

# Aggregate versus disaggregate information in dynamic factor models\*

Rocio Alvarez	Maximo Camacho <sup>†</sup>	Gabriel Perez-Quiros
Universidad de Alicante	Universidad de Murcia	Banco de España and CEPR
rocio@merlin.fae.ua.es	mcamacho@um.es	gabriel.perez@bde.es

## Abstract

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

**Keywords:** Business Cycles, Output Growth, Time Series.

**JEL Classification:** E32, C22, E27.

---

\*Camacho acknowledges financial support from MICINN (ECO2010-19830 and ECO2013-45698). We thank the editor, the associate editor, and the two referees for their comments and suggestions. The Appendix is available at <http://www.um.es/econometria/Maximo/>. The views in this paper are those of the authors and do not represent the views of Bank of Spain or the EuroSystem.

<sup>†</sup>Corresponding Author: Universidad de Murcia, Facultad de Economía y Empresa, Departamento de Métodos Cuantitativos para la Economía y la Empresa, 30100, Murcia, Spain. E-mail: mcamacho@um.es.

# 1 Introduction

Empirical macroeconomists face a peculiar data structure. Due to recent advances in information technologies, data are becoming increasingly available with 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 sectorial splits for industrial production and labor, detailed information on prices and disaggregations of surveys into sectors.

To deal with these large data sets, factor models have received a growing attention for 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. Mainly, the factor models rely on different sophistications of the works of Forni and Reichlin (1996) and Stock and Watson (2002a). Recent examples are Forni, Hallin, Lippi and Reichlin (2005), Giannone, Reichlin and Small (2008), Angelini et al. (2011) and Banbura and Modugno (2014). The Chicago Fed National Activity Index (CFNAI), released by the Federal Reserve Bank of Chicago, is an index computed by following this approach.

Although the natural choice would be to use all the sectorial information in factor models, extracting information from such large data sets could be suboptimal. Boivin and Ng (2006) for the US and Caggiano, Kapetanios, and Labhard (2009) for some euro area countries show that including sectorial information could lead to model mis-specification in small samples since it increases the idiosyncratic cross-correlation. Poncela and Ruiz (2015) show that, when model parameters have to be estimated, 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 contribution to this literature is 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 the simulations to evaluate the precision in estimating the space spanned by the common factors as well as in forecasting a target time series under two empirical scenarios. In the *disaggregated* scenario, the factors are estimated and the forecasts are computed from a large data set, which is generated by including additional series in 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 by using several statistical criteria.

Our Monte Carlo experiment has been designed to focus on the effects on these two scenarios of across-category and within-category correlations, serial correlation of factors and idiosyncratic components, sample sizes, oversampled categories, and ragged edges. Although there is no unambiguous evidence in favor of using aggregate or disaggregate data, our results show the cases in which using disaggregate information is advisable and the cases in which it is not. In particular, we find that using aggregate information outperforms using disaggregate information in factor models when the cross-correlation across the series of 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 mitigate substantially when the target series are persistent. In addition, we show that our results are not qualitatively affected by the presence of missing data, which typically appear in 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 using large cross sections could deteriorate the performance of factor models, the choice of the data set from which to extract the factors and to perform the forecasts remains partly unaddressed. To address the problem of selecting the representative indicators from the small set of separate economic categories, we propose the following criterion: select one representative of each category, the time series with largest averaged correlation with

the series of 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 randomly picking one series from each category. We show that the forecasts computed from our selection process outperform the forecasts computed from these two alternatives.

The empirical performance of aggregate versus disaggregate factor models is examined by using the balanced set of US monthly macroeconomic indicators suggested by Stock and Watson (2002b). The time series included in the data set are classified by these authors 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 of the complete set of indicators versus a factor model that uses only the indicators that exhibit the highest averaged correlation with the series of the same category. For this purpose, we analyze the accuracy to forecast four key macroeconomic variables at different short-term horizons. The empirical results obtained from actual data are in concordance with those obtained from generated data. For highly persistent target series, we do not find 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 and Modugno (2014) and Banbura, Giannone and Reichlin (2011). Remarkably, our variable selection method clearly outperforms the statistical method suggested by Boivin and Ng (2006) and the random method.

The empirical analysis helps us 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). By contrast, some of variables selected from the alternative procedures are finer disaggregations of these headline concepts, which complicates the interpretation of empirical economic applications. In addition, the variables selected from the Boivin and Ng (2006) method generally belong to only a reduced

number of economic categories while some key categories become unrepresented. This could be problematic since some of these categories are routinely monitored by users of factor models and including them in the analysis is usually important not only to eventually improve the forecasts but also for interpreting the forecasts.

These results could help in formalizing the variable selection of those factor models that use aggregate information from small sets of indicators. Although they typically focus on different enlargements of the four-variable single-index factor model suggested by Stock and Watson (1991), the data selection is sometimes arbitrary and does not rely on statistical criteria. Examples are Mariano and Murasawa (2003), Nunes (2005), Aruoba, Diebold and Scotti (2009), Aruoba and Diebold (2010), and Camacho and Perez Quiros (2010). The ADS index, released by the Federal Reserve Bank of Philadelphia, also follows this approach.

This paper proceeds as follows. Section 2 presents the details of the simulation exercise and shows how the series of each category and their finer disaggregations are generated. Section 3 shows the main findings in the comparison of factor models that use aggregate and disaggregate information for different parameter's values. Section 4 describes our empirical application. Section 5 concludes and points out some lines of further research.

## 2 Designing the simulation study

In this section, we perform Monte Carlo simulations to assess the extent to which using all the disaggregate time series that belong to a small set of broad economic categories might deteriorate the performance of the dynamic factor models in finite samples.<sup>1</sup> For this purpose, the analysis accounts for different degrees of across-category and within-category correlation, for different degrees of serial correlation of factors and idiosyncratic components, for different sample sizes, for oversampled categories and for the case of dealing with ragged edges.

---

<sup>1</sup>Moench, Ng and Potter (2013) develop an interesting analysis by using dynamic hierarchical factor models. The comparison between these models and the factor models used in this paper is left for further research.

## 2.1 Dynamic factor models

Let  $y_t$  be a scalar stationary time series variable to be forecasted and let  $X_t = (X_{1t}, \dots, X_{Nt})'$ , with  $t = 1, \dots, T$ , be the observed  $N$  stationary time series which are candidate predictors of  $y_t$ . The baseline model to compute one-step-ahead predictions could be stated as

$$y_{t+1} = \pi_0 + \pi_1' X_t + \sum_{j=1}^p \pi_{2j} y_{t-j+1} + \epsilon_{yt+1}, \quad (1)$$

where  $\pi_1 = (\pi_{11}, \dots, \pi_{1N})'$ , and  $\epsilon_{yt+1}$  is a zero mean white noise.

Since estimating this expression becomes impractical as the number of predictors increases, it is standard to assume that each predictor  $X_{it}$  admits a factor structure:

$$X_{it} = \lambda_i' F_t + \xi_{it}, \quad (2)$$

for the  $i$ -th cross-section unit at time  $t$ ,  $i = 1, \dots, N$ ,  $\lambda_i = (\lambda_{i1}, \dots, \lambda_{ir})'$ , and  $t = 1, \dots, T$ . In this framework the  $r \times 1$  vector  $F_t$  contains the  $r$  common factors,  $\lambda_i$  the  $r$  factor loadings,  $\chi_{it} = \lambda_i' F_t$  the common components, and  $\xi_{it}$  the idiosyncratic errors. In vector notation the model can be written as

$$X_t = \Lambda F_t + \xi_t, \quad (3)$$

where  $\Lambda = (\lambda_{ij})$  is the  $N \times r$  matrix of factor loadings and  $\xi_t$  is the vector of  $N$  idiosyncratic shocks. We assume that the vectors  $F_t$  and  $\xi_t$  are cross-sectionally uncorrelated processes.<sup>2</sup> The dynamics of the common factors are supposed to follow autoregressive processes. Although it is very easy to generalize, let us assume that the factors follow a simple  $VAR(1)$  process

$$F_t = A F_{t-1} + u_t, \quad (4)$$

where  $A$  is the  $r \times r$  matrix of coefficients and  $u_t$  is serially uncorrelated, with  $E[u_t] = 0$  and  $E[u_t u_t'] = \Sigma_u$ . In addition,  $\xi_t$  is also assumed to follow a simple stationary  $VAR(1)$  process

$$\xi_t = C \xi_{t-1} + v_t, \quad (5)$$

---

<sup>2</sup>In this framework the common factor is supposed to generate most of the cross-correlation between the series of the data set  $\{X_{it}\}_{i=1}^N$ .

where  $v_t$  is serially uncorrelated, with  $E[v_t] = 0$  and  $E[v_t'v_t] = \Sigma_v$ .<sup>3</sup> Then, the target variable  $y_t$  can be forecasted through the common factors by using the expression<sup>4</sup>

$$y_{t+1} = \beta_0 + \beta' F_t + \sum_{j=1}^p \gamma_j y_{t-j+1} + e_{yt+1}. \quad (6)$$

## 2.2 Generating the data

To simplify the analysis, we assume in this section that the time series are generated from only one common factor.<sup>5</sup> To start with the simulations, we generate the common factor  $\{F_t\}_{t=1}^T$  by using expression (4). In this case,  $\{u_t\}_{t=1}^T$  are random numbers which are drawn from a normal distribution with zero mean and variances  $\Sigma_u = 1$ . To examine the dependence of the results on the persistence of the factor, we allow for different values for the parameter  $A = 0.1, 0.5, \text{ and } 0.75$ . In the empirical applications,  $F_t$  usually represents the “state of the economy” or the “business cycle”.

The next step is to generate 10 different categories of data. For this purpose, we generate  $\{\xi_{i,t}\}_{i,t=1}^{10,T}$  from (5), where the errors  $v_t = (v_{1t}, \dots, v_{10t})'$ , are random numbers which are drawn from a normal distribution with zero mean and covariance matrix  $\Sigma_v$ . To simplify simulations, the autoregressive coefficients matrix  $C$  is diagonal with two possible values,  $c = 0.1$  and  $c = 0.75$ , in all the elements of the main diagonal. In addition, to examine the effects of the across-category correlation, the covariance matrix  $\Sigma_v$  takes different values across the simulations. In particular, for a given value for the parameter

---

<sup>3</sup>Although assuming  $VAR(p)$  dynamics for the factors and the idiosyncratic components is straightforward, it would complicate notation.

<sup>4</sup>In practice, the factors are replaced by factor estimates.

<sup>5</sup>Using more than one factor would require identifying assumptions that complicate the analysis of all the possible identification schemes. Nevertheless, we address the possibility of estimating more than one factor when showing the results.

$\rho_s$ , the covariance is the Toeplitz matrix

$$\Sigma_v = \begin{pmatrix} 1 & \rho_s & \rho_s^2 & \dots & \rho_s^9 \\ \rho_s & 1 & \rho_s & \dots & \rho_s^8 \\ \rho_s^2 & \rho_s & 1 & \dots & \rho_s^7 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_s^9 & \rho_s^8 & \rho_s^7 & \dots & 1 \end{pmatrix}. \quad (7)$$

Now, we use (3) to generate  $\{X_{it}^s\}_{i,t=1}^{10,T}$ , where  $\Lambda$  is assumed to be a column vector of  $N$  ones. Therefore, the parameter  $\rho_s$  governs the correlation across categories of data. In the simulations, the values of this parameter is  $\rho_s = 0, 0.1, 0.5$ , and  $0.75$ .

Let us focus now on generating the 10 indicators that belong to each of the 10 different categories, i.e., the data set  $\{X_{i,kt}^l\}_{i,k,t=1}^{10,10,T}$ . For this purpose, we add an error term representing the idiosyncratic error of the specific series of each category to each of the ten time series  $\{X_{it}^s\}_{i,t=1}^{10,T}$ . We denote these errors as  $\{w_{i,kt}\}_{i,k,t=1}^{10,10,T}$ , where  $i$  represents the category, and  $k$  represents each of the series that belongs to the  $i$ -th category. We assume that the errors are correlated with the series of the  $i$ -th category and serially correlated. Hence, we generate the 100 indicators from

$$X_{ikt}^l = X_{it}^s + w_{ikt}, \quad (8)$$

where  $i = 1, \dots, 10$ ,  $k = 1, \dots, 10$ , and  $w_{it} = (w_{i1t}, \dots, w_{i10t})'$  is the vector of idiosyncratic errors, which is generated by

$$w_{ikt} = Dw_{ikt-1} + e_{ikt}^l. \quad (9)$$

In this expression,  $\{e_{ikt}^l\}_{i,k,t=1}^{10,10,T}$  are random numbers drawn from a normal distribution with zero mean and covariance matrix  $\Sigma_w$ , which is the Toeplitz matrix constructed from  $\rho_l$  as in (7), where  $\rho_l = 0, 0.1, 0.5$ , and  $0.9$ . Therefore, the parameter  $\rho_l$  controls the correlation that appear within each of the 10 categories of data. The matrix  $D$  of autoregressive coefficients is diagonal, with values of  $d = 0.1$  and  $d = 0.75$  in the main diagonal.

According to expressions (3), (8), and (9), each series of the 100 time series that forms the data set can be decomposed as follows

$$X_{ikt}^l = \lambda_i F_t + \xi_{ikt}^l, \quad (10)$$



where  $\xi_{ikt}^l = \xi_{it} + w_{ikt}$ . Then, the idiosyncratic components  $\xi_{ikt}^l$  are composed by a common error inside the categories,  $\xi_{it}$ , which could be correlated across different categories and a specific error term,  $w_{ikt}$ , which could be correlated with series of the same category. Finally, putting together the series along all the categories, the data set becomes

$$X_t^l = \left( X_{1,1,t}^l, X_{1,2,t}^l, \dots, X_{1,10,t}^l, X_{2,1,t}^l, X_{2,2,t}^l, \dots, X_{2,10,t}^l, \dots, X_{10,1,t}^l, X_{10,2,t}^l, \dots, X_{10,10,t}^l \right)'. \quad (11)$$

These time series can be interpreted as economic indicators that have been generated as the sum of two components: the common factor,  $F_t$ , and the idiosyncratic component,  $\xi_{ikt}^l$ . Therefore, the idiosyncratic components depend on the within-category cross correlation (measured by  $\rho_l$ ), on the across category correlation (measured by  $\rho_s$ ) and on the serial correlation of the idiosyncratic components (measured by  $c$  and  $d$ ).

### 2.3 Using disaggregate information

As suggested by Banbura and Modugno (2014), who extend the quasi-maximum likelihood approach proposed by Doz et al. (2012) to allow for serially correlated idiosyncratic components on a dataset with an arbitrary pattern of missing data, the estimates of the parameters the model stated in expressions (3) to (5) are obtained via the EM algorithm, which consists on an iterative two-step estimator.<sup>6</sup> In the first step, the algorithm computes an estimate of the parameters given an initial estimate of the common factor. In the second step, the algorithm uses the estimated parameters to approximate the common factor by the Kalman smoother. Iterating this process is an algorithm where in every step the likelihood increases, and hence, under regularity conditions, it converges to the maximum likelihood solution.

In our simulations, the  $(i + 1)$ -th iteration of the algorithm is defined as follows. Let  $X_t^l$  be the vector that collects the observations at time  $t$  of  $\{X_{j,kt}^l\}_{i,k,t=1}^{10,10,T}$ . Let us assume that  $\hat{\Lambda}^i$ ,  $\hat{A}^i$ ,  $\hat{\Sigma}_u^i$ ,  $C^i$  (whose elements are  $C_j^i$ ) and the diagonal matrix  $\hat{\Sigma}_v^i$  (whose elements are  $\hat{\Sigma}_{jv}^i$ ) are estimates from the previous iteration. Let  $F_t^i$  be the common factor, which

---

<sup>6</sup>Alternative approaches to account for serial correlation in the idiosyncratic components of large models are proposed by Jungbacker et al. (2011), Pinheiro et al. (2013) and Bai and Wang (2015). Pinheiro et al. (2013) show that it improves the accuracy.

is the output of the Kalman filter from the  $i$ -st iteration. The updated estimates of  $\Lambda$ ,  $A$ ,  $C$  and  $\Sigma_v$  can be obtained from

$$\hat{\Lambda}^{i+1} = \left( \frac{1}{T} \sum_{t=1}^T \left( E[\widehat{X_t^l F_t^l}] - E[\widehat{\xi_t F_t^l}] \right) \right) \left( \frac{1}{T} \sum_{t=1}^T E[\widehat{F_t F_t^l}] \right)^{-1}, \quad (12)$$

$$\hat{A}^{i+1} = \left( \frac{1}{T} \sum_{t=1}^T E[\widehat{F_t F_{t-1}^l}] \right) \left( \frac{1}{T} \sum_{t=1}^T E[\widehat{F_{t-1} F_{t-1}^l}] \right)^{-1}, \quad (13)$$

$$\hat{C}_j^{i+1} = \left( \frac{1}{T} \sum_{t=1}^T E[\widehat{\xi_{jt} \xi_{jt-1}}] \right) \left( \frac{1}{T} \sum_{t=1}^T E[\widehat{\xi_{jt-1} \xi_{jt-1}^l}] \right)^{-1}, \quad (14)$$

$$\hat{\Sigma}_u^{i+1} = \frac{1}{T} \sum_{t=1}^T E[\widehat{F_t F_t^l}] - \frac{1}{T} \hat{A}^{i+1} \sum_{t=1}^T E[\widehat{F_{t-1} F_t^l}], \quad (15)$$

$$\hat{\Sigma}_{jv}^{i+1} = \frac{1}{T} \sum_{t=1}^T E[\widehat{\xi_{jt}^2}] - \frac{1}{T} \hat{C}_j^{i+1} \sum_{t=1}^T E[\widehat{\xi_{jt-1} \xi_{jt}}], \quad (16)$$

with  $j = 1, \dots, N$ . The estimates of the expectations can be obtained from

$$E[\widehat{X_t^l X_t^l}] = X_t^l X_t^l, \quad (17)$$

and

$$E[\widehat{X_t^l F_t^l}] = X_t^l E[F_t^l]. \quad (18)$$

The estimates of the moments of the latent factors,  $E[\widehat{F_t}]$ ,  $E[\widehat{F_t F_t^l}]$ ,  $E[\widehat{F_{t-1} F_{t-1}^l}]$ ,  $E[\widehat{F_t F_{t-1}^l}]$ , and the moments involving  $\xi_t$  can be obtained from the Kalman smoother for the following state space representation. Let  $I_j$  be the identity matrix of dimension  $j$ , let  $0_{a \times b}$  be the  $(a \times b)$  matrix of zeroes, and let  $h_t$  be the  $(r + N)$  vector  $h_t = (F_t^l, \xi_t^l)'$ , where  $N = 100$ . Hence, the measurement equation can be defined as

$$X_t^l = H h_t + e_t, \quad (19)$$

where

$$H = \begin{pmatrix} \Lambda & I_N \end{pmatrix}, \quad (20)$$

and  $e_t$  is a vector of zeroes. In addition, the transition equation can be stated as

$$h_{t+1} = F h_t + w_t, \quad (21)$$

where the  $(r + N \times r + N)$  matrix  $F$  is

$$F = \begin{pmatrix} A & 0_{r \times N} \\ 0_{N \times r} & C \end{pmatrix}, \quad (22)$$

and  $w_t = (u'_t, v'_t)'$  with zero mean and covariance matrix

$$Q = \begin{pmatrix} \Sigma_u & 0 \\ 0 & \Sigma_v \end{pmatrix}, \quad (23)$$

where  $\Sigma_u$  and  $\Sigma_v$  are diagonal matrices that refer to the variances of  $u_t$  and  $v_t$ .

Then,  $\hat{\Lambda}^{i+1}$ ,  $\hat{A}^{i+1}$ ,  $\hat{\Sigma}_u^{i+1}$ ,  $C^{i+1}$  and  $\hat{\Sigma}_v^{i+1}$  are used again in the Kalman filter to compute the factors of the  $(i + 1)$ -th iteration. The algorithm, which starts with the static principal components estimates of the common factors  $F_t^0$  and their factor loadings  $\Lambda^0$ , is repeated until the quasi-maximum likelihood estimates of the parameters are obtained.<sup>7</sup> Finally, the forecasts of  $y_{t+1}$  are estimated by OLS regressions on (6). Let us denote the dynamic factor model that uses all the disaggregate information as *DDFM* ( $N$  and  $T$  are large).

## 2.4 Using aggregate information

Initiated by the four-variable single-index factor model suggested by Stock and Watson (1991), there is a strand of literature on factor models that focuses on constructing indexes of coincident indicators from the aggregate information provided by a subset of indicators of the total economy. Typically, the analyses focus on different enlargements of the four variables used by Stock and Watson, i.e., industrial production, retail sales, unemployment and income. However, the data used in these analyses are usually designed for a specific problem and does not rely on statistical criteria, which complicates the implementation to other countries or data sets. Examples are Mariano and Murasawa (2003), Nunes (2005), Aruoba, Diebold and Scotti (2009), Aruoba and Diebold (2010), and Camacho and Perez Quiros (2010). The ADS index, released by the Federal Reserve Bank of Philadelphia, also follows this approach.

---

<sup>7</sup>The algorithm requires small number of iterations to converge. In our simulations, we only required 3 or 4 iterations to converge, which occurs whenever the difference between two consecutive log-likelihood values is lower than a  $10^{-4}$ .

With the aim of analyzing the effects of estimating the factors using an small set of informative indicators, we contribute to the literature by proposing a data selection procedure that benefits from the peculiar data structure typically faced by economists: choose the time series from each category that exhibit the highest averaged correlation with respect to the other series included in the same category. Let  $\{X_{i,*t}^l\}_{i,t=1}^{10,T}$  be this set of  $N = 10$  time series with the highest average within-category correlation, whose observations at time  $t$  are collected in the vector  $X_t^*$ .

The model used in this strand of the literature is typically cast in the state space form described in (19)-(23). The parameters of  $\Lambda$ ,  $A$ ,  $C$  and  $\Sigma_v$  are estimated by direct optimization via the Kalman filter, where  $X_t^l$  is replaced by  $X_t^*$  and  $N = 10$ .<sup>8</sup> In the standard way, the Kalman filter also produces filtered and smoothed inferences of the common factor:  $\{F_{t|t}^s\}_{t=1}^T$  and  $\{F_{t|T}^s\}_{t=1}^T$ . These inferences can be used in the prediction equation (6) to compute OLS forecasts of the variable  $y_{t+h}$ . Let us denote the dynamic factor model that uses the aggregate information in the way that is described in this section as *ADFM* ( $N$  is fixed and small and  $T$  is large).

Only for comparative purposes, we also apply this method to the set of 10 time series used to generate the categories, i.e.,  $\{X_{it}^s\}_{i,t=1}^{10,T}$ , which are collected at time  $t$  in the vector  $X_t^s$ . Since these time series are less noisy than any of the time series included in the disaggregate and aggregate analyses, we know that the results obtained from this method are, by construction, better.<sup>9</sup> However, we consider these results as “lower bounds” with which the results of the disaggregate and aggregate approaches are compared. Finally, we compare our selection criterion with the selection procedure used by Boivin and Ng (2006), who select the subset of indicators by removing those with highest correlation across the idiosyncratic components. Collecting the selected data,  $\{X_{i,b,t}^l\}_{i,t=1}^{10,T}$ , at time  $t$  in the vector  $X_t^b$ , the analysis of this section is also applied.

---

<sup>8</sup>For identification purposes,  $\Sigma_u$  is usually assumed to be the identity matrix.

<sup>9</sup>Therefore, the less noisy series are not included in the set of 100 time series from which we performed the disaggregate and aggregate information approaches.

## 2.5 Generating the target series

We generate the series to be predicted in a simple scenario. To simplify simulations, we consider that the forecasts are computed from the factors and only one lag of the time series. Hence, the series  $y_t$  is generated from (6), where  $\beta_0 = 0$ ,  $\beta_1 = 1$ ,  $p = 1$  and  $e_{yt}$  is a white noise process with  $\sigma_{e_y} = 1$ . The parameter  $\gamma$ , which measures the autocorrelation of the target series, is assumed to take on the values of 0, 0.3, and 0.8.

## 3 Simulation results

In each replication,  $j$ , we estimate the factor models that use aggregate and disaggregate information and compute the accuracy of these models to infer the factor by using the Mean Squared Error over the  $J = 1000$  replications

$$MSE^i = \frac{1}{J} \sum_{j=1}^J \frac{1}{T} \sum_{t=1}^T (F_{jt} - F_{jt|T}^i)^2, \quad (24)$$

for  $i = a$  in the case of the models that use aggregate information and  $i = d$  in the case of the models that use disaggregate information. In addition, we compare the out-of-sample forecasting accuracy of *ADFM* and *DDFM* by computing the errors in forecasting one step ahead the generated target series. Let  $\hat{\beta}$  and  $\hat{\gamma}$  be the OLS estimates of the parameters that appear in the prediction equation using the common factor series and the values of  $y$  up to period  $T$ . Then, we construct the one-step-ahead forecast of  $y_{jT+1}$  by using the relation  $\hat{y}_{jT+1}^i = \hat{\beta} F_{jt|T}^i + \hat{\gamma} y_{jT}$ . In this way, one can define the Mean Squared one-step-ahead Forecast Errors of model  $i$  as

$$MSFE^i = \frac{1}{J} \sum_{j=1}^J (y_{jT+1} - \hat{y}_{jT+1}^i)^2. \quad (25)$$

According to the forecasting scenarios described above, we call  $MSE_p^a$ ,  $MSE_r^a$ ,  $MSE^d$ ,  $MSE_b^a$ ,  $MSFE_p^a$ ,  $MSFE_r^a$ ,  $MSFE^d$  and  $MSFE_b^a$  the mean across replications of the *MSE* and *MSFE*. They are computed from an *ADFM* that uses the 10 series used to generate the categories (superscript  $a$ , subscript  $p$ ),  $X_t^s$ ; from an *ADFM* that uses the 10 indicators with highest averaged correlation with the series of each category (superscript  $a$ , subscript  $r$ ),  $X_t^*$ ; from a *DDFM* that uses the 100 time series of the simulation exercise

(superscript  $d$ ),  $X_t^l$ ; and from a *ADFM* that uses a reduced number of series that are chosen from the large pool by using the Boivin and Ng (2006) statistical criteria (superscript  $a$ , subscript  $b$ ),  $X_t^b$ .<sup>10</sup>

### 3.1 Factor estimates

In this section, we investigate the benefits of computing inferences from aggregate and disaggregate data. Let us start the analysis of the simulations by using *MSEs* to examine the relative accuracy of the models to infer the factors. To facilitate understanding, let us describe how the results are presented in the tables. First, the results in Tables 1 to 3 are classified according to different values of the autoregressive coefficient of the common factor (coefficient  $A$ ). This coefficient takes on the value of 0.1 (low correlation) in Table 1, the value of 0.5 (medium correlation) in Table 2 and the value of 0.75 (high correlation) in Table 3.

Second, each of these tables shows the accuracy of the models for different values of the cross-correlation of the errors within (measured by  $\rho_l$ ) and across (measured by  $\rho_s$ ) categories. The first block of the results refers to the case when the cross-correlation present in the idiosyncratic components is only due to series that belong to the same category, which occurs when  $\rho_s = 0$ . The following blocks of results examine the effects of progressively increasing the correlation across categories to 0.1, 0.5 and 0.75. Within each of these blocks, the tables report the models accuracy to infer the common factor when the within-category correlation, which is measured by  $\rho_l$ , increases from 0 to 0.1, 0.5 and 0.9.

Third, the first four columns of the tables refer to MSEs from dynamic factor models which either use the set of 10 time series that were used to generate the categories (results labelled as  $MSE_p^a$ ); or use the 10 indicators that exhibits the highest averaged correlation with the series of their respective categories (results labelled as  $MSE_r^a$ ); or use the complete set of 100 indicators (results labelled as  $MSE^d$ ); or use the 10 series selected with the Boivin-Ng statistical criteria (results labelled as  $MSE_b^a$ ), respectively.<sup>11</sup>

<sup>10</sup>We iterate the algorithm proposed by Boivin and Ng (2006) five times. This method provides a number of time series that average 10.

<sup>11</sup>Notice that  $MSE_p^a$  is valid for comparison but not relevant for practical purposes since it reaches the

Fourth, it is a common practice in large scale factor models that each category is represented by different number of time series and frequently some categories might be over represented.<sup>12</sup> We address the effects of over sampling on the performance of the models in the last three columns of these tables. For this purpose, we simulate ten categories of data but including 20 series instead of 10 in the first category, using 5 series instead of 10 in the second and third categories, and using 10 series in each of the other 7 categories.<sup>13</sup>

Fifth, in Tables 1 to 3, we assume that the idiosyncratic components and the errors of the same category have low serial correlation (values of  $c = d = 0.1$ ), that the sample is small ( $T = 50$ ), and that there is only one common factor in the estimation.<sup>14</sup> The robustness of the results in the cases of serial correlation in errors, larger samples, and ragged edge are analyzed in Tables A1 to A9 in the Appendix.<sup>15</sup>

A summary of the main results is the following. As expected, the tables show that the factor model that uses the 10 time series that generate the categories reaches the smallest *MSE*. Notably, the tables also show that there is valuable gains by selecting those series of each category that exhibit the highest correlation with the other series of the same category. The relative performance of *ADFM* with the representative highly correlated series and the *DDFM* shows that the former improves upon the latter ( $MSE_r^a < MSE^d$ ) when the persistence of the factor and the within categories correlation increase.

These results are in line with some recent findings proposed in the related literature. First, our results are in line with the findings of Stock and Watson (2002b), although these

---

lowest value by construction.

<sup>12</sup>Typically, the number of series of disaggregated industrial production indicators is quite higher than the number of time series included in other categories. Significant examples are Stock and Watson (2002a, 2002b), Giannone et al. (2008), and Angelini et al. (2011).

<sup>13</sup>The accuracy of the factor model that relies on the 10 time series used to generate the categories does not depend on the number of series that are included in each category. Hence, the last columns of the tables only show  $MSE_r^a$ ,  $MSE^d$  and  $MSE_b^a$ .

<sup>14</sup>In their simulations, Stock and Watson (2002b) consider that  $T$  is large when it is greater than 100, that  $T$  is small when it is smaller than 50, and that  $T$  is very small when it is equal to 25.

<sup>15</sup>The Appendix, which shows all the expressions of this section, is available at <http://www.um.es/econometria/Maximo>. However, since the method proposed by Boivin and Ng (2006) does not provide significant gains with respect to the other proposals, this method is omitted in the tables that appear in the Appendix.

authors do not model the serial correlation. These authors find some deterioration on the quality of the factor estimates when the degree of serial correlation in the factor and in the idiosyncratic errors is high even when the number of variables and observations is large. This coincides with the finding that we show in Tables 2 and 3, which report the results of increasing the inertia in the simulated common factor, with  $A$  ranging from 0.1 (almost no serial correlation) in Table 1 to 0.5 (moderate correlation) in Table 2 and to 0.75 (high correlation) in Table 3. Although our results confirm the deterioration in factor estimation from all the factor models, the relative losses are not uniformly distributed across the models. When the serial correlation of the factor increases, the relative performance of the factor model that uses aggregate information increases significantly, except for the case of very large correlation across categories, where the relative gains attenuate. In fact, when  $A = 0.75$  the factor model that use only the series of each category with highest within-category correlation outperforms the factor model that uses disaggregate information in all scenarios.<sup>16</sup>

Second, our results are in concordance with those of Boivin and Ng (2006) who suggest that the factor estimates are adversely affected by cross-correlation in the errors and by oversampling. The *MSEs* displayed in Tables 1 to 3 suggest that the factor model that uses disaggregated information is usually beaten by the factor model that uses aggregate information when the correlation of the variables of the same category is high. In addition, the effect of using oversampled categories in factor analysis are analyzed in the last three columns of the tables, which report the *MSEs* of estimating the factor from models that use the 10 unbalanced sets of indicators. Overall, the factor model that uses the disaggregated information performs worse than in the case of balanced categories, especially when the correlation across categories is small. Again, the relative better accuracy of the factor model that uses aggregate information is more evident when the low correlation across categories is combined with high within-category correlation and high persistence of the factor.

---

<sup>16</sup>The scenario of high persistence in the common factor is very realistic in practice. For example, regressing the Chicago Fed National Activity Index (CFNAI) and the Philadelphia Fed ADS Business Conditions Index (converted to monthly by averaging) on their past values lead to slope coefficients of 0.86 and 0.93, respectively.



Notably, Tables 1 to 3 also show that the statistical selection criterion proposed by Boivin and Ng (2006) leads to poor relative performance in most cases. Only in the case of very low correlation across categories, very low within-category correlation and extremely low autocorrelation of the factor, the results obtained by using their selection criterion are comparable with the results obtained by using ours. Even in this case, it is also remarkable that the good performance of their statistical criterion are not robust to oversampling. This result suggests that, when data sets are drawn from a small number of broad categories, using our variable selection criterion is appropriate.

In the online Appendix, Tables A1 to A4 examine the effects of increasing the serial correlation of the idiosyncratic components on the factor models. In particular, the effects of having higher autocorrelation of the series specific shock (measured by  $d$ ) are analyzed in Tables A1 and A2 whereas the effects of assuming higher autocorrelation of the category specific shocks (measured by  $c$ ) are analyzed in Tables A3 and A4.

Tables A1 and A2 show the *MSEs* of the models when the serial correlation of the series specific shock is assumed to grow from  $d = 0.1$  to  $d = 0.75$  in two scenarios, when the serial correlation of the factor is low ( $A = 0.1$  in Table A1) and when it is high ( $A = 0.75$  in Table A2). Tables A3 and A4 analyze the role of the serial correlation of category specific shocks, which is measured by the parameter  $c$ . This parameter is allowed to increase from  $c = 0.1$  to  $c = 0.75$  when the serial correlation of the factor is low ( $A = 0.1$  in Table A3) and when it is high ( $A = 0.75$  in Table A4).

The *MSEs* reported in the tables show that increasing the serial correlation in the idiosyncratic components contributes to deteriorate the overall performance of the models. However, the tables show a better accuracy of the factor models that use the 10 representative series compared to the factor models that use the disaggregate information of the 100 series, especially when there is high serial correlation in the idiosyncratic components. This result reveals that factor models that use large panels could be more negatively affected by serial correlation than factor models that use a reduced number of selected indicators. Finally, the tables also show that the relatively larger negative effects of increasing the correlation of the idiosyncratic components on large panels are magnified in the case of oversampled categories.

The role of the number of observations in the performance of factor models is examined in Tables A5 and A6. According to the theory, in absence of the typical data problems which are accounted for by our simulations and that usually appear in empirical applications, the larger the sample size of the time series the better expected performance of factor models that use the largest panels over the factor models that uses only a subset of indicators. This theoretical result is documented in Table A5 where the reported  $MSEs$  show that under low serial correlation of the factor and low correlation of the idiosyncratic errors, the accuracy of the factor model that uses the ten representative indicators that exhibit the largest correlation with the series of each category with respect to the factor model that uses disaggregate information diminishes. However, the tables also show that when the serial correlation of the factor increases, the former clearly outperforms the latter. Interestingly, the tables also reveal that the relative losses in accuracy due to oversampling in factor models that use disaggregate information are still large when the sample size increases even in absence of data problems.

It is worth noting that the number of factors has been restricted to be one according to the data generating process. However, since the time series are generated in different categories with high within-category and across-category correlation, this assumption might be too restrictive in practice. To evaluate the effect of this potential restriction that could affect the accuracy of factor models that use disaggregate information, we leave the models to select the number of factors according to the procedure described in Bai and Ng (2002), where the maximum number of factor is 11.<sup>17</sup> Table A7 and A8 report the  $MSE^d$  and the averaged number of estimated factors across the 1000 replications both in the case of balanced sets of categories and in the case of oversampled categories. According to the previous discussion, we find that the higher the within-categories correlation, the larger the number of estimated factors. The intuition is that the high correlation in each category is interpreted by the model as if the series belonging to the same category would share a common factor. Notably, although selecting the number of factors increases the accuracy of factor models that use disaggregated information, the gains are not sufficiently large to

---

<sup>17</sup>We also selected the factors by using BIC. However, although this method selected lower factors, it typically exhibited a deterioration in MSE.

qualitatively alter the results obtained in this section, especially when the autocorrelation of the factor is high and there are oversampled categories.

Finally, factor models are an appropriate framework to handle in an automatic manner data sets with series that are available with different publication lags.<sup>18</sup> Following Banbura and Modugno (2014), the idea is to write the likelihood as if the data were complete and to discard in the Kalman filter the rows that correspond to missing observations. Table A9 examines the effects of missing data on factor models when 20% of time series end at  $T-2$ , 40% of the time series end at  $T-1$  and 40% of the time series end at  $T$ .<sup>19</sup> In line with these authors, we find that although the accuracy decreases when missing data appear in the data, the losses are not large. Interestingly, our results suggest that the deterioration in factor estimation is uniformly distributed across the models. This suggests that our results are not qualitatively affected by the presence of missing data.

### 3.2 Forecasting accuracy

This section examines how close the one-step-ahead out-of-sample forecasts based on the estimated factors are to the target series, which has been generated from (6). Part of the forecast performance analysis has already been developed in the previous section since, in absence of autocorrelation in the target series (measured by  $\gamma$ ), the forecast performance is expected to increase when the discrepancy between the actual and the estimated factors diminishes.<sup>20</sup> Accordingly, this section examines the effects of different values of  $\gamma$  ranging from 0 (no inertia) to 0.8 (high degree of time series dependence) on forecast performance. In addition, the section also addresses the effects of the data problems outlined above on the relative forecast performance of factor models that use aggregate and disaggregate information.

Tables 4 to 6 evaluate the ability of factor models in forecasting.<sup>21</sup> As in the case of

---

<sup>18</sup>For a deeper analysis of the effects of missing data on factor models, the readers are referred to Banbura and Modugno (2014) and the references therein.

<sup>19</sup>We tried with other combinations of missing data and the results were qualitatively similar.

<sup>20</sup>Note that the variance of the errors has been normalized  $\sigma_{e_y} = 1$ .

<sup>21</sup>To save space, the tables that show the in-sample forecast analysis were omitted. In addition, the tables that show the forecast analysis has been simplified. More complete versions of these tables are available from the authors upon request.

factor estimates, the relative forecasting accuracy of factor models that use aggregate and disaggregate information is examined under different scenarios, in which the simulations allow for different degrees of cross-correlation across ( $\rho_s$  from 0 to 0.5) and within ( $\rho_l$  from 0 to 0.9) categories. Table 4 shows the *MSFEs* of the models when the factor exhibits low serial correlation ( $A = 0.1$ ) while Tables 5 and 6 display the *MSEs* of the models when the factor autocorrelation increases to medium ( $A = 0.5$ ) and to high ( $A = 0.75$ ), respectively.

The robustness analysis has been conducted through Tables A10 to A15 in the online Appendix. Tables A10 and A11 display the *MSFEs* of the models when the autocorrelation of the series specific shock increases to  $d = 0.75$ , Tables A12 and A13 show the effects of increasing the sample size to  $T = 150$  and Tables A14 and A15 analyze the forecasting accuracy when the number of common factors is selected as Bai and Ng (2002) describe.<sup>22</sup>

Overall, the tables show that the data problems lead to similar effects on the forecasting ability of the models than those observed on the analysis of factor estimation. When the time series are correlated with the indicators already included in some categories, the factor or the idiosyncratic components are persistent, or some categories are oversampled, forecasting with disaggregate data performs worse than forecasting from a representative series data set, especially when the categories are not highly correlated. Again, the tables show the better accuracy of our proposed criterion of picking the series with the highest average within category correlation with respect to the selection criterion suggested by Boivin and Ng (2006), which is based on picking the time series with highest correlation across the idiosyncratic components.

As expected, the results depend on the magnitude of the autocorrelation of the target variable since it tends to mitigate the forecasts losses of those models which are more contaminated with data problems. That is, the models that exhibited larger deteriorations in factor estimation due to data problems present smaller increases in *MSFE* when the autocorrelation of the target variable increases. The intuition is clear: the larger the

---

<sup>22</sup>The tables that examine the effects of higher category-specific autocorrelation, measured by  $c$  are omitted to save space. The results are similar to those obtained when the series-specific autocorrelation, measured by  $d$ , increases.

autocorrelation of the target variable the smaller the weights of the factor in forecasting the time series and the lower the effect on forecasting from mis-specified models.

For example, Tables 1 and 3 show that if the set of parameters that measure the data problems change from  $\rho_s = 0, \rho_l = 0, A = 0.1$  to  $\rho_s = 0.5, \rho_l = 0.9, A = 0.75$ , the accuracy of the factor estimation reduces from  $MSE^d = 0.14$  to  $MSE^d = 0.59$ , which implies a 321% increase in  $MSE$ . However, under the same change in the set of parameters, Tables 4 and 6 show that the forecast accuracy moves from  $MSFE^d = 1.12$  to  $MSFE^d = 1.52$  when  $\gamma = 0$ , which implies a 36% increase in  $MSE$ , and to  $MSFE^d = 1.46$  when  $\gamma = 0.8$ , which implies only a 30% increase in  $MSE$ .

In sum, in this section we find that adding data that bear little additional information about the factor components does not necessarily lead factor models that use all the disaggregate information to improve upon the forecasts of factor models that use the aggregate information that appear in the selected indicators. When the additional data are too correlated with the data of some categories that are already included in factor estimation, forecasting with disaggregated information tends to perform worse than forecasting from aggregated data. Our results are stronger in the case of high persistence of the common factor, in the case of high serial correlation of the idiosyncratic components and in the case of oversampled categories. However, the relative good performance of models that use aggregate information mitigates when the across-category correlation increases. In addition, the gains in forecasting reduce substantially when the target series are persistent.

## 4 Empirical example

This section examines the forecasting accuracy of small versus large scale dynamic factor models by using the data set of monthly macroeconomic indicators used in the balanced panel factor estimation by Stock and Watson (2002a) for the US economy.<sup>23</sup> The variables, which are available over the sample 1959:01-1998:12, are standardized and transformed to induce stationarity following their instructions.

---

<sup>23</sup>Although the unbalanced panel proposed by Stock and Watson (2002a) included 215 time series, we concentrate on the time series that form the balanced panel.

#### 4.1 Preliminary analysis of data

According to Stock and Watson (2002a), Table 7 classifies the data in 13 different categories: (1) real output and income (series 1-19); (2) employment and hours (series 20-44); (3) retail and manufacturing trade (series 45-53); (4) consumption (series 54-58); (5) housing starts and sales (series 59-65); (6) inventories (series 66-76); (7) orders (series 77-92); (8) stock prices (series 93-99); (9) exchange rate (series 100-104); (10) interest rates (105-119); (11) money and credit (series 120-126); (12) price indexes (series 127-144); (13) Average hourly earnings (series 145-146).<sup>24</sup> This table displays the name of the categories in column 1 and the number of the series included in each category in column 2. Notably, some categories contain more indicators than others. Therefore, the problem of oversampling outlined in the simulations applies in this empirical analysis.

According to the motivation of the paper, much of the time series included in each category are disaggregations of the headline concept and, therefore, are expected to exhibit large correlations with each others. To gauge this fact in our example, Table 7 also shows in the third column the averaged correlation across the series of each category. Overall, the categories contains very collinear indicators, which exhibit averaged correlations of more than 0.5 in the cases of housing starts and sales and exchange rates and of more than 0.4 in the cases of real output and income, consumption, stock prices, and interest rates.

In order to gain better insights on the set of indicators selected from the selection method proposed in the paper, the fourth column of Table 7 shows the name of the indicators that exhibit the largest averaged correlation with the series of each category and the last column of Table 7 reports the magnitudes of these averaged correlations. Overall, the representative indicators exhibit averaged correlations with the series of the same category of more than 0.5, and in some cases the correlations rise up to 0.70 in the case of exchange rates and to 0.74 in the case of housing starts. Interestingly, when finer disaggregations of sectorial data are included in a category, the representative indicator of the category in the aggregate information scenario usually refers to the headline concept.

---

<sup>24</sup>The last category labelled as *miscellaneous* has been omitted from the empirical analysis since it included only one series.

In this sense, the proposed criterion of picking the series with the highest average within category correlation selects economically meaningful sets of representative indicators from each category.

## 4.2 Out-of-sample forecasting

We consider two real (industrial production and non agricultural employment) and two nominal (consumer and producer price indexes) target series, which are called  $Y_t$ . We investigate the accuracy of the different specifications of dynamic factor models to forecast the target series using the multi-step ahead forecasting procedure described in Stock and Watson (2002a)

$$y_{t+h}^h = \alpha_0 + \sum_{i=0}^m \beta_i' \widehat{F}_{t-i} + \sum_{j=0}^v \gamma_j z_{t-j} + \varepsilon_{t+h}^h. \quad (26)$$

In this equation,  $y_{t+h}^h$  is the  $h$ -step ahead covariance stationary transformation of the original series  $Y_t$ , where  $y_{t+h}^h = \ln(Y_{t+h}/Y_t)$ ,  $\widehat{F}_{t-i}$  is the  $i$ -lagged ( $i = 0, 1, \dots, m$ ) value of the ( $r \times 1$ ) vector of estimated factors, and  $z_{t-j}$  is the  $j$ -lagged ( $j = 0, 1, \dots, v$ ) value of the 1-step ahead covariance stationary transformation of  $Y_t$ , where  $z_t = \ln(Y_t/Y_{t-1})$ . Expressions  $\beta_j'$  and  $\gamma_j$  refer to the standard parameters of autoregressive processes. The term  $\varepsilon_{t+h}^h$  is a homoskedastic martingale difference sequence with respect to the set of information at time  $t$ . Finally, in line with previous studies in forecasting with empirical factors, our model is allowed to choose values of  $m$  lying between 1 and 6 and  $v$  lying between 1 and 12 based upon the BIC selection criterion. In large scale factor specifications,  $r$  is either imposed to be just one or selected as Bai and Ng (2002) describe.

The out-of-sample forecasting exercise begins with data from 1959:3-1970:1. Using this sample,  $m$ ,  $v$ , and (in some cases)  $r$  are chosen, and a  $h$  period ahead forecast is formed by using values of the regressors at 1970:1 to give  $y_{1970:1+h}^h$ . Then, the sample is updated by one period, the factors and the forecasting models (including  $m$ ,  $v$ , and, in some cases,  $r$ ) are re-estimated, and a  $h$ -month forecast for 1970:1+ $h$  is computed (for  $h = 1$  it would be 1970:2, for  $h = 6$  1970:7 and for  $h = 12$  1971:1). The forecasting procedure continues iteratively until the final forecast  $y_{1998:12}^h$  which is made using data until 1998:11 for  $h = 1$ , 1998:6 for  $h = 6$  and 1997:12 for  $h = 12$ . In each iteration, the

root of the squared deviation of  $h$ -ahead forecasts from actual data are computed and the average of these figures is labeled as *RMSFE*.

To investigate the benefits of forecasting from aggregate and disaggregate data, Tables 8 and 9 consider a forecast competition of different factor models. The first competitor is a simple autoregressive model which is obtained when  $\beta_i = 0$  for all  $i = 1, \dots, m$  in (26). The second set of competitors are autoregressive models that are enlarged with the factors obtained from factor models applied to reduced sets of indicators, which have been selected by using three different alternatives. In the first alternative, the 13 representative indicators are the series of each category that exhibit the highest averaged cross-correlation. In the second alternative, the representative indicators are selected as in Boivin and Ng (2006) by removing the time series with highest correlation across the idiosyncratic components.<sup>25</sup> In the third scenario, the 13 indicators are selected by randomly picking one series from each category. The last competitor is a factor model that uses the disaggregate information of the balanced set of the 146 indicators. In some cases, the number of factors is restricted to be  $r = 1$  while in others the optimal number of factors  $r^*$  is determined by using BIC in the case of *ADFM* and by using Bai and Ng (2002) in the case of *DDFM*.

To facilitate comparisons, the tables report the *RMSFEs* relative to the autoregressive models. Hence, an entry less than one indicates that the factor model forecast is superior to the autoregressive univariate forecast. According to Stock and Watson (2002a), regardless of the factor model and the forecasting horizon used in the analysis, the forecasts of the factor models generally improve over the benchmark univariate forecasts.

In addition, we find that the forecasting accuracy depends on the number of factors included in the analysis. In the case of real variables, we find significant improvements in the forecast performance of factor models whose factors are determined by statistical criteria over the forecasts computed from only one factor, especially when the forecasting horizon becomes large. For example, Table 8 shows that when only one factor is included in the factor forecasts, most of the relative *RMSFEs* are greater than 0.9, which implies that the factor forecasts are only slightly more accurate than the univariate autoregressive forecasts. However, Table 9 shows that in the case of nominal variables, we failed to find

---

<sup>25</sup>To allow for balanced comparisons, the method is iterated until the final dataset contains 13 indicators.



the forecast deterioration when the number of factors is restricted to one.

According to our simulation results, which model do we expect to perform better? We find that the factors are highly persistent and some categories are over-represented. Therefore, the aggregate model should perform better, especially if there is high within-category correlation. However, if the across-category correlations are high, forecasting from the aggregate scenario could suffer from a significant deterioration. In addition, the nominal target series are very persistent, which implies that the potential benefits of the two alternatives could become negligible in these cases.<sup>26</sup> In line with Banbura and Modugno (2014) and Banbura, Giannone and Reichlin (2011), in Tables 8 and 9, we find that including more disaggregated data does not improve the accuracy of the forecasts. For less persistent target series of economic activity, the factor model that uses aggregate information yields satisfactory forecasting results with respect to those of the factor model that uses disaggregate information. In the case of nominal variables, which are very persistent, we do not find substantial differences between the two approaches. This suggests that using the series with highest within-category averaged correlation can reliably extract the relevant signals from the large data set suggested by Stock and Watson (2002a), which relies on finer disaggregations of the headline concepts of a small number of broad categories.

Remarkably, Tables 8 and 9 show that the variable selection criterion proposed in this paper, which is based on picking the series of each category that exhibits the largest within-category correlation uniformly performs better than the variable selection suggested by Boivin and Ng (2006) of removing the time series with highest correlation across the idiosyncratic components and than the criterion of randomly picking one series from each category. Therefore, we consider that our proposal could be the basis of factor models that use aggregated data whose variable selection process does not follow statistical criteria. Among others, examples are Mariano and Murasawa (2003), Nunes (2005), Aruoba, Diebold and Scotti (2009), Aruoba and Diebold (2010), Camacho and Perez Quiros (2010), and the ADS index, released by the Federal Reserve Bank of Philadelphia.

---

<sup>26</sup>An additional explanation is that all the categories are assumed to be equally informative in the Monte Carlo analysis, while this might not be the case in practice.

## 5 Conclusions

Two versions of dynamic factor models have received a growing attention in the recent forecasting literature, the dynamic factors that use disaggregate information from large data sets and the dynamic factors that use the aggregate information from a small number of representative indicators. In this paper, we propose simulations that mimic different scenarios of empirical forecasting when the large data sets rely on finer disaggregations of the headline concepts of a small number of broad categories.

As in economic applications, we assume that some categories are oversampled, that cross correlation and serial correlation among idiosyncratic components may appear, that the factors exhibit different degrees of correlation and that there could be missing data. In this context, we ask whether it is sufficient to use only some key selected indicators of each category (those with largest averaged within-category correlation) or it would be more efficient to use the disaggregated (sometimes redundant) information of each category.

We find that adding data that bear little additional information about the factor components does not necessarily lead factor models that use all the disaggregate information to improve upon the forecasts of factor models that use the aggregate information. When the within-category idiosyncratic components are correlated, forecasting with disaggregated data tends to perform worse than forecasting with aggregated data. Our results are stronger in the case of high persistence of the common factor, in the case of high serial correlation of the idiosyncratic components and in the case of oversampled categories. However, the relative good performance of models that use aggregate information mitigates when the across-category correlation increases. In addition, the gains in forecasting reduce substantially when the target series are persistent. Therefore, our paper contributes to the existing literature by adding some new lights on the existing black box decisions in the use of disaggregate and aggregate approaches.

All the data problems outlined above are very frequent in empirical applications, as in the example proposed in the paper, which focuses in the Stock and Watson (2002a) large database. Therefore, in line with Banbura and Modugno (2014), we failed to find substantial differences between the aggregate and disaggregate approaches. This suggests

that using the series with highest within-category averaged correlation could be a good alternative to using large disaggregated data sets.

Therefore, the paper offers a formal method to select the series for factor models once the categories are fixed, which is a typical data structure faced by empirical macroeconomists. However, three issues are left for further research. First, the analysis of which categories to choose when the data structure described in this paper is not available. Second, the issue of ragged edges, where apart from signal to noise ratio also timeliness matters. Third, the comparison between dynamic hierarchical factor models and the factor models used in this paper.

## References

- [1] Angelini, E., Camba-Mendez, G., Giannone, D., Reichlin, L., and Rünstler, G. 2008. Short-term forecasts of euro area GDP growth. *Econometrics Journal* 14: 25-44.
- [2] Aruoba, B., Diebold, F., and Scotti, C. 2009. Real-time measurement of business conditions. *Journal of Business and Economic Statistics* 27: 417-427.
- [3] Aruoba, B., and Diebold, F. 2010. Real-time macroeconomic monitoring: Real activity, inflation, and interactions. *American Economic Review* 100: 20-24.
- [4] Bai, J., and Ng, S. 2002. Determining the number of factors in approximate factor models. *Econometrica* 70: 191-221.
- [5] Bai, J., and Wang, P. 2015. Identification and bayesian estimation of dynamic factor models. *Journal of Business and Economics Statistics*, forthcoming.
- [6] Banbura, M., Giannone, D, and Reichlin, L. 2011. Nowcasting. In *Oxford Handbook on Economic Forecasting*, edited by M. Clements and D. Hendry. Oxford University Press.
- [7] Banbura, M., and Rünstler, G. 2011. A look into the model factor model black box: Publication lags and the role of hard and soft data in forecasting GDP. *International Journal of Forecasting* 27 333-346.
- [8] Banbura, M., and Modugno, M. 2014. Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics* 29: 133-160.
- [9] Boivin, J., and Ng, S. 2006. Are more data always better for factor analysis? *Journal of Econometrics* 132: 169-194.
- [10] Caggiano, G., Kapetanios, G., and Labhard, V. 2011. Are more data always better for factor analysis? Results for the Euro area, the six largest Euro area countries and the UK. *Journal of Forecasting* 30: 736-752.

- [11] Camacho, M., and Perez Quiros, G. 2010. Introducing the Euro-STING: Short Term Indicator of Euro Area Growth. *Journal of Applied Econometrics*, 25: 663-694.
- [12] Doz, C., Giannone, D., and Reichlin, L. 2012. A quasi-maximum likelihood approach for large approximate dynamic factor models. *The Review of Economics and Statistics* 94: 1014-1024.
- [13] Forni, M., and Reichlin, L. 1996. Dynamic common factors in large cross-sections. *Empirical Economics* 21: 27-42.
- [14] Forni, M., Hallin, M., Lippi, M., and Reichlin, L. 2005. The generalized dynamic factor model: one-sided estimation and forecasting. *Journal of the American Statistical Association* 100: 830-40.
- [15] Giannone, D., Reichlin, L., and Small, D. 2008. Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics* 55: 665-676.
- [16] Jungbacker, B., Koopman, S., and van der Wel M. 2011. Maximum likelihood estimation for dynamic factor models with missing data. *Journal of Economic Dynamics and Control* 35: 1358-1368.
- [17] Mariano, R., and Murasawa, Y. 2003. A new coincident index os business cycles based on monthly and quarterly series. *Journal of Applied Econometrics* 18: 427-443.
- [18] Moench, E., Ng, S., and Potter S. 2013. Dynamic hierarchical factor models. *Review of Economics and Statistics* 95: 1811-1817.
- [19] Nunes, L. 2005. Nowcasting quarterly GDP growth in a monthly coincident indicator model. *Journal of Forecasting* 24: 575-592.
- [20] Pinheiro, M., Rua, A., and Dias, F. 2013. Dynamic factor models with jagged edge panel data: Taking on board the dynamics of the idiosyncratic components. *Oxford Bulletin of Economics and Statistics* 75: 80-102.
- [21] Poncela, P., and Ruiz, E. 2015. More is not always better: back to the Kalman filter in dynamic factor models. In *Un-observed Components and Time Series Econometrics*, edited by S. Koopman, and N. Shephard. Oxford University Press.

- [22] Stock, J., and Watson, M. 1991. A probability model of the coincident economic indicators. In *Leading Economic Indicators: New Approaches and Forecasting Records*, edited by K. Lahiri and G. Moore. Cambridge University Press.
- [23] Stock, J., and Watson, M. 2002a. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics* 20 147-162.
- [24] Stock, J., and Watson, M. 2002b. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97: 1167-1179.

Table 1. Common factor estimation ( $T=50$ ,  $c=0.1$ ,  $d=0.1$ ,  $A=0.1$ )

Correlation within categories $\rho_l$	Same number of series in each category				Over sampling one category		
	$MSE_p^a$	$MSE_r^a$	$MSE^d$	$MSE_b^a$	$MSE_r^a$	$MSE^d$	$MSE_b^a$
Correlation across categories $\rho_s=0$							
0	0.101	0.195	0.144	0.172	0.191	0.173	0.210
0.1	0.101	0.192	0.148	0.176	0.191	0.175	0.217
0.5	0.101	0.196	0.143	0.188	0.190	0.174	0.225
0.9	0.101	0.195	0.195	0.231	0.192	0.330	0.337
Correlation across categories $\rho_s=0.1$							
0	0.116	0.207	0.143	0.190	0.204	0.164	0.218
0.1	0.116	0.205	0.145	0.191	0.204	0.166	0.221
0.5	0.116	0.205	0.158	0.201	0.203	0.181	0.229
0.9	0.116	0.206	0.207	0.240	0.202	0.321	0.339
Correlation across categories $\rho_s=0.5$							
0	0.223	0.289	0.238	0.284	0.285	0.236	0.283
0.1	0.223	0.286	0.240	0.285	0.284	0.238	0.283
0.5	0.223	0.286	0.250	0.288	0.284	0.248	0.289
0.9	0.223	0.287	0.287	0.309	0.284	0.313	0.335
Correlation across categories $\rho_s=0.75$							
0	0.350	0.383	0.348	0.399	0.382	0.342	0.392
0.1	0.350	0.380	0.349	0.398	0.383	0.343	0.391
0.5	0.350	0.381	0.356	0.392	0.376	0.350	0.389
0.9	0.350	0.377	0.381	0.396	0.378	0.379	0.394

Notes. The values of  $\rho_s$  determine the cross-correlation of the idiosyncratic shocks across categories, and the values of  $\rho_l$  determine the cross-correlation of the idiosyncratic shocks across the series of the same category.  $T$  is the sample size. Parameters  $A$  and  $c$  measure the serial correlation of the factor and the idiosyncratic shocks, respectively. The Mean Squared Errors of the model that uses the 10 series that generate the categories, the model that uses the 10 series with higher correlation with the series of the same category, the model that uses all the 100 series, and the model that uses the series selected by following Boivin and Ng (2006), are denoted by  $MSE_p^a$ ,  $MSE_r^a$ ,  $MSE^d$ , and  $MSE_b^a$ , respectively.

Table 2. Common factor estimation ( $T=50, c=0.1, d=0.1, A=0.5$ )

Correlation within categories $\rho_l$	Same number of series in each category				Over sampling one category		
	$MSE_p^a$	$MSE_r^a$	$MSE^d$	$MSE_b^a$	$MSE_r^a$	$MSE^d$	$MSE_b^a$
Correlation across categories $\rho_s=0$							
0	0.100	0.191	0.185	0.229	0.190	0.211	0.262
0.1	0.100	0.190	0.187	0.231	0.188	0.213	0.263
0.5	0.100	0.192	0.200	0.243	0.188	0.229	0.271
0.9	0.100	0.191	0.251	0.274	0.187	0.370	0.362
Correlation across categories $\rho_s=0.1$							
0	0.115	0.204	0.200	0.244	0.201	0.220	0.269
0.1	0.115	0.203	0.202	0.244	0.201	0.222	0.270
0.5	0.115	0.204	0.215	0.255	0.200	0.237	0.278
0.9	0.115	0.203	0.263	0.284	0.199	0.363	0.360
Correlation across categories $\rho_s=0.5$							
0	0.227	0.293	0.301	0.344	0.290	0.300	0.341
0.1	0.227	0.291	0.303	0.345	0.288	0.301	0.342
0.5	0.227	0.291	0.314	0.347	0.290	0.312	0.345
0.9	0.227	0.291	0.353	0.363	0.288	0.374	0.385
Correlation across categories $\rho_s=0.75$							
0	0.372	0.399	0.425	0.473	0.403	0.418	0.466
0.1	0.372	0.400	0.426	0.472	0.405	0.419	0.464
0.5	0.372	0.407	0.434	0.467	0.402	0.427	0.461
0.9	0.372	0.400	0.463	0.096	0.402	0.459	0.095

Notes. See notes of Table 1.



Table 3. Common factor estimation ( $T=50, c=0.1, d=0.1, A=0.75$ )

Correlation within categories $\rho_l$	Same number of series in each category				Over sampling one category		
	$MSE_p^a$	$MSE_r^a$	$MSE^d$	$MSE_b^a$	$MSE_r^a$	$MSE^d$	$MSE_b^a$
Correlation across categories $\rho_s=0$							
0	0.097	0.182	0.411	0.461	0.180	0.435	0.506
0.1	0.097	0.182	0.413	0.455	0.180	0.437	0.481
0.5	0.097	0.183	0.427	0.464	0.181	0.453	0.486
0.9	0.097	0.182	0.477	0.489	0.180	0.566	0.541
Correlation across categories $\rho_s=0.1$							
0	0.112	0.195	0.427	0.468	0.192	0.446	0.488
0.1	0.112	0.195	0.429	0.469	0.193	0.448	0.488
0.5	0.112	0.196	0.443	0.477	0.194	0.463	0.496
0.9	0.112	0.195	0.491	0.500	0.191	0.567	0.545
Correlation across categories $\rho_s=0.5$							
0	0.230	0.290	0.538	0.575	0.289	0.536	0.572
0.1	0.230	0.291	0.539	0.578	0.286	0.537	0.573
0.5	0.230	0.291	0.551	0.577	0.288	0.550	0.574
0.9	0.232	0.289	0.593	0.561	0.286	0.608	0.569
Correlation across categories $\rho_s=0.75$							
0	0.406	0.425	0.682	0.729	0.432	0.674	0.721
0.1	0.406	0.425	0.684	0.729	0.430	0.675	0.721
0.5	0.406	0.425	0.693	0.722	0.426	0.685	0.713
0.9	0.406	0.425	0.727	0.717	0.427	0.722	0.712

Notes. See notes of Table 1.

Table 4. Forecasting accuracy ( $T=50, c=0.1, d=0.1, A=0.1$ )

Correlation within categories $\rho_l$	Persistency of the target series $\gamma$	Same number of series in each category				Oversampling one category		
		$MSFE_p^a$	$MSFE_r^a$	$MSFE^d$	$MSFE_b^a$	$MSFE_r^a$	$MSFE^d$	$MSFE_b^a$
Correlation across categories $\rho_s = 0$								
0	0	1.107	1.215	1.212	2.097	1.171	1.238	2.099
	0.3	1.101	1.202	1.215	2.095	1.218	0.301	2.097
	0.8	1.086	1.172	1.200	2.074	1.178	1.229	2.075
0.9	0	1.107	1.354	1.301	2.094	1.378	1.388	2.104
	0.3	1.101	1.146	1.304	2.093	1.166	1.388	2.100
	0.8	1.086	1.273	1.295	2.069	1.318	1.374	2.076
Correlation across categories $\rho_s = 0.5$								
0	0	1.197	1.280	1.294	2.011	1.371	1.296	2.008
	0.3	1.200	1.248	1.298	2.006	1.324	1.299	2.003
	0.8	1.154	1.222	1.286	1.989	1.288	1.287	1.986
0.9	0	1.197	1.324	1.362	2.001	1.248	1.386	2.004
	0.3	1.200	1.320	1.366	1.996	1.441	1.388	1.999
	0.8	1.154	1.320	1.358	1.978	1.394	1.375	1.982

Notes. The estimated model is  $y_{t+1} = \beta F_t + \gamma y_t + e_{y,t+1}$ . See notes of Table 1.

Table 5. Forecasting accuracy ( $T=50, c=0.1, d=0.1, A=0.5$ )

Correlation within categories $\rho_l$	Persistency of the target series $\gamma$	Same number of series in each category				Oversampling one category		
		$MSFE_p^a$	$MSFE_r^a$	$MSFE^d$	$MSFE_b^a$	$MSFE_r^a$	$MSFE^d$	$MSFE_b^a$
Correlation across categories $\rho_s = 0$								
0	0	1.169	1.299	1.260	2.097	1.279	1.281	2.099
	0.3	1.121	1.203	1.253	2.090	1.289	1.274	2.091
	0.8	1.222	1.363	1.225	2.178	1.411	1.248	2.179
0.9	0	1.169	1.300	1.347	2.086	1.302	1.429	2.090
	0.3	1.121	1.313	1.341	2.079	1.306	1.418	2.082
	0.8	1.222	1.328	1.318	2.167	1.204	1.408	2.170
Correlation across categories $\rho_s = 0.5$								
0	0	1.220	1.290	1.342	2.104	1.291	1.343	2.102
	0.3	1.299	1.357	1.335	2.097	1.433	1.336	2.094
	0.8	1.218	1.382	1.313	2.186	1.374	1.313	2.182
0.9	0	1.220	1.259	1.416	2.089	1.238	1.435	2.091
	0.3	1.299	1.395	1.410	2.082	1.397	1.426	2.084
	0.8	1.218	1.357	1.396	2.170	1.293	1.415	2.172

Notes. See notes of Tables 1 and 4.

Table 6. Forecasting accuracy ( $T=50, c=0.1, d=0.1, A=0.75$ )

Correlation within categories $\rho_l$	Persistency of the target series $\gamma$	Same number of series in each category				Oversampling one category		
		$MSFE_p^a$	$MSFE_r^a$	$MSFE^d$	$MSFE_b^a$	$MSFE_r^a$	$MSFE^d$	$MSFE_b^a$
Correlation across categories $\rho_s = 0$								
0	0	1.133	1.275	1.385	2.607	1.213	1.400	2.613
	0.3	1.132	1.216	1.350	2.489	1.212	1.365	2.495
	0.8	1.151	1.345	1.301	2.773	1.27	1.321	2.776
0.9	0	1.133	1.201	1.462	2.286	1.273	1.515	2.284
	0.3	1.132	1.212	1.425	2.228	1.205	1.477	2.225
	0.8	1.151	1.243	1.387	2.511	1.27	1.461	2.504
Correlation across categories $\rho_s = 0.5$								
0	0	1.329	1.371	1.455	2.296	1.414	1.457	2.295
	0.3	1.373	1.449	1.419	2.238	1.498	1.420	2.237
	0.8	1.315	1.379	1.389	2.522	1.454	1.390	2.520
0.9	0	1.329	1.401	1.521	2.606	1.378	1.531	2.613
	0.3	1.373	1.410	1.485	2.487	1.428	1.494	2.494
	0.8	1.315	1.326	1.468	2.772	1.231	1.485	2.776

Notes. See notes of Tables 1 and 4.

Table 7. Data description

Category name	Number of series	Averaged cross-correlation	Representative series of the category	Highest averaged cross-correlation
1. Real output and income	19	0.422	Industrial production: total index	0.570
2. Employment and hours	25	0.323	Employees on nonagricultural Payrolls: total	0.475
3. Real retail, manufacturing and trade sales	9	0.381	Manufacturing & trade: total	0.623
4. Consumption	5	0.403	Personal consumption expend, total	0.640
5. Housing starts and sales	7	0.559	Housing starts: total farm & nonfarm	0.740
6. Real inventories and inventory-sales ratios	11	0.272	Manufacturing & trade inventories: total	0.426
7. Orders and unfilled orders	16	0.363	Mfg new orders: mfg industries with unfilled orders	0.435
8. Stock prices	7	0.476	S&P's common stock price index: composite	0.635
9. Exchange rates	5	0.515	United States effective exchange rate	0.701
10. Interest rates	15	0.427	Spread US treasury bills, secondary market 10-years and federal fund rate	0.517
11. Money and credit quantity aggregates	7	0.286	Money stock: M2	0.345
12. Price indexes	18	0.214	Cpi-u: all items	0.288
13. Average hourly earnings	2	0.313	Average hourly earnings of production workers: manufacturing	0.313
Total	146		13	

Notes. The dataset, the definition of the thirteen categories, and the distribution of the indicators across these categories follows the Stock and Watson (2002a). The representative series of each category is the economic indicator that exhibits the largest averaged correlation with the series of the same category. The last column reports these correlations.

Table 8. Forecasting real variables

	Industrial production			Nonagricultural employment		
	Forecast horizon			Forecast horizon		
	$h=1$	$h=6$	$h=12$	$h=1$	$h=6$	$h=12$
Forecast method	<i>RMSFE</i>			<i>RMSFE</i>		
<i>AR</i>	0.007	0.031	0.049	0.002	0.009	0.017
	Relative (to the <i>AR</i> ) <i>RMSFE</i>			Relative (to the <i>AR</i> ) <i>RMSFE</i>		
	Aggregate information					
Largest corr., $r=1$	0.96	0.96	0.92	0.92	0.89	0.86
Boivin-NG, $r=1$	0.98	0.99	1.03	0.98	1.28	1.21
Random, $r=1$	1.35	1.09	1.10	1.33	1.31	1.24
Largest corr., $r^*$ BIC	<b>0.87</b>	<b>0.73</b>	<b>0.52</b>	0.91	<b>0.78</b>	<b>0.63</b>
Boivin-NG, $r^*$ BIC	0.97	0.89	1.33	0.94	1.35	0.93
Random, $r^*$ BIC	1.47	1.09	1.15	1.42	1.31	1.29
	Disaggregate information					
$r=1$	0.88	0.81	0.85	<b>0.76</b>	0.88	0.88
$r^*$ Bai-Ng	<b>0.87</b>	0.79	0.75	0.76	0.79	0.76

Notes. The sample period is 1959:03-1998:12 and the out-of-sample forecast period is 1971:01-1998:12. The competing models are the autoregressive model, and the autoregressive model extended with factors as follows. First, the factors use the aggregate information of the 13 indicators that exhibit the largest average autocorrelation with the series of the same category, of the 13 indicators selected as in Boivin and Ng (2006), and of randomly picking one series from each category. Second, the factors use all the disaggregate information of the 146 indicators. In some cases, the number of factors is restricted to be  $r=1$  while in others the optimal number of factors  $r^*$  is determined by using BIC or Bai and Ng (2002).

Table 9. Forecasting nominal variables

	Consumer price index			Producer price index		
	Forecast horizon			Forecast horizon		
	<i>h=1</i>	<i>h=6</i>	<i>h=12</i>	<i>h=1</i>	<i>h=6</i>	<i>h=12</i>
Forecast method	<i>RMSFE</i>			<i>RMSFE</i>		
<i>AR</i>	0.002	0.010	0.021	0.008	0.026	0.046
	Relative (to the <i>AR</i> ) <i>RMSFE</i>			Relative (to the <i>AR</i> ) <i>RMSFE</i>		
	Aggregate information					
Largest corr., <i>r=1</i>	0.99	0.80	0.75	0.94	0.91	0.90
Boivin-NG, <i>r=1</i>	1.31	2.22	2.45	0.97	1.36	1.47
Random, <i>r=1</i>	1.87	1.87	1.80	1.01	1.15	1.22
Largest corr., <i>r*</i> BIC	0.99	0.92	0.86	1.00	0.95	0.88
Boivin-NG, <i>r*</i> BIC	1.43	1.79	2.21	0.99	1.22	1.28
Random, <i>r*</i> BIC	1.77	1.88	1.76	0.98	1.15	1.31
	Disaggregate information					
<i>r=1</i>	0.94	0.96	0.87	0.86	0.84	0.91
<i>r*</i> Bai-Ng	0.98	0.96	0.90	0.87	0.82	0.91

Notes. See notes of Table 9.