

In-sample Inference and Forecasting in Misspecified Factor Models

Marine Carrasco*
Université de Montréal
CIREQ, CIRANO

Barbara Rossi†
ICREA-Univ. Pompeu Fabra,
Barcelona GSE and CREI

December 9, 2015

Abstract

This paper considers in-sample prediction and out-of-sample forecasting in regressions with many exogenous predictors. We consider four dimension reduction devices: principal components, Ridge, Landweber Fridman, and Partial Least Squares. We derive the rate of convergence of the prediction error for two representative models: a mildly sparse model and an approximate factor model. The theory is developed for a large cross-section and a large time-series. We also propose data-driven selection methods based on cross-validation and establish their optimality. Monte Carlo simulations and an empirical application to forecasting inflation and output growth in the U.S. show that data-reduction methods outperform conventional methods in several relevant settings, and might effectively guard against instabilities in predictors' forecasting ability.

Keywords: Forecasting, regularization methods, factor models, Ridge, partial least squares, principal components, sparsity, large datasets, variable selection, GDP forecasts, inflation forecasts.

J.E.L. Codes: C22, C52, C53.¹

*Address: Université de Montréal, Département de Sciences Economiques, CP 6128, succ Centre Ville, Montreal, QC H3C3J7, Canada. E-mail: marine.carrasco@umontreal.ca.

†Address: Universitat Pompeu Fabra, C/ Ramon Trias Fargas 25-27, 08005 Barcelona, Spain. E-mail: barbara.rossi@upf.edu