International Journal of Forecasting 32 (2016) 680–694

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Aggregate versus disaggregate information in dynamic factor models

Rocio Alvarez [a], Maximo Camacho [b,*], Gabriel Perez-Quiros [c,d]

[a] Universidad de Alicante, Spain
[b] Universidad de Murcia, Spain
[c] Banco de España, Spain
[d] CEPR, Spain

ARTICLE INFO

Keywords:
Business cycles
Output growth
Time series

ABSTRACT

We examine the finite-sample performances of dynamic factor models that use either aggregate or disaggregate data, where the latter rely on finer disaggregations of the headline concepts of a small set of economic categories. Our Monte Carlo analysis reveals that the use of the series with the largest averaged within-category correlations outperforms the use of disaggregate data for factor estimation and forecasting in several cases. This occurs for high levels of cross-correlation across the idiosyncratic errors of series that belong to the same category, for oversampled categories, and especially for high levels of persistence in either the common factor or the idiosyncratic errors. However, the forecasting gains are reduced considerably when the target series are persistent. This could potentially explain why there is no clear ranking between the aggregate and disaggregate approaches when using the constituent balanced panel of the Stock-Watson factor model, which classifies the US data into 13 economic categories.

1. Introduction

Empirical macroeconomists face peculiar data structures. Recent advances in information technologies mean that, increasingly, data are becoming available with an unprecedented degree of disaggregation. However, in practice, the data sets typically rely on finer disaggregations of the headline concepts of a small number of broad economic categories. For example, the data sets usually contain sectoral splits for industrial production and labor, detailed information on prices, and disaggregations of surveys into sectors.

Factor models have received a growing amount of attention for dealing with these large data sets, due to their ability to summarize the information contained in lots of series in a small number of unobserved common factors that may capture the comovements across the series, which are usually used to forecast some key economic aggregates. Factor models generally rely on different sophistications of the work of Forni and Reichlin (1996) and Stock and Watson (2002a). Recent examples include studies by Angelini, Camba-Mendez, Giannone, Reichlin, and R ünstler (2011), Banbura and Modugno (2014), Forni, Hallin, Lippi, and Reichlin (2005) and Giannone, Reichlin, and Small (2008). The Chicago Fed National Activity Index (CFNAI), released by the Federal Reserve Bank of Chicago, is an index that is computed by following this approach.

* Correspondence to: Universidad de Murcia, Facultad de Economia y Empresa, Departamento de Metodos Cuantitativos para la Economiay la Empresa, 30100, Murcia, Spain.
E-mail addresses: rocio@merlin.fae.ua.es (R. Alvarez), mcamacho@um.es (M. Camacho), gabriel.perez@bde.es (G. Perez-Quiros).

http://dx.doi.org/10.1016/j.ijforecast.2015.10.006
0169-2070/

Although the natural choice would be to use all of the sectoral information in the factor models, extracting information from such large data sets could be suboptimal. Boivin and Ng (2006) for the US and Caggiano, Kapetanios, and Labhard (2011) for some euro area countries show that including sectoral information could lead to model misspecification in small samples, since it increases the idiosyncratic cross-correlation. Poncela and Ruiz (2015) show that, when model parameters have to be estimated, the parameter and total uncertainties could increase when the number of indicators increases. Banbura and Runstler (2011) show that forecast weights are concentrated among a relatively small set of euro area indicators. Banbura, Giannone, and Reichlin (2011) and Banbura and Modugno (2014) find that including disaggregated information does not improve the accuracy of the euro area forecasts.

Our contributions to this literature are twofold. Our first contribution is to design a Monte Carlo experiment that allows us to document the conditions under which adding more disaggregated data, which rely on the particular structure described above, could be undesirable. Within this context, we develop simulations for evaluating the precision of estimating the space spanned by the common factors, as well as of forecasting a target time series under two empirical scenarios. In the *disaggregated* scenario, the factors are estimated and the forecasts computed from a large data set, which is generated by including additional series in each of a small set of broad categories, under the assumption that the additional series in each category are finer disaggregations of one broad indicator with which they could be correlated. In the *aggregated* scenario, the factors and the forecasts come from a factor model that uses only a small number of time series from the large data set. In this case, the subset of time series is selected from each category using several statistical criteria.

Our Monte Carlo experiment has been designed to investigate the effects on these two scenarios of across-category and within-category correlations, serial correlations of factors and idiosyncratic components, differing sample sizes, oversampled categories, and ragged edges. Although there is no unambiguous evidence in favor of using either aggregate or disaggregate data, our results show the cases in which using disaggregate information is advisable and those in which it is not. In particular, we find that the use of aggregate information outperforms the use of disaggregate information in factor models when the cross-correlation across series in the same category is high, when the factor is persistent, when some categories are overrepresented, and when the serial correlation of the idiosyncratic errors is high. However, we find that these gains in forecasting are reduced substantially when the target series are persistent. In addition, we show that our results are not affected qualitatively by the issue of missing data, which is typical of real-time applications, since data are released in a non-synchronous manner and with different reporting lags.

Our second contribution has to do with the fact that, in spite of the empirical evidence that the use of large cross sections could cause the performances of factor models to deteriorate, the selection of the data set from which to extract the factors and produce forecasts has still not been addressed fully. To address the problem of selecting representative indicators from the small set of separate economic categories, we propose the following criterion: select one representative from each category, the time series with the largest averaged correlation with the series in the same category. We compare our selection criterion with that used by Boivin and Ng (2006), who select the variables by removing the time series with highest correlation across the idiosyncratic components, and with a random selection of one series from each category. We show that the forecasts computed based on our selection process outperform those computed using these two alternatives.

The empirical performances of aggregate versus disaggregate factor models are examined using the balanced set of US monthly macroeconomic indicators introduced by Stock and Watson (2002b). These authors classified the time series included in the data set into 13 economic categories, such as real output, prices, and employment. In an out-of-sample exercise, we examine the performance of a factor model that uses the disaggregate information from the complete set of indicators with that of a factor model that uses only the indicators that exhibit the highest averaged correlations with the series in the same category. For this purpose, we analyze the accuracy for forecasting four key macroeconomic variables at different short-term horizons. The empirical results obtained from actual data are in accord with those obtained from the generated data. For highly persistent target series, we do not find any substantial differences between aggregate and disaggregate factor models. For less persistent target series, the factor model that uses aggregate information yields satisfactory forecasting results with respect to those of the factor model that uses disaggregate information, which agrees with the findings of Banbura et al. (2011) and Banbura and Modugno (2014). Remarkably, our variable selection method clearly outperforms both the statistical method suggested by Boivin and Ng (2006) and the random method.

The empirical analysis helps to show an additional advantage of our variable selection criterion. We find that picking the series with the highest average within-category correlation leads to economically meaningful sets of representative indicators, in the sense that the indicators picked from each category typically coincide with the aggregate headline concepts (such as total industrial production). In contrast, some of the variables selected using the alternative procedures are finer disaggregations of these headline concepts, which complicates the interpretation of the empirical economic applications. In addition, the variables selected using the Boivin and Ng (2006) method generally belong to only a reduced number of economic categories, with some key categories remaining unrepresented. This could be problematic, since some of these categories are routinely monitored by the users of factor models and it is usually important to include them in the analysis, not only to eventually improve the forecasts, but also to assist in interpreting the forecasts.

These results could help in formalizing the variable selection of factor models that use aggregate information from small sets of indicators. Although they typically focus on different enlargements of the four-variable single-index