International Journal of Forecasting 31 (2015) 587–597

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Testing for multiple-period predictability between serially dependent time series

Chris Heaton

Macquarie University, Department of Economics, Balaclava Road, North Ryde, New South Wales 2109, Australia

ARTICLE INFO

Keywords:
Covariance estimation
Causality
Statistical tests
Inflation forecasting
Macroeconomic forecasting

ABSTRACT

This paper reports the results of a simulation study that considers the finite-sample performances of a range of approaches for testing multiple-period predictability between two potentially serially correlated time series. In many empirically relevant situations, but not all, most of the test statistics considered are significantly oversized. In contrast, both an analytical approach proposed in this paper and a bootstrap are found to have accurate empirical sizes. In a small number of cases, the bootstrap is found to have a superior power. The test procedures considered are applied to an empirical analysis of the predictive power of a Phillips curve model during the 'great moderation' period, which illustrates the practical importance of using test statistics with accurate empirical sizes.

1. Introduction

Tests of multiple-period-ahead predictability are complicated by the fact that the prediction errors necessarily have a moving average structure, due to the overlap of successive prediction periods. In a theoretically ideal large sample setting, this does not present any problem for applied researchers, since there exist many covariance matrix estimators that are consistent in the presence of this type of autocorrelation. However, for samples of the sorts of magnitudes that are encountered most frequently in economics, test statistics constructed using many of the well-known alternative heteroscedasticity and autocorrelation consistent (HAC) covariance estimators may be severely oversized. While this issue has received some attention in the literature (see Ang & Bekaert, 2007; Hodrick, 1992; Kilian, 1999; Nelson & Kim, 1993; Richardson & Smith, 1991; Smith & Yadav, 1996; Wei & Wright, 2009), the context has often been the prediction of asset returns, and so the null model has usually been a martingale difference sequence (MDS). Consequently, while this body of literature has found evidence of significant size distortions when using well-known techniques for dealing with autocorrelation, and has suggested some superior methods, these methods are not usually directly applicable to cases in which the predicted variable is serially correlated under the null, as would be expected for macroeconomic series and many other applications of interest. Previous work that has considered a serially correlated predicted variable includes that of Lütkepohl and Burda (1997), who consider the Wald test in the context of a vector autoregression (VAR); Dufour, Pelletier, and Renault (2006), who use a parametric bootstrap to circumvent the technical difficulties of the Wald test; Pesaran, Pick, and Timmermann (2011), who pool non-overlapping regressions and propose a SURE estimator for a factor-augmented VAR; and Britten-Jones, Neuberger, and Nolte (2011), who propose a transformation to account for the serial correlation induced by the construction of the overlapping dependent variable, and deal with serial correlation in the variable from which it is constructed using the Newey–West estimator.

In this paper, the covariance estimator that was proposed by Hodrick (1992) for the multiple-period prediction

---

E-mail address: chris.heaton@mq.edu.au.

http://dx.doi.org/10.1016/j.ijforecast.2014.09.004
0169-2070/

regression of a variable that is an MDS under the null is generalized to cover cases in which the predicted variable is serially correlated. A simulation study is then conducted that compares the small-sample performance of a test for multiple-period predictability based on this estimator to those of a selection of other approaches that might be considered by applied researchers. The simulation study considers a range of prediction horizons, sample sizes, and degrees of serial correlation in both the predictor and predicted variables, and the results provide clear guidance for researchers who are interested in testing multiple-period predictability.

West (1997) also proposed a generalization of the Hodrick (1992) estimator for dealing with regression models with moving average errors. However, his approach differs from that proposed in this paper. Furthermore, the simulation study reported in West's paper considers only moving average orders of one and two, which is of little interest for applications in which predictions are being made more than three periods ahead.[1] Other authors have reported simulation studies that consider longer horizons for test statistics based on different estimators. Simulation studies by Ang and Bekaert (2007), Britten-Jones et al. (2011), Dufour et al. (2006), and Smith and Yadav (1996) found the kernel-based estimators of Andrews (1991) and Newey and West (1987) to provide test statistics for multiple-period predictability that are significantly oversized. Smith and Yadav (1996) found that the statistics were still oversized when the prewhitening procedure of Andrews and Monahan (1992) was used. Similar results were found for the Hansen and Hodrick (1980) estimator by Ang and Bekaert (2007) and Smith and Yadav (1996). In contrast, Dufour et al. (2006) and Kilian (1999) found that a bootstrapped statistic may have an accurate empirical size, and Ang and Bekaert (2007) found a similar result for the method of Hodrick (1992).

The simulation study reported in this paper extends this literature in several ways. Firstly, in contrast to the studies by Ang and Bekaert (2007) and Smith and Yadav (1996), this study considers serial correlation in the predicted variable, and, in contrast to Britten-Jones et al. (2011) and Dufour et al. (2006), a range of different strengths of this correlation are considered. Like Smith and Yadav (1996), I consider different strengths of the serial correlation in the predictor variable. Secondly, the present study considers a wider range of prediction horizons and sample sizes than has been considered in previous studies. Thirdly, this study considers nine different alternative test procedures, in contrast to the studies by Ang and Bekaert (2007), Dufour et al. (2006) and Smith and Yadav (1996), which consider seven, three and two, respectively. As a consequence, it provides a comparison of a wide range of test statistics within a single study design.

The remainder of this paper is structured as follows. In Section 2, the generalization of the Hodrick (1992) covariance estimator is derived, and compared to the alternative generalization due to West (1997). In Section 3, the Monte Carlo simulations are presented. Section 4 presents a brief application of all of the test procedures considered to the question of whether an expectations-augmented Phillips curve model was able to predict inflation over horizons of four to 12 quarters during the 'great moderation' period between the 1980s and the start of the financial crisis in 2008. Section 5 provides some concluding comments.

2. The estimator and test statistic

Suppose that we wish to test the null hypothesis that a vector w_t does not predict the change in a scalar variable y_t over h time periods. Let $y_t^{(1)} = y_{t+1} - y_t$ be an observable variable. The change in y_t over h time periods is then $y_t^{(h)} = \sum_{k=0}^{h-1} y_{t+k}^{(1)}$. It is assumed that, under the null hypothesis, $y_t^{(1)}$ may be approximated well by a stable, finite-ordered autoregression

$$y_t^{(1)} = \beta_0 + \sum_{j=1}^{p} \beta_j y_{t-j}^{(1)} + \varepsilon_{t+1}, \quad t = p+1, \ldots, T-1. \tag{1}$$

The technical assumptions for what follows are that (under the null hypothesis) $E(\varepsilon_{t+1}|\mathscr{F}_t) = 0$, where $\mathscr{F}_t = \sigma(\varepsilon_t, w_t, \varepsilon_{t-1}, w_{t-1}, \ldots)$; that the fourth moments of ε_t and w_t are finite; and that the characteristic roots of Eq. (1) lie outside the unit circle. Autoregressions are used widely in applied prediction problems, often quite successfully. Nonetheless, not all economic variables can be represented adequately by an autoregression. The working assumption made in this paper is that a stable, finite-order autoregression approximates the process of interest sufficiently well to provide errors that, at most, differ from an MDS only negligibly. Standard statistical tools exist for estimating the order and assessing the fit of an autoregression, and these should be applied prior to the application of the procedures proposed in this paper.

What follows may be understood more easily by first considering a simple special case. Set the order (p) in Eq. (1) to one, set $\beta_0 = 0$, and set the prediction horizon h to two. Since $y_t^{(2)} = y_{t+1}^{(1)} + y_t^{(1)}$, it is simple to show that

$$y_t^{(2)} = b y_{t-1}^{(1)} + \eta_t, \quad \text{where } \eta_t = (\beta+1)\varepsilon_{t+1} + \varepsilon_{t+2} \text{ and } b = \beta(1+\beta). \tag{2}$$

Let $\hat{b}$ be the OLS estimator of b. Then

$$T(\hat{b}-b)^2 = S_T \left(\frac{1}{T} \sum_{t=2}^{T-2} y_{t-1}^{(1)2} \right)^{-2},$$

$$\text{where } S_T = \frac{1}{T} \sum_{t=2}^{T-2} \sum_{s=2}^{T-2} \eta_t \eta_s y_{t-1}^{(1)} y_{s-1}^{(1)}, \tag{3}$$

and the estimation of the variance of $\hat{b}$ requires the construction of an estimator of $E(S_T)$. This task is the main topic of the present paper. Popular approaches include the kernel-based estimators of Andrews (1991), Andrews and Monahan (1992) and Newey and West (1987). The new approach proposed here is related to, but distinct from, the approach taken by West (1997). Let

$$g_t = \varepsilon_{t+1}((\beta+1) y_{t-1}^{(1)} + y_{t-2}^{(1)})$$

---

[1] As is well-known, and as Section 2 shows, an h-step-ahead prediction regression has an error that follows an MA($h-1$) process.