Journal of Econometrics 177 (2013) 250–264

Contents lists available at ScienceDirect

Journal of Econometrics

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

Predictive regression under various degrees of persistence and robust long-horizon regression

Peter C.B. Phillips [a,b,c,d,*], Ji Hyung Lee [a]

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

ARTICLE INFO

Article history:
Available online 17 April 2013

JEL classification:
C22

Keywords:
Asymptotic theory
Balanced regression
Endogeneity
Instrumentation
IVX methods
Local power
Mild integration
Mildly explosive
Predictive regression
Robustness

ABSTRACT

The paper proposes a novel inference procedure for long-horizon predictive regression with persistent regressors, allowing the autoregressive roots to lie in a wide vicinity of unity. The invalidity of conventional tests when regressors are persistent has led to a large literature dealing with inference in predictive regressions with local to unity regressors. Magdalinos and Phillips (2009b) recently developed a new framework of extended IV procedures (IVX) that enables robust chi-square testing for a wider class of persistent regressors. We extend this robust procedure to an even wider parameter space in the vicinity of unity and apply the methods to long-horizon predictive regression. Existing methods in this model, which rely on simulated critical values by inverting tests under local to unity conditions, cannot be easily extended beyond the scalar regressor case or to wider autoregressive parametrizations. In contrast, the methods developed here lead to standard chi-square tests, allow for multivariate regressors, and include predictive processes whose roots may lie in a wide vicinity of unity. As such they have many potential applications in predictive regression. In addition to asymptotics under the null hypothesis of no predictability, the paper investigates validity under the alternative, showing how balance in the regression may be achieved through the use of localizing coefficients and developing local asymptotic power properties under such alternatives. These results help to explain some of the empirical difficulties that have been encountered in establishing predictability of stock returns.

1. Introduction

Predictive regression typically encounters the problem of predicting some noisy stationary variable using a highly persistent regressor. A leading practical example is stock return predictability in finance and the empirical puzzles associated with these regressions that have emerged in the financial literature. The traditional form of the efficient market hypothesis supports the idea of martingale behavior in stock prices and stock return unpredictability. But empirical evidence on the predictability of returns shows mixed results on the explanatory power of various economic fundamentals such as the dividend–price ratio, leading to what has become known as the stock return predictability puzzle. Some researchers have even characterized stock return predictability as a new stylized fact in finance.

Forecasting stock returns has been a longstanding interest of Hashem Pesaran. His well cited paper with Alan Timmerman (1995) is an early contribution in the field that highlighted the need to robustify econometric procedures to the time varying predictive power of economic factors on stock returns. The present paper explores a related theme and investigates robust predictive regressions in the presence of multivariate nonstationary regressors developing results that have direct application to stock return forecasting.

The empirical model employed in stock return predictive regressions commonly involves a linear regression of returns on economic fundamentals. The regressors typically manifest a high, but imprecisely determined, degree of persistence. This uncertainty in the degree of regressor persistence is usually modeled in terms of an autoregressive coefficient with an unknown local to unity parameter (the localizing coefficient) that measures the (sample size normalized) departure of this autoregressive coefficient from unity. The localizing coefficient is not consistently estimable and this characteristic leads to nonstandard and nonpivotal inference problems.

Other commonly occurring predictive regressions include forward premium regressions in international finance and consumption growth regressions in macroeconomics. These models share the same problem of nonstandard and nonpivotal inference that

* Corresponding author at: Yale University, United States.
E-mail addresses: peter.phillips@yale.edu (P.C.B. Phillips), jihyung.lee@yale.edu (J.H. Lee).

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.04.011

originates in the unknown degree of persistence in the predictive regressors. These common difficulties in predictive regressions have kindled a widespread search for robust inference methods. There have been extensive efforts and some procedures such as Bonferroni type methods have received much attention. At present, Bonferroni type procedures represent the state-of-the-art in this literature but they do have some undesirable and limiting properties. In particular, simulations are required to compute critical values to perform inference and confidence interval construction since the limit distribution employed in the calculations is nonstandard. A second limitation is that the method is very difficult to extend beyond the scalar regressor case. In practical work, there are often a selection of variables representing various economic fundamentals which need to be investigated in applied work on predictive regression and it is delimiting for empirical procedures to be restricted to a single regressor.

Magdalinos and Phillips (2009b, MP henceforth) recently developed a novel extended IV procedure (called IVX regression) and established some attractive asymptotic features of this method that apply in quite general cointegrating regression models. In particular, the IVX method has some useful and somewhat surprising features such as standard chi square testing without any precise knowledge about the degree of persistence in the regressors, straightforward extension of the methods to multivariate models, and great generality in terms of the permissible persistence (or vicinity of unity) space. This method has very recently been applied to a predictive regression model and was shown to inherit all these advantages in this context (Kostakis et al., 2010).

In empirical research short horizon predictive regressions have shown generally inconclusive findings. In response, researchers have been studying prediction over longer horizons for such variables as stock returns and the forward premium. Long-horizon predictive regressions share the same problems as their short horizon counterparts (viz., nonstandard and nonpivotal inference with persistent regressors) and similar solutions such as Bonferroni techniques have been applied with the same limitations noted above. It is therefore natural to explore whether IVX methodology has potentially beneficial applications in long-horizon regressions.

That question forms the focus of the present paper. We study long-horizon predictive regressions and propose a novel inference procedure which is based on an extended version of MP and Kostakis et al. (2010). The long-horizon version of IVX is analyzed, and shown to be applicable to even much wider parameter region near unity: from boundary of stationary/unit root side (mildly integrated regressors) to mildly explosive regressors. This is the most extensive treatment of the parameter region near unity not only within this setting of predictive regression but also in more general time series regressions such as cointegrating regressions. All the attractive features of IVX regression such as standard asymptotic chi square inference with possible multivariate regressor and regressand are shown to apply.

A further contribution of the paper is to investigate validity of the predictive model specification under the alternative. In particular, we show how balance in the regression may be achieved through the use of localizing coefficients and that non trivial local asymptotic power applies under such alternatives. These results partly explain some of the practical difficulty that has been encountered in establishing empirical evidence of predictability. In effect, the departures from the null that deliver predictability are necessarily small in order to preserve the observed character of the dependent variable, thereby making detection difficult. Nonetheless, as we show here, long horizon IVX regression provides a simple and effective machinery for testing predictability that has non-trivial asymptotic power against local alternatives.

The paper is organized as follows. Section 2 overviews existing results on predictive regression literature and the various limitations of the methods currently used in empirical research. Section 3 develops a limit theory for the extended IVX approach in long horizon predictive regression. Section 4 concludes and proofs of the main results are given in the Appendix. The Appendix also contains a discussion of balancing predictive regression and an analysis of local asymptotic power.

A supplement (Phillips and Lee, 2012a) is available online and provides supporting lemmas and further technical arguments that are used in the paper.

2. Predictive regressions: literature review and motivation

This Section reviews key results in the predictive regression literature and identifies the source of the difficulties encountered by existing methods. The basic linear predictive model can be characterized as:

$$y_t = \beta x_{t-1} + u_{0t}, \tag{2.1}$$

$$x_t = \rho x_{t-1} + u_{xt}. \tag{2.2}$$

We impose a simple but widely used structure of martingale difference sequence (mds) innovations for $u_t := [u_{0t}, u_{xt}]$ with conditional variance

$$\mathbb{E}_{\mathcal{F}_{t-1}}\left[u_t u_t'\right] = \begin{bmatrix} \Sigma_{00} & \Sigma_{0x} \\ \Sigma_{x0} & \Sigma_{xx} \end{bmatrix},$$

where $\mathcal{F}_t$ is the natural filtration. This framework allows only for contemporaneous correlation between the components of the model. More general dependence structures will be permitted later but for the purpose of this overview we retain the simple mds structure. Full details of the conditions and notation used in the paper are provided in Appendix A.1.

2.1. Existing problems

2.1.1. Finite sample bias with stationary regressors

The centered OLS coefficient estimator in (2.1) has the form:

$$\hat{\beta} - \beta = \frac{\sum_{t=1}^{n} x_{t-1}u_{0t}}{\sum_{t=1}^{n}(x_{t-1})^2} = \frac{\sum_{t=1}^{n} x_{t-1}u_{0.xt}}{\sum_{t=1}^{n}(x_{t-1})^2} + \left(\frac{\Sigma_{0x}}{\Sigma_{xx}}\right)\frac{\sum_{t=1}^{n} x_{t-1}u_{xt}}{\sum_{t=1}^{n}(x_{t-1})^2}$$

$$= \frac{\sum_{t=1}^{n} x_{t-1}u_{0.xt}}{\sum_{t=1}^{n}(x_{t-1})^2} + \left(\frac{\Sigma_{0x}}{\Sigma_{xx}}\right)\left(\hat{\rho} - \rho\right), \tag{2.3}$$

where $u_{0.xt} = u_{0t} - \frac{\Sigma_{0x}}{\Sigma_{xx}}u_{xt}$ and $\hat{\rho} = \left(\sum_{t=1}^{n} x_{t-1}^2\right)^{-1}\sum_{t=1}^{n} x_{t-1}x_t$. Under normality and with a stationary regressor ($|\rho| < 1$), Stambaugh (1999) gave a bias expansion for $E\left(\hat{\beta} - \beta\right)$ using the well known bias expansion for a fitted AR(1) (Kendall, 1954), which in the case of a fitted intercept has the form

$$E\left(\hat{\beta} - \beta\right) = -\left(\frac{\Sigma_{0x}}{\Sigma_{xx}}\right)\left(\frac{1+3\rho}{n}\right) + O\left(\frac{1}{n^2}\right), \tag{2.4}$$

which has come to be known as the "Stambaugh bias".[1] Accordingly, the first order bias adjusted estimator has been used by

[1] As indicated, the bias formula given in (2.4) is for a stationary AR(1) process with a fitted intercept. Unlike the stationary case, fitting an intercept affects asymptotics in both the nonstationary and explosive regressor cases. For the subsequent development, which focuses on persistent and explosive regressors, it is convenient to keep to the no-intercept case in the generating mechanism for x_t. On the other hand, introducing an intercept in the predictive regression (2.1), so that $y_t = \mu_y + \beta x_{t-1} + u_{0t}$, is easily handled even with nonstationary regressors — see Kostakis et al. (2010) — and this is the primary case of interest in practice.