Improving Forecast Performance with Reduced Parameter, Large Order AR Models

Octavio A. Ramirez and Jeffrey H. Dorfman
Department of Agricultural & Applied Economics
The University of Georgia
Athens, GA 30602 USA


Copyright 2010 by Octavio A. Ramirez and Jeffrey H. Dorfman. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.
We examine the role of time series model specification on forecast performance, focusing on the impact of the maximum model order and on full versus reduced parameter models. Forecasters should reconsider their practices and search over a wider range of model specifications, including longer lags and reduced parameter models.

Braun and Mittnik (1993) establish that misspecified ARMA models can cause significant errors in estimation of impulse response functions and variance decompositions. In fact, they find that an under-specified model is generally worse than using an AR model when the true data generating process is an ARMA as long as the AR model has a sufficiently high lag order. Their only caveat to greater use of higher order models is the increased sample error and the limitation in the number of parameters to be estimated versus the number of observations that are available for estimation.

A recent Bayesian method in a spirit similar to what we propose here is found in Jochmann, Koop, and Strachan (2010) who utilize a Bayesian stochastic search variable selection technique to effectively zero out numerous parameters in a VAR model. They do not, however, examine long lags, but restrict their model to a maximum of four lags.

To test the just discussed findings against empirical experience, we conducted an expansive experiment measuring out-of-sample forecast performance for nine commonly-studied monthly data series across several model specification strategies. We confined our experiments to AR models since the very long maximum lag considered in our largest models should capture any ignored MA component. Models were estimated with maximum lag lengths of 12, 24, and 60. Reduced parameter models were obtained using a simple stepwise procedure to remove parameters. A probability (p-value) threshold of 0.075 was set for a parameter to stay in the model for the AR(12). Then, the 12 additional coefficients needed to complete an AR(24) were subjected to the same step-wise procedure and probability ceiling. Finally, the 36 extra lags required for an AR(60) were tested using a 0.025 threshold. The rationale for this procedure is to give the lower lag parameters a higher probability of staying the model.

The forecasting performance results are summarized in Table 2. Averaged over the nine different series, the maximum lag lengths, reduced models yield moderate (≈4%) out-of-sample forecast MSE for the AR(12) specifications, substantial (>7%) gains for AR(24) models, and tremendous (>25%) benefits for AR(60) models.

Braun and Mittnik (1993) pointed out the potential benefits of forecasting time series with autoregressive models of high order, even when the true data generating process included a moving average component. They added a caveat about the sampling error and degree of freedom limitations associated with a model having a large number of parameters. Here, we propose a way around that caveat by employing reduced models in which many parameters are eliminated using a simple stepwise procedure. For a diverse set of nine data series, we find meaningful reductions in MSE for out-of-sample forecasting associated with using the reduced models in general, and the high order reduced models in particular. We also confirmed that model specification is sensitive to the forecasting horizon as we find some difference in model rankings as the forecast horizon changes. Our recommendation, based on this reasonably broad experiment, is that if the goal is forecasting performance, one should use reduced models and very long maximum lag lengths in model specification searches.

The middle section of Table 2 highlights the results for the individual cases where the reduced models performed the worst and the best. For the longest out-of-sample forecast period being examined (OS=120), the reduced model exhibits a lower MSE than the corresponding full model for all nine data series. So the reduced models are not only better on average, but outperform the full models across a wide range of series. Also note that the reduced AR(60) model substantially outperforms both the full and reduced smaller order models (12F, 24F, 12R, and 24R). This suggests that, if using reduced models, researchers should set the maximum lag length as high as reasonably can be justified given the sample size. In our experiments, sample sizes are at least 6 times the number of parameters in the full models.