## Econometric Reviews >

Volume 29, 2010 - Issue 5-6: The Link Between Statistical Learning Theory and Econometrics: Applications in Econometrics, Finance, and Marketing

65348
0
Views CrossRef citations to date Altmetric
Original Articles

# Bagging or Combining (or Both)? An Analysis Based on Forecasting U.S. Employment Growth 

David E. Rapach D\& Jack K. Strauss
Pages 511-533 | Published online: 15 Sep 2010
6S Cite this article $\boldsymbol{\sigma}$ https://doi.org/10.1080/07474938.2010.481550
Sample our
Mathematics \& Statistics
Journals
$\gg$ sign in here to start your access
to the latest two volumes for 14 days

自 Full Article Figures \& data E Beferences Citations Metrics
Reprints \& Permissions
Read this article

## Abstract

Forecasting a macroeconomic variable is challenging in an environment with many potential predictors whose predictive ability can vary over time. We compare two approaches to forecasting U.S. employment growth in this type of environment. The first approach applies bootstrap aggregating (bagging) to a general-to-specific

## procedu <br> About Cookies On This Site

 predicto from 30 square bagging incorpor

We and our partners use cookies to enhance your website experience, learn how our site is used, offer personalised features, measure the effectiveness of our services, and tailor content and ads to your interests while you navigate on the web or interact with us across devices. You can choose to accept all of these cookies or only essential cookies. To learn more or manage your preferences, click "Settings". For further information about the data we collect from you, please see our Privacy Policy.

principal components often leads to further gains in forecast accuracy.
r

[^0]
## ACKNOWLEDGMENTS

The authors thank two anonymous referees for very helpful comments. The authors also gratefully acknowledge financial support from the Simon Center for Regional Forecasting at Saint Louis University. The results reported in this article were generated Forecasting at Saint Louis University. The results reported in
using GAUSS 6.1. The GAUSS program files are available at http://pages.slu.edu/faculty/rapachde/ Research.htm.

## Notes

Lee and Yang (2006) use bagging techniques to develop binary and quartile forecasts of financial variables.

See Timmermann (2006) for a comprehensive review of forecast combining methods.
For example, employment growth-in the context of the so-called "jobless" recovery from the 2001 U.S. recession-received considerable attention during the 2004 presidential election, arguably more than any other economic variable. Employment growth is also viewed as a key indicator of labor market activity and crucial in Federal Reserve formulations of interest rate policy, and hence forecasts of employment growth For example, employment growth-in the context of the
from the 2001 U.S. recession-received considerable at
presidential election, arguably more than any other eco
growth is also viewed as a key indicator of labor market
Reserve formulations of interest rate policy, and hence
are important in ruminations about Fed policy actions.

Note that Innuo and Kilian (One) ale rnmnara hanging fnraracte af 11 c inflation to
combine About Cookies On This Site
on
h on a lag


We and our partners use cookies to enhance your website experience, learn how our site is used, offer personalised features, measure the effectiveness of our services, and tailor content and ads to your interests while you navigate on the web or interact with us across devices. You can choose to accept all of these cookies or only essential cookies. To learn more or manage your preferences, click
"Settings". For further information about the data we collect from you, please see our Privacy Policy. from you, please see our Privacy Policy. le電

位
$\square$
$+$$-$

ER10yment



$\qquad$


#### Abstract

$\checkmark$


Inoue and Kilian (2008) consider a range of critical values. We obtain similar results using other conventional critical values such as 1.96.

Following Inoue and Kilian (2008), we use $m=h$ and $B=100$.
"Recursive" indicates that the forecasts are generated using an expanding estimation window. The out-of-sample forecasts are "simulated"-as opposed to "real-time"because they are based on revised data and not on the data actually available at the time of forecast formation. Real-time forecasting exercises are not feasible in the present article, as data vintages are not readily available for all of the variables we consider in our forecasting exercise. We follow much of the macroeconomic forecasting
literature, including Stock and Watson (1999, 2003, 2004), in analyzing simulated outconsider in our forecasting exercise. We follow much of the macroeconomic forecasting
literature, including Stock and Watson (1999, 2003, 2004), in analyzing simulated out-of-sample forecasts.

Our results are not very sensitive to the maximum lag lengths.
Observe that the number of clusters serves to define the size of the first cluster, as none of the other clusters are used in generating the forecast. The greater the number of clusters, the smaller the size of the first cluster.

Using the taxonomy in Huang and Lee (2007), all of the combining methods we consider are classified as "combination of forecasts," while the bagging forecastswhich are formed by including all of the variables in a single general model-are classified as "combination of information." Another combination of information procedure that could be used is the Stock and Watson (2002) diffusion index, which involves extracting principal components from the potential predictors themselves (instead of extracting principal components from the individual ARDL forecasts). We experimented with diffusion index forecasts and found that the PC combining method forecasts generally perform better. It would be interesting in future research to consider
the appr
$(2006) \mathrm{t}$
for estin
The
treatme

A word
$\qquad$

Bai and Ng About Cookies On This Site
We and our partners use cookies to enhance your website
experience, learn how our site is used, offer personalised
features, measure the effectiveness of our services, and About Cookies On This Site
We and our partners use cookies to enhance your website
experience, learn how our site is used, offer personalised
features, measure the effectiveness of our services, and About Cookies On This Site
We and our partners use cookies to enhance your website
experience, learn how our site is used, offer personalised
features, measure the effectiveness of our services, and About Cookies On This Site
We and our partners use cookies to enhance your website
experience, learn how our site is used, offer personalised
features, measure the effectiveness of our services, and tailor content and ads to your interests while you navigate on the web or interact with us across devices. You can choose to accept all of these cookies or only essential cookies. To learn more or manage your preferences, click "Settings". For further information about the data we collect from you, please see our Privacy Policy.

Essential Onl

Settings a textbook
Accept All e as proxies
Essential Onl
Settings a model.
a textbook -
research demonstrates that a number of issues-such as the size of the in-sample period relative to the out-of-sample period, type of estimation window (for example, fixed, rolling, or recursive), and whether the models are nested or non-nested-can affect the asymptotic distribution of the test statistics; see Corradi and Swanson (2006) for an informative review of these issues. We recognize that, strictly speaking, all of the conditions required for the validity of the asymptotic distribution may not be met in our applications, so that inferences based on the MHLN $n$ statistic serve as a rough guide to statistical significance.

Vendor performance is an index that measures how quickly companies receive deliveries from their suppliers. An increase in the index means that it is taking longer for companies to receive deliveries.

Notes: The entries for the AR benchmark model report the MSFE; the other entries report the MSFE for the ARDL model forecasts indicated on the left to the MSFE for the AR benchmark model forecasts.

Notes to Tables 2-5: The MSFE ratio reports the ratio of the MSFE for the BA model or combination forecasts indicated on the left to the MSFE for the AR benchmark model forecasts. $\mathrm{H}_{0}$ : BA encompasses $\mathrm{CB}\left(\mathrm{H}_{0}: C B\right.$ encompasses BA$)$ corresponds a test of the null hypothesis that the BA model (combination) forecasts encompass the combination (BA model) forecasts against the one-sided, upper-tail alternative hypothesis that the BA model (combination) forecasts do not encompass the combination (BA model) forecasts; () is the OLS estimate of the weight on the combination (BA model) forecast in the optimal convex combination forecast given by (7); MHLN ${ }_{n}$ is the test statistic corresponding to the null hypothesis; 0.00 indicates less than 0.005 ; ${ }^{\dagger}, *, * *$ indicate significance at the $10 \%, 5 \%$, and $1 \%$ levels, respectively.

Nevertheless, all of the MSFE ratios for the combination forecasts are below unity, indicatin the MSF ARDL m for thes

## About Cookies On This Site

We and our partners use cookies to enhance your website experience, learn how our site is used, offer personalised features, measure the effectiveness of our services, and tailor content and ads to your interests while you navigate on the web or interact with us across devices. You can choose to accept all of these cookies or only essential cookies. To learn more or manage your preferences, click oting that
for "Settings". For further information about the data we collect e individual from you, please see our Privacy Policy.

```
Essential Onl
```

Settings eriod at the the BA (PC combination) torecasts encompass the PL combination (BA model) torecasts
according to the results in Tables 2-5 using a 10\% significance level. The entries for the average rows report the ratio of the MSFE for a forecast formed as a simple average of the BA model and PC combination forecasts to the MSFE for the AR benchmark model forecasts.

For the cases where both of the MHLN ${ }_{\mathrm{h}}$ statistics are significant, the weights on the BA and PC forecasts in Tables 2-5 are close to 0.50, so taking the mean of the two forecasts is reasonable. This procedure also avoids having to estimate the weights, making it easier to implement in practice.

We examined the robustness of our results along a number of dimensions and obtained similar results. For example, the results are very similar when we use the AIC instead of the SIC to select the lag lengths in (1) and (2). We also computed combination forecasts for a set of potential predictors that excludes manufacturing capital orders and manufacturing and trade sales, two variables that are available with a one-month lag relative to the other potential predictors (and so are not "coincident" with the other predictors). We again obtain very similar results. The complete results for these robustness checks are available at http://pages.slu.edu/faculty/rapachde/Research.htm.

Related research ©

Cited by
48

## About Cookies On This Site

We and our partners use cookies to enhance your website experience, learn how our site is used, offer personalised features, measure the effectiveness of our services, and
tailor content and ads to your interests while you navigate on the web or interact with us across devices. You can choose to accept all of these cookies or only essential cookies. To learn more or manage your preferences, click

## Information for

## Authors

R\&D professionals
Editors

## Librarians

## Societies

## Opportunities

Reprints and e-prints
Advertising solutions
Accelerated publication
Corporate access solutions

Open access
Overview
Open journals
Open Select
Dove Medical Press
F1000Research
Help and information
Help and contact
Newsroom
All journals
Books

Keep up to date
Register to receive personalised research and resources by email


Sign me up
f $x$ in
用

## About Cookies On This Site

Accept All
We and our partners use cookies to enhance your website experience, learn how our site is used, offer personalised features, measure the effectiveness of our services, and tailor content and ads to your interests while you navigate on the web or interact with us across devices. You can choose to accept all of these cookies or only essential cookies. To learn more or manage your preferences, click


[^0]:    - 

