Diverging Tests of Equal Predictive Ability

Michael W. McCracken*

March 13, 2020

Abstract

We investigate claims made in Giacomini and White (2006) and Diebold (2015) regarding the asymptotic normality of a test of equal predictive ability. A counterexample is provided in which, instead, the test statistic diverges with probability one under the null.

JEL Nos.: C12, C52, C53

Keywords: prediction, out-of-sample, inference

*Research Division; Federal Reserve Bank of St. Louis; P.O. Box 442; St. Louis, MO 63166; michael.w.mccracken@stls.frb.org. We are grateful to Ken West, Tatevik Sekhposyan, Rafaela Giacomini, Minchul Shin, the Editor, and three anonymous referees for helpful comments. Additional thanks for comments from participants at the 2019 International Symposium of Forecasting, the 2019 International Association for Applied Econometrics, and the 2019 Federal Reserve System Meeting on Econometrics. The views expressed herein do not reflect the official positions of the Federal Reserve Bank of St. Louis or the Federal Reserve System.
1 Introduction

We describe a data generating process and forecasting exercise that yields a Diebold-Mariano (1995) test statistic \( \Theta_T \) that diverges under the null of equal predictive ability. We do so to address two claims made in the literature on tests of this null. The first is made in Giacomini and White (2006). The authors claim their Theorem 4 applies when model parameters are estimated using a fixed and finite window of observations and, subsequently, \( \Theta_T \) is asymptotically standard normal under the null. The second is made in Diebold (2015). The author claims that, as long as the loss differentials are covariance stationary, \( \Theta_T \) must be asymptotically standard normal under the null. We show with a counterexample that both claims are incorrect.

Consider an application in which the accuracy of a no-change point forecast of a scalar \( y_{t+1} \) is compared with the accuracy of a forecast based on a location model that is estimated using a fixed and finite window of observations \( R < \infty \). That is, \( \hat{y}_{1,t+1} = 0 \) while \( \hat{y}_{2,t+1} = \tilde{y}_t \) where \( \tilde{y}_t = \tilde{y}_R = R^{-1} \sum_{s=1}^{R} y_s \) for all forecast origins \( t = R, \ldots, R + P - 1 = T - 1 \). Under quadratic loss, the loss differential takes the form \( \hat{d}_{t+1} = (y_{t+1} - 0)^2 - (y_{t+1} - \tilde{y}_t)^2 \). Straightforward algebra reveals that if \( y_t = \mu + \varepsilon_t \) with \( \varepsilon_t \sim i.i.d. N(0, \sigma^2) \), the null of equal predictive ability \( E\hat{d}_{t+1} = 0 \) holds for all \( t \) when \( \mu = \sigma / \sqrt{R} \).

To test the null hypothesis we use the statistic \( \Theta_T = P^{-1/2} \sum_{t=R}^{T-1} \hat{d}_{t+1} / \hat{\omega} \), where \( \hat{\omega}^2 = P^{-1} \sum_{t=R}^{T-1} \hat{d}_{t+1}^2 \). We are interested in the behavior of this statistic as \( P \) diverges. Rearranging terms we find that

\[
\Theta_T = \frac{2(P^{-1/2} \sum_{t=R}^{T-1} \hat{d}_{t+1}) \tilde{y}_R + P^{1/2}(\mu^2 - (R^{-1} \sum_{s=1}^{R} \varepsilon_s)^2)}{\sqrt{4\tilde{y}_R^2(P^{-1} \sum_{t=R}^{T-1} \varepsilon_t^2) + 4\tilde{y}_R^2(2\mu - \tilde{y}_R)(P^{-1} \sum_{t=R}^{T-1} \varepsilon_t) + (\mu^2 - (R^{-1} \sum_{s=1}^{R} \varepsilon_s)^2)^2}}.
\]

As \( P \) diverges, the denominator is \( O_p(1) \). In contrast, while the first part of the numerator is \( O_p(1) \), the second part diverges to \( \pm \infty \) with probability one. Together this implies that \( \Theta_T \) is not asymptotically standard normal under the null and, in fact, diverges with probability one. Unreported simulations reinforce the analytical example: for reasonable values of \( R \) and \( P \), the test exhibits severe size distortions.

The root cause of the problem is not whether the loss differentials are covariance stationary. The problem is that the fixed and finite estimation window implies loss differentials that do not exhibit short memory. In the example, the loss differentials are covariance stationary but the autocovariances never die out.

Although the example suffices to counter the claims made in Giacomini and White (2006) and Diebold (2015), it does not necessarily overturn any existing applications that use \( \Theta_T \) to
conduct a test of equal predictive ability. In many such applications, model parameters are estimated using a rolling window of observations, in which case Theorem 4 of Giacomini and White (2006) remains valid. Nevertheless, the potential for future applications in which the loss differential is covariance stationary, and yet $\Theta_T$ is not asymptotically normal, remains.

In particular, it is easy to find extensions and applications of Giacomini and White (2006) that continue to claim the fixed estimation scheme is a viable option. An early example includes Giacomini and Komunjer (2005), but later examples can be found in our own work (Clark and McCracken, 2013) as well as very recent work by Rossi and Sekhposyan (2019).

References


