Computational Statistics and Data Analysis ■ (■■■■) ■■■–■■■

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

A bootstrap approximation for the distribution of the Local Whittle estimator

Josu Arteche, Jesus Orbe *

Department of Econometrics and Statistics, University of the Basque Country UPV/EHU, Bilbao 48015, Spain

ARTICLE INFO

Article history:
Received 25 March 2014
Received in revised form 24 February 2015
Accepted 16 March 2015
Available online xxxx

Keywords:
Long memory
Bootstrap
Local Whittle estimation

ABSTRACT

The asymptotic properties of the Local Whittle estimator of the memory parameter d have been widely analysed and its consistency and asymptotic distribution have been obtained for values of $d \in (-1/2, 1]$ in a wide range of situations. However, the asymptotic distribution may be a poor approximation of the exact one in several cases, e.g. with small sample sizes or even with larger samples when $d > 0.75$. In other situations the asymptotic distribution is unknown, as for example in a noninvertible context or in some nonlinear transformations of long memory processes, where only consistency is obtained. For all these cases a bootstrap strategy based on resampling a (perhaps locally) standardised periodogram is proposed. A Monte Carlo analysis shows that this strategy leads to a good approximation of the exact distribution of the Local Whittle estimator in those situations where the asymptotic distribution is not reliable.

1. Introduction

Long memory is a common feature of many time series in areas as diverse as economics, finance, hydrology, climatology, politics and network traffic. It means that observations which are far apart maintain a significant relationship such that the autocorrelations are not summable. The non summability of the autocorrelations implies that the spectral density function $f(\lambda)$ diverges at the origin. In fact, the most common definition of long memory is established by the behaviour of the spectral density function around the origin such that it satisfies

$$f(\lambda) \sim C\lambda^{-2d} \quad \text{as } \lambda \to 0, \tag{1}$$

for a finite positive constant C, where $a \sim b$ means that $a/b \to 1$. The memory parameter d governs the persistence of the series. If $d = 0$ the series has short memory, whereas a value of $d > 0$ implies long memory or strong dependence such that $f(\lambda)$ diverges at $\lambda = 0$ and the autocovariances are not summable. Finally, the antipersistent case $d < 0$ entails a zero in the spectral density function at the origin, usually caused by overdifferencing.

Knowledge of d provides useful information on many characteristics of the series and is of particular interest in a wide range of situations. For example, a value $d \geq 1/2$ implies that the series is nonstationary since $f(\lambda)$, which in this case has a pseudo spectral density interpretation, is not integrable, although mean reversion is possible as long as $d < 1$. Also, the value of d of the cointegrating errors determines the degree of cointegration by comparison with the memory parameter of the series in the cointegrating regression. As a third example, d can be interpreted as the degree of differentiation necessary to reach a weak dependent stationary and invertible process in a context of fractional integration.

* Corresponding author. Tel.: +34 946013740.
E-mail address: jesus.orbe@ehu.es (J. Orbe).

http://dx.doi.org/10.1016/j.csda.2015.03.014
0167-9473/

Please cite this article in press as: Arteche, J., Orbe, J., A bootstrap approximation for the distribution of the Local Whittle estimator. Computational Statistics and Data Analysis (2015), http://dx.doi.org/10.1016/j.csda.2015.03.014

Estimating d is thus of inherent importance. Parametric methods such as (quasi) maximum likelihood or the Whittle approximation entail a risk of inconsistency if the model is misspecified, even if that misspecification occurs at frequencies not affected by d. In order to avoid that risk, semiparametric or local techniques have been proposed, which only restrict the behaviour of the spectral density around the pole as in (1). One of the most popular is the Local Whittle (LW) estimator, which is the one with which we are concerned in this paper. It was first proposed by Künsch (1987) but it was not until Robinson (1995) that its nice asymptotic properties were proved in the stationary and invertible case $-1/2 < d < 1/2$. In particular Robinson showed consistency, pivotal asymptotic normal distribution and higher efficiency than other rival techniques (such as log periodogram regression) under very mild conditions, allowing for non Gaussian series. Velasco (1999), Phillips and Shimotsu (2004) and Shao and Wu (2007) extended the asymptotic properties of the LW estimator to the nonstationary case, obtaining consistency for $d \leq 1$ and asymptotic normality for $d < 3/4$. For larger values of d the asymptotic distribution is non normal and the estimator is inconsistent for $d > 1$. These limitations have been overcome by the Exact LW estimator of Shimotsu and Phillips (2005) and the Extended LW of Abadir et al. (2007), which are extensions of the original LW estimation that preserve the asymptotic properties of the LW estimator with $d < 3/4$ for any value of d with no loss of efficiency.

The standard asymptotic distribution of the LW estimator and its pivotal nature (at least for $d < 3/4$) makes it very simple to implement asymptotic inference on the value of d. However the asymptotic distribution may be a poor approximation of the exact one in moderately-sized samples or even in large samples if $d \in (0.75, 1]$ (see Phillips and Shimotsu, 2004), rendering inference based on asymptotic results rather unreliable. In other situations inference on the memory parameter is not a feasible possibility because the asymptotic distribution of the LW estimator is unknown despite its consistency. This is the case in noninvertible fractionally integrated processes (Shimotsu and Phillips, 2006) and in some nonlinear transformations of long memory series (Dalla et al., 2005). In all these cases we propose to use a bootstrap method to approximate the exact distribution of the LW estimator, which makes it possible to implement reliable inference on d.

The bootstrap was originally designed for samples of independent observations, but some refinements have been proposed to deal with dependent data. In this context there are basically two approaches. One is based on describing the dependence through a parametric model with independent disturbances. After the model is estimated the bootstrap is implemented in the residuals, which are assumed to be close to being independent. The sieve bootstrap follows this spirit but instead of identifying the correct model, an AR approximation of sufficiently high order is estimated to capture the relevant dependence of the series. The second approach does not rely on a model but attempts to retain the structure of dependence by resampling overlapping or nonoverlapping blocks of observations. This is the block bootstrap, designed to maintain dependence inside the block while assuming independence between blocks.

The applicability of these traditional bootstrap methods to long memory series is influenced by the strong persistence of the series and the absence of mixing conditions, as analysed for example by Poskitt (2008) and Kreiss et al. (2011) for the sieve bootstrap and by Lahiri (1993) and Kim and Nordman (2011) for the block bootstrap (see also Murphy and Izzeldin, 2009). To avoid these problems Kapetanios and Papailias (2011) propose fractionally differencing the series using a consistent estimator of d prior to the application of a bootstrap strategy. If the estimate is close to the true value of d the bootstrap is finally applied to a series that is close to being weak dependent. But an estimate far from the true d invalidates the procedure since the bootstrap would have to be finally implemented in a long memory series regardless of the (incorrect) prior differencing.

In the paper presented here we follow a different approach based on a bootstrap strategy in the frequency domain, which is the context in which the LW estimator is defined. We take advantage of the fact that bootstrap approximations to the distribution of the Local Whittle estimator do not need bootstrap samples of the original series; only resampling of the periodogram is necessary. This implies that the problems originated by the strong dependence of the data are partially avoided, since the transformation that leads to the periodogram entails a significant modification in the structure of dependence. For example, the periodogram ordinates of weak dependent series are asymptotically independent, regardless of the dependence of the original series. However, periodogram ordinates of long memory series are not asymptotically independent around the spectral pole and they show a marked structure (far from the periodogram of a white noise) that would have to be replicated by the bootstrap samples. With that purpose, but in a weak stationary context, Paparoditis and Politis (1999) adapt a local bootstrap suggestion introduced by Shi (1991) and set out to resample near periodogram ordinates locally, thus retaining the global structure of the periodogram. Considering blocks of near periodogram ordinates makes this strategy resemble the more traditional block bootstrap. But, whereas the block bootstrap is designed to maintain the local structure by resampling blocks of observations while assuming independence between blocks, the local bootstrap resamples periodogram ordinates within a block of neighbouring frequencies, keeping the global structure of the periodogram unaltered. Thus, a local strategy seems to be more adequate under long memory where the periodogram shows a marked structure. However, as noted by Silva et al. (2006), replication of this structure in the bootstrap samples compels the use of a very narrow interval around the frequency of interest, which affects the performance of the bootstrap (in fact, Silva et al., 2006 propose resampling within a neighbourhood of only one or two frequencies).

In order to avoid this limitation we follow Franke and Härdle (1992) and Dahlhaus and Janas (1996) who, dealing with weak dependent series, propose resampling Studentised periodogram ordinates obtained by dividing the periodogram by an estimate of the spectral density function, which implies smoothing the structure of the raw periodogram such that the Studentised periodogram is closer to the periodogram of a white noise. In this context, the main challenge with long memory series is to obtain an estimator of the spectral density that is consistent over the whole band of frequencies used

Please cite this article in press as: Arteche, J., Orbe, J., A bootstrap approximation for the distribution of the Local Whittle estimator. Computational Statistics and Data Analysis (2015), http://dx.doi.org/10.1016/j.csda.2015.03.014