Journal of Empirical Finance 28 (2014) 261–272

Contents lists available at ScienceDirect

Journal of Empirical Finance

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

# A frequency-domain alternative to long-horizon regressions with application to return predictability

Natalia Sizova *

*Department of Economics, Rice University, Houston, TX 77251, USA*

ARTICLE INFO

*Article history:*
Received 28 September 2012
Received in revised form 5 November 2013
Accepted 10 March 2014
Available online 20 March 2014

*JEL classification:*
C12
C14
G12
G14
E47

*Keywords:*
Predictive regression
Semiparametric method
Local-to-unity
Long memory
Long-horizon regression
Subsampling

ABSTRACT

This paper aims at improved accuracy in testing for long-run predictability in noisy series, such as stock market returns. Long-horizon regressions have previously been the dominant approach in this area. We suggest an alternative method that yields more accurate results. We find evidence of predictability in S&P 500 returns even when the confidence intervals are constructed using model-free methods based on subsampling.



## 1. Introduction

Define the long-horizon return regression as the regression of the aggregated quantity of interest, here the return, $r_{t+1} + \ldots + r_{t+H}$, on a predictive variable, $x_t$. Three facts emerge from the vast literature on long-horizon regressions. First, despite the appealing empirical results, long-horizon regressions appear to carry little or no information about the relation of $r_{t+1}$ to $x_t$ aside from what is already conveyed by the simple regression of $r_{t+1}$ on $x_t$ (e.g., Boudoukh et al., 2007). Second, when correct confidence intervals are used, both simple regressions and long-horizon regressions only weakly reject the hypothesis of no predictability (e.g., Goyal and Welch, 2008; Valkanov, 2003; Wolf, 2000). Third, if $x_t$ is persistent, then both simple regressions and long-horizon regressions often yield biased estimates owing to the contemporaneous correlation of $r_t$ and $x_t$. Because the contemporaneous correlation is usually of opposite sign to the estimated regression slope for the return regressions, it follows from Stambaugh (1999) that the predictability in returns is systematically overestimated.

In this paper, we demonstrate that long-horizon regressions do carry additional information when the short-run dynamics of returns differ from their long-run dynamics. Furthermore, we show that there exist methods, defined in the frequency domain, that provide a superior fit for testing the long-run predictability compared to that provided by long-horizon regressions. In

* Tel.: +1 713 348 5613; fax: +1 713 348 5278.
*E-mail address:* Natalia.Sizova@rice.edu.

http://dx.doi.org/10.1016/j.jempfin.2014.03.002
0927-5398/

particular, we focus on two estimators: the local Whittle estimator and the frequency-domain least squares (FDLS) estimator by Robinson (1994). The corresponding confidence intervals for the tests are constructed in two ways: first, using a simple parametric approach by Boudoukh et al. (2007), and, second, using a method based on subsampling by Wolf (2000).

Applications of frequency-domain methods to the estimation of variable interactions have been extensively studied in the literature (see Granger, 1969; Granger and Hatanaka, 1964). The current paper is also related to the literature on the estimation of co-integrating systems of fractionally integrated processes (e.g., Bandi and Perron, 2006; Christensen and Nielsen, 2006; Lobato, 1997; Marmol and Velasco, 2004; Nielsen and Frederiksen, 2011; Robinson, 1994; Robinson and Marinucci, 2003). To the best of our knowledge, none of these papers has investigated the problem of testing for predictability in series resembling white noise, such as asset returns.

The paper is structured as follows. Section 1 introduces the main concepts to be applied in this study. Section 2 presents the results of our simulations. An empirical application of our results is demonstrated in Section 3. Section 4 presents empirical findings matching our tests with subsampled confidence intervals. The results are summarized in the Conclusion.

## 2. Basic concepts

Suppose that the return dynamics are governed by the following linear model:

$$r_{t+1} = \mu + \beta_x x_t + \varepsilon_{t+1},$$

where $\mathrm{cov}(x_t, \varepsilon_{t+1}) = 0$, for which we will test that $\beta_x = 0$. In the usual situation, which is also the situation of concern, $x_t$ is persistent. The most widely studied case is that in which $x_t$ follows an autoregressive process,

$$x_{t+1} = \rho x_t + u_{t+1}, \tag{1}$$

where $\rho$ is less than but close to 1. Here, $\varepsilon_{t+1}$ and $u_{t+1}$ are significantly less persistent than $x_t$. Once $r_t$ is aggregated, the following relation emerges:

$$\sum_{i=1}^{H} r_{t+i} = \mu H + \beta_x \sum_{i=1}^{H} x_{t+i-1} + \sum_{i=1}^{H} \varepsilon_{t+i}.$$

The higher the persistence of $x_t$, the more rapid the increase in the variance of the sum $\sum_{i=1}^{H} x_{t+i-1}$ with the horizon, $H$. At high frequencies, the process $r_t$ may appear noisy when $\beta_x$ is sufficiently small. However, if $x_t$ is persistent and $H$ is large, then the variation of the "predictable" portion, $\beta_x \sum_{i=1}^{H} x_{t+i-1}$, dominates the variation of the unpredictable portion, $\sum_{i=1}^{H} \varepsilon_{t+i}$, in effect removing the noise. This is a mechanism through which the long-horizon regressions extract information from return series (e.g., Fama and French, 1988).

Although long-horizon regressions are a straightforward generalization of simple regressions, they are not guaranteed to yield efficient estimators under any assumptions. In this paper, the estimation is instead based on the Gaussian likelihood maximization, which yields quasi-maximum likelihood estimates (QMLEs) for non-normally distributed $x_t$ and $r_t$. The maximization of the Gaussian likelihood function with respect to the model parameters $\theta$ can be replaced by the minimization of the Whittle approximation for a class of processes,[1]

$$l_T(\theta) = \sum_{j=1}^{T-1} \left[ \log\left| s^{xr}\left(\theta, \omega_j\right)\right| + tr\left(s^{xr}\left(\theta, \omega_j\right)^{-1} I_T^{xr}\left(\omega_j\right)\right)\right],$$

where $T$ is the number of observations, $I_T^{xr}(\omega_j)$ is the sample periodogram of the pair $(x_t, r_t)$ calculated at natural frequencies, $\omega_j = 2\pi j/T$. The procedure minimizes the distance from $I_T^{xr}(\omega_j)$ to its population counterpart, $s^{xr}(\theta, \omega_j)$. The function $s^{xr}(\theta, \omega)$ is the theoretical spectrum of $(x_t, r_t)$, in which the (1, 1) element $s_{xx}^{xr}(\theta, \omega)$ is the spectrum of $x_t$. Its (2, 2) element $s_{rr}^{xr}(\theta, \omega)$ is the spectrum of $r_t$ and $s_{xr}^{xr}(\theta, \omega)$ and $s_{rx}^{xr}(\theta, \omega)$ are off-diagonal elements that characterize the cross-spectrum, i.e., the interrelatedness of $x_t$ and $r_t$ at different frequencies.

The exact formula for the spectrum depends on the model. For example, for an autoregressive $x_t$ as in model (1), the spectrum is represented by the following matrix:

$$s^{xr}(\theta, \omega) = \begin{bmatrix} \dfrac{s_{uu}^{u\varepsilon}(\theta', \omega)}{|1-\rho e^{-i\omega}|^2} & \beta_x e^{\omega i} \dfrac{s_{uu}^{u\varepsilon}(\theta', \omega)}{|1-\rho e^{-i\omega}|^2} + \dfrac{s_{u\varepsilon}^{u\varepsilon}(\theta', \omega)}{(1-\rho e^{-i\omega})} \\ \beta_x e^{-\omega i} \dfrac{s_{uu}^{u\varepsilon}(\theta', \omega)}{|1-\rho e^{-i\omega}|^2} + \dfrac{s_{\varepsilon u}^{u\varepsilon}(\theta', \omega)}{(1-\rho e^{+i\omega})} & \beta_x^2 \dfrac{s_{uu}^{u\varepsilon}(\theta', \omega)}{|1-\rho e^{-i\omega}|^2} + 2Re\left(\beta_x e^{-\omega i} \dfrac{s_{u\varepsilon}^{u\varepsilon}(\theta', \omega)}{1-\rho e^{-i\omega}}\right) + s_{\varepsilon\varepsilon}^{u\varepsilon}(\theta', \omega) \end{bmatrix},$$

[1] E.g., the results of Dzhaparidze (1986) for Gaussian processes rely on the conditions that should hold for the Fourier coefficients $\beta(\tau)$ of the spectral density, $\sum_{\tau=1}^{\infty} \tau \|\beta(\tau)\|^2 < \infty$. Hannan (1973) provides the results for moving average (not necessarily Gaussian) processes with square-summable moving average coefficients and driven by innovations with constant variances.