Available online at www.sciencedirect.com
ScienceDirect

Journal of Econometrics 138 (2007) 104–124

JOURNAL OF Econometrics

www.elsevier.com/locate/jeconom

Unit root log periodogram regression

Peter C.B. Phillips[a,b,c,*]

[a]*Cowles Foundation for Research in Economics, Yale University, Box 208281 Yale Station, New Haven, CT 06520, USA*
[b]*University of York, UK*
[c]*University of Auckland, New Zealand*

Available online 5 July 2006

Abstract

Log periodogram (LP) regression is shown to be consistent and to have a mixed normal limit distribution when the memory parameter $d = 1$. Gaussian errors are not required. The proof relies on a new result showing that asymptotically infinite collections of discrete Fourier transforms (dft's) of a short memory process at the fundamental frequencies in the vicinity of the origin can be treated as asymptotically independent normal variates, provided one does not include too many dft's in the collection.

JEL classification: C22

Keywords: Asymptotic independence; Discrete Fourier transform; Fractional integration; Log periodogram regression; Long memory parameter; Nonstationarity; Semiparametric estimation; Unit root

1. Introduction

For the last two decades a primary focus in econometric research has been on the long-run properties of economic time series, including the intrinsic memory properties displayed by individual series and the existence of long-run relationships between series. Many economic time series, such as inflation and interest rates, display long memory in the sense that their temporal autocorrelations decay slowly (if at all) and characterizing this property

*Corresponding author at: Cowles Foundation for Research in Economics, Yale University, Box 208281 Yale Station, New Haven, CT 06520, USA. Tel.: +1 203 421 4708; fax: +1 203 432 5429.
E-mail address: peter.phillips@yale.edu.

0304-4076/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2006.05.017

empirically has presented many econometric challenges. A particularly attractive econometric approach is semiparametric, in which the parameter (d) that characterizes memory in the data is estimated without making any delimiting assumptions about the short memory components in the data generating process. Accordingly, semiparametric estimation of the parameter d in fractionally integrated ($I(d)$) time series has received much recent study.

In applied work, $I(d)$ processes with fractional $d>0$ have been found to provide good empirical models for financial time series and volatility measures, as well as certain macroeconomic time series like inflation, money stock, and interest rates. Robinson (1994a) and Baillie (1996) reviewed aspects of this work relevant to econometrics up to the mid 1990s and there has been much work in the field since then. Growing evidence in applied work indicates that fractionally integrated processes can describe certain long range characteristics of economic data rather well, including the volatility of financial asset returns, forward exchange market premia, interest rate differentials, and inflation rates.

Two commonly used semiparametric estimators are log periodogram (LP) regression (Geweke and Porter-Hudak, 1983) and local Whittle (LW) estimation (Künsch, 1986). LW estimation involves numerical optimization of the LW likelihood and is attractive because it is asymptotically efficient. LP regression is popular because of its convenience, which stems from the simplicity of its construction as a linear regression estimator, and it has been extensively used in applied econometric research.

Let X_t be a fractional process satisfying

$$(1-L)^d X_t = u_t, \quad t \geqslant 0, \;\; X_0 = u_0 = 0, \tag{1}$$

where u_t is stationary with zero mean, finite moments to order p and continuous spectral density $f_u(\lambda)>0$. The parameter d in (1) measures the extent of the memory or long range dependence in X_t. The present paper concentrates on the special case where $d=1$, which corresponds to the important unit root case.

The LP estimator $\widehat{d}$ is obtained from the least squares regression

$$\log(I_X(\lambda_s)) = \widehat{c} - \widehat{d} \log|1-\mathrm{e}^{\mathrm{i}\lambda_s}|^2 + \text{residual} \tag{2}$$

taken over fundamental frequencies $\{\lambda_s = 2\pi s/n : s=1,\ldots,m\}$ for some $m<n$. Setting $a_s = \log|1-\mathrm{e}^{\mathrm{i}\lambda_s}|$ and $x_s = a_s - \overline{a}$, where $\overline{a} = m^{-1}\sum_{s=1}^m a_s$, we have

$$\widehat{d} = -\frac{1}{2}\frac{\sum_{s=1}^m x_s \log I_X(\lambda_s)}{\sum_{s=1}^m x_s^2}, \tag{3}$$

where $I_X(\lambda_s) = w_X(\lambda_s)w_X(\lambda_s)^*$ is the periodogram and $w_X(\lambda_s)$ is the discrete Fourier transform (dft), $w_X(\lambda_s) = (1/\sqrt{2\pi n})\sum_{t=1}^n X_t \mathrm{e}^{\mathrm{i}t\lambda_s}$ of the time series X_t. The regression (2) is motivated by the form of the log spectrum of X_t and has appeal because of its nonparametric treatment of u_t and the convenience of linear least squares. Under Gaussian assumptions and in the stationary case, where $d \in \left(-\frac{1}{2}, \frac{1}{2}\right)$, Robinson (1995) developed consistency and asymptotic normality results for a version of $\widehat{d}$ which trims out low frequencies periodogram ordinates (i.e. takes $s \geqslant l$, for some $m>l>1$), as suggested by Künsch (1986). Hurvich and Beltrao (1993) have developed data-driven criteria for the selection of m; and Hurvich et al. (1998) extend Robinson's (1995) results to include low frequencies ordinates and find an optimal choice of the number of periodogram ordinates