/S}) + \gamma_{-Sh-s}(X_{\delta/S})\right\}$$
$$= \sum_{h=0}^{HS} k_S\left(\frac{h-1}{HS}\right)\tilde{\gamma}_{Sh+s}(X_{\delta/S}).$$

So a subsampled realised kernel is a realised kernel simply operating on a higher frequency (ignoring end effects). The implied kernel weights, $k_S\left(\frac{h}{HS}\right)$, $h = 1, \ldots, SH$, are convex combinations of neighboring weights of the original kernel,

$$k_S\left(\frac{hs}{HS}\right) = \frac{S-s}{S}k\left(\frac{h}{S}\right) + \frac{s}{S}k\left(\frac{h+1}{S}\right),$$
$$h = 0, \ldots, H,\ s = 1, \ldots, S. \quad (1)$$

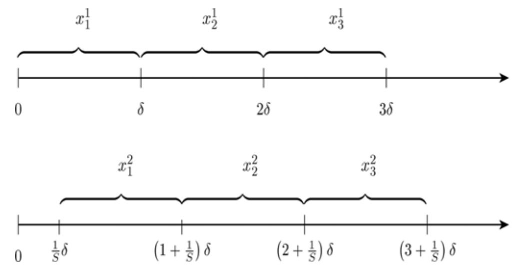

Fig. 1. x_j^1 are the usual returns. The bottom series are the offset returns x_j^s, $s = 2, \ldots, S$.

In Fig. 2 we trace out the implied kernel weights for three subsampled realised kernels. The left panels display the original kernel functions and right panels display the implied kernel functions. For the truncated kernel ($H = 1$) subsampling leads to a substantially different implied kernel function—the trapezoidal kernel of Politis and Romano (1995). For the kinked Bartlett kernel, subsampling leads to the same kernel function. For a smooth kernel function, the original and implied kernel functions are fairly similar; however subsampling does impose some piecewise linearity which is the reason that subsampling of smooth kernels increases the asymptotic variance.

The connection between subsampled realised kernels and realised kernels is perhaps not too surprising, because Bartlett (1950) motivated his kernel with the subsampling idea. The conventional form of subsampling is based on subseries that consist of consecutive observations. This is different from the case for our subsamples that consist of every Sth observation. Such ones are called subgrids in Zhang et al. (2005). While the two types of subsampling are different, they can result in identical estimators. For instance, the sparsely sampled realised variance, $\gamma_0^1(X_\delta)$, is identical to Carlstein's subsample estimator (of the variance of a sample mean when the mean is zero); see Carlstein (1986). Carlstein's estimator is based on non-overlapping subseries and Künsch (1989) analysed the closely related estimator on the basis of overlapping subseries. Interestingly, the (overlapping) subsample estimator of Künsch (1989) is identical to the average sparsely sampled realised variance called "second best" in Zhang et al. (2005), so the TSRV and MSRV estimators, by Zhang et al. (2005), Aït-Sahalia et al. (forthcoming), and Zhang (2006), can be expressed as linear combinations of two or more subsample estimators of the overlapping subseries type by Künsch (1989). For additional details on the relation between the Bartlett kernel and various subsample estimators; see Anderson (1971, p. 512), Priestley (1981, pp. 439–440), and Politis et al. (1999, pp. 95–98).

This paper has the following structure. We present the basic framework in Section 2 along with some known results. In Section 3 we present our main results. Here we derive the limit theory for subsampled realised kernels and show that subsampling cannot improve realised kernels within a very broad class of estimators. In Section 4, we given some specific recommendations on empirical implementation of subsampled realised kernels and how to conduct valid inference in this context. We present results from a small simulation study in Section 5 and an empirical application in Section 6. We conclude in Section 7 and present all proofs in an Appendix.

2. Notation, definitions and background

2.1. Semimartingales and quadratic variation

The fundamental theory of asset prices says that the log-price at time t, Y_t, must, in a frictionless arbitrage free market, obey a *semimartingale* process (written as $Y \in \mathcal{SM}$) on some filtered