Available online at www.sciencedirect.com
ScienceDirect
International Journal of Forecasting 27 (2011) 496–511
www.elsevier.com/locate/ijforecast

Forecasting national activity using lots of international predictors: An application to New Zealand

Sandra Eickmeier[a,*], Tim Ng[b]

[a] *Deutsche Bundesbank, Germany*
[b] *Reserve Bank of New Zealand, New Zealand*

Abstract

We assess the marginal predictive content of a large international dataset for forecasting GDP in New Zealand, an archetypal small open economy. We apply "data-rich" factor and shrinkage methods to efficiently handle hundreds of predictor series from many countries. The methods covered are principal components, targeted predictors, weighted principal components, partial least squares, elastic net and ridge regression. We find that exploiting a large international dataset can improve forecasts relative to data-rich approaches based on a large national dataset only, and also relative to more traditional approaches based on small datasets. This is in spite of New Zealand's business and consumer confidence and expectations data capturing a substantial proportion of the predictive information in the international data. The largest forecasting accuracy gains from including international predictors are at longer forecast horizons. The forecasting performance achievable with the data-rich methods differs widely, with shrinkage methods and partial least squares performing best in handling the international data.

Keywords: Forecasting; Factor models; Shrinkage methods; Principal components; Targeted predictors; Weighted principal components; Partial least squares; Ridge regression; Elastic net; International business cycles

1. Introduction

Exploiting information from large datasets has been shown to improve forecasts. Certain "data-rich" methods such as principal components (PC) are now widely used in economic forecasting and analysis by both academics and practitioners (Eickmeier & Ziegler, 2008; Stock & Watson, 2002a,b, 2004). In this paper, we explore PC and other lesser-known data-rich methods using the case of New Zealand GDP growth forecasting. We focus in particular on whether data-rich methods are a useful way of effectively handling the large amount (hundreds of series) and diversity (many countries and types) of available international data for forecasting purposes.

We use New Zealand as an example because it is highly exposed to the international economy.

* Corresponding address: Economic Research Centre, Deutsche Bundesbank, Wilhelm-Epstein-Str. 14, Frankfurt/M., Germany.
E-mail address: s_eickmeier@gmx.net (S. Eickmeier).

0169-2070/$ - see front matter
doi:10.1016/j.ijforecast.2009.10.011

Table 1
International dataset — number of series in each class.

	North Amer.	Europe	Asia	Australia	World	Sum
Real	50	41	95	20	–	206
Prices	28	28	41	11	–	108
Monetary + financial	23	34	62	11	–	130
Commodity prices	–	–	–	–	38	38
Sum	101	103	198	42	38	482

New Zealand has many trading partners across the world, with a potentially very large set of relevant predictor variables. Structural vector-autoregressive (SVAR) analyses of New Zealand's macroeconomic fluctuations (e.g. Buckle, Kim, Kirkham, McLellan, & Sharma, 2007; Dungey & Fry, 2009) find that international shocks contribute substantially to New Zealand GDP fluctuations. Furthermore, there are many potential channels for these shocks. In addition to a large proportion of New Zealand's real demand being sourced externally, New Zealand is a "commodity economy" (Chen & Rogoff, 2003), exposed to the international commodity price cycle, and is highly dependent on foreign funding (Bedford, 2008).

Most data-rich applications use datasets that include some international variables, but not very many relative to the number available. Surprisingly few explicitly assess the marginal predictive content of international data. We make three contributions.

First, we look at the potential improvement in forecasting accuracy when a large number of international variables are added to a predictor dataset containing national (in this case, New Zealand) data only. We also look at the forecasting performances of models based on international data only. Second, we compare different data-rich approaches to capturing the predictive information in our international dataset. Third, we assess the extent to which New Zealand survey data capture international information.

Certain data-rich approaches have been shown to be useful in the context of essentially single-country datasets, but the dimensionality of the problem is an order of magnitude greater in the case of international datasets. The dataset we use here contains several hundred macroeconomic time series from 12 of New Zealand's major trading partner economies, which together account for about three-quarters of New Zealand's total trade in goods (exports and imports). These economies represent the three major world economic regions – North America and Europe (which each account for about 15% of goods trades) and Asia-Pacific (about half).[1] The dataset is also roughly balanced by macroeconomic type, i.e. real activity, prices, monetary and financial indicators, and international commodity prices (Table 1).

Data-rich methods deal with large dimensional datasets by shrinking the variance of the parameters in estimated equations, or by summarising the information contained in many data series into a few common factors, or using both methods. These approaches make forecasting using many predictors feasible. In addition to the familiar PC, we also look at other data-rich methods which address some of the known limitations of PC when applied to forecasting in practice. Variants of PC include targeted predictors (TP) and weighted principal components (WPC). We also look at partial least squares (PLS), elastic net (EN) and ridge regression (RR; a special case of EN).

We compare the forecasting accuracy using these data-rich approaches to the international data with that from using trade-weighted aggregates. Trade-weighting is a much simpler approach to summarising international data and is often used by practitioners due to its simplicity and intuitive appeal. Obvious advantages of the data-rich methods, though, are that they can easily cope with a much greater quantity and diversity of data; also, related to this, they can (implicitly) capture international transmission channels beyond direct trade linkages, such as indirect trade, financial markets, commodity prices, and confidence; and they (generally) derive weights from an explicit statistical optimisation problem, which is not usually the case for trade-weights.

[1] North America is represented by the US and Canada; Europe by the euro area (measured as a single economy) and the UK; and Asia-Pacific by Australia, Japan, China, Singapore, Korea, Malaysia, Taiwan and Hong Kong.