Econometric Reviews ›
Volume 33, 2014 - Issue 5-6: Special Issue in Honor of Les Godfrey

Original Articles

Robustify Financial Time Series Forecasting with Bagging

Sainan Jin, Liangjun Su & Aman Ullah

Pages 575-605 | Published online: 27 Nov 2013

Cite this article https://doi.org/10.1080/07474938.2013.825142 Check for updates

Abstra

In this pa ... ng as a
forecast ... models. The
revised ... data and
can be u ... red forecast
error wh ... show that
the ... ahead linear
forecast ... he in-
sample ... based on
misspec ... arametric
models, ... of out-of-
sample ... ables
suggeste ... . We find
that, co... stock return

forecasts may beat other predictor variables in the literature when we apply traditional one-step linear forecast and the nonparametric forecasting methods. However, when using the bagging method or its revised version, which help to improve the mean squared forecast error for “unstable” predictors, the predictive variables have a better forecasting power than the historical mean.

Keywords:

Bagging | Combined forecasts | Nonparametric models | Predictability | Time series

JEL Classification:

C14 | C53

ACKNOWLEDGMENTS

The authors are grateful to a referee for useful comments and suggestions. They are also thankful to Tae-Hwy Lee and Yundong Tu for discussions on the subject matter of this paper. The first author gratefully acknowledges financial support from the SMU research grant (Grant number: C244/MSS10E006).

Notes

In practi ... nd employ a weighte ... weight function

We a ... d −0.95. The resu ... tively. To save spa

We try d ... ry different out-of-sa ... similar and not repo

As suggested by one referee, The choice of equal weights is optimal in the iid case but not necessarily with dependent series like the ones considered in the paper. We also try to compute the weights by using Bayesian model averaging (BMA) technique as introduced in Lee and Yang (2006). The BMA gives a large weight to the bth bootstrap predictor at each period t when it has forecasted well over the past k periods and a small weight to the predictor at period t when it forecasted poorly over the past k periods. We have set k = 1, 5 and R. The results are similar to those based on equal weights and not reported here for brevity.

As one referee remarks, the choice of h affects the results and may not be optimal as chosen. We are dealing with dependent series and cross validation methods require blocking here too. We follow Hart and Vieu (1990) and set different leave-out sequences to take care of the dependence structure of the time series. We set $l_n = 0, 1, 2, 3, 4, 5$ as in Hart and Vieu (1990), where $l_n = 0$ corresponds to the ordinary leave-one-out cross validation, $l_n > 0$ corresponds to leave $2l_n + 1$ observations out, and the leave-out sequence is $\{X_j\}$ with $|j - t| \leq l_n$. The results for $l_n > 0$ are similar to those based on the usual leave-one-out least squares cross validation and thus not reported here.

We also try to choose the bandwidth by the “rule of thumb”: $h_l = c_0 s_l n^{-1/(4+q)}$, where s_l stands for the sample standard deviations of $X_{it,l}$, the lth regressor in X_{it}. We set $c_0 = 0.5, 1$, and 2 to examine the sensitivity of our test to the choice of bandwidth. It turns o ... are robust to differe ... local constant

Note: In ... E averaged over 200 ... pectively; the out-o ... f bootstrap resampl

Note: ... 5 M12. The benchm

Note: Sa ... 5 M12. The benchm

Note: Sa ... 5 M12. The benchm

As we discussed in the previous section, the one-step-ahead local constant and local linear predictors are sensitive to the bandwidth choice. If we try to choose the bandwidth by the “rule of thumb”: $h_l = c_0 s_l n^{-1/(4+q)}$, and set different values of $c_0 = 0.5$, 1, and 2, the results are quite different.

The calculation is based on Eq. (13) in Campbell and Thompson (2008). For illustration, the paper considered an investor with a single-period horizon and mean-variance preference and calculated the expected excess return when the investor observes the predicting variable and the expected excess return when the investor does not observe the predicting variable. The difference between these two expected excess returns is , where γ is the coefficient of relative risk aversion, and S is the unconditional Sharpe ratio of the risky asset.