Journal of Econometrics 185 (2015) 468–494

Contents lists available at ScienceDirect

Journal of Econometrics

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

Nonparametric predictive regression☆

Ioannis Kasparis [a,*], Elena Andreou [a,b], Peter C.B. Phillips [c,d,e,f]

[a] University of Cyprus, Cyprus
[b] CERP
[c] Yale University, United States
[d] University of Auckland, New Zealand
[e] University of Southampton, United Kingdom
[f] Singapore Management University, Singapore

ARTICLE INFO

Article history:
Received 7 September 2012
Received in revised form
9 February 2014
Accepted 14 May 2014
Available online 25 June 2014

JEL classification:
c22
C32

Keywords:
Fractional Ornstein–Uhlenbeck process
Functional regression
Nonparametric predictability test
Nonparametric regression
Stock returns
Predictive regression

ABSTRACT

A unifying framework for inference is developed in predictive regressions where the predictor has unknown integration properties and may be stationary or nonstationary. Two easily implemented nonparametric F-tests are proposed. The limit distribution of these predictive tests is nuisance parameter free and holds for a wide range of predictors including stationary as well as non-stationary fractional and near unit root processes. Asymptotic theory and simulations show that the proposed tests are more powerful than existing parametric predictability tests when deviations from unity are large or the predictive regression is nonlinear. Empirical illustrations to monthly SP500 stock returns data are provided.

1. Introduction

The limit distributions of various estimators and tests are well known to be non-standard in the presence of stochastic trends (e.g., Phillips, 1987a,b; Chan and Wei, 1987). For instance, least squares cointegrating regression does not produce mixed-normal limit theory or pivotal tests unless strong conditions of long run orthogonality hold. Several early contributions (among others, Phillips and Hansen, 1990, Saikkonen, 1991, Phillips, 1995) developed certain modified versions of least squares for which mixed normality and standard methods of inference apply. While these approaches are now in widespread use in empirical research, some important obstacles to valid inference remain. First, modified statistics require for their validity some prior information about integration properties in order to choose appropriate tests. In consequence, the use of unit root and stationarity tests prior to parametric inference is common practice in applied work, exposing this approach to pre-test difficulties. Second, inference based on modified techniques is not robust to local deviations from the unit root model (Elliott, 1998) and modified tests can exhibit severe size distortions when there are local deviations from unity and significant correlations between the covariates and the equation error. Both of these problems arise in cointegrating and predictive regressions.

To address the second difficulty, several inferential methods that are robust to local deviations from unity have been proposed, including Wright (2000), Lanne (2002), Torous et al. (2004), Campbell and Yogo (2006), Jansson and Moreira (2006), and Magdalinos and Phillips (2009). The methods have attracted particular attention in the predictive regression literature. Some of the techniques proposed are technically complicated and difficult to implement in practical work, which in part explains why some

☆ First version of the paper: September 2012. The authors thank the Editor, the Associate Editor and two referees for helpful comments and suggestions. Further thanks go to Timos Papadopoulos for substantial assistance with the simulations and to Eric Ghysels and Tassos Magdalinos for useful comments and suggestions. Andreou acknowledges support of the European Research Council under the European Community FP7/2008-2013 ERC grant 209116. Phillips acknowledges partial support from the NSF under Grant Nos. SES 09-56687 and SES 12-58258.

* Correspondence to: Department of Economics, University of Cyprus, University Avenue 1, P.O. Box 20537, 1678 Nicosia, Cyprus. Tel.: +357 22893717; fax: +357 22895028.

E-mail address: kasparis@ucy.ac.cy (I. Kasparis).

http://dx.doi.org/10.1016/j.jeconom.2014.05.015
0304-4076/

methods have never been used in empirical work. Most of these approaches also focus on regressions with nearly integrated (NI) covariates and some are invalid for stationary regressors. Implementation of the Campbell and Yogo (2006) method, for instance, typically imposes bounds on the near-to-unity parameter that rule out stable autoregressions. Further, if those bounds are relaxed, it has recently been shown that confidence intervals produced by this method have zero coverage probability in the limit when the predictive regressors are stationary (Phillips, 2014), so there is complete failure of robustness in this case. It is also unknown whether these techniques are valid when the regressors involve fractional processes or other types of nonstationarity. Extension of valid inference to fractional processes is particularly important. Unlike NI processes, fractional processes directly bridge the persistence gap between $I(0)$ and $I(1)$ processes, so that partial sums have a range of magnitudes of the form

$$\sum_{t=1}^{n} x_t = O_p(n^{\alpha}), \quad \text{for some } \alpha \in (1/2, 3/2). \tag{1}$$

The approach of Magdalinos and Phillips (2009) holds for moderately integrated processes, whose partial sums are of the general form (1).

All of these methods are parametric and may not be robust to functional form misspecification. Functional form affects the power of predictive tests under nonstationarity. For instance, fully modified t-tests are based on linear regression and for a near integrated predictor, the test statistic has divergence rate $O_p(n)$ under a linear alternative but may be inconsistent for certain nonlinear alternatives, as we discuss in the paper. In a related vein, Wang and Phillips (2012) found that nonparametric nonstationary specification tests have divergence rates under local alternatives that depend explicitly on the functional form and may be inconsistent for certain functional forms.

The present paper contributes to this literature in several ways. First, we adopt a nonparametric approach using recent theory for nonparametric regression in nonstationary settings by Wang and Phillips (2009a), hereafter WP). Nonparametric F-tests are proposed which have limit distributions that are invariant to integration order. The tests are easy to implement, rely on simple functionals of the Nadaraya–Watson kernel regression estimator, and have limit distributions that apply for a wide range of predictors including stationary as well as non-stationary fractional and near unit root process. In this sense the proposed tests provide a unifying framework for inference. Further, the tests are robust to functional form. The limit distribution of the tests, under the null hypothesis (no predictability), is determined by functionals of independent χ^2 variates. Under the alternative hypothesis (predictability), asymptotic power rates are obtained. The power rates of the nonparametric tests are affected by the bandwidth parameter and are slower than that of parametric tests against linear alternatives. Interestingly, however, the nonparametric tests may attain faster divergence rates than those of parametric tests in cases where parametric fits are misspecified in terms of functional form.

Simulation results suggest that in finite samples the proposed nonparametric tests have stable size properties and can be more powerful than existing parametric predictability tests even when the latter are based on correctly specified models. An empirical illustration of the proposed tests evaluates the predictability of the monthly S&P 500 excess returns using the Earnings Price and Dividend Price ratios as predictors over the period 1926–2010 and various subperiods.

The remainder of the paper is organized as follows. Section 2 provides the model, assumptions and some preliminary results. The nonparametric tests and limit theory is given in Section 3. Section 4 considers power. Simulations results are reported in Section 5. The empirical illustration is given in Sections 6 and 7 concludes. Proofs are given in Appendices A and B.

Notation is standard. For instance, for two sequences a_n, b_n the notation $a_n \sim b_n$ denotes $\lim_{n\to\infty} a_n/b_n = 1$, and $=_d$ represents distributional equality. We use $\lfloor \cdot \rfloor$ to denote integer part, $1\{A\}$ as the indicator function of A, and $i = \sqrt{-1}$. For any sequence X_t, $\overline{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$ and $\underline{X}_t := X_t - \overline{X}$. Similarly, for any functions f_r, $\overline{f} := \int_0^1 f_r dr$ and $\underline{f}_r := f_r - \overline{f}$. Integrals of the form $\int_0^1 G_r dr$ and $\int_0^1 G_r dV_r$ are often written as $\int_0^1 G$ and $\int_0^1 G dV$.

2. Model and assumptions

We consider predictive regressions of the (possibly nonlinear) form

$$y_t = f(x_{t-\nu}) + u_t, \qquad f(x) = \mu + g(x), \tag{2}$$

where g is some unknown regression function, $\nu \geq 1$ is an integer valued lag term and u_t is a martingale difference term whose properties are specified below. When x_t is a stationary weakly dependent process, the limit theory of nonparametric regression estimators for models such as (2) is well known from early research (e.g. Robinson, 1983) and overviews in the literature (e.g. Li and Racine, 2007). The limit theory of the nonparametric tests proposed here follows readily from the standard theory in such cases.

The present work focuses on cases where x_t is nonstationary. We are particularly interested in models where $\{x_t\}_1^n$ is generated as a NI array of the commonly used form

$$x_t = \rho_n x_{t-1} + v_t, \qquad x_0 = 0, \tag{3}$$

with $\rho_n = 1 + \frac{c}{n}$, for some constant c. The error v_t may be a short-memory (SM) time series or an $ARFIMA(d)$, $d \in (-1/2, 1/2)$, process with either long memory (LM) or anti-persistence (AP). Both x_t and u_t are defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with a filtration specified below. The regression function f in (2) is estimated by the Nadaraya–Watson estimator

$$\hat{f}(x) = \frac{\sum_{t=\nu+1}^{n} K_h(x_{t-\nu} - x)\, y_t}{\sum_{t=\nu+1}^{n} K_h(x_{t-\nu} - x)}, \tag{4}$$

where $K_h(.) = K(./h)$, $K(.)$ is a kernel function and h is a bandwidth with $h = h_n \to 0$ as $n \to \infty$.

To fix ideas and for subsequent analysis we introduce the following technical conditions. Assumptions 2.1 and 2.2 below are largely based on WP (2009a), to which we refer readers for discussion. The WP notation is used here for ease of cross-reference. First, it is convenient to standardize x_t in array form as $x_{t,n} = x_t / d_n$ for some suitable sequence $d_n \to \infty$ so that $x_{\lfloor ns \rfloor, n}$ is compatible with a functional law as $n \to \infty$. It is also convenient to introduce a standardizing array $d_{l,k,n}$, $1 \leq k \leq l \leq n$ with $d_{l,k,n} \sim C d_{l-k}/d_n$ for some constant C. We note that $\left(x_{l,n} - x_{k,n}\right)/d_{l,k,n}$ has a limit distribution as $n, l-k \to \infty$. As in WP, it is convenient to use the set notation.

$$\Omega_n(\eta) = \{(l,k) : \eta n \leq k \leq (1-\eta)n, k + \eta n \leq l \leq n\},$$
$$0 < \eta < 1/2.$$

Assumptions 2.1 and 2.2 deal with the density function properties of x_t and their relation to the function f.

Assumption 2.1. For all $0 \leq k < l \leq n, n \geq 1$, there exist a sequence of σ-fields $\mathcal{F}_{n,k-1} \subseteq \mathcal{F}_{n,k}$ such that, (u_k, x_k) is adapted to $\mathcal{F}_{n,k}$ and conditional on $\mathcal{F}_{n,k}$, $\left(x_{l,n} - \rho_n^{l-k} x_{k,n}\right)/d_{l,k,n}$ has density