



Contents lists available at ScienceDirect

International Journal of Forecasting

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

Factor models for large and incomplete data sets with unknown group structure[☆]

Maximo Camacho^{a,*}, German Lopez-Buenache^b

^a Department of Quantitative Analysis, University of Murcia, Campus de Espinardo 30100 Murcia, Spain

^b Department of Business Administration, Catholic University of Murcia, Avenida de los Jeronimos, 135, 30107 Murcia, Spain



ARTICLE INFO

Keywords:

Big data
Factor models
Clustering
Missing data
Business cycles
Forecasting

ABSTRACT

Most economic applications rely on a large number of time series, which typically have a remarkable clustering structure and they are available over different spans. To handle these databases, we combined the expectation–maximization (EM) algorithm outlined by Stock and Watson (JBES, 2002) and the estimation algorithm for large factor models with an unknown number of group structures and unknown membership described by Ando and Bai (JAE, 2016; JASA, 2017). Several Monte Carlo experiments demonstrated the good performance of the proposed method at determining the correct number of clusters, providing the appropriate number of group-specific factors, identifying error-free group membership, and obtaining accurate estimates of unobserved missing data. In addition, we found that our proposed method performed substantially better than the standard EM algorithm when the data had a grouped factor structure. Using the Federal Reserve Economic Data FRED-QD, our method detected two distinct groups of macroeconomic indicators comprising the real activity indicators and nominal indicators. Thus, we demonstrated the usefulness of our group-specific factor model for studies of business cycle chronology and for forecasting purposes.

© 2022 The Authors. Published by Elsevier B.V. on behalf of International Institute of Forecasters. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Most economic empirical analyses are based on unusual data structures. Improvements in information technology allow empirical economists to deal with data sets containing hundreds of economic indicators, which can frequently lead to information overload without a clear method for organizing the data. These data are usually highly collinear, so factor analysis has become one of

the most appealing econometric approaches for managing these large dimensional data sets because they condense the common dynamics of the time series into a relatively small number of unobserved factors. This class of models has been successfully applied in economics, as shown in surveys by Bai and Ng (2008) and Stock and Watson (2011).

A second characteristic of economic data sets is that they are usually collected with incomplete statistical information. In these data sets, some time series are available over a diminished time span and observations may be missing at the beginning of the sample because data collection from their sources started on different dates. This so-called ragged edge problem also arises at the end of the sample due to differences in the publication lags among the variables characterizing the flow of economic information in real time.

Stock and Watson (2002) introduced an important computational method for handling unbalanced panels by

[☆] M. Camacho is grateful for the support from grant PID2019-107192GB-I00 funded by MCIN/AEI/10.13039/501100011033. We thank the two reviewers, associate editor, and editor in charge of our manuscript for their valuable comments. All remaining errors are our responsibility. Data and codes that replicate our results are available from the authors' websites.

* Corresponding author.

E-mail addresses: mcamacho@um.es (M. Camacho), glopez@ucam.es (G. Lopez-Buenache).

extending the factor model with an iterative expectation–maximization (EM) algorithm. Briefly, the algorithm is initialized with an estimate of the factors from a balanced panel where missing data are replaced by initial guesses. The factors are then used to provide estimates of the missing observations. The process is iterated until the estimates do not change substantially. In practice, the EM algorithm can also handle outliers if they are replaced by missing values in the data set. In empirical applications, [Bernanke and Boivin \(2003\)](#), [Angelini et al. \(2006\)](#), [Schumacher and Breitung \(2008\)](#), and [Marcellino and Schumacher \(2010\)](#) dealt satisfactorily with missing observations by applying an EM algorithm in a principal components framework.¹

The final characteristic of the economic data sets is that they typically have a remarkable clustering structure. For example, group structures appear in macroeconomic indicators because they usually either contain sectoral splits of different categories or they are classified into real and nominal data. Group structures also appear in international data sets, which usually contain country region-specific data that exhibit strong cross-correlation among series in the same region and they are separated from the cross-correlations observed in other regions.

Within this framework, it is reasonable to consider that a model that uses group-specific factors will describe group co-movements better than a model that only uses global common factors. Factor models with block structures were first studied based on data sets organized into blocks using a priori information. Typically, these methods consider a panel of variables with a set of common pervasive factors that affect all of the variables in all groups, and group-specific pervasive factors that only affect the variables in a specific group. Examples of these methods were proposed by [Gregory et al. \(1997\)](#), [Kose et al. \(2003\)](#), [Crucini et al. \(2011\)](#), [Hallin and Liška \(2011\)](#), [Ando and Bai \(2015\)](#), [Breitung and Eickmeier \(2016\)](#), [Rodríguez-Caballero and Ergemen \(2017\)](#), and [Choi et al. \(2018\)](#). However, [Ando and Bai \(2016\)](#) developed a new procedure designed to explore group structures in large factor models with distinctive features comprising endogenously assigning variables to groups, determining the number of groups, and estimating the group-specific factors.² [Ando and Bai \(2017\)](#) extended their method to include global common factors and group-specific factors.³

Despite the development of new methods for handling economic data sets mentioned above, a comprehensive approach is not available for dealing with large panels with group structures, unknown membership, and

missing observations, which is the norm rather than the exception in empirical applications. This deficiency considerably limits the potential use of factor models for characterizing the co-movement of economic variables. However, considering the missing observations in standard factor models by using an EM algorithm, as described by [Stock and Watson \(2002\)](#), suggests that the iterative updated estimates of the missing observations for a variable i that belongs to group g are provided by the expectation conditional only on global factors, which is clearly misleading under the premise of group heterogeneity.

In addition, the algorithms proposed by [Ando and Bai \(2016, 2017\)](#) require data sets where all of the panel members are observed in every time period. Thus, considering group membership implies removing observations without a full set of data, thereby leading to the computation of inferences based on shorter cross sections (when complete columns are removed from the panel) and/or shorter time series (when complete rows are removed from the panel). This approach could potentially lead to high costs due to a loss of efficiency when omitting data with highly valuable information, or due to bias when the data are not missing at random. In addition, handling missing data is particularly challenging for real-time forecasters because real-time data routinely exhibit end-of-sample ragged edge problems as some variables are observed with longer delays than others.

Due to these limitations, we aimed to improve factor models by integrating the two approaches described above. In particular, we developed an algorithm for dealing with unbalanced panels of data with cluster structures and potentially containing outliers. Our proposed method uses the Stock–Watson EM algorithm to consider missing data while the Ando–Bai algorithm automatically determines the number of groups, group membership, number of global factors, and number of group-specific factors. Our method only applies a C_p -type criterion for model selection, which is minimized by implementing a recursive algorithm for endogenously dealing with missing data by using a group-specific EM algorithm. Based on this criterion, our method selects from three factor model alternatives: only global factors, as in the method of [Stock and Watson \(2002\)](#), global and group-specific factors, as in the method of [Ando and Bai \(2017\)](#), and only group-specific factors, as in the method of [Ando and Bai \(2016\)](#).

In Monte Carlo experiments designed to capture some basic data problems that characterize economic data sets, we evaluated the ability of the proposed algorithm for selecting models with the most appropriate number of groups, true number of group-specific factors, convenient group membership, and suitable inference of the missing data. Among the data generating processes, our method includes non-homoskedastic errors with cross-sectional dependence, errors with some serial correlations, and factors with temporal dependence. Overall, our results indicate that despite the inclusion of missing observations to simulate unbalanced data sets, no significant differences were found between the clustering abilities of our proposed method and the group factor models of [Ando and Bai \(2016, 2017\)](#), which require balanced data sets. In addition, we found that our group-specific inference method

¹ In recent studies, [McCracken and Ng \(2021\)](#), and [Jin et al. \(2021\)](#) proposed extensions of the EM algorithm based on principal component frameworks. Comparisons of these approaches with the method described by [Stock and Watson \(2002\)](#) are left for further research.

² The method proposed by [Ando and Bai \(2016\)](#) does not consider common factors that affect all of the variables. However, their method permits a large number of potential observable factors.

³ [Alonso et al. \(2020\)](#) proposed an alternative procedure for building dynamic factor models with a cluster structure that considers common global factors and group-specific factors, but this method was not investigated in the present study.

for missing data performed better than the imputation of missing data conditional on the full set of common factors in the standard EM algorithm described by Stock and Watson (2002).

Furthermore, we implemented an empirical application to illustrate how our proposed method can be used by practitioners to identify a model for an unbalanced large set of observed variables in terms of a smaller number of underlying latent factors characterized by group structures and unknown membership. We employed the novel quarterly Federal Reserve Economic Data FRED-QD data set described by McCracken and Ng (2021), which comprises a set of 248 economic indicators from the United States. This data set provides researchers with access to a regularly updated version of the data used by Stock and Watson (2012), which comprise the quarterly version of the monthly data set used by Stock and Watson (2002). McCracken and Ng (2021) acknowledged that the data set contains missing values in 38 out of 248 cases, and thus about 15% of the indicators are incomplete because their samples started late. In addition, McCracken and Ng (2021) noted that 30 outliers are present in the FRED-QD data set, mostly among the bank reserves variables.

Our results indicated that the quarterly FRED-QD data set could be characterized better by a factor model that admits only group-specific factors rather than by factor models that also include some pervasive factors common to all variables or by a factor model with only common factors. In particular, we identified two distinctive groups of economic indicators, which we interpreted as the real activity group and the nominal group. The selected factors explained 50 percent and almost 40 percent of the variance in the indicators that belonged to the real activity group and nominal group, respectively. We demonstrated the high empirical reliability of the group-specific EM algorithm at filling in the gaps in missing observations for both the nominal and real activity indicators. In addition, our results suggest a very promising role for the estimated factors in studies of the US business cycle chronology.

Finally, we evaluated the forecasting performance of the proposed method based on an out-of-sample forecasting exercise from 2006:1 to 2019:4. In this experiment, the group-specific factor model obtained meaningful forecasting improvements compared with the Stock–Watson factor model. The improvements were uniformly observed over the forecasts for 248 economic indicators and they were substantially large for many real economic activity indicators. These out-of-sample results are promising but they do not allow data revisions; however, the forecasting performance could be reassessed in the future as more real-time data sets become available and real-time assessments become feasible.

The remainder of this paper is organized as follows. In Section 2, we describe the model selection algorithm and estimation process. In Section 3, we provide a description of the Monte Carlo experiments conducted to analyze the ability of the proposed method to select the correct number of groups, group membership, and number of factors, and to conduct inference based on missing data. In Section 4, we present the results of studies conducted

to examine the ability of the new algorithm to transform the information content in the FRED-QD data set into a smaller number of group-specific factors, which could be used for business cycle inference and to perform forecasts. In Section 5, we give our conclusions and suggestions for future research.

2. Grouped factor models with missing data

2.1. Model formulation

Let $t = 1, \dots, T$ be the time index, S is the number of groups, N_1, \dots, N_S is the number of indicators in each group observed for a sample of T observations, and $N = \sum_{s=1}^S N_s$ is the total number of indicators. Let us assume that the number of groups is finite and independent of N and T , and that each group contains a proportion of indicators, $p_s = N_s/N$, with $s = 1, \dots, S$, which is bounded between a minimum p_{min} and a maximum p_{max} .

The value of the i th indicator x_{it} observed at time t belonging to group $g_i \in \{1, \dots, S\}$ from a collection of stationary time series data $\{x_{it}\}_{i=1, \dots, N}$ that admits a grouped factor representation is expressed as follows

$$x_{it} = f'_{ct} \lambda_{ci} + f'_{g_i t} \lambda_{g_i i} + \epsilon_{it}. \tag{1}$$

The $r \times 1$ vector λ_{ci} collects the factor loadings that measure the unknown sensitivity of x_i to the unobservable common factors r factors collected in the $r \times 1$ vector f_{ct} , which affect all indicators in all groups. The $r_{g_i} \times 1$ vector $\lambda_{g_i i}$ collects the factor loadings that measure the unknown sensitivity of x_i to the unobservable group-specific r_{g_i} factors collected in the $r_{g_i} \times 1$ vector $f_{g_i t}$, which only affect the indicators for group g_i .⁴ As suggested by Choi et al. (2018), we assume that the common factors and group-specific factors are orthogonal, but correlations between factors in different groups are allowed, although they cannot be perfectly correlated. In this expression, ϵ_{it} is the unit-specific error, which is independent of f_{ct} and $f_{g_i t}$ for all i and t . First, we assume that $E(\epsilon_{it}) = 0$ and $var(\epsilon_{it}) = \sigma_i^2$, and that it is independent over i and t .⁵

The group membership is unknown but the time series are labeled through the unobservable state indicator g_i in the whole sequence of realizations, which are collected in $G = (g_1, \dots, g_N)$. For example, $g_i = s$ indicates that the time series x_i belongs to group s . It is useful to represent the conditional factor representation of the T observations of this i th indicator in matrix notation

$$x_i = F_c \lambda_{ci} + F_s \lambda_{si} + \epsilon_i, \tag{2}$$

where $x_i = (x_{i1}, x_{i2}, \dots, x_{iT})'$; $F_c = (f_{c1}, f_{c2}, \dots, f_{cT})'$ and $\lambda_{ci} = (\lambda_{ci}^1, \lambda_{ci}^2, \dots, \lambda_{ci}^r)'$ are the $T \times r$ matrix of common factors and the $r \times 1$ vector of common factor loadings; $F_s = (f_{s1}, f_{s2}, \dots, f_{sT})'$ and $\lambda_{si} = (\lambda_{si}^1, \lambda_{si}^2, \dots, \lambda_{si}^{r_s})'$ are the $T \times r_s$ matrix of group-specific factors and the $r_s \times 1$

⁴ To eliminate scale effects, the time indicators are standardized with zero mean and unit variance.

⁵ In the simulations, we examined the potential effects of cross-sectional dependent and serially correlated errors as well as the effects of serially correlated factors.

vector of group-specific factor loadings for group s ; and $\epsilon_i = (\epsilon_{i1}, \epsilon_{i2}, \dots, \epsilon_{iT})'$.

Given the group membership G and common factor structures $F_c \lambda_{ci}$, we define the variable $z_i^s = x_i - F_c \lambda_{ci}$ for each group $s = 1, \dots, S$. If we collect the N_s variables z_i^s that belong to group s in the $T \times N_s$ matrix $Z_s = (z_1, \dots, z_{N_s})$, we can obtain the estimates based on the principal components for the group-specific factors, \hat{F}_s , subject to the normalization $F_s' F_s / T = I_{r_s}$, as \sqrt{T} times the eigenvectors corresponding to the r_s largest eigenvalues of the $T \times T$ matrix $Z_s Z_s'$. If we collect the group-specific factor loadings of this group in the $r_s \times N_s$ matrix $\Lambda_s = (\lambda_{s1}, \dots, \lambda_{sN_s})$, the estimates of the factor loadings can be obtained as $\hat{\Lambda}_s = \hat{F}_s' Z_s / T$. In addition, given the group membership G and group-specific factor structures $F_s \lambda_{si}$, we define the variable $z_i^c = x_i - F_s \lambda_{si}$ for each group $s = 1, \dots, S$. If we collect the N variables z_i^c in the $T \times N$ matrix $Z_c = (z_1, \dots, z_N)$, we can obtain the estimates based on the principal components for the common factors, \hat{F}_c , subject to the normalization $F_c' F_c / T = I_r$, as \sqrt{T} times the eigenvectors corresponding to the r largest eigenvalues of the $T \times T$ matrix $Z_c Z_c'$. If we collect the common factor loadings in the $r \times N$ matrix $\Lambda_c = (\lambda_{c1}, \dots, \lambda_{cN})$, the estimates of the factor loadings can be obtained as $\Lambda_c = \hat{F}_c' Z_c / T$.

To consider missing values in the data, we assume that not all the observations are available in x_i . Then, we define x_i^{obs} as the $T^{obs} \times 1$ vector that contains the observations available for variable i , which is only a subset of x_i when $T > T^{obs}$. As shown in the following, a key step during the EM iterations is mapping from the full set x_i to the observed data x_i^{obs} . For a known $T^{obs} \times T$ matrix A_i , this relationship is defined by

$$x_i^{obs} = A_i x_i, \tag{3}$$

where A_i is the identity matrix of size T with the $T - T^{obs}$ rows corresponding to the missing observations in x_i removed. When x_i is fully observed, A_i is simply the identity matrix.

2.2. Model estimation

In this section, we present an estimation method for the model by combining the EM algorithm proposed by [Stock and Watson \(2002\)](#) to deal with missing observations and the algorithm described by [Ando and Bai \(2017\)](#) for identifying group membership and estimating common and group-specific factors.⁶ Given the number of groups, S , the number of common factors, r , and the number of group-specific factors, r_s , with $s = 1, \dots, S$, the least squares estimator of $F = \{F_c, F_1, \dots, F_S\}$ and $\Lambda = \{\Lambda_c, \Lambda_1, \dots, \Lambda_S\}$ is defined as the minimizer of the objective function:

$$V(F, \Lambda) = \sum_{i=1}^N \sum_{t=1}^T I_{it} (x_{it} - f_{ct}' \lambda_{ci} - f_{g_i t}' \lambda_{g_i i})^2, \tag{4}$$

⁶ It should be noted that in contrast to [Ando and Bai \(2017\)](#), we do not allow observed factors.

where $I_{it} = 1$ if x_{it} is observable at t , but 0 otherwise.

Algorithm 1 Model estimation: minimization of $V(F, \Lambda)$

Set $S, r, \{r_s\}_{s=1}^S$ and a naive guess of the missing values

Initialization $x_{it}^{(0)}, G^{(0)}, \{F_c^{(0)}, \Lambda_c^{(0)}\}$, and $\{F_s^{(0)}, \Lambda_s^{(0)}\}_{s=1}^S$

while $m \leq M$ **do** (*)

Update membership $G^{(m)}$ given $x_{it}^{(m-1)}, \{F_c^{(m-1)}, \Lambda_c^{(m-1)}\}$ and $\{F_s^{(m-1)}, \Lambda_s^{(m-1)}\}_{s=1}^S$ using (5)

for $j \leq J$ **do** (*)

Expectation of $x_{it}^{(m)}$ given $G^{(m)}, \{F_c^{(m-1)}, \Lambda_c^{(m-1)}\}$ and $\{F_s^{(m-1)}, \Lambda_s^{(m-1)}\}_{s=1}^S$ using (6)

Maximization 1 PC of $\{F_c^{(m)}, \Lambda_c^{(m)}\}$ given $G^{(m)}, x_{it}^{(m)}$ and $\{F_s^{(m-1)}, \Lambda_s^{(m-1)}\}_{s=1}^S$

Maximization 2 PC of $\{F_s^{(m)}, \Lambda_s^{(m)}\}_{s=1}^S$ given $G^{(m)}, x_{it}^{(m)}$ and $\{F_c^{(m)}, \Lambda_c^{(m)}\}$

end for

end while

(*) Stop when convergence is achieved. PC denotes the principal components.

Our algorithm is described in Algorithm 1. The first step requires initial (naive) guesses of the missing observations.⁷ These guesses together with the available data comprise a balanced data set for estimating the initial factors and loadings using standard principal components for a fixed number of common factors. The standard version of the EM algorithm can then be applied to the balanced panel to provide initial estimates of the missing values in the time series and a first guess of the balanced data set $x^{(0)}$ while ignoring the group-specific factor structures. Next, we require an initialization of the group membership G . For the sake of simplicity, we use the K -means algorithm, which separates the data set into S clusters and provides an initial membership $G^{(0)} = (g_1^{(0)}, \dots, g_N^{(0)})$. Finally, we obtain starting values of $\{F_c^{(0)}, \Lambda_c^{(0)}\}$ based on the principal components on x_{it} , and $\{F_s^{(0)}, \Lambda_s^{(0)}\}$ for $s = 1, \dots, S$ based on the group-specific principal components on $x_{it} - F_c \lambda_{ci}$.

In a given iteration (m), the second step involves providing an optimal assignment for each indicator $G^{(m)} = (g_1^{(m)}, \dots, g_N^{(m)})$ given the common and group-specific factor structure of the previous iteration, $\{F_c^{(m-1)}, \Lambda_c^{(m-1)}\}$ and $\{F_s^{(m-1)}, \Lambda_s^{(m-1)}\}$ for $s = 1, \dots, S$. For this purpose, we assign each indicator to the group with the best in-sample

⁷ Initially, we replace missing observations by random draws from the standard normal distribution.

fit

$$g_i^{(m)} = \underset{s \in \{1, \dots, S\}}{\operatorname{argmin}} \sum_{i=1}^N \sum_{t=1}^T \left(x_{it} - f_{ct}^{(m-1)} \lambda_{ci}^{(m-1)} - f_{st}^{(m-1)} \lambda_{si}^{(m-1)} \right)^2, \tag{5}$$

which implies that the time series x_i is assigned to group s if it minimizes the sum of squared residuals among the S possible groups.

In the third step, given the group membership $G^{(m)}$, the common factor structures, $\{F_c^{(m)}, \Lambda_c^{(m)}\}$, and group-specific factor structures, $\{F_s^{(m)}, \Lambda_s^{(m)}\}$ for $s = 1, \dots, S$, are updated with an iterative group-specific EM algorithm by alternating the expectation of $x_i^{(m)}$ with respect to $\{F_c^{(m)}, \Lambda_c^{(m)}\}$ and $\{F_s^{(m)}, \Lambda_s^{(m)}\}$ and the optimization of $\{F_c^{(m)}, \Lambda_c^{(m)}\}$ and $\{F_s^{(m)}, \Lambda_s^{(m)}\}$ given $x_i^{(m)}$. In the expectation step of the j th iteration, missing observations for each variable are updated by the expectation of x_{it} conditional on the observations available for i and the common and group-specific factors and loadings from the previous iteration. If we define $F \lambda_i^{(m(j-1))} = F_c^{(m(j-1))} \lambda_{ci}^{(m(j-1))} + F_s^{(m(j-1))} \lambda_{si}^{(m(j-1))}$, the missing observations are updated by using

$$\hat{x}_i^{(mj)} = F \lambda_i^{(m(j-1))} + A_i' (A_i A_i')^{-1} \left(X_i^{obs} - A_i F \lambda_i^{(m(j-1))} \right), \tag{6}$$

where $s = 1, \dots, S$. Then, the elements of the estimated balanced panel are constructed as $\hat{x}_{it}^{(mj)} = x_{it}$ if x_{it} is observed, but $\hat{x}_{it}^{(mj)} = f_{ct}^{(m(j-1))} \lambda_{ci}^{(m(j-1))} + f_{git}^{(m(j-1))} \lambda_{gi}^{(m(j-1))}$ otherwise.

After obtaining the updated balanced data set, the maximization step involves re-estimating the common factors and loadings, $\{F_c^{(mj)}, \lambda_{ci}^{(mj)}\}$, based on the principal components on $\hat{x}_{it}^{(mj)} - f_{git}^{(m(j-1))} \lambda_{gi}^{(m(j-1))}$. In addition, the maximization step re-estimates the group-specific factor and loadings, $\{F_s^{(mj)}, \lambda_{si}^{(mj)}\}$ based on the principal components on $\hat{x}_{it}^{(mj)} - f_{ct}^{(m(j-1))} \lambda_{ci}^{(m(j-1))}$ for each group $s = 1, \dots, S$. The expectation and maximization steps are iterated until the maximum percentage changes in the estimates of the variables are smaller than a convergence tolerance δ_{EM} . This algorithm provides estimates of the missing values in the time series and estimates of the common and group-specific factors and loadings.

To obtain the minimizer of $V(F, \Lambda)$, we iterate the second and third steps until convergence. In practical implementations, we stop the iterations once

$$100 * \left(V(F^{(m)}, \Lambda^{(m)}) - V(F^{(m-1)}, \Lambda^{(m-1)}) \right) / V(F^{(m-1)}, \Lambda^{(m-1)}) < \delta_V, \tag{7}$$

for some absolute convergence tolerance δ_V .

Although we do not consider observed factors, this multi-level factor structure that admits common and group specific factors generalizes the method of Ando and Bai (2017) to deal with missing observations. In a similar manner, our proposed method can be readily simplified to generalize the method of Ando and Bai (2016), who advocated a model with completely separated group-specific factor structures in different clusters without pervasive factors common to all variables. In this simplified algorithm, which we call Algorithm 1', the term $F_c \lambda_{ci}$ does

not appear in expression (4). In addition, the estimation method outlined in Algorithm 1 is still valid, although Algorithm 1' does not include $\{F_c, \Lambda_c\}$ or maximization step 1.

2.3. Model selection

We assume that the number of groups S , number of common factors, r , and number of group-specific factors, $\{r_1, \dots, r_S\}$, are known, but in practice, we are faced with the problem of estimating these quantities from the data. Thus, we adapt the approach proposed by Ando and Bai (2017) to deal with this model specification uncertainty.

Based on the method described by Ando and Bai (2017), we propose selecting these unknown parameters as minimizers for the panel information criterion (PIC)

$$PIC(s, r, r_1, \dots, r_S) = \frac{1}{TN} \sum_{i=1}^N \sum_{t=1}^T I_{it} \left(x_{it} - f_{ct} \lambda_{ci} - f_{git} \lambda_{gi} \right)^2 + \sigma^2 r \left(\frac{T+N}{TN} \right) \log(TN) + \sigma^2 \sum_{s=1}^S r_s \left(\frac{N_s}{N} \right) \left(\frac{T+N_s}{TN_s} \right) \log(TN_s). \tag{8}$$

Thus, model selection involves choosing minimizers of the distances between the actual data and factor estimation measured based on the sum squared error over the observed data, $V(F, \Lambda)$, which is penalized by redundant model flexibility. In particular, the second and third terms penalize overestimations of the number of common and group-specific factors, where σ^2 provides the scaling for the penalty term.⁸

In practice, the number of groups, number of common factors, and number of group-specific factors can be determined with the model search algorithm described by Algorithm 2. The first step requires fixing a maximum number of groups, S_{max} , maximum number of common factors, r_{max} , and maximum number of group-specific factors, $r_{max,j}$. Thus, the potential number of groups is $\{S_1, \dots, S_{max}\}$, the potential number of common factors is $\{r_1, \dots, r_{max}\}$, and the potential number of group-specific factors for group s is $\{r_{1,s}, \dots, r_{max,s}\}$, with $s = 1, \dots, S$ and $S = S_1, \dots, S_{max}$.⁹ For the first value of the number of groups, S_1 , we optimize the number of common and group-specific factors by minimizing PIC . Next, we repeat the second step for $\{S_2, \dots, S_{max}\}$ and choose the number of groups (and its corresponding number of common and group-specific factors) that minimizes the value of PIC .

This algorithm generalizes the model selection algorithm proposed by Ando and Bai (2017) to deal with missing observations. Our proposed algorithm can also be

⁸ In practical implementations, σ^2 can be obtained as the mean squared error based on the maximum number of groups and the maximum number of common and group-specific factors.

⁹ In the applications developed in this study, we set $S_1 = 1$, which implies that $S \in \{1, 2, \dots, S_{max}\}$.

Algorithm 2 Model selection: minimization of $PIC(s, r, r_1, \dots, r_s)$

```

Set  $S_{max}, r_{max}, \{r_{max,s}\}_{s=1}^S$  for  $S = S_1, \dots, S_{max}$ 
while  $S \leq S_{max}$  do
    for  $r \leq r_{max}$  and  $r_s \leq r_{max,s}$  do
        Run Algorithm 1
    end for
    Select  $r^*$  and  $\{r_s^*\}_{s=1}^S$  that minimize (8). Store the
    resulting  $PICs$  for each  $S \leq S_{max}$ 
end while
Choose  $S^*$  that minimizes the stored  $PICs$ 
    
```

readily simplified to generalize the model selection algorithm described by Ando and Bai (2016), which was designed for a factor model that admits only group-specific factors. The simplified algorithm, which we call Algorithm 2', omits the second term in the right-hand side of expression (8) and expressions r_{max} and r^* are dropped from the algorithm.

In empirical applications, the framework described above is useful for choosing among three alternative factor structures: only common factors as described by Stock and Watson (2002), common and group-specific factors as described by Ando and Bai (2017), or only group specific-factors as described by Ando and Bai (2016). For a given data set, the discrimination method involves running Algorithm 1 and Algorithm 2 to estimate a factor model that admits common and group-specific factors, thereby leading to an optimal value of the model selection criterion, $PIC_c(S_c^*, r_c^*, r_{c1}^*, \dots, r_{cS_c^*}^*)$. In addition, the method involves running the modifications of Algorithm 1' and Algorithm 2', which do not allow for common factors, thereby obtaining an optimal value of the model selection criterion, $PIC_g(S_g^*, r_{g1}^*, \dots, r_{gS_g^*}^*)$.

If the optimal number of groups is $S_c^* = 1$, then we can infer that the data generation process does not admit a grouped factor structure because the factors are common to all o the time series. If $S_c^* > 1$, then we can infer that the data generation process admits common and group-specific factors when $PIC_c(S_c^*, r_c^*, r_{c1}^*, \dots, r_{cS_c^*}^*) < PIC_g(S_g^*, r_{g1}^*, \dots, r_{gS_g^*}^*)$. However, we propose a factor structure with factors that are common only within some clusters of variables and excluding pervasive factors common to all variables when $PIC_g(S_g^*, r_{g1}^*, \dots, r_{gS_g^*}^*) < PIC_c(S_c^*, r_c^*, r_{c1}^*, \dots, r_{cS_c^*}^*)$.

3. Simulation experiments

We conducted several Monte Carlo experiments to examine the small sample performance of our proposed method at determining the correct number of clusters, providing the appropriate number of common and group-specific factors, achieving error-free group membership, and producing accurate estimates of unobserved missing data compared with the Stock and Watson (2002) EM algorithm. To facilitate comparisons, we designed the data

generation processes according to Ando and Bai (2016, 2017).

In our simulation experiments, we generated a total of 500 sets of N_s idiosyncratic components ϵ_i of length T in each group $s = 1, 2, \dots, S$. According to Ando and Bai (2017), we set the number of groups $S = 3$, only one common factor, $r = 1$, the same number of group-specific factors in each group, $r_1 = r_2 = r_3 = 3$, and the same number of time series in each group, $N_1 = N_2 = N_3 = N$. The common factor f_{ct} was drawn from a uniform distribution on $[0,1]$ and each element of the factor-loading matrix Λ_c followed the uniform distribution $[-2,2]$. Each element of the $r_s \times 1$ group-specific factors f_{st} was drawn from $N(0, 1)$, and in order to generate group heterogeneity, each element of the $r_s \times 1$ vector of group-specific factor loadings λ_{si} was drawn from $N(0.5 \times s, 1)$, for $s = 1, 2, 3$.

To examine the relative performance of the model according to the number of variables and number of observations, we generated times series with a grouped factor structure $x_i = F_c \lambda_{ci} + F_s \lambda_{si} + \epsilon_i$ for several combinations of (N, T) : (50, 150), (100, 150), (50, 250), (100, 250). These settings are similar to the standard data sets used in economics and they include combinations similar to the cross-section and time-series dimensions of our empirical application.

First, we generated the data according to the processes described above but we replaced a specific fraction of the data with missing values. To examine the effect of the magnitude of these missing values on the model's performance, we considered that fractions of the time series comprising $P = 10\%$ and $P = 20\%$ had a certain amount of missing data. For these time series, we set amounts of $M = 50$ and $M = 83$ missing data (out of a total of T observations).¹⁰ Again, these settings replicated the standard situations in economic applications, including our empirical example.

For the sets of simulations with different data generation processes, we applied the grouped factor model with missing data described in this study by allowing for a maximum number of four groups and a maximum number of eight common factors and eight group-specific factors. The iterative method for computing successive approximations to the final model estimates stopped when the magnitude of the percentage difference between two consecutive values of $V(F, \Lambda)$ was below the convergence tolerance $\delta_V = 10^{-3}$. Similarly, the group-specific EM algorithm stopped when the maximum percentage change in the estimates of the variables was smaller than $\delta_{EM} = 10^{-3}$.

Our Monte Carlo experiments were designed to focus on the effects of heteroskedasticity, cross-correlated and

¹⁰ In the (N, T, P, M) combinations (50, 150, 10, 50), (100, 150, 10, 50), (50, 250, 10, 83), and (100, 250, 10, 83), the missing data represented 3.3% of the total data. This percentage increased to 6.7% in the combinations (50, 150, 20, 50), (100, 150, 20, 50), (50, 250, 20, 83), and (100, 250, 20, 83). Following a reviewer's suggestion, we enlarged the proportion of missing values and found a slight deterioration with up to 20% missing values and a moderate deterioration with 25% missing values. The detailed results are available from the authors upon request.

Table 1
Simulation: accuracy of the model.

N	T	P	M	S		r		r ₁		r ₂		r ₃		N ₁		N ₂		N ₃	
				U	C	U	C	U	C	U	C	U	C	U	O	U	O	U	O
Panel A: Errors are homoskedastic and serially uncorrelated																			
50	150	10	50	5.8	86.8	0	100	0	82.2	0	84.8	0	82.0	1.4	1.3	1.2	1.3	1.4	1.3
50	150	20	50	6.5	84.8	0	100	0	81.8	0	83.0	0	77.4	2.0	2.1	1.8	2.0	2.1	1.8
100	150	10	50	1.6	94.2	0	100	0	92.8	0	92.4	0	93.4	0.8	0.9	0.7	0.7	0.8	0.7
100	150	20	50	2.2	93.2	0	100	0	91.2	0	90.4	0	90.4	1.3	1.1	1.2	1.3	1.0	1.2
50	250	10	84	6.8	84.8	0	100	0	89.1	0	83.2	0	81.8	1.9	1.7	2.1	1.5	1.9	1.7
50	250	20	84	8.6	82.2	0	100	0	80.1	0	80.6	0	80.1	2.1	2.2	2.0	1.9	2.2	2.1
100	250	10	84	1.8	93.2	0	100	0	93.8	0	92.8	0	91.8	0.9	0.9	0.8	0.8	0.9	0.8
100	250	20	84	3.2	91.0	0	100	0	89.4	0	90.2	0	90.0	1.2	1.2	1.2	1.2	1.1	1.3
Panel B: Errors are nonhomoskedastic and cross-sectionally correlated																			
50	150	10	50	6.8	81.6	0	100	0	81.8	0	80.6	0	80.6	1.8	1.7	2.2	2.1	1.9	1.8
50	150	20	50	7.2	80.2	0	100	0	80.4	0	78.8	0	80.1	1.9	2.1	2.1	1.9	2.1	1.9
100	150	10	50	2.8	90.6	0	100	0	91.6	0	91.8	0	90.2	1.2	0.9	0.8	1.1	0.9	1.1
100	150	20	50	3.2	89.4	0	100	0	89.4	0	89.6	0	89.8	1.6	1.1	1.3	1.5	1.2	1.3
50	250	10	84	7.1	79.8	0	100	0	80.1	0	80.6	0	80.1	1.9	1.8	2.1	2.1	1.9	1.9
50	250	20	84	8.6	76.8	0	100	0	76.4	0	78.6	0	78.8	2.1	2.1	1.9	2.1	2.1	1.9
100	250	10	84	4.6	87.8	0	100	0	89.6	0	90.2	0	89.6	1.3	1.2	1.1	0.8	1.1	0.2
100	250	20	84	6.2	84.8	0	100	0	87.8	0	88.6	0	86.8	1.3	1.3	1.2	1.9	1.3	1.6
Panel C: Errors have some serial and cross-sectional correlations																			
50	150	10	50	7.2	79.8	0	100	0	80.8	0	78.8	0	80.1	1.9	1.9	2.1	2.2	1.9	2.0
50	150	20	50	8.1	77.8	0	100	0	78.6	0	78.2	0	77.5	2.1	2.2	2.3	2.1	2.2	1.9
100	150	10	50	3.6	87.2	0	100	0	90.1	0	90.4	0	90.6	1.1	1.1	0.9	0.8	1.2	1.1
100	150	20	50	4.1	83.4	0	100	0	87.6	0	88.2	0	90.2	1.3	1.2	1.4	1.5	1.1	1.4
50	250	10	84	8.6	78.8	0	100	0	78.8	0	81.2	0	80.2	1.9	2.0	2.0	2.1	2.1	1.9
50	250	20	84	9.4	72.8	0	100	0	75.2	0	74.8	0	76.2	2.2	2.1	2.0	2.2	1.9	2.1
100	250	10	84	5.2	83.6	0	100	0	90.1	0	89.2	0	89.8	1.2	1.1	1.2	0.9	1.9	1.1
100	250	20	84	6.4	81.2	0	100	0	86.8	0	88.6	0	84.6	1.4	1.4	1.1	1.1	1.4	1.8
Panel D: Dynamic factors																			
50	150	10	50	7.8	79.4	0	100	0	80.2	0	79.4	0	80.0	1.8	2.1	2.0	2.1	1.9	1.9
50	150	20	50	8.2	76.4	0	100	0	77.8	0	77.2	0	77.4	2.2	2.1	2.4	2.2	2.1	2.0
100	150	10	50	3.8	86.8	0	100	0	90.4	0	90.4	0	90.6	0.9	1.2	1.8	0.9	1.3	1.0
100	150	20	50	4.2	82.2	0	100	0	87.6	0	89.4	0	87.2	1.2	1.4	1.5	1.6	1.2	1.3
50	250	10	84	8.6	78.6	0	100	0	80.2	0	80.1	0	80.0	1.8	2.1	2.1	1.9	1.9	1.9
50	250	20	84	9.6	72.4	0	100	0	77.2	0	74.2	0	74.6	2.3	2.2	2.1	1.9	1.8	2.1
100	250	10	84	5.8	82.2	0	100	0	89.6	0	88.2	0	88.8	1.1	1.3	1.3	1.2	0.9	1.3
100	250	20	84	6.2	80.6	0	100	0	87.2	0	87.2	0	87.6	1.5	1.6	1.2	1.1	1.3	1.4

Notes. In the first four columns, *N* is the number of time series in each group, *T* is the time dimension, *P* is the percentage of time series with missing data, and *M* is the number of missing data in each series. In the next eight columns, the figures refer to the percentages of under- (U) or correct (C) identification of the number of groups (*S*) and the number of common (*r*) and group-specific factors (*r_i*, *i* = 1, 2, 3). The last six columns refer to underclassification (U) or overclassification (O) rates.

serially correlated errors, and the effects of allowing for dynamic changes in the factors. However, for comparative purposes, the noise term was homoskedastic and serially uncorrelated in the first data generation process. Thus, each element ϵ_{it} was drawn independently from a normal distribution with mean 0 and variance $\sigma^2 = 1$. Panel A in Table 1 shows the results of the simulations for the combinations of (*N*, *T*, *P*, *M*), which are presented in the first four columns.

The next two columns in Table 1 show the percentages of under- (U) and correct (C) identification of the number of groups. The figures reported in Table 1 indicate that the clustering ability of the panel data model with a factor structure did not deteriorate substantially when the data sets contained missing observations because the high percentages of times that the model determined the correct number of clusters was comparable with the results obtained by Ando and Bai (2016, 2017). Remarkably, the accuracy of the model decreased only slightly when the number of missing data increased from 3.3% to 6.7%

of the total observations, although the proposed method continued to provide very good results.

The figures in the next six columns in Table 1 show the percentages of under- and correct identification of the true number of factors, *r_s*, with *s* = 1, 2, 3, which were calculated under the condition that the number of groups was selected correctly. As shown in Table 1, the proposed method provided accurate results by selecting the true numbers of common and group-specific factors when observations were missing. Again, we found only a small deterioration when the percentage of time series with missing data increased, but the results showed that the true number of group-specific factors was still estimated very well.¹¹

¹¹ Following a reviewer's suggestion, we also tested the performance of the method when missing data were placed randomly and we found that the results were robust. These results are available from the authors upon request.

The last six columns in Table 1 show the accuracy of the model at correctly classifying the generated time series into the three groups. The metric used to assess the effectiveness of the model was the proportion of time series for each group classified incorrectly. Thus, for a given group, we computed the over (under) classification rate as 100 times the number of time series that exceeded (were lower than) the correct number of series for this group throughout the 500 simulations over the total number of time series generated for this group. The results in Table 1 indicate that all of the misclassification rates were within a very narrow range from 0.7% to 2.2%, thereby demonstrating that our method identified group membership with high accuracy.

Before investigating more complex data generation processes, we assessed the performance of our proposed model estimation and model selection algorithms at inferring from data when the data generation process admitted common and group-specific factors or only group-specific factors. We set $S = 3$ and generated 500 data sets from $x_i = F_c \lambda_{ci} + F_s \lambda_{si} + \epsilon_i$ and 500 data sets from $x_i = F_s \lambda_{si} + \epsilon_i$. For these two groups of data sets, we implemented the two pairs of algorithms, Algorithm 1 and Algorithm 2 (common and group-specific factors), and Algorithm 1' and Algorithm 2' (only group-specific factors), as described in Section 2. We found that $PIC_c(3, r_c^*, r_{c1}^*, r_{c2}^*, r_{c3}^*) < PIC_g(3, r_{g1}^*, r_{g2}^*, r_{g3}^*)$, thereby showing that we selected a factor model with common and group-specific factors in 91% of the simulations that allowed for common and group-specific factors. In addition, for all of the data generation processes that allowed for factor models with only group-specific factors, we obtained $PIC_g(3, r_{g1}^*, r_{g2}^*, r_{g3}^*) < PIC_c(3, r_c^*, r_{c1}^*, r_{c2}^*, r_{c3}^*)$, which implies that we selected a factor model with only group-specific factors. These results suggest that our procedure could select the true data generation process from the observed data with very high accuracy.

Next, we investigated the performance of our proposed method when the noise term was non-homoskedastic and cross-sectionally correlated. In this case, the data generation process for the errors was defined as follows: $\epsilon_{it} = 0.9e_{it}^1 + \delta_t 0.9e_{it}^2$, where $\delta_t = 1$ if t is odd and zero if t is even. The $N \times 1$ vectors $e_t^1 = (e_{t1}^1, \dots, e_{tN}^1)'$ and $e_t^2 = (e_{t1}^2, \dots, e_{tN}^2)'$ were independent and followed multivariate normal distributions $N(0, \Sigma)$, with $\sigma_{ij} = 0.3^{|i-j|}$. The figures shown in Panel B in Table 1 suggest that non-homoskedastic errors and cross-sectional correlations led to some deterioration in performance for all combinations (N, T, P, M) , although the performance losses were not significant, especially with reasonably large numbers of time series.

The third data generation process involved idiosyncratic errors with some serial and cross-sectional correlations. The noise term was generated as: $\epsilon_{it} = e_{it} + 0.2\epsilon_{it-1}$, where $t = 1, \dots, T$. The $N \times 1$ vector $e_t = (e_{t1}, \dots, e_{tN})'$ followed a multivariate normal distribution $N(0, \Sigma)$, with $\sigma_{ij} = 0.3^{|i-j|}$. Panel C in Table 1 shows that the performance also deteriorated slightly, although the results in the table confirmed that our method continued to perform well regardless of whether we focused on

determining the number of groups, factor extraction, or group membership.

The fourth data generation process focused on allowing for the temporal dependence of factors. In particular, the common factor was assumed to be an autoregressive process of order one, $f_{ct} = \phi f_{ct-1} + u_t$, where u_t is an independent Gaussian error term with mean 0 and variance 1. In addition, each element of the $r_s \times 1$ group-specific factors was assumed to follow an autoregressive processes of order one, $f_{st}^j = \phi f_{st-1}^j + u_{st}^j$, where u_{st}^j is an independent Gaussian error term with mean 0 and variance 1, and $j = 1, \dots, r_s$. In the simulations, we set $\phi = 0.3$. The results in Panel D in Table 1 suggest that serially correlated factors led to slight finite-sample performance deterioration in a similar manner to the idiosyncratic errors with serial correlations.¹²

In summary, these results suggest that our proposed method for dealing with unbalanced panels of data sets with a group factor structure was robust to the data problems that typically occur in empirical applications, especially for reasonably large data sets.

Finally, we conducted simulation experiments to assess the performance of our proposed method at estimating unobserved missing data. For each combination (N, T, P, M) , we used the data generation model with homoskedastic and serially uncorrelated noise terms to simulate data from a factor model with a group structure and replaced M observations with missing values in a proportion P of the series. We then imputed the unobserved data using the standard EM method of Stock and Watson (2002) and our proposed group-specific extension of the EM algorithm. For each simulation, we computed the mean squared differences between these two estimates of the missing values and the true data removed before the simulation.

For each combination (N, T, P, M) , the last column in Table 2 labeled as *better* shows the percentage of times (among the total simulations) that Algorithm 1 and Algorithm 2 obtained lower mean squared errors than the standard EM algorithm of Stock and Watson (2002). These results indicate that our proposed method performed substantially better than the standard EM algorithm when the data had grouped factor structures. Thus, our algorithm exploited the group similarity when imputing the values of the unobserved data points.¹³

4. Empirical application

We assessed the effectiveness of our proposed method at handling a data set of time series over different spans,

¹² For reasonably larger values of ϕ that covered the standard empirical applications, the method continued to perform well. Factors with near to unit-root dynamics ($\phi = 0.9$) are very rare in practice, but we found that they resulted in some deterioration in performance, as also shown by Stock and Watson (2002).

¹³ In this section, we focused on group factor models that allow for common and group-specific factors. We repeated the analysis for group factor models that only allow for group-specific factors and found that the results were qualitatively similar. These results are available from the authors upon request.

Table 2
Comparison with the standard EM algorithm.

<i>N</i>	<i>T</i>	<i>P</i>	<i>M</i>	<i>Better</i>
50	150	10	50	88.2
50	150	20	50	82.4
100	150	10	50	96.8
100	150	20	50	95.4
50	250	10	84	85.4
50	250	20	84	81.2
100	250	10	84	94.7
100	250	20	84	92.8

Notes. See notes for Table 1. The last column shows the percentage of times (out of the total simulations) that our grouped factor model with the EM algorithm obtained a lower mean squared error compared with the standard EM algorithm of Stock and Watson (2002) when filling in missing data.

containing outliers, and with remarkable clustering structures. We focus on FRED-QD, which updates the data set used by Stock and Watson (2012) in real time. The data comprised 248 US quarterly economic indicators compiled in an easily downloadable manner by the Research division of the Federal Reserve Bank of St Louis, which are available from its website.¹⁴

4.1. In-sample analysis

The effective sample of the latest available vintage ranged from the third quarter of 1959 to the last quarter of 2019. However, 38 of the 248 indicators (15.3%) were not available for the entire sample, thereby implying 2,102 missing observations out of 60,264 potential figures (3.5%). In contrast to Ando and Bai (2016, 2017), we did not remove indicators with missing observations or shorten the sample period. In addition, McCracken and Ng (2021) documented 30 outliers that we also treated as missing.¹⁵

Before implementing the model, each series was transformed by approximately integrating to order zero using the transformation codes provided by McCracken and Ng (2021). In general, real activity variables were transformed into growth rates, interest rates were transformed into first differences, and prices were transformed into first differences of rates of inflation. We then standardized all of the series to zero mean and unit variance.

We applied the proposed model selection criteria in the pairs of algorithms, Algorithm 1 and Algorithm 2, and Algorithm 1' and Algorithm 2', to decide whether the data admitted a standard factor structure, as described by Stock and Watson (2002), a grouped factor with common and group-specific factors, or a grouped factor with only group-specific factors. In addition, the algorithms allowed us to choose the number of groups from a maximum of five groups and the number of factors from a maximum of eight factors.

¹⁴ The data are available at <https://research.stlouisfed.org/econ/mccracken/fred-databases/>

¹⁵ Similar to Stock and Watson (2002), and McCracken and Ng (2021), we defined an outlier as an observation that deviated from the sample median by more than ten interquartile ranges.

Fig. 1 shows the behavior of *PIC* as a function of the number of groups when the grouped factor model admitted common and group-specific factors (red line, Algorithm 1 and Algorithm 2), and when the grouped factor model admitted only group-specific factors (black line, Algorithm 1' and Algorithm 2'). The two pairs of algorithms showed that the model selection criterion was minimized for only two groups. However, the minimum *PIC* obtained for the model that admitted common and group-specific factors was 0.797, whereas that for the model that admitted only group-specific factors was 0.785, thereby suggesting that a factor model that allowed for only group-specific factors was more suitable for FRED-QD than a factor model that allowed for both common and group-specific factors.

The numbers of time series in the two clusters were 103 and 145. The estimated number of group-specific factors was three for the first group and four for the second group. In this context, it is standard practice to analyze the extent to which the factors can capture the variability in the original series using the estimated eigenvalues. In our study, we found that the three factors estimated for the first group accounted for almost 40 percent of the variance in the indicators that belonged to that group. In the second group, the four selected factors accounted for more than 50 percent of the variance in the data.

To provide economic meaning to group membership, it should be noted that McCracken and Ng (2021) classified the indicators into the following 14 categories: national income and product accounts (NIPA); industrial production; employment and unemployment; housing; inventories, orders, and sales; prices; earnings and productivity; interest rates; money and credit; household balance sheets; exchange rates; other; stock markets; and non-household balance sheets. The classification of the economic indicators into one of the two groups is clearly shown in Fig. 2. The two membership groups are represented as floating dots on the ordinate axis against the 248 economic indicators, which are grouped on the abscissa axis by category and ordered as described in the Appendix of the study by McCracken and Ng (2021).

Overall, our method classified categories 4, 6, 10, 11, 12, 13, and the monetary indicators in category 9 into the first group. These categories were related to housing, prices, household balance sheets, exchange rates, other, and stock markets. According to this classification, we referred to this group of economic indicators as the nominal group.

We also analyzed the relationship between this group and the nominal economic conditions by regressing the 248 indicators against the first factor for this group, which accounted for almost 20 percent of the variance in the first group. The top panel in Fig. 3 show the coefficients of determination for these regressions. Clearly, this factor was loaded mainly on prices (category 6). The bottom panel in the figure shows the strong correlation between this factor and the stationary transformation of the consumer price index (CPI).

By contrast, categories 1, 2, 3, 5, 7, and 8 were classified in the second group. These categories include real

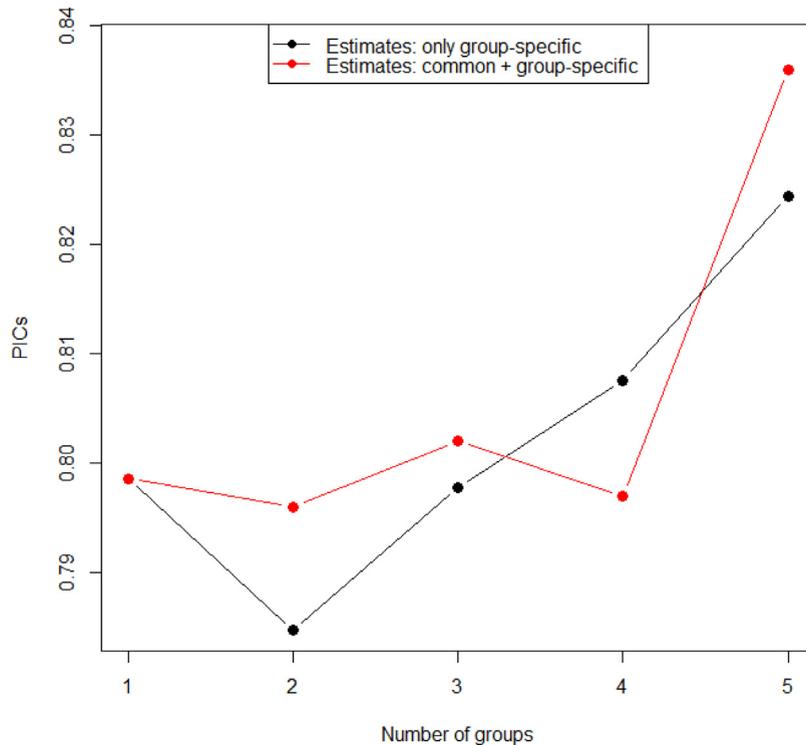


Fig. 1. Model selection. (Notes. The figure shows the value of PICs as a function of the number of groups. The red line refers to a group factor model that allows for both common and group-specific factors, and the black line refers to a group factor model that allows for only group-specific factors.)

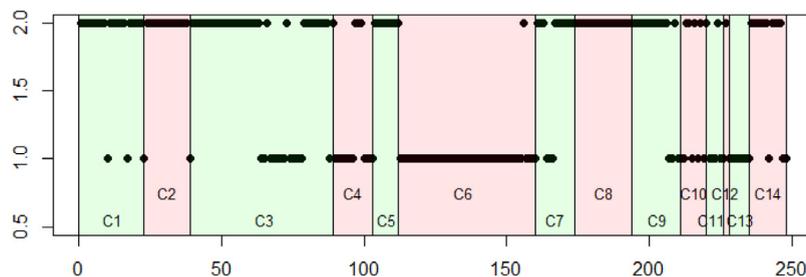


Fig. 2. Classification $G = (g_1, \dots, g_N)$. (Notes. Classification of the $N = 248$ indicators provided by FRED-QD in the two identified groups. According to McCracken and Ng (2021), the indicators were grouped by categories labeled as follows: national income and product accounts (C1), industrial production (C2), employment and unemployment (C3), housing (C4); inventories, orders, and sales (C5); prices (C6), earnings and productivity (C7), interest rates (C8), money and credit (C9), household balance sheets (C10), exchange rates (C11), other (C12), stock markets (C13), and non-household balance sheets (C14).)

output, employment, real earnings, and productivity. Interestingly, interest rate and the credit variables in the ninth category also belonged to this group. These indicators typically capture the broad movements in economic activity, so we referred to this group as the real activity group.

Following the same reasoning applied for the nominal group, the top panel in Fig. 4 shows the coefficients of determination for the regressions of the first factor in the real activity group against the entire panel of indicators. The results show the strong correlations between this factor and the NIPA indicators, industrial production, employment, unemployment, inventories, orders, and sales.

The bottom panel in this figure shows the strong correspondence between the first factor in the second group and the GDP growth rate, thereby providing a broad assessment of US economic performance.

This result agrees with that obtained by McCracken and Ng (2016) who suggested reorganizing the factor estimates into two groups, with one for real activity and one for nominal activity.¹⁶ Before factor estimation, McCracken and Ng (2016) constructed a diffusion index for real activity by computing the common variations at low

¹⁶ McCracken and Ng (2016) employed the monthly version of the database used in the present study (FRED-MD).

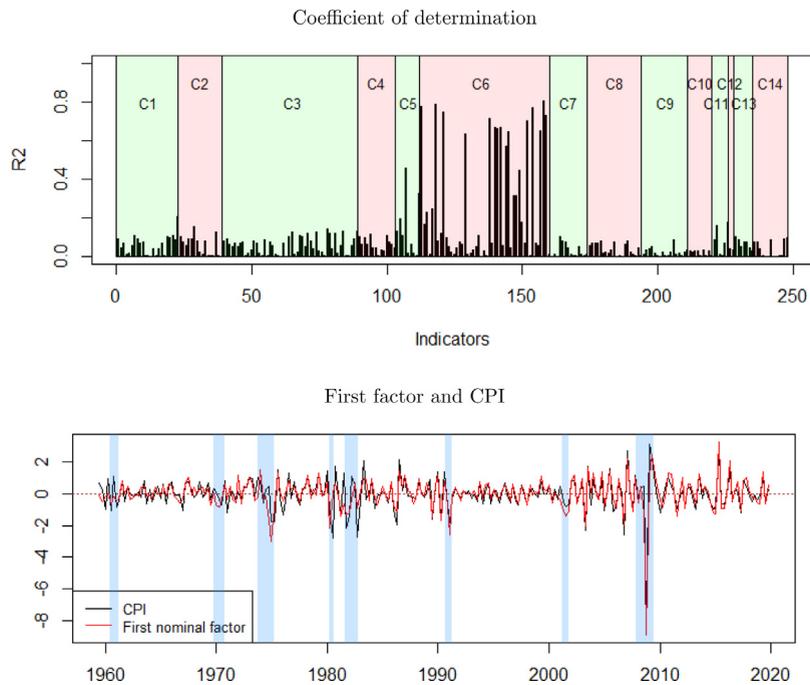


Fig. 3. Nominal group: first factor and economic indicators. (Notes. The top panel shows the coefficients of determination calculated between the first factor of the nominal group and the economic indicators grouped by categories (C1 to C14), as described by McCracken and Ng (2021) (see the notes for Fig. 2). The factor and stationary transformation of CPI are plotted (normalized with zero mean and unit variance) in the bottom panel. The shaded areas denote the recession documented by the National Bureau of Economic Research (NBER).)

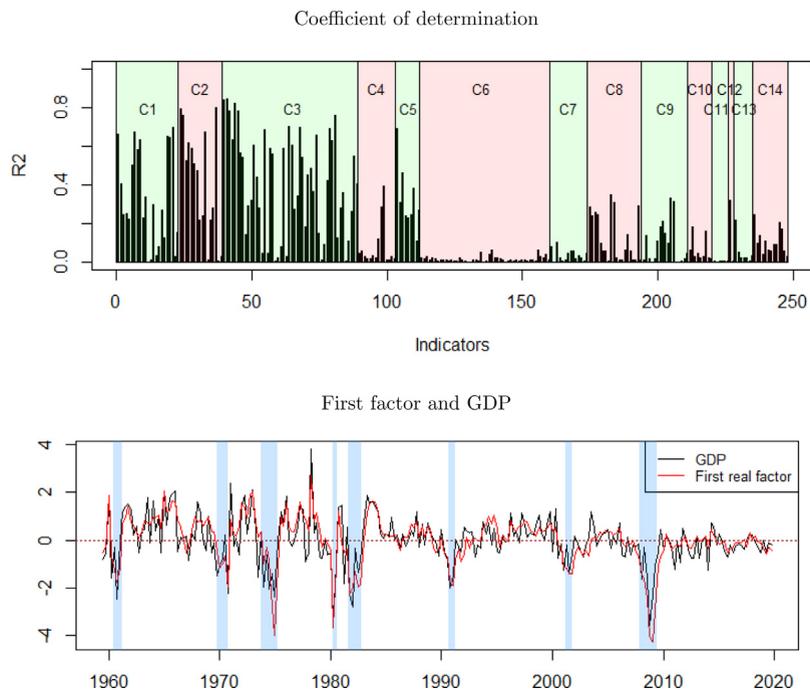


Fig. 4. Real activity group: first factor and economic indicators. (Notes. The top panel shows the coefficients of determination calculated between the first factor of the real activity group and the economic indicators grouped by categories (C1 to C14), as described by McCracken and Ng (2021) (see the notes for Fig. 2). The factor and GDP growth rate are plotted (normalized with zero mean and unit variance) in the bottom panel. The shaded areas denote the NBER-referenced recessions.)

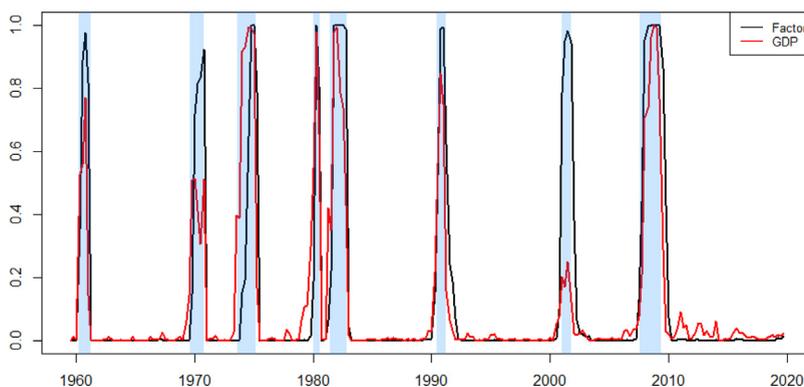


Fig. 5. Smoothed probabilities. (Notes. The figure shows the smoothed probability estimates for the recession regime based on the real activity factor estimates (solid line) and GDP (dashed line). The shaded areas denote the NBER-referenced recessions.)

frequencies for the factor associated with the industrial production and employment series. They showed that the evolution of this index was closely related to the NBER-referenced business cycle. In the same manner, they constructed a diffusion index for the nominal activity from the factors related to price, spread, and interest rate variables, and found that the index appeared to align with price pressure inflation expectations in the past five decades.

In a similar manner, previous empirical studies pre-classified economic data sets into nominal and real data to extract indexes for US inflation and real activity from two separate factor models. [Aruoba and Diebold \(2010\)](#) did not claim that their study was exhaustive but they used five indicators of real activity and six inflation indicators to compute real activity and inflation indexes, which were then applied to examine real activity and inflation together over the cycle. [Leiva-Leon \(2014\)](#) extended this approach by considering a unified factor model where the two separate factors were extracted from the same two sets of real and nominal indicators. Our results are consistent with this classification but we selected the two separate groups and the indicators in our large data set assigned to each group in a systematic and statistically optimal manner.

Similar to [McCracken and Ng \(2016\)](#), we also examined the business cycle information that could be extracted from the first factor of the real activity group. We assumed a regime switch in the index f_{1t}^1 itself and the switching mechanism at time t was controlled by an unobservable state variable, v_t , which followed a first-order Markov chain. Thus, we specified a simple switching model ([Hamilton, 1989](#)) as:

$$f_{1t}^1 = \mu_{v_t} + u_t, \tag{9}$$

where u_t is an independent Gaussian perturbation with zero mean and variance σ_u^2 . The nonlinear behavior of the time series was governed by the mean of the process, which was allowed to change within each of the two distinct regimes $v_t = 0$ and $v_t = 1$. We required that $\mu_0 > \mu_1$, so the first regime referred to expansions and the second regime to recessions. The Markov switching assumption implied that the transition probabilities were

independent of the information set at $t - 1$, $I_{t-1} = \{f_{11}^1, f_{12}^1, \dots, f_{1t-1}^1\}$, and of the business cycle states prior to $t - 1$

$$p(v_t = i | v_{t-1} = j, v_{t-1} = l, \dots, I_{t-1}) = p(v_t = i | v_{t-1} = j) = p_{ij}. \tag{10}$$

The black line in [Fig. 5](#) shows the smoothed probabilities of recession inferred from the first real factor, $p(v_t = 1 | I_t)$, and the shaded areas correspond to the NBER-referenced recessions. The results indicate that the probabilities of recession were in good agreement with the professional consensus regarding the history of US business cycles. The probabilities were close to zero or one, thereby demonstrating that the first factor of the real group captured the underlying pattern of the dichotomous shifts between expansions and recessions. Remarkably, NBER recessions failed to match with high recession probabilities in no cases. In addition, we found that our factor model could be used to date the reference cycle because the spikes in the probabilities of recession occurred near the NBER turning points.

We also assessed the accuracy of the first real activity factor at capturing the fluctuations in the US business cycle by comparing the smoothed probabilities inferred from the estimated factor with those inferred from observable information. In particular, we used the smoothed probabilities inferred from GDP because it is a meaningful observable measure of the overall economic conditions. These probabilities are also plotted as a red line in [Fig. 5](#). Comparisons of both estimates indicated clear gains in the characterization of the US business cycle recessions based on the higher informational content of the factor summarizing the real activity data.¹⁷ Using the simple metric suggested by [Hamilton \(1989\)](#) for identifying a recession, which implies that $p(v_t = 1 | I_t) > 0.5$, only the factor-based recession probabilities would have identified the recession at the beginning of the 1970s and the mild recession in 2001.

Finally, we assessed the reliability of the values used to fill in the gaps in the missing observations with the

¹⁷ Within recessions, the Brier score loss is 0.06 for probabilities inferred using the factor and 0.17 for probabilities obtained using GDP.

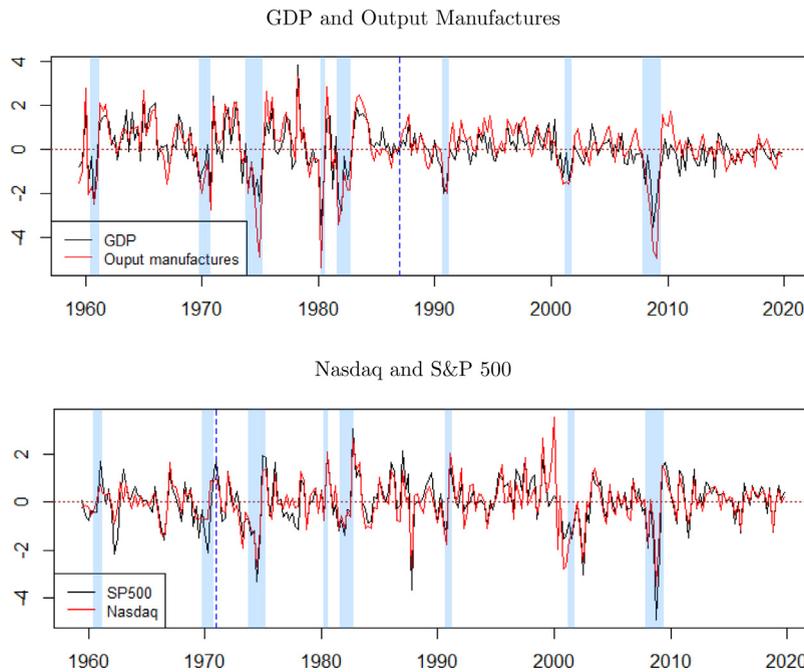


Fig. 6. Filling in missing observations. (Notes. The top panel shows the growth rates in the GDP and real output from the manufacturing sector (first 111 observations are missing). The bottom panel shows the stationary transformations of the NASDAQ composite index and the S&P common stock price index (first 47 observations are missing). The vertical lines indicate the dates of the last missing observations. To facilitate comparisons, the series were normalized with zero mean and unit variance. The shaded areas denote the NBER-referenced recessions.)

group-specific EM algorithm. For illustrative purposes, we compared the dynamic changes in some indicators available for the whole sample period with other indicators from the same category that contained missing values, which were imputed by the model. The top panel in Fig. 6 shows the evolution in the growth rates of real output in the manufacturing sector, which were not available from 1959:3 to 1987:1 and only observable since 1987:2. Thus, this series was extended back to 1959:3 by using the EM estimates for the first 111 missing observations and the cutoff is indicated by a vertical dotted line in the figure.

The figure shows that the dynamic changes in the real output in the manufacturing sector in the unobserved period were similar to those in the observed period, where the reduction in the variance characterized the economic activity in the US since the mid-1980s (the Great Moderation). To facilitate comparisons, the figure also shows the rate of growth in the GDP, which was available for the whole sample period. As expected, the two indicators were highly synchronized in the second part of the observed sample. Interestingly, we also found a very similar pattern between the imputed values for the real output in the manufacturing sector and the observed GDP values in the first part of the sample.

As a comparable example with indicators for the nominal group, the bottom panel in Fig. 6 shows the S&P Common Stock Price Index, which was not available from 1959:3 to 1971:1 (missing the first 47 observations), where the cutoff is indicated by a vertical dotted line. Again, the imputed values for the time series in the first

part of the sample had the same dynamic patterns found in the observed data in the second part of the sample. For comparison, the figure also shows the NASDAQ composite index for the whole sample period. Irrespective of the part of the sample considered, the figure shows that the S&P Common Stock Price Index was highly synchronized with the observed NASDAQ composite index.

4.2. Out-of-sample analysis

The model was evaluated in terms of its ability to obtain one- and four-quarters ahead forecasts, particularly by evaluating the improvements in forecasting by applying a group-specific factor model compared with the forecasts obtained using the Stock–Watson factor model.

In each case, in order to compute the h -step-ahead forecast of the indicator $x_{i,t}$, we used the direct multistep forecasting equation

$$x_{i,t}^h = \alpha_{i,h} + \sum_{j=0}^{q-1} \beta'_{i,hj} \hat{f}_{gi,t-h-j} + \sum_{l=0}^{p-1} \gamma_{i,hl} x_{i,t-h-l} + \epsilon_{i,t}^h, \quad (11)$$

where $\hat{f}_{gi,t-j}$ is the $(h+j)$ -lagged value of the $r_{gi} \times 1$ vector of estimated group-specific factors and $\epsilon_{i,t}^h$ is an independent error term for the i th indicator. According to McCracken and Ng (2016), we fixed the number of autoregressive lags p at 4. In addition, we considered the same number of lags $q = 4$ for the group-specific factors.

It would have been interesting to use real-time vintages but they were only available from May 2018, thereby making real-time forecasting analysis unfeasible at present.

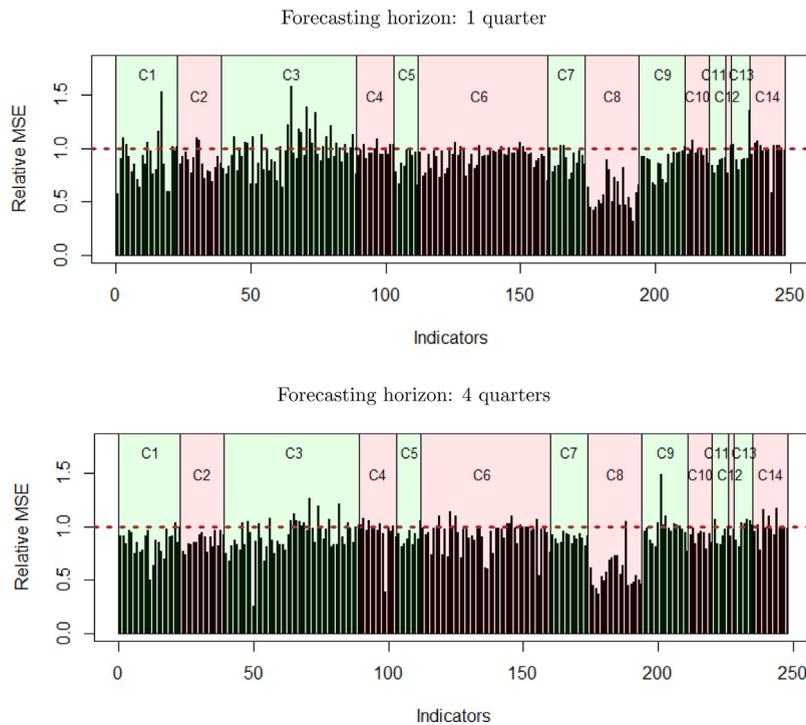


Fig. 7. Relative MSE for one- and four-quarter horizons. (Notes. The figure shows the relative MSEs of the forecasts by the group-specific factor model compared with the Stock–Watson factor model across the 248 macroeconomic indicators grouped by categories (C1 to C14), as described by McCracken and Ng (2021) (see the notes for Fig. 2).)

Instead, we performed an out-of-sample exercise in order to mimic real-time analysis. This method involved computing forecasts from successive enlargements of a partition of the latest available data set. We started with data from the beginning of the sample until 2005:4. Using this sample, the data were screened for outliers and standardized, before performing classification and choosing the group-specific number of factors. Based on these choices, the forecasting equation was used to compute forecasts $x_{2005:4+h}^h$ (for $h = 1$, the forecast was for 2006:1; and for $h = 4$, the forecast was for 2006:4).

The sample was then updated by one period, classified, and the factors were estimated, before the forecasting models were re-estimated and forecasts were computed for $x_{2006:1+h}$. The forecasting procedure continued iteratively until the forecast for 2019:4 was computed using data until 2019:3 for $h = 1$ and data until 2018:4 for $h = 4$. In each iteration, we computed the squared deviation for the h -ahead forecasts based on actual data and stored the averages of these figures for each indicator, i.e., the mean squared errors (MSEs).

For comparison, similar h -ahead forecasts were obtained using the factors extracted as described by Stock and Watson (2002) and the MSEs were also stored. Fig. 7 shows the results of this forecasting exercise based on the 248 series, where the MSEs for the group-specific factor model are plotted relative to the MSE for the forecasts using the Stock–Watson factor model. Hence, bars smaller than one indicate that group-specific factors produced more accurate forecasts than common factors.

Regardless of the forecasting horizon, the figure shows that the improvements in the forecasts using group-specific factor models compared with the forecasts using the standard Stock–Watson factor model were spread over all categories. In particular, the group-specific model performed better for 75% of the one-quarter-ahead forecasts and this figure increased to almost 80% in the case of the four-quarter-ahead forecasts. It should be noted that the highest performance gains were found in interest rates, although substantial improvements were also found in NIPA, industrial production, and employment.

Finally, we focused on the one- and four-quarter-ahead forecasts for the eight target variables selected by McCracken and Ng (2021). In particular, they considered four real activity indicators comprising the real GDP, industrial production (INDPRO), unemployment rate (UNRATE), and federal funds rate (FFR), and four nominal indicators comprising the CPI, personal consumption expenditures price index (PCEPI), GDP deflator (GDPPI), and production price index (PPI). To compare the forecasting performance of our proposed method, Table 3 shows the MSEs computed for our grouped factor model relative to the MSEs for several forecasting alternatives. Thus, the forecasts obtained using our model had a relative MSE of 1.00 in the rows labeled as M1, and figures lower than 1.00 indicate that our model performed better than the alternatives.

The relative MSEs achieved by various forecasting methods are included in the table. Rows labeled as M2 show the relative MSEs with the Stock–Watson factor model, which is a standard factor model that handles missing

Table 3
Out-of-sample forecasting accuracy.

Model	GDP	INPRO	UNRATE	FFR	CPI	PCEPI	GDPPI	PPI
1-quarter-ahead forecasts								
M1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M2	0.57	0.85	1.23	0.64	0.93	0.97	0.77	0.96
M3	0.69	0.80	0.92	0.94	0.73	1.03	0.98	1.00
M4	0.81	0.99	1.24	0.94	0.93	0.95	0.95	1.00
M5	0.73	0.83	0.93	0.91	1.00	0.73	0.95	0.96
M6	0.53	0.86	1.13	0.69	0.90	0.74	0.88	0.93
M7	0.55	0.95	0.94	0.75	0.96	0.96	0.97	0.95
M8	0.59	0.80	1.18	0.95	0.89	0.90	0.83	0.84
4-quarter-ahead forecasts								
M1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
M2	0.91	0.77	1.06	0.61	0.95	0.98	0.94	0.98
M3	0.92	1.06	0.90	0.86	0.96	0.98	0.97	1.01
M4	0.87	1.01	0.90	0.86	0.96	0.97	0.87	0.93
M5	0.89	0.88	1.03	0.64	0.96	1.03	0.95	0.99
M6	0.78	0.80	0.99	0.68	0.91	0.95	0.92	0.97
M7	0.86	1.05	0.86	1.06	1.00	0.96	0.91	0.91
M8	0.85	1.02	0.85	1.00	0.94	0.93	0.96	0.84

Notes. The figure shows the relative MSEs for the forecasts using the group-specific factor model (M1, first row) over several alternatives that are described in the text for the following eight target variables selected by McCracken and Ng (2021): real GDP (GDP), industrial production (INDPRO), unemployment rate (UNRATE), federal funds rate (FFR), consumer price index (CPI), personal consumption expenditures price index (PCEPI), GDP deflator (GDPPI), and production price index (PPI).

data with an EM algorithm. PIC suggests that a common factor is not necessary but rows labeled as M3 show the relative forecasting performance with a grouped factor model that allows for both common and group-specific factors. Rows labeled as M4 refer to the relative MSEs with a model such as M1 but using a less automatic predictive regression equation that focuses only on the factors that are statistically significant at 5%. The relative performance results with a grouped factor model that handles missing using with the naive procedure based on imputing the last valid observations are shown in the rows labeled as M5. Rows labeled as M6 (standard factor model) and M7 (grouped factor model) refer to models that handle missing data by excluding the indicators with missing data. Finally, the rows labeled as M8 show the relative MSEs based on the forecasts with AR(1) processes.

Overall, the results in Table 3 show that the forecasts obtained with a grouped factor that allowed for only group-specific factors and that handled missing data with an EM algorithm were generally better than those produced with the univariate and multivariate methods, regardless of the forecasting horizon. Except for unemployment, comparisons with the model of Stock and Watson (2002) indicated that the grouped factor model performed better than the standard factor model, where the forecasting gains in terms of GDP and FFR were very large. The improvements were robust to forecasting with a less automatic predictive regression equation that focused only on statistically significant factors. Similar to the results obtained using the C_p -type criterion for model selection, allowing for factors common to all variables did not improve the forecasts.

In addition, the relative MSEs shown in Table 3 demonstrate that using the EM algorithm to handle missing data performed better than the other naive alternatives, such as replacing missing data by the last valid observations or removing indicators with missing observations from the data set. Finally, similar to the results reported by Stock and Watson (2002), and McCracken and Ng (2021), we obtained substantial evidence that the factors could provide additional predictive information compared with that utilized by univariate autoregressive processes.

5. Conclusion

To deal with large sets of unbalanced panels with cluster structures and potential outliers, we merged two previously proposed econometric methods comprising the EM algorithm combined with the factor estimator-based principal component analysis, as introduced by Stock and Watson (2002), and grouped factor structures with unknown group membership and numbers of groups, as proposed by Ando and Bai (2016, 2017).

We assessed the benefits of the proposed method for factor modeling based on Monte Carlo simulations. The experimental results demonstrated that the proposed method was highly effective. In particular, we found no significant differences between the clustering abilities of our method, which deals with missing data in a simple manner, and those of the factor models with unknown group membership proposed by Ando and Bai (2016, 2017), which require balanced data sets. The performance of the model did not deteriorate substantially when the simulations included non-homoskedastic errors with cross-sectional dependence, errors with serial correlations, and factors with temporal dependence. In addition, the simulation results indicated that our group-specific EM algorithm performed better than the standard EM algorithm proposed by Stock and Watson (2002) and it could accurately infer the missing data.

Using the novel repository FRED-QD maintained by the St Louis Federal Reserve Bank and documented by McCracken and Ng (2021), we examined the empirical performance of our proposed method at determining the number of clusters, providing group-specific factors, obtaining accurate estimates of unobserved missing data, and producing out-of-sample forecasts for US economic indicators. According to the results, the factor model that allows for only group-specific factors was more suitable for this data set than a factor model that allows for both common and group-specific factors. We identified two distinctive groups of economic indicators comprising a group of nominal indicators and a group of real activity indicators. The estimated factors could be applied for identifying the business chronology in the US. Finally, we found substantial improvements in the forecasting results obtained with group factor models compared with forecasting using factor models that do not deal with grouped structures, as described by Stock and Watson (2002), or forecasting with factor models that require the removal of indicators with missing observations, as described by Ando and Bai (2016).

According to our results, the proposed framework is also a very promising tool for handling large and unbalanced data sets with group structures. In fact, we plan to address the following issues in future research. First, a natural extension of our proposed factor model for large and incomplete data sets with unknown group structure involves considering additional explanatory variables. This extension may consider homogeneous coefficients over all cross-sectional units (Ando & Bai, 2016), heterogeneous group-specific coefficients (Ando & Bai, 2016), or coefficients that vary over all cross-section units (Ando & Bai, 2017).

Second, the research division of the Federal Reserve Bank of St Louis also provides a monthly frequency companion to the data set used in our empirical application called FRED-MD. Thus, further empirical gains could be achieved by extending our proposed factor method to handle mixed frequencies. We did not consider this feature in the present study, but mixed frequencies require further investigation because they could combine the economic information provided by FRED-MD and FRED-QD at the same time.

Third, the real-time vintages for the data set used in the empirical application are only available since 2018. According to this data limitation, our assessment of the out-of-sample forecasting performance was conducted based on fractions of the finally revised data. New vintages will become available and real-time reassessments should provide novel insights into the forecasting accuracy of our model.

Fourth, our approach focuses on dealing with missing data using the EM algorithm of Stock and Watson (2002). However, some modifications of this method have been applied in recent studies. For example, McCracken and Ng (2021) initialized the algorithm to the unconditional sample mean based on the non-missing values, and the mean and variance of the data were re-calculated in each iteration. In a recent study, Jin et al. (2021) replaced the missing observations by zeros and conducted the usual PC analysis for a scaled version of the data matrix, where the scale was determined by the percentage of observed values in the data. Based on the EM algorithm, they replaced the missing observations with initial estimators of the common components and obtained updated PC estimators, before iterating this procedure until convergence. Comparisons of these two approaches could be conducted in future research.

Finally, the performance of the estimation procedure could be specifically evaluated by applying it to data sets with different structures of practical relevance. In particular, we could extend our approach to handle seasonal patterns or long-term memory in sets of economic indicators.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Alonso, A., Galeano, P., & Peña, D. (2020). A robust procedure to build dynamic factor models with cluster structure. *Journal of Econometrics*, 216, 35–52.
- Ando, T., & Bai, J. (2015). Asset pricing with a general multifactor structure. *Journal of Financial Econometrics*, 13, 556–604.
- Ando, T., & Bai, J. (2016). Panel data models with grouped factor structure under unknown group membership. *Journal of Applied Econometrics*, 31, 163–191.
- Ando, T., & Bai, J. (2017). Clustering huge number of financial time series: A panel data approach with high-dimensional predictors and factor structures. *Journal of the American Statistical Association*, 112, 1182–1198.
- Angelini, E., Henry, J., & Marcellino, M. (2006). Interpolation and back-dating with a large information set. *Journal of Economic Dynamics & Control*, 30, 2693–2724.
- Aruoba, S., & Diebold, F. (2010). Real-time macroeconomic monitoring: Real activity, inflation, and interactions. *American Economic Review*, 100, 20–24.
- Bai, J., & Ng, S. (2008). Large dimensional factor analysis. *Foundations and Trends in Econometrics*, 3, 89–163.
- Bernanke, B., & Boivin, J. (2003). Monetary policy in a data-rich environment. *Journal of Monetary Economics*, 50, 525–546.
- Breitung, J., & Eickmeier, S. (2016). Analyzing international business and financial cycles using multi-level factor models: a comparison of alternative approaches. In *Dynamic Factor Models*. Emerald Group Publishing Limited.
- Choi, I., Kim, D., Kim, Y., & Kwark, N. (2018). A multilevel factor model: Identification, asymptotic theory and applications. *Journal of Applied Econometrics*, 33, 355–377.
- Crucini, M., Kose, M., & Otrok, C. (2011). What are the driving forces of international business cycles? *Review of Economic Dynamics*, 14, 156–175.
- Gregory, A., Head, A., & Raynauld, J. (1997). Measuring world business cycles. *International Economic Review*, 38, 677–701.
- Hallin, M., & Liška, R. (2011). Dynamic factors in the presence of blocks. *Journal of Econometrics*, 163, 29–41.
- Hamilton, J. (1989). A new approach to the economic analysis of nonstationary time series and the business cycles. *Econometrica*, 57, 357–384.
- Jin, S., Miao, K., & Su, L. (2021). On factor models with random missing: EM estimation, inference, and cross validation. *Journal of Econometrics*, 222, 745–777.
- Kose, M., Otrok, C., & Whiteman, C. (2003). International business cycles: World, region, and country-specific factors. *American Economic Review*, 93, 1216–1239.
- Leiva-Leon, D. (2014). Real vs nominal cycles: a multistate markov-switching Bi-factor approach. *Studies in Nonlinear Dynamics and Econometrics*, 18, 557–580.
- Marcellino, M., & Schumacher, Ch (2010). Factor MIDAS for nowcasting and forecasting with ragged-edge data: A model comparison for german GDP. *Oxford Bulletin of Economics and Statistics*, 72, 518–550.
- McCracken, M., & Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34, 574–589.
- McCracken, M., & Ng, S. (2021). FRED-QD: A quarterly database for macroeconomic research. *Federal Reserve Bank of St. Louis Review*, 103, 1–44.
- Rodriguez-Caballero, C., & Ergemen, Y. (2017). *Estimation of a Dynamic Multilevel Factor Model with Possible Long-Range Dependence: Technical report*, Universidad Carlos III de Madrid. Departamento de Estadística.
- Schumacher, Ch, & Breitung, J. (2008). Real-time forecasting of german GDP based on a large factor model with monthly and quarterly data. *International Journal of Forecasting*, 24, 386–398.
- Stock, J., & Watson, M. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20, 147–162.
- Stock, J., & Watson, M. (2011). Dynamic factor models. In M. Clements, & D. Hendry (Eds.), *The oxford handbook of economic forecasting*. Oxford: Oxford University Press.
- Stock, J., & Watson, M. (2012). Disentangling the channels of the 2007–2009 recession, brookings papers on economic activity. *Spring*, 8, 1–156.