Dynamic factor models: A review of the literature

Karim Barhoumi, Olivier Darné, Laurent Ferrara

Dynamic Factor Models:
A Review of the Literature*

Karim BARHOUMI,** **Olivier DARNɆ**
and Laurent FERRARA‡

July 2013

1 Introduction

In recent years, the volume of available financial and economic data has led econometricians to develop or adapt methods to efficiently summarize the information contained in these large data bases. In applied macroeconomics, it is frequently the practitioner who has the tricky task of identifying, from among the large number *N* of variables available to him, the few variables of interest that will enable him to best solve his problem.

For example, economic growth and inflation forecasts are made in national and international institutions that have access to large volumes of data from surveys of households and businesses and various series on prices and real activity, such as the industrial production index (IPI), household consumption, the unemployment rate, etc. Similarly, central banks conduct monetary policy in a data-rich environment, looking at macroeconomic activity and the different financial markets on a regular basis and tracking numerous monetary aggregates.

A number of econometric methods have been proposed in the literature for working in such data-rich environments. For example, to explain the changes in a particular variable using a vast

* This article was submitted to *Économie et Prévision*.
** International Monetary Fund.
† LEMNA, University of Nantes, and Banque de France.
‡ Banque de France, International Macroeconomics Division. The authors would like to thank Éric Dubois and Hélène Erkel-Rousse for their comments and suggestions, as well as two anonymous reviewers. This article does not necessarily reflect the opinions of the Bank of France or the IMF.

set of exogenous variables N in a linear regression model, the so-called "general-to-specific" method (Krolzig and Hendry, 2001) proposes an algorithm that makes it possible to select just a few variables from among those variables N. Similarly, vector autoregressive (VAR) models are recognized as allowing for simultaneous modeling of variables in a multivariate context. Traditionally, VAR models use a small number of variables to avoid inflating the number of parameters to be estimated. To remedy this problem, Bayesian approaches have been proposed to estimate VAR models with a high number of variables N by imposing restrictions (see, for example, De Mol, Giannone and Reichlin, 2008). Finally, if we consider the problem of predicting a particular variable when we have a large number of variables N that are potentially very relevant, we can imagine estimating N linear regressions, which then provide N forecasts that we will seek to combine (see, for example, Newbold and Harvey, 2002, for forecast combination methods). Eklund and Kapetanios (2008) also provide a review of the literature on the various forecasting techniques using large data sets.

Among the different methodologies proposed in the literature, dynamic factor models have grown significantly in popularity since the early 2000s and have been shown to be very useful in macroeconomic analysis and forecasting in a data-rich environment. These models can be used to summarize the information contained in a large number of economic variables into a small number of factors common to the set of variables. In this type of model, the N variables (x_{it}), for $i = 1, \dots, N$ and $t = 1, \dots, T$, where t refers to the time index, are each assumed to be the sum of two unobservable orthogonal components: one component resulting from the factors that are common to the set of variables, (χ_{it}), and an idiosyncratic component (ξ_{it}).The component (χ_{it}) is obtained by extracting a small number $r \geq 1$ of common factors (F_{jt}), $j = 1, \dots, r$ from all of the variables present in the data set. Often, by extension, this component (χ_{it}) is identified by the term "common component," which we will also use in this article. The idiosyncratic component (ξ_{it}) covers the shocks specific to each of the variables. Thus, in a factor model of dimension $(N \times 1)$, each element of the vector $X_t = (x_{1t}, \dots, x_{Nt})'$, assumed to be zero mean, can be written as follows:

$$x_{it} = \chi_{it} + \xi_{it},$$

or:

$$x_{it} = \lambda_{i1}F_{1t} + \cdots + \lambda_{ir}F_{rt} + \xi_{it},$$

For $i = 1,\ldots, N$ and $t = 1,\ldots, T$. The loadings (λ_{ij}) for $i = 1,\ldots, N$ and $j = 1,\ldots, r$, represent the contributions of the variable i to the common factor (F_t) of dimension $(r \times 1)$ such as $F_t = (F_{1t}, \ldots, F_{rt})'$. The vector $(\xi_t) = (\xi_{1t}, \ldots, \xi_{Nt})'$ of dimension $(N \times 1)$ is a vector consisting of N idiosyncratic components. The vectorial form of the model is presented as follows, for all $t = 1, \ldots, T$:

$$X_t = \Lambda F_t + \xi_t, \tag{1}$$

Where Λ is the weighting matrix of dimension $(N \text{ x } r)$. The matrix version is given as:

$$X = F\Lambda' + \xi, \tag{2}$$

Where X is of dimension $(T \times N)$, F is of dimension $(T \times r)$, Λ is of dimension $(N \times r)$, and ξ is of dimension $(T \times N)$.

Given the rapid development of dynamic factor models in applied macroeconomics, we felt that the time was right to propose a review of the literature on these models so as to recap the current situation for practitioners. In this article, we begin by presenting the so-called traditional or classical factor models, which were developed initially for a small number of variables with common movements. We distinguish between static and dynamic approaches for these models. Then, we describe approximate factor models, which can take a large number of variables into account, again in a static or dynamic context. Next, we present some estimation models proposed in the literature. A crucial aspect of these models is the selection of the number of common factors r to use in the analysis and, so, we provide a review of the various information criteria developed to select the optimal number of factors. There are many applications of factor models in the empirical economic literature, including, for example, asset pricing models (Ross, 1976), consumer theory (Gorman, 1981; Lewbel, 1991), performance assessment and risk measurement in finance (Campbell et al., 1997). In the final section, we focus on some recent applications that underscore the interest of this approach for macroeconomists, particularly (i) for the construction of short-term economic indicators, (ii) for macroeconomic forecasting, and (iii) for international macroeconomics and monetary policy analysis.

2 Factor models for a small number of variables (small *N*)

In this section, we present factor models used to model a small number of variables *N*, where, in practice, *N* is generally lower than 6 or 7 variables. We begin with the simplest non-dynamic models (static factors) and, then, look at dynamic models, ending with a few recent extensions of this type of model.

2.1 Static factor models (SFM)

In this type of model, a small number of unobservable variables *r* provides a linear explanation of a small number of observed variables *N* so that $r < N$. In the applications presented in the final section of our article, the number of variables is such that $N \leq 7$ and a single factor can generally explain most of the variance, *i.e.* $r = 1$. The series are assumed to be stationary, to have finite variance, and to be standardized. We put forward the following hypotheses, which could subsequently be abandoned:

(SH1) The factors (F_t) are centered, $E(F_t) = 0$, and are mutually orthogonal for all *t*, i.e.: $\forall t,\ E(F_{jt}F_{j't}) = 0$ for $j \neq j'$. Consequently, the variance-covariance matrix of (F_t), $\Sigma_F = E(F_tF_t')$, where F_t' is the transpose of F_t, is a diagonal matrix.

(SH2) The idiosyncratic processes (ξ_{it}) and $(\xi_{i't})$ are mutually orthogonal for all $i \neq i'$, with $E(\xi_t) = 0$. Consequently the variance-covariance matrix (ξ_t) is a diagonal matrix: $\Sigma_\xi = E(\xi_t\xi_t') = diag(\sigma_1^2, \dots, \sigma_N^2)$.

(SH3) The factors (F_t) and idiosyncratic noise $(\xi_{it})_{i=1,\dots,N}$ are not correlated, i.e.: $\forall i, j, t, t'$, we have: $E(F_{jt}\xi_{it'}) = 0$.

(SH4) The variables are assumed to be independent and identically distributed over time (the so-called IID hypothesis), so that, in particular, for $t \neq t'$, $E(F_{jt}F_{jt'}) = 0$ and $E(\xi_{it}\xi_{it'}) = 0$.

The model given by equation (1) represents the static factor model (SFM) in which the factors (F_t) do not possess their own dynamic and the relationship between the factors and variables is linear with constant weights over time. This model can be estimated either by assuming that the variables are IID (hypothesis SH4), or by assuming that there is a time dynamic within the variables (SH4 is abandoned).

Assuming that (F_t) and (ξ_t) are not correlated and are zero mean, then the variance-covariance matrix for the static factor model, denoted $\Sigma_X = E(X_t X_t')$, is given by:

$$\Sigma_X = \Lambda \Sigma_F \Lambda' + \Sigma_\xi \tag{3}$$

By normalizing the variance-covariance matrices of (F_t), $\Sigma_F = I_r$, and by assuming that the diagonal elements of the variance-covariance matrix Σ_ξ of (ξ_t) are bounded, we obtain:

$$\Sigma_X = \Lambda\Lambda' + \Sigma_\xi \tag{4}$$

For additional details, we refer to Lawley and Maxwell (1971) and Anderson (1984). The static factor model can, thus, be identified and estimated. The factorial analysis method is used for the static estimation of the factors. The weighting matrix Λ can be estimated by minimizing the sum of the squared residuals as follows:

$$\sum_{t=1}^{T} (X_t - \Lambda F_t)' (X_t - \Lambda F_t) \tag{5}$$

subject to the constraint $\Lambda'\Lambda = I_r$.

In this context, Doz and Lenglart (1999) establish the asymptotic properties of the estimator. Specifically, they show that this method produces convergent estimators even when the data used are autocorrelated, as is the case with time series. Moreover, they show empirically that this method provides a very good approximation of the dynamic method, while being easier to establish, which is an essential quality for forecasters who regularly estimate the model.

2.2 Exact or strict dynamic factor models (DFM)

Static factor models (SFM) are different from exact or strict dynamic factor models (DFM) in the sense that the latter incorporate a time dynamic. Thus, in the DFM, the common component can be seen as a sum of common shocks, whether contemporaneous or lagged. The model is, then, defined as follows:

$$x_{it} = \chi_{it} + \xi_{it}, \tag{6}$$

where:

$$\begin{aligned}\chi_{it} = b_{i1}^{0} u_{1t} + \cdots + b_{i1}^{s} u_{1,t-s} + b_{i2}^{0} u_{2t} + \cdots + b_{i2}^{s} u_{2,t-s} \\ + \cdots + b_{iq}^{0} u_{qt} + \cdots + b_{iq} u_{q,t-s},\end{aligned} \tag{7}$$

where (u_t), of dimension $(q \times 1)$, is the vector of common shocks such as $u_t = (u_{1t}, u_{2t}, \dots, u_{qt})'$, with $q \leq N$, and where s is the number of lags included in the model. The parameters (b_{il}^{τ}), for $\tau = 0, \dots, s$, $i = 1, \dots, N$ and $l = 1, \dots, q$, represent the weights of the finite dynamic factors s. We speak of a "restricted" DFM when s is finite and a "generalized" DFM when s is infinite.[1]

Equation (7) can be rewritten as follows:

$$\chi_{it} = \sum_{l=1}^{q} b_{il}(L) u_{lt,} \tag{8}$$

Where $b_{il}(z) = b_{il}^{0} + b_{il}^{1}.z + \ldots + b_{il}^{s}.z^{s}$ are polynomials of degree s and where L is the lag operator so that, for all s, $L^{s} u_t = u_{t-s}$. In a matrix form, equation (7) can be rewritten as:

$$\chi_{it} = B_i(L) u_t, \tag{9}$$

Where $B_i(L) = (b_{i1}(L), \dots, b_{iq}(L))$ is a q-vector of polynomials with degree s.

[1] For a discussion of the relationship between restricted DFMs and generalized DFMs, see Giannone et al. (2006) or Forni et al. (2009).

Moreover, we assume (Bai and Ng, 2007) that the vector $B_i(L)$ can be decomposed as follows:

$$B_i(L) = \lambda_i^*(L)C(L) \tag{10}$$

Where $\lambda_i^*(L)$ is a r-vector of polynomials with degree s and where $C(L)$ is a matrix of dimension $(r \times q)$. Using equations (9) and (10), we can then write the common component in the following manner:

$$\chi_{it} = \lambda_i^*(L)F_t^*, \tag{11}$$

Where $F_t^* = C(L)u_t$ is a vector of dimension r that refers to the *static factors*, and the common shocks u_t of dimension $(q \times 1)$ to the *dynamic factors*. A model with q dynamic factors can, thus, be considered as a model with $r = q(s+1)$ static factors.

In the context of small-dimension dynamic factor models, the estimation is generally done in the time domain by likelihood maximization, as proposed by Dempster et al.(1977), Shumway and Stoffer (1982), Watson and Engle (1983), and Stock and Watson (1989).[2]

To estimate the DFM when N is small, the following hypotheses are generally put forward:

(DH1) The factors (F_{jt}) and $(F_{j't})$ are mutually orthogonal, but the factors (F_{jt}) can be autocorrelated and are variance-covariance stationary, i.e.: $\forall j \neq j'$, $\tau \neq 0$, $E(F_{j,t}) = 0$, $cov(F_{j,t}, F_{j',t-\tau}) = 0$, and $cov(F_{j,t}, F_{j,t-\tau})$ depends only on τ.

(DH2) The idiosyncratic processes (ξ_{it}) and $(\xi_{i't})$ are mutually orthogonal, but the processes (ξ_{it}) can be autocorrelated and covariance-stationary, i.e.: $\forall i \neq i'$, $\tau \neq 0$, $E(\xi_{i,t}) = 0$, $cov(\xi_{i,t}, \xi_{i',t-\tau}) = 0$, and $cov(\xi_{i,t}, \xi_{i,t-\tau})$ depends only on τ.

(DH3) The factors (F_{jt}) and the idiosyncratic processes (ξ_{it}) are orthogonal for all i, j.

[2] Another method of estimating this type of model has been proposed by Sargent and Sims (1977) and Geweke (1977) in the frequency domain, based on a spectral analysis. We come back to this type of frequency domain estimation in the [fourth] section, entitled "Estimation of factor models for large N."

Based on these hypotheses, we can, then, attempt to estimate a dynamic factor model by likelihood maximization in the time domain, with the additional hypothesis of Normality for the model residuals. The maximum likelihood estimator is calculated by, first, placing the model in a space-state form and, then, using a Kalman-type recursive filter.

The DFM can be written in a space-state form, assuming that the common factors follow a VAR process of order p such as:

$$\Phi(L)F_t = \varepsilon_t \qquad \Leftrightarrow \qquad F_t = \sum_{\tau=1}^{p} \Phi_\tau F_{t-\tau} + \varepsilon_t, \tag{12}$$

and, for a given index i, the idiosyncratic process (ξ_{it}) follows an AR process of order p' in the following form:

$$\psi_i(L)\xi_{it} = \eta_{it} \qquad \Leftrightarrow \qquad \xi_{it} = \sum_{\tau=1}^{p'} \psi_{i\tau}\xi_{i,t-\tau} + \eta_{it}, \tag{13}$$

Where (ε_t) and (η_{it}) are innovations of (F_t) and (ξ_{it}), respectively, so that (ε_{it}) and (η_{it}) are independent. $\Phi(.)$ and $\psi_i(.)$ are polynomials of order p and p', respectively, so that:

$$\Phi(L) = I - \Phi_1 L - \ldots - \Phi_p L^p \quad \text{and} \quad \psi_i(L) = I - \psi_{i1} L - \ldots - \psi_{ip'} L^{p'} .$$

The hypothesis of Normality is put forward for (ε_t) and (η_{it}). In practice, the orders p and p' of the lag polynomials must be selected prior to the estimation stage. This selection is generally done by minimizing an AIC-type information criterion (Akaike information criterion) or a BIC-type information criterion (Bayesian information criterion) or by using the Doz and Lenglart (1999) test. In empirical studies, $p = 2$ and $p' = 1$ are often shown to be sufficient to whiten the residuals.

This type of model shown by equations (1), (12) and (13) allows a space-state representation as follows:

$$X_t = c_t\beta_t + m_t Z_t + w_t \tag{14}$$

Where (Z_t) is a vector of *n* explanatory variables, for example, the lagged values of the observed variables (X_t), and where:

$$\beta_t = a_t \beta_{t-1} + v_t \tag{15}$$

Equation (14) is the measure equation, which describes the relations between the unobservable states, of dimension *r*, and the observable variables, of dimension *n*, where β_t represents the state vector:

$$\beta_t = \begin{bmatrix} F_t \\ \vdots \\ F_{t-p+1} \\ \xi_t \\ \vdots \\ \xi_{t-q+1} \end{bmatrix}$$

Equation (15) represents the state or transition equation, which describes the development of unobservable states. We see that a_t, c_t and m_t are matrices that can depend on time, of the dimensions $((p \times r + q \times N) \times (p \times r + q \times N))$, $(N \times (p \times r + q \times N))$ and $(N \times n)$, respectively, and where v_t is a Gaussian white noise vector of dimension $(p \times r \times q \times N)$, w_t is a Gaussian white noise vector of dimension *N*, of the variance-covariance matrices Q_t and R_t, respectively. In practice, the system is generally assumed to be invariant over time, i.e., a_t, c_t and m_t are constant. It is also assumed that for all *t*, $t' \neq t$, $E\ (v_t w'_t) = 0$.

The model in its space-state form can then be estimated by maximum likelihood using a filtering method such as the Kalman filter. We refer, for example, to Hamilton (1994) for a description of the filtering algorithm. The maximum likelihood estimation algorithm can take a great deal of time as it requires inversion of a large dimensional matrix, even when *N* is small. In general, in the case of numerical optimization, the expectation-maximization (EM) algorithm is used, as proposed by Dempster et al. (1977) or Shumway and Stoffer (1982).[3]

[3] Alternatively, the Fisher scoring algorithm is used by Watson and Engle (1983).

2.3 Recent extensions of factor models with small N

Several extensions of factor models with a small number of variables have recently been proposed to take certain data characteristics into account. We present two of these extensions below: Markov regime-switching models and mixed frequency models.

2.3.1 Markov regime-switching models

These models are directly linked to the Markov regime-switching processes introduced by Hamilton (1989) and assume that the common unobservable factors have their own dynamics governed by a two-regime Markov chain, denoted (S_t), with for all t, $S_t \in \{1, 2\}$. The idea of these models is to assume that the factors are related to the state of the economy, which itself evolves cyclically but non-periodically based on two economic phases that follow one another. We, then, assume that, for example, the first regime ($S_t = 1$) corresponds to the low phase of the business cycle and the second regime ($S_t = 2$) to the high phase of the business cycle. The model can easily be extended to a larger number of regimes, but the estimation problems then become tricky in that the model contains two latency levels, i.e., the common factors and the Markov chain.

An initial model of this type was proposed by Diebold and Rudebusch (1996), but the theoretical and empirical aspects were more broadly considered by Kim and Yoo (1995) and Kim and Nelson (1998). At the same time, Chauvet (1998) independently proposed a similar model. For example, in the case of a single factor (*i.e.*, $r = 1$) for N centered stationary variables, the Markov regime-switching model can be defined as follows for $i = 1, \dots, N$ and $t = 1, \dots, T$:

$$x_{it} = \lambda_i' F_t + \xi_{it} \tag{16}$$

With:

$$\phi(L) F_t = \mu(S_t) + \epsilon_t, \tag{17}$$

where the $\lambda_i = (\lambda_{i1}, \dots, \lambda_{ir})$ are the loadings, where, for each i, (ξ_{it}) follows a Gaussian autoregressive process of order one (AR(1)) of finite variance σ_i^2, where (ϵ_t) is a Gaussian white noise of unit variance and where $\phi(L) = I - \phi_1 L - \cdots - \phi_p L^p$. If we assume that ($S_t$) is a first-order two-regime Markov chain, this means that the probability of S_t belonging to a regime at date t depends only on the probability of being in a certain regime at date $t-1$, or:

$$P(S_t \mid S_{t-1}, S_{t-2}, S_{t-3}, \dots) = P(S_t \mid S_{t-1}) \tag{18}$$

The transition probabilities $p_{12} = P(S_t = 2 \mid S_{t-1} = 1)$ and $p_{21} = P(S_t = 1 \mid S_{t-1} = 2)$ measure the probability of switching from one regime to the other. Similarly, the probabilities $p_{11} = 1 - p_{12}$ and $p_{22} = 1 - p_{21}$ measure the probability of remaining in the same regime, thus reflecting the degree of persistence of each regime. The estimation stage allows for the estimation for each date t of the expected, filtered and smoothed probabilities of being in a particular regime, respectively given by $P(S_t \mid \hat{\theta}, X_{t-1}, \dots, X_1)$, $P(S_t \mid \hat{\theta}, X_t, \dots, X_1)$ and $P(S_t \mid \hat{\theta}, X_T, \dots, X_1)$, where $\hat{\theta}$ represents the set of estimated parameters of the model, which includes the N autoregressive parameters of the AR(1)models, the N idiosyncratic variances and the p parameters of the polynomial $\phi(.)$.

The parameters of this model can be estimated simultaneously by maximum likelihood, as proposed by Kim and Nelson (1998), or in two steps, by first estimating the common factor (F_t) in the time or spectral domain (see above), and then adjusting an autoregressive regime-switching process on the estimated factor (see, on this point, Diebold and Rudebusch, 1996). From a theoretical standpoint, simultaneous estimation is preferable but, empirically, the maximization algorithm is often difficult to converge, in particular if the variables are very volatile. The two-stage estimation is more practical (see Darné and Ferrara, 2011, for an application), but the second equation (17) then includes an error in the measure of the estimated factors that is not explicitly integrated in the model, which can thus create statistical inference problems.

2.3.2 Mixed frequency models

Numerous macroeconomic series are available to forecasters but do not necessarily have the same sampling frequency (or periodicity). The national accounts in particular, which most economists seek to predict, are available only on a quarterly basis while many economic indicators, such as the industrial production index (IPI), consumer expenditure by households, or opinion surveys are monthly. To be able to simultaneously handle these two periodicities in a single model, Mariano and Murasawa (2003) proposed a dynamic factor model in a space-state form that considers quarterly series as monthly series with missing values.

The aim of this type of model is to estimate a common factor with *N* variables, some of which are quarterly and some of which are monthly. Thus, we write $(Y_{1,t})$ a vector of N_1 quarterly variables that are observable only in the third month t of the quarter and $(Y_{2,t})$ a vector of N_2 monthly variables, so that $N_1 + N_2 = N$. We assume here that these series (in logarithms) are integrated of order one. We also assume that there is a vector $(Y^*_{1,t})_t$ of N_1 unobservable monthly variables so that for all t, $Y_{1,t}$ is the geometric mean of $Y^*_{1,t}$ over the three months in a given quarter, i.e.:

$$\log\left(Y_{1,t}\right) = \frac{1}{3}\left[\log(Y^*_{1,t}) + \log\left(Y^*_{1,t-1}\right) + \log(Y^*_{1,t-2})\right] \tag{19}$$

We note that this identity (19) differs from the arithmetical mean generally used in the quarterly accounts but makes it possible to implement a linear space-state form, in contrast to the arithmetical identity, that requires a non-linear form.

Mariano and Murasawa (2003), then, show that:

$$y_{1,t} = \frac{1}{3}y^*_{1,t} + \frac{2}{3}y^*_{1,t-1} + y^*_{1,t-2} + \frac{2}{3}y^*_{1,t-3} + \frac{1}{3}y^*_{1,t-4} \tag{20}$$

Where $y_{1,t} = Y_{1,t} - Y_{1,t-3}$ and $y^*_{1,t} = Y^*_{1,t} - Y^*_{1,t-1}$, $(y_{1,t})$ being observable in all three periods and $(y^*_{1,t})$ being unobservable.

Under all of the standard hypotheses of dynamics on factors and idiosyncratic errors, and the Normality of residuals (see Mariano and Murasawa, 2003, p. 430), one can show that the model with a factor (F_t) is written in the following form for all *t*:

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} a_1(\frac{1}{3}F_t + \frac{2}{3}F_{t-1} + F_{t-2} + \frac{2}{3}F_{t-3} + \frac{1}{3}F_{t-4}) \\ a_2F_t \end{pmatrix}$$
$$+ \begin{pmatrix} \frac{1}{3}u_{1,t} + \frac{2}{3}u_{1,t-1} + u_{1,t-2} + \frac{2}{3}u_{1,t-3} + \frac{1}{3}u_{1,t-4} \\ u_{2,t} \end{pmatrix} \tag{21}$$

Where $a = (a_1', a_2')'$ is the weight vector of dimension *N*, (F_t) is the scalar common factor, $u_t = (u_{1,t}', u_{2,t}')'$ is the idiosyncratic component of dimension *N*, and where $y_{2,t} = Y_{2,t} - Y_{2,t-1}$. The model is, then, put in a space-state form (see Mariano and Murasawa, 2003, p. 431) and estimated by maximum likelihood with the help of a Kalman filter.

Other approaches have been proposed to manage data with different periodicities simultaneously in factor models. For example, Aruoba, Diebold and Scotti (2009) also propose a factor model containing four variables of different periodicities (daily, weekly, monthly and quarterly) to estimate U.S. GDP with a high periodicity. This indicator is actually updated weekly by the Philadelphia Federal Reserve on its internet site. Similarly, Camacho and Perez-Quiros (2010, 2011) propose a factor model that simultaneously treats variables with different periodicities and regime-switching in factors to estimate GDP growth in the eurozone and Spain. For French data, Cornec and Desperraz (2006) construct a composite activity indicator for services based on monthly and quarterly survey data. Cornec (2006) also develops an indicator based on a mixed frequency factor model (see the last section on applications). More generally, we should point out that this type of approach makes it possible to treat the problem of missing data in series in an econometric model.

3 Approximate factor models (large *N*)

Although the concept of factor models is attractive, the traditional approach presented in the previous section has a number of limitations that are both theoretical and practical in nature.

1. The number of variables (*N*) is often larger than the number of observations (*T*) in economic data series. Consequently, potentially important information is lost when a small number of variables must be selected to respect the constraint that *N* be small;
2. Asymptotic convergence of estimators is assured when *T* tends to infinity and *N* is fixed, but not when *N* also tends to infinity;

3. IID hypotheses and hypotheses on the diagonality of the variance-covariance matrix of the idiosyncratic component Σ_ξ, which prohibit the cross correlation, are often too strong for economic data. This can result in a risk of misspecification;
4. The maximum likelihood estimation (MLE) is generally considered unachievable for factor models of large dimensions because the number of parameters to be estimated is too large (Bai, 2003; Bai and Ng, 2002);
5. The traditional approach makes it possible to consistently estimate the coefficients of the weighting factors (λ_i) by MLE when T is large, but not the common factors (F_t), for which only the estimated value can be obtained (Steiger, 1979). Meanwhile, in most economic problems, it is these common factors that are of greatest interest since they represent the common shocks, the diffusion indices, etc., for example.

To respond to a number of these limitations, the idea of factor models was generalized to allow for the manipulation of less strict hypotheses on the variance-covariance matrix of the idiosyncratic components by proposing an approximate factor structure. Non-parametric estimators of common factors based on the principal components have been suggested (Forni et al., 2000; Stock and Watson, 2002), their asymptotic properties being known when N is large. New methodologies have, consequently, been proposed.

3.1 Approximate static factor models (SFM)

Chamberlain and Rothschild (1983) are the first to introduce the so-called "approximate" factor structure concept by abandoning the hypothesis that idiosyncratic disturbances are not mutually correlated within the so-called "strict" factor structure, *i.e.*, allowing for idiosyncratic errors to be weakly correlated. This concept makes it possible to obtain a non-diagonal variance-covariance matrix $\Sigma_\xi = E(\xi_t \xi_t')$. Moreover, Chamberlain and Rothschild (1983) show that a principal components analysis (PCA) is equivalent to a factor analysis (or to a maximum likelihood analysis under the hypothesis of Normality of the idiosyncratic component (ξ_{it})) when N increases to infinity. However, they assume that the variance-covariance matrix of the population, Σ_X, of dimension$(N \times N)$, is known. Connor and Korajczyk (1986, 1988, 1993) study the case of an unknown variance-covariance matrix Σ_X and suggest that, when N is larger than T, the factor model can be estimated by applying a PCA to the variance-covariance matrix Σ_X, of dimension $(T \times T)$.

Connor and Korajczyk (1986) establish the coherence of factors estimated by PCA when T is fixed and N tends to infinity in the context of approximate factor models but they provide no formal argument when N and T simultaneously tend to infinity.[4] Stock and Watson (1999) study the uniform coherence of estimated factors and derive convergence rates for large T and N. The convergence rate is also studied by Bai and Ng (2002). Finally, Bai (2003) shows that the PCA estimator of the common component is asymptotically Gaussian, converging to a rate equal to $\min(N^{1/2}, T^{1/2})$, even when the idiosyncratic component is serially correlated and/or heteroskedastic when N and T are large.[5]

3.2 Approximate dynamic factor models (DFM)

Forni and Lippi (1997), Forni and Reichlin (1998) and Forni et al. (2000, 2004) extend approximate factor models by considering dynamic factor models of large dimensions and introduce different methods for the estimation of this type of model. These models are referred to as "generalized" because they combine both dynamic and approximate structures, *i.e.*, they generalize exact dynamic factor models by assuming that the number of variables N tends to infinity and by allowing idiosyncratic processes to be mutually correlated.

Forni et al. (2000, 2004) expand the dynamic principal components analysis introduced by Brillinger (1981) when N is large. The estimation proposed by Brillinger (1981) generalizes the static PCA by placing the analysis in the frequency domain. First, the spectral density of the vector X_t is estimated using a consistent spectral density estimator, denoted $\hat{S}(\omega)$, for a frequency $\omega \in]0,2\pi]$. Then, the eigenvectors corresponding to the largest q eigenvalues of this

---

[4] Ding and Hwang (1999) obtain results on the coherence for PCA estimation of traditional exact factor models when N and T tend to infinity.

[5] Jones (2001) and Boivin and Ng (2005) propose weighted PCA estimators by considering the problem of non-linear generalized least squares as follows: $\min_{F_1,\ldots,F_T,\Lambda} \sum_{t=1}^{T} (X_t - \Lambda F_t)' \sum_{\xi}^{-1} (X_t - \Lambda F_t)$, with (ξ_t) being IID and of the normal distribution $N(0, \Sigma_{\xi})$. Stock and Watson (2005a) extend the weighted PCA approach by assuming an autoregressive structure of weak order for (ξ_t).

spectral matrix are calculated. Finally, one returns to the time domain by applying the inverse Fourier transform to these eigenvectors, to recover the estimators of the time series in dynamic principal components (see next section).

Brillinger (1981) obtains distributional results when N is fixed and T tends to infinity. Forni et al. (2000) show that the dynamic PCA provides a consistent estimate of the common component when both N and T increase. Forni et al.(2004) discuss the coherence conditions and convergence rates. It has been shown that the principal components are convergent estimators of factors, both in the static context (Bai and Ng, 2002; Stock and Watson, 2002; Bai, 2003) and in the dynamic context (Forni et al., 2000, 2004).

Approximate factor models have several advantages over strict models. They are flexible and appropriate under general hypotheses on measurement errors and, usually, on the cross-correlation of idiosyncratic components. The misspecification error resulting from the approximate structure of the idiosyncratic component disappears when N and T are large, as long as the cross-correlation of the idiosyncratic processes is relatively small and that of the common components increases across the transverse dimension when N increases. These conditions are introduced in Chamberlain and Rothschild (1983) and used, reinterpreted and expanded in Connor and Korajczyk (1986, 1988, 1993), Forni and Lippi (1997), Forni et al. (2000) and Stock and Watson (2002), respectively. In short, approximate factor models have two important advantages over traditional factor models:

1. The idiosyncratic components can both be weakly mutually correlated and show little heteroskedasticity. This can reflect the condition in which all the eigenvalues of the idiosyncratic variance-covariance matrix $\Sigma_{\xi} = E(\xi_t \xi_t')$ are bounded. Thus, the absolute mean of the covariances is bounded, *i.e.*, $\lim_{N\to\infty} N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \left| E(\xi_{it}\xi_{jt}) \right| < \infty$ (Stock and Watson, 2002);[6]

[6] The precise technical condition allowing for a weak correlation of idiosyncratic terms varies from one study to another but, in general, this condition limits the contribution of the idiosyncratic covariances to the variance-covariance matrix of X_t when N is large.

2. In this type of model, it is possible to have a weak correlation between the factors (F_t) and the idiosyncratic components (ξ_t).

3.3 Recent extensions of approximate factor models

Among the recent extensions of dynamic factor models when N is large, we would like to mention FAVAR models, models whose parameters vary over time and mixed frequency models. The applications of these different types of models are presented in the last section on applications. We will see that these models assume that the variables are stationary. For the development of dynamic factor models on data that are non-stationary, we refer, for example, to Peña and Poncela (2006a, 2006b). Moreover, Banerjee and Marcellino (2009) have extended the FAVAR models to the factor-augmented error correction model, which makes it possible to integrate variables that are non-stationary.

3.3.1 FAVAR models

To remedy the problem of missing variables, generally encountered in traditional VAR and SVAR (structural VAR) modeling, Bernanke, Boivin and Eliasz (2005) propose using factor-augmented VAR (FAVAR) models, particularly in the context of monetary policy analysis.
The FAVAR model can be described by the following equation:

$$X_t = \Lambda F_t + BX_{t-1} + \xi_t \tag{22}$$

where (X_t) represents the endogenous variables of a traditional VAR model, such as in Bernanke et al. (2005) and Boivin et al. (2009), (F_t) the common factor, Λ the weighting matrix, and (ξ_t) the idiosyncratic component. In Stock and Wilson (2005a), B is a diagonal matrix $D(L)$ with a lag polynomial $\delta_i(L)$ on the i^{th} diagonal. It is also conceivable to specify a short-term dynamics, of the first-order autoregressive type, for example, on the common factor (F_t) and on the idiosyncratic component (ξ_t).
Stock and Watson (2005a) propose using an iterative procedure to estimate the FAVAR model given by equation (22). This procedure begins with an initial estimation of the static factor $\hat{F}_t$ using a PCA. Then, the weighting matrix $\hat{\Lambda}$ and the coefficients $\hat{B}$ are estimated by ordinary least

squares. Finally, the $\hat{F}_t$ factors are re-estimated by the principal components of $X_t - \hat{B}X_{t-1}$ and this procedure is iterated until convergence. Boivin et al. (2009) also use this iterative procedure with $B = 0$ as the initial estimation. Bernanke et al. (2005) propose estimating the unobservable factors in two steps: (1) the principal components of the informational variables are, first, calculated by ignoring the presence of the observable variables; (2) equation (22) is, then, estimated by integrating the factors estimated in the previous step. Other works on monetary policy analysis involve the estimation of factors using a dynamic factor model, these estimated factors then being introduced into a VAR model as additional regressors (see, for example, Giannone et al., 2004, and Favero et al., 2005).

3.3.2 Time-varying parameter models

Some authors are also interested in factor models for which, for example, the weights grouped in the matrix Λ of equation (1) vary over time (see, for example, Motta, Hafner and von Sachs, 2011). This type of approach is promising since it makes it possible to integrate structural changes in respect of the source and amplitude of the shocks into the modeling, as well as their channels of transmission to the economy. As well, this type of modeling incorporates non-linearity into relations. In particular, this makes it possible to assess whether behaviors change over the course of a business cycle.

In the area of FAVAR models, some recent papers incorporate a dynamic structure by allowing a change over time (i) of the weights, (ii) of the autoregressive dynamics of the factors, or (iii) of the variance of innovations. Thus, a possible specification of a time-varying FAVAR (TV-FAVAR) model is as follows:

$$X_t = \Lambda_t F_t + B_t X_{t-1} + \xi_t \tag{23}$$

With $E(\xi_t) = 0$ and $E(\xi_t \xi_t') = \Sigma_{\xi_t}$.

Generally, the FAVAR model given by equation (23) is estimated in a Bayesian context, as is done, for example, in the articles of Del Negro and Otrok (2008), Mumtaz and Surico (2009) or Baumeister, Liu and Mumtaz (2010). Similarly, Kose, Otrok and Whiteman (2003) propose an estimation of dynamic factor models based on an MCMC (Monte Carlo Markov-Chain)[7] approach in a Bayesian context.

Eickmeier, Lemke and Marcellino (2011) propose an alternative approach by developing a two-step standard estimation model by maximum likelihood. The first step consists in estimating the factors in a static context (see next section). Then, in a second step, the model with non-constant parameters is estimated equation by equation. Thus, each univariate regression equation is put in a space-state form, then estimated in the traditional fashion using a Kalman filter. This approach requires assuming that FAVAR model equations are conditionally independent.

3.3.3 Mixed frequency models

Finally, in the case of a large number N of available variables in the data set, variables with different periodicities can be managed using a MIDAS (MIxed DAta Sampling) regression proposed by Ghysels and his co-authors (see, for example, Ghysels, Sinko and Valkanov, 2007, for a presentation). The MIDAS approach can be used to explain a sampled variable with a certain periodicity (such as annual or quarterly) by variables with a higher periodicity (such as monthly or daily) without having to, first, aggregate the higher periodicity data. This approach is based on a standard linear regression equation but involves the estimation of a weight function depending on a hyper-parameter with a smaller dimension compared with the initial dimension of the problem.

Let us assume, for example, that we are seeking to estimate the quarterly GDP growth rate of an economy, denoted (y_t), assumed to be stationary, for a number of quarters T, the index t here designating the quarter. Let us also assume that, with the help of one of the methods shown in the next section, we have estimated a single monthly stationary factor $(\hat{F}_t^{(m)})$, (*i.e.*, $r = 1$) based on a large monthly data set. We, thus, observe m = 3 times $(\hat{F}_t^{(m)})$ over the period $[t - 1, t]$. The standard MIDAS equation makes it possible to link the quarterly variable to be explained to the monthly estimated factor as follows:

[7] MCMC methods are numerical methods that create long Markov chains, making it thus possible to obtain samples distributed asymptotically according to a certain distribution.

$$y_t = c_0 + c_1\, G(\theta)\hat{F}_t^{(m)} + \epsilon_t, \tag{24}$$

Where c_0 and c_1 are parameters to be estimated and where $(\epsilon_t)\epsilon$ is assumed to be a Gaussian white noise of finite variance that will also need to be estimated. The term $G(\theta)$ controls the polynomial weights, which allows the mixing of frequencies. In fact, the MIDAS specification consists in smoothing the past values of the $\left(\hat{F}_t^{(m)}\right)$ using the polynomial $G(\theta)$ of the following form:

$$G(\theta) = \sum_{k=1}^{K} g_k(\theta) L^{(k-1)/m} \tag{25}$$

Where K is the number of points on which the smoothing operates, L is the lag operator, so that, for any monthly variable $x_t^{(m)}$, $L^{\tau/m} x_t^{(m)} = x_{t-\tau/m}^{(m)}$ and $g_k(.)$ is the weight function, which can take various forms. As in Ghysels et al. (2007), we generally use a two-parameter Almon polynomial $\theta = (\theta_1, \theta_2)$ so that:

$$g_k(\theta) \equiv g_k(\theta_1, \theta_2) = \frac{\exp\left(\theta_1 k + \theta_2 k^2\right)}{\sum_{k=1}^{K} \exp\left(\theta_1 k + \theta_2 k^2\right)} \tag{26}$$

The θ parameter is part of the estimation problem. It is influenced by the information contained in the last K values of $\left(\hat{F}_t^{(m)}\right)$, the size of the window K being an exogenous parameter. Other specifications can be considered in the literature for equation (24), particularly by adding monthly explanatory variables or autoregressive terms for the target variable (y_t). Similarly, other weight functions can be considered.

In terms of application, Marcellino and Schumacher (2010) propose an approach in which they, first, estimate the monthly factors based on a data set of 111 representative variables of the German economy. They, then, use these factors to forecast the quarterly German GDP using a MIDAS regression (so-called factor-MIDAS approach). The authors show the usefulness of such an approach in making better use of the most recent data for short-term macroeconomic forecasting purposes. As for the MIDAS approach itself, some examples of applications, in forecasting, for example, are found in the articles by Clements and Galvao (2008) and Ferrara and Marsilli (2013).

4 Estimation of factor models for large *N*

In this section, we present the main estimation methods of factor models, whether static or dynamic, when the number of variables is high (large *N*). In this case, the usual methods based on maximizing likelihood run into the problem of the dimension of the parameter to be estimated.

4.1 Static factor models: the Stock and Watson (2002) approach

One of the first approximate factor models is the one proposed by Stock and Watson (SW) (2002), which is based on a static PCA. The PCA is used since it allows for the estimation of both the parameters and the factors of the model given by equation (1) by maximizing the variance explained by the initial variables, for a small number *r* of static factors (F_t). The main aim of the SW approach is to approximate the factors by a linear combination of the data $\hat{F}_{j,t} = \widehat{W}_j' X_t$, for $j = 1, \dots, r$, that maximizes the variance of the estimated factors $\widehat{W}_j' \hat{\Sigma}_x \widehat{W}_j$, where $\hat{\Sigma}_x = (1/T) \sum_{t=1}^{T} X_t X_t'$ is the empirical variance-covariance matrix of the vector of the initial standardized data X_t.

Under the following normalization assumption: $\widehat{W}_j' \widehat{W}_{j'} = 1$ for $j = j'$ and $\widehat{W}_j' \widehat{W}_{j'} = 0$ for $j \neq j'$, the maximization problem can, then, be transformed into the solution of a eigenvalues' problem:

$$\hat{\Sigma}_x \widehat{W}_j = \hat{\mu}_j \widehat{W}_j \tag{27}$$

Where $\hat{\mu}_j$ is the j^{th} eigenvalue and $\widehat{W}_j$ is the associated eigenvector of dimension $(N \times 1)$. Once they have been calculated, the highest *N* eigenvalues are classified in decreasing order. Then, the eigenvectors are, in turn, classified in decreasing order with respect to the highest *r* eigenvalues. The factors proposed by SW are, then, written as follows:

$$F_t^{SW} = \widehat{W}' X_t, \tag{28}$$

Where $\widehat{W}$ is the matrix of dimension $(N \times r)$ of the stacked eigenvectors $\widehat{W} = (\widehat{W}_1, \dots, \widehat{W}_r)$.[8]

[8] Stock and Watson (1998) develop theoretical results for this methodology.

However, the Stock and Watson approach does not allow for use of the different dynamics that may exist between the variables used. To take account of this dynamic structure in factor models, several alternatives to the static factor model have been proposed in the literature. Specifically, there are two main types of dynamic factor models or approaches. Developed by Doz, Giannone and Reichlin (2011, 2012), the first one is based on a space-state representation of models in the time domain. Proposed by Forni, Hallin, Lippi and Reichlin (2004, 2005), the second one is based on the spectral domain. We, now, present the estimation strategies of these different dynamic factor models.

4.2 Dynamic factor models

4.2.1 Time domain approach

Doz, Giannone and Reichlin (DGR) (2011, 2012) propose a dynamic factor model that can be represented in a space-state form. Specifically, DGR (2011, 2012) estimate their dynamic factor model using two different approaches. The first one is the so-called two-step approach (DGR, 2011). The second one is based on the quasi maximum likelihood (DGR, 2012).

According to DGR (2011), for a number r of factors and q of dynamic shocks, the estimation is carried out in two steps. In the first step:

1. $\hat{F}_t$ is estimated using a PCA, as an initial estimate;
2. Then, equations (6) and (11) are estimated using the estimated factor from the previous step, $\hat{F}_t$, to obtain both $\hat{\lambda}_i^*(L)$ and the variance-covariance matrix of the residuals $\hat{\xi^*}$, denoted $\hat{\Sigma}_{\xi^*}$. To obtain an estimate of $C(L)$, appearing in equation (10), DGR (2011) apply a decomposition of eigenvalues to the matrix $\hat{\Sigma}_{\xi^*}$ by taking into account the number of dynamic shocks q. Let us introduce the matrix M of dimension $(r \times q)$ corresponding to the largest q eigenvalues and the matrix P of dimension $(q \times q)$ containing the largest q eigenvalues in its diagonal and zeros elsewhere. The estimate of $C(L)$ is, then, obtained by $\hat{C}(\mathrm{L}) = M \times P^{-1/2}$.

In a second step, the coefficients and parameters of the system described by equations (6) and (11) are considered to be known and provided by the first step. The model is, then, written in a space-state form and the Kalman filter is applied to obtain new estimates of the factors.

In their alternative approach, DGR (2012) estimate an approximate dynamic factor model using the quasi maximum likelihood method.[9] The main aim of this approach is to consider the strict factor model as a misspecification of the approximate factor model and to analyze the properties of the maximum likelihood indicator of the factors under this misspecification. This estimator is called the *quasi* maximum likelihood in the sense of White (1982). By analyzing the properties of the maximum likelihood estimator under several sources of misspecifications, such as an omitted serial correlation of the observations or a cross-sectional correlation of the idiosyncratic components, DGR (2012) show that these misspecifications do not affect the robustness of the common factors, particularly for fairly large N and T. More specifically, this estimator is a valid parametric alternative for the estimator resulting from a PCA. The model defined by means of equations (6) and (11) can be put in a space-state form, with a number of states equal to the number of common factors r. It is noteworthy that the estimation of the parameters of the model, particularly the common factors, by the quasi maximum likelihood can be approximated by their anticipated values, using the Kalman filter.[10]

These dynamic factor models have also been called restricted dynamic factor models, since the r static factors are caused by a number q of dynamic factors, with $q \leq r$ (Forni et al., 2005; Hallin and Liska, 2007).

Kapetanios and Marcellino (2004) also propose an approach based on a space-state representation. Their approach is based on the use of specific subspaces in which the factors are estimated. This subspace algorithm can be used to estimate factors without having to specify or identify the model entirely in its space-state form.

---

[9] Jungbacker and Koopman (2008) propose new results for the estimation of a dynamic factor model using the maximum likelihood method and a Bayesian method based on Markov chains. Jungbacker et al. (2011) adapt this approach in the context of missing data.

[10] The likelihood can be maximized by means of the EM algorithm, which requires the use of the Kalman filter for each iteration.

4.2.2 Frequency domain approach

In a series of articles, Forni, Hallin, Lippi and Reichlin (2000, 2003, 2004, 2005) (FHLR) propose a dynamic PCA in the frequency domain, also called a generalized dynamic factor model, to estimate dynamic factors.[11] The purpose of their model is to identify the dynamic structure of a factor model. The dynamic factor model is given by equations (6) and (7). The method proposed by FHLR makes it possible to estimate dynamic factors in a first step and, then, obtain the static factors from the estimated dynamic factors in a second step. The approach proposed by FHLR aims to estimate both the dynamic factors and their covariances. This estimation is performed to maximize the variance of the common component under certain orthogonality restrictions. The optimization program is likened to a problem to determine the dynamic eigenvalues of the spectral density matrix of the variables observed. The spectral matrix $I_x(\omega)$ of X_t is estimated using a representation of time series in the frequency domain for each frequency ω in the interval $[0,2\pi[$. The estimated spectral matrix contains information on both the cross-correlation between variables and their dynamic relations. Thus, we write $\widehat{\Sigma}_x(\tau)$ the estimated autocovariance matrix between X_t and $X_{t-\tau}$ for a particular lag τ. The estimated spectral density of the vector of observed variables is given by:

$$\widehat{I}_x(\omega_h) = \sum_{\tau=-H}^{H} \widehat{\Sigma}_x(\tau)\left(1 - \frac{|\tau|}{H+1}\right) e^{-i\tau\omega_h} \tag{29}$$

for each Fourier frequency $\omega_h = 2\pi h/(2H+1)$ and for each $h = 0, \dots, 2H$, with i representing the imaginary number such as $i^2 = -1$. For each frequency ω_h, the dynamic eigenvalues and eigenvectors resulting from $\hat{I}_x(\omega_h)$ are calculated. The eigenvectors are classified in decreasing order. More specifically, the eigenvectors $\hat{P}_l(\omega_h)$, of dimension $(N \times 1)$, are collected for $l = 1, \dots, q$ (the highest q eigenvalues). To return to the time domain, the eigenvectors are obtained based on the inverse Fourier transform:

[11] FHLR generalize the dynamic factor model of Sargent and Sims (1977) and Geweke (1977) by raising the hypothesis of the orthogonality of the idiosyncratic factors (see also Forni, Giannone, Lippi and Reichlin, 2009). Hallin and Liska (2011) recently adapted these models to estimate common factors specific to "blocks" of data, *i.e.*, large subpanels of variables.

(30) $\hat{P}_l(L) = \sum_{\tau=-H}^{H} \hat{P}_{l,\tau} L^{\tau}$, with $\hat{P}_{l,\tau} = \frac{1}{2H+1} \sum_{h=0}^{2H} \hat{P}_l(\omega_h) e^{i\tau\omega_h}$,

for $\tau = -H, \ldots, H$ and $j = 1, \ldots, q$. The j^{th} dynamic principal component $\hat{F}_{j,t}$, is, then, given by the j^{th} component of $\sum_{l=1}^{q} \hat{P}_l' (L) \hat{P}_l(L) X_t$.

Thus, the dynamic principal components are obtained from a decomposition of the spectral density matrix into dynamic eigenvalues and eigenvectors. This breakdown also makes it possible to divide the spectral density matrix into a spectral density matrix of common components $I_{\chi}(\omega)$ and a spectral density matrix of idiosyncratic components $I_{\xi}(\omega)$.

Moreover, the estimator of the frequency domain is reduced to a symmetric filter. This presents problems at the end of samples, particularly when future observations are useful to estimate the principal components. To remedy this problem, FHLR (2005) suggest a refinement of their procedure that maintains the advantages of the dynamic approach, while basing the estimation of the common components on an asymmetric filter.[12] With this procedure, the space of the factors is approximated by adding the *r* static factors rather than the *q* dynamic principal components. However, the mean resulting from the *r* contemporary static factors is based on information from the dynamic approach. The estimation of the model, then, consists in maximizing the variance of the common components or minimizing the variance of the idiosyncratic components. Thus, a convergent estimator of the spectral density matrix of the common component is given by:

$$\hat{I}_{\chi}(\omega) = \hat{P}(\omega) \Omega(\omega) \hat{P}'(\omega), \tag{31}$$

Where $\Omega(\omega)$ is a diagonal matrix of dimension $(q \times q)$ containing the largest q dynamic eigenvalues on the diagonal and $\hat{P}(\omega) = (\hat{P}_1(\omega), \ldots, \hat{P}_q(\omega))$ is a matrix of dimension $(N \times q)$ containing the eigenvectors corresponding to the frequency ω. We, then, deduce the spectral density matrix of the idiosyncratic component:

$$\hat{I}_{\xi}(\omega) = \hat{I}_x(\omega) - \hat{I}_{\chi}(\omega) \tag{32}$$

This estimation in the frequency domain is carried out in two steps. The first one is based on the dynamic approach, by which we obtain the variance-covariance matrices of the common

[12] See also Forni and Lippi (2011).

components $\hat{I}_{\chi}(\omega)$ and idiosyncratic components $\hat{I}_{\xi}(\omega)$, respectively, estimated by an inverse Fourier transform. Thus, the variance-covariance matrix of the common components is estimated as follows:

$$\widehat{\Sigma}_{\chi}(\tau) = \frac{1}{2H+1} \sum_{h=0}^{2H} \widehat{I}_{\chi}(\omega_h) e^{i\tau\omega_h} \tag{33}$$

for $\tau = -H, \ldots, H$. The variance-covariance matrix of the idiosyncratic components is estimated in the same way:

$$\widehat{\Sigma}_{\xi}(\tau) = \frac{1}{2H+1} \sum_{h=0}^{2H} \widehat{I}_{\xi}(\omega_h) e^{i\tau\omega_h} \tag{34}$$

In a second step, this information is used to construct the factor space using the *r* aggregated means. More specifically, the variables are weighted in terms of the common-to-idiosyncratic variance ratio, obtained by means of the variance-covariance matrices estimated in the first step. These *r* aggregated means are defined as the solutions to a generalized principal components problem and they have the advantage of minimizing the idiosyncratic quadratic errors of the common factors by selecting only those variables with the highest common-to-idiosyncratic variance ratio. The number of these aggregated means is equal to $r = s(p+1)$, which represents the rank of the spectral density matrix of the common factors, where *s* indicates the number of lags for $\lambda_i^*(L)$ in equation (10). FHLR (2005) show that to determine the number of aggregated means *r*, the problem of maximization can be converted into a problem of generalized eigenvalues:

$$\widehat{\Sigma}_{\chi}(0)\hat{Z}_j = \hat{\mu}_j \widehat{\Sigma}_{\xi}(0)\hat{Z}_j \tag{36}$$

Where $\hat{\mu}_j$ is the j^{th} generalized eigenvalue, $\hat{Z}_j$ its associated eigenvector of dimension $(N \times 1)$, and $\widehat{\Sigma}_{\chi}(0)$ and $\widehat{\Sigma}_{\xi}(0)$ are the contemporaneous variance-covariance matrices $(\tau = 0)$ of the common and idiosyncratic components, respectively. Moreover, FHLR (2005) impose the following normalization $\hat{Z}_j' \widehat{\Sigma}_{\xi}(0) \hat{Z}_{j'} = 1$ for $j = j'$ and $\hat{Z}_j' \widehat{\Sigma}_{\xi}(0) \hat{Z}_{j'} = 0$ for $j \neq j'$. Then, the eigenvalues are classified in decreasing order and the factors obtained correspond to the product of the *r* eigenvectors corresponding to the highest eigenvalues and the vector X_t. The estimator proposed by FHLR (2005) is written as follows:

$$F_t^{FHLR} = \widehat{Z}' X_t \tag{36}$$

Where $\hat{Z} = (\widehat{Z_1}, \dots, \widehat{Z_r})$ is a matrix of dimension $(N \times r)$ of the stacked eigenvectors.

To conclude this section, we will point out that the estimation methods proposed are relatively recent and the literature does not yet have sufficient perspective to systematically choose one method over another. Because it is easy to use, the Stock and Watson (2002) approach is naturally attractive, and the empirical results, particularly in a forecasting context, show that the results yielded by this approach are not significantly poorer than the other approaches in terms of forecasting error (on this point, see D'Agostino and Giannone, 2012, or Barhoumi, Darné and Ferrara, 2013).

The recent literature has also looked at the estimation of factor models in a Bayesian context. This approach makes it possible to reduce uncertainty regarding the parameters by first applying hypotheses on the distributions of these parameters. In this regard, we refer interested readers to Kose, Otrok and Whiteman (2003, 2008) or Lopes and West (2004), for example.

Finally, the asymptotic properties of the estimators presented above are proved under the simple hypothesis "*N* and *T* tend to infinity," the interpretation of which is sometimes rather vague. Bai (2003) and Forni et al. (2004) emphasize that asymptotic properties, such as convergence, hold along specific trajectories $\{(N, T(N));\ N \in \mathbb{N}\}$. For example, a property that holds for $\min(N, T)$ holds along the entire trajectory $(N, T(N))$, while a property that holds for $N = O(T^k)$ requires that the number of observations *T* be at least on the order of $N^{1/T}$. In fact, three concepts of limits exist: (i) sequential, (ii) pairwise, and (iii) simultaneous. Let $g(N, T)$ be a function that one wishes to study. A sequential limit stretches *N* and *T* to infinity, one after the other. A pairwise limit stretches (N, T) to infinity only along a particular trajectory, which can be denoted $lim_{N,T\to\infty} g\big(N, T(N)\big)$. A simultaneous limit authorizes (N, T) to increase along all possible trajectories: $\lim_{N,T\to\infty} g(N, T)$. It is noteworthy that the existence of a simultaneous limit implies the existence of a pairwise limit and a sequential limit, but the reverse is not true. Another approach using the theory of random matrices postulates that *N* and *T* tend to infinity with $N/T \to c \in (0, \infty)$, where c is a constant. For a more detailed discussion, see, for example, Bai and Ng (2008b) and Harding (2009).

5 Selection of the number of factors

An important step in the statistical analysis of static and dynamic factor models is the preliminary identification of the number of factors. A number of papers focus on the problem of determining the number of factors. For example, Forni and Reichlin (1998) suggest a graphic approach to identify the number of factors when $N \to \infty$ and T is fixed but no theory is proposed. Stock and Watson (1998) modify the BIC criterion to select the optimal number of factors in forecasting when $N, T \to \infty$ with $\sqrt{N}/T \to \infty$. However, their criterion is restrictive since it requires that $N \gg T$ and it is appropriate only in a forecasting context. Forni et al. (2000) consider a multivariate version of the AIC criterion but no theoretic or empirical property is known for their criterion.

In this section, we present the criteria most used in the empirical literature, *i.e.*, the criteria of Bai and Ng (2002) and Alessi et al. (2010) for static factor models and those of Stock and Watson (2005a), Amengual and Watson (2007), Bai and Ng (2007), Hallin and Liska (2007) and Breitung and Pigorsch (2013) for dynamic factor models. Note that those criteria have been compared in a forecasting framework by Barhoumi, Darné and Ferrara (2013).

5.1 Selection of the number of factors for static factor models

To specify the number of factors, Bai and Ng (2002) suggest using information criteria to select the optimal number of static factors r when N and T tend to infinity. Bai and Ng (2002) propose information criteria based on the quality of adjustment of the model to the data measured by the variance $V(j, F)$ such that:

$$V(j,F) = (NT)^{-1} \sum_{t=1}^{T} \left(X_t - \widehat{\Lambda} \widehat{F}_t \right)^2, \tag{37}$$

Where j is a given number of factors such as $\hat{F}_t = (\hat{F}_{1t}, \dots, \hat{F}_{jt})'$. Thus, if the number of factors j increases, the variance of the factors increases mechanically and the sum of the squares of the residuals decreases in turn. Bai and Ng (2002), then, suggest introducing a penalty function in the criterion to be optimized and propose the following three criteria, corresponding to different penalty functions:

$$IC_1(j) = \ln\left(V(j,F)\right) + j\left(\frac{N+T}{NT}\right)\ln\left(\frac{NT}{N+T}\right), \tag{38}$$

$$IC_2(j) = \ln\left(V(j,F)\right) + j\left(\frac{N+T}{NT}\right), \tag{39}$$

$$IC_3(j) = \ln\ (V(j,F)) + j(\ln C_{NT}^2/C_{NT}^2), \tag{40}$$

where $C_{NT} = \min\{\sqrt{N},\sqrt{\mathrm{T}}\}$ and ln denotes the natural logarithm. The estimation of the number of factors r is obtained by minimization of the information criteria for $j = 0, \ldots, r_{max}$, where r_{max} is the maximum number of static factors. These criteria reflect the trade-off between the quality of the adjustment and the risk of overadjustment.[13] Bai and Ng (2002) show that their criteria are robust to the presence of a heteroskedastic component in the time and cross-section dimensions between variables, but also in the presence of weak serial and cross-section dependence.

Subsequently, Alessi et al. (2010) extend this criterion by modifying the strength of the penalty function that appears in the preceding three criteria given by equations (38), (39), and (40). Alessi et al. (2010) propose an alternative to the criteria proposed by Bai and Ng (2002) by multiplying the penalty function by a positive constant c, suggested originally by Hallin and Liska (2007), representing the strength of the penalty function. The authors, thus, propose the following two criteria:

$$IC_1^*(j) = \ln\left(V(j,F)\right) + c.j.\left(\frac{N+T}{NT}\right)\ln\left(\frac{NT}{N+T}\right) \tag{41}$$

$$IC_2^*(j) = \ln\left(V(j,F)\right) + c.j.\left(\frac{N+T}{NT},\right) \tag{42}$$

Where $V(j,F)$ is given by equation (37). The estimation of the number of factors r is obtained by minimization of the information criteria IC_1^* and IC_2^* for $j = 0, \ldots, r_{max}$, where r_{max} is the maximum number of static factors. The procedure for the selection of the number of static

[13] Bai and Ng (2002) also propose another class of information criteria for which the variance $V(j,F)$ replaces $\ln(V(j,F))$ in equations (38), (39) and (40). Bai and Ng (2002, Theorem 2) give the results of the convergence of these criteria when N and T tend to infinity.

factors depends both on the variance of the number of estimated factors $V_c(r)$ (for N and T tending to infinity) and on the constant $c \in [0, c_{max}]$. Alessi et al. (2010) suggest estimating this variance $V_c(r)$ by reiterating the procedure for estimating r for a finite number of subsets of the initial N variables, also making the number of observations T vary.

Kapetanios (2010) proposes a concurrent method to the information criterion to estimate the number of static factors, based on the random matrix theory. His approach is based on a series of tests on the largest eigenvalues of the variance-covariance matrix of the initial data, which we have denoted Σ_x. Other procedures have been suggested by Yao and Pan (2008) and Onatski (2010).

5.2 Selection of the number of factors for dynamic factor models

5.2.1 The Bai and Ng (2007) criterion

In the context of dynamic factor models, the number of dynamic shocks q (for the estimation of factors in dynamic principal components and their space-state form) can be determined using the Bai and Ng (2007) information criterion. This criterion is obtained by considering the r estimated static factors as given and, then, estimating a VAR model of order p on these factors, where the order p is selected using the BIC criterion. Next, a spectral decomposition of the variance-covariance matrix of the estimated residuals of the VAR model, denoted $\hat{\Sigma}_\varepsilon$, of dimension $(r \times r)$, is calculated. Then, the j^{th} ordered eigenvalue $\hat{c}_j$, where $\hat{c}_1 \geq \hat{c}_2 \geq \cdots \geq \hat{c}_j \geq \cdots \geq \hat{c}_r \geq 0$, is recovered. Finally, for $l = 1, \dots, r-1$, Bai and Ng (2007) propose the following two quantities:

$$\begin{aligned} \widehat{D}_{1,l} &= \left(\frac{\widehat{c}_{l+1}}{\sum_{j=1}^{r} \widehat{c}_j} \right)^{1/2} \\ \widehat{D}_{2,l} &= \left(\frac{\sum_{j=l+1}^{r} \widehat{c}_j}{\sum_{j=1}^{r} \widehat{c}_j} \right)^{1/2} \end{aligned}$$

Where $\widehat{D}_{1,l}$ represents a measure of the marginal contribution of the l + 1th eigenvalue and $\widehat{D}_{2,l}$ represents a measure of the cumulative contribution of the eigenvalues, under the hypotheses that $\widehat{\Sigma}_{\varepsilon}$ is the unit matrix of dimension $(r \times r)$ and that $c_l = 0$ for $l > q$.[14]

Thus, according to the selected marginal contribution measure, the number of dynamic factors q is obtained by minimizing:

$$\left\{ l \text{ } such \text{ } that\colon \widehat{D_{1,l}} \leq \frac{c}{\min(N^{\frac{2}{5}}, T^{\frac{2}{5}})} \right\}$$

or:

$$\left\{ l \text{ } such \text{ } that\colon \widehat{D_{2,l}} \leq \frac{c}{\min(N^{\frac{2}{5}}, T^{\frac{2}{5}})} \right\}$$

Bai and Ng (2007) suggest using $c = 1$ based on Monte Carlo simulations.

In practice, these different criteria are used at three stages:

1. First, one of the Bai and Ng (2002) criteria is used to determine the optimal number of factors $r \in \{1,\ldots, r_{max}\}$ in a static context;[15]

2. Then, a VAR(p) is estimated on these r estimated factors and the order p of the VAR is selected to minimize the BIC criterion;

3. Finally, the Bai and Ng (2007) criteria are applied to the variance-covariance matrix or correlation matrix of the residuals (ε_t) of the VAR(p) to obtain the optimal number of dynamic factors q.

[14] Bai and Ng (2007) show that $\widehat{D}_{1,l}$ and $\widehat{D}_{2,l}$ converge toward zero when $l > q$.
[15] The criterion IC_2 is used more often in practice.

5.2.2 The Stock and Watson (2005a) and Amengual and Watson (2007) criteria

Stock and Watson (2005a) and Amengual and Watson (2007) show that the Bai and Ng (2002) estimator can be used to estimate the number of dynamic factors. To do this, they propose applying this estimator to the errors resulting from the projection of observed data on the lagged values of static factors, *i.e.*, on $v_t = X_t - \sum_{\tau=1}^{p} \Lambda\, \Phi(L) F_{t-\tau}$. They propose two ways of estimating the errors $(\hat{v}_t)$:

$$\hat{\nu}_t^A = X_t - \sum_{\tau=1}^{p} \hat{\Lambda}\hat{\phi}_\tau \hat{F}_{t-\tau}$$
$$\hat{\nu}_t^B = X_t - \sum_{\tau=1}^{p} \hat{\Pi}_\tau \hat{F}_{t-\tau}$$

Where $(\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_p)$ are the ordinary least squares estimators of the regression of $\hat{F}_t$ on $(\hat{F}_{t-1}, \dots, \hat{F}_{t-p})$ and $(\hat{\Pi}_1, \hat{\Pi}_2, \dots, \hat{\Pi}_p)$ are the ordinary least squares estimators of the regression of X_t on $(\hat{F}_{t-1}, \dots, \hat{F}_{t-p})$.

5.2.3 The Breitung and Pigorsch (2013) criteria

Breitung and Pigorsch (2013) also propose two information criteria to select the number of dynamic factors. Their criteria are based on an analysis of the canonical correlations of static factors (obtained by a principal components analysis) and depend on the estimation of a $VAR(p)$ model on these factors, where the order p is selected by the BIC criterion.

The first criterion is based on the following statistic:

$$\zeta(q^*) = \tilde{C}_{NT}^{2-\delta} \sum_{j=1}^{r-q^*} (1 - \tilde{\lambda}_j)$$

where $\tilde{C}_{NT}^{2-\delta} = (2-\delta)N^{-1} + (2-\delta)T^{-1}$, with $0 < \delta < 2$, and λ_j are values resulting from the solution to the following problem $\left|\tilde{\lambda}_j \tilde{S}_{00} - \tilde{S}_{01}\tilde{S}_{11}^{-1}\tilde{S}_{01}'\right| = 0$, with $\tilde{S}_{00} = \sum_{t=\tau+1}^{T} \hat{F}_t \hat{F}_t'$,, $\tilde{S}_{01} = \sum_{t=\tau+1}^{T} \hat{F}_t\, \hat{G}_{t-1}'$, $\tilde{S}_{11} = \sum_{t=\tau+1}^{T} \hat{G}_{t-1}\, \hat{G}_{t-1}'$, and $\hat{G}_{t-1} = \left[\hat{F}_{t-1}', \dots, \hat{F}_{t-\tau}'\right]$. The number of dynamic factors can be estimated using a large number of $q*$ resulting from this sequence $q^* = r-1, r-2, \dots, 0$, where the statistic $\zeta(q^*)$ is larger than the level of the threshold κ, or:

$$q = \max\{q^* \quad \text{such that: } \zeta(q^*) > \kappa\}$$

Breitung and Pigorsch (2013) suggest using the following values of the parameters: $\tau = 1$, $\delta = 0.5$, and $\kappa = 1$.

The second criterion is based on the following statistic:

$$LR(q^*) = T \sum_{j=r-q^*+1}^{r} \tilde{\lambda}_j$$

where the null hypothesis is H_0: $q^* = q$ as against the alternative H_1: $q^* < q$.

5.2.4 The Hallin and Liska (2007) criterion

Hallin and Liska (2007) develop an information criterion for generalized dynamic factor (GDF) models. This criterion is based on the spectral density matrix of the observations. It is written as follows:

$$IC_{2,N}^{T}(j) = \ln\left[\frac{1}{N}\sum_{s=j+1}^{N}\frac{1}{2M_T+1}\sum_{h=-M_T}^{M_T}\lambda_{N_s}^{T}(\omega_h)\right] + c.j.p(N,T), \tag{43}$$

With $0 \le j \le q_{max}$, $\omega_h = \frac{\pi h}{M_T + \frac{1}{2}}$, for $h = -M_T, \dots, M_T$ with the truncation parameter $M_T > 0.$, c is a positive constant such that $c = [0.01, 0.02, \dots, 3.00].$, $N_s < N$ is the number of variables contained in a given subset and $p(N,T)$ is a penaly function[16] such that:

$$p(N,T) = \left(-M_T^{-2} + M_T^{1/2}T^{-1/2} + N^{-1}\right) \times \ln\left[\min(N, M_T^2, M_T^{-\frac{1}{2}}T^{\frac{1}{2}})\right] \quad)$$

The eigenvalues $\lambda_{N_s}^{T}$ result from $\hat{I}_x(\omega)$, which represents the spectral density matrix estimator of X_t with $\omega \in [-\pi, \pi]$. The number of estimated factors is, then, given by:

$$q = \text{argmin } IC_{2,N}^{T}(j)$$

[16] Hallin and Liska (2007) also propose two other penalty functions. Onatski (2009) suggests alternative tests in the context of approximate dynamic factor models. Jacobs and Otter (2008) propose a test based on a canonical correlation procedure to determine simultaneously the number of dynamic factors q and the order of the lag p in a dynamic factor model, but for fixed N and T.

$$0 \leq j \leq q_{max}$$

The procedure for selecting the number of dynamic factors is similar to that used by Alessi et al. (2010), *i.e.*, by examining the variance of the number of estimated factors, $V_c(r)$ for N and T tending to infinity and for an interval of values for the constant c. In their numerical illustration, Hallin and Liska (2007) propose retaining $M_T = [0.75\sqrt{T}]$ and $q_{max} = 13$.

6 Recent results in the empirical literature

Applications of dynamic factor models abound in the empirical economic literature. A few examples are asset pricing models (Ross, 1976), consumer theory (Gorman, 1981; Lewbel, 1991), and the assessment of performance and risk measurement in finance (Campbell et al., 1997). In this section, we present some recent applications of these models that underscore the interest of this approach for (i) the construction of short-term economic indicators, (ii) macroeconomic forecasting, and (iii) international macroeconomics and monetary policy analysis.

6.1 Tools for short-term economic monitoring

Dynamic factor models are useful for developing economic activity indicators based on the mass of data available to short-term forecasters. These models can be used to synthesize large datasets into a composite indicator that reflects the most relevant data available on a given date.

One of the most common indicators to which forecasters working on the U.S. economy refer is the Chicago Fed National Activity Indicator (CF-NAI) developed using the Stock and Watson (1999) approach. This indicator is based on 85 monthly series representative of the U.S. economy, covering production, income, employment, personal consumption, housing, sales, inventories and orders. The CF-NAI corresponds to the first factor estimated in a principal components analysis. A value close to zero means that activity is close to its long-term trend. As well, the CF-NAI can be used to detect recessions in the U.S. by using historical estimated thresholds. The Philadelphia Federal Reserve, for its part, each week publishes a daily economic activity indicator that is based on the mixed frequency model presented in the article by Aruoba, Diebold and Scotti (2009) and includes daily, weekly, monthly and quarterly data. Given its high frequency, this indicator is interesting for forecasting since it makes it possible to provide an early signal very rapidly.

In the euro zone, the Centre for Economic Policy Research (CEPR) has for some years now been disseminating the EuroCoin indicator developed at the Bank of Italy by Altissimo et al. (2001, 2010). The purpose of this indicator is to estimate a monthly GDP growth in the euro zone for the coincident quarter using smoothing to remove very short-term effects (frequency less than one year). The model used to synthesize the information is a generalized dynamic factor model proposed by Forni, Hallin, Lippi and Reichlin (2000) applied to a very large number of variables. From an economic standpoint, this measure corresponds to a kind of medium-term quarterly growth, but does not aim to precisely estimate the quarterly accounts figures provided by Eurostat. The first version of EuroCoin has a smaller variance than quarterly GDP growth. The economic outlook that it provides deviates considerably from the national accounts. CEPR has recently tried to construct a more effective new version of EuroCoin (Altissimo et al., 2010). The aim of the new indicator is the same but the number of input variables used has been reduced from 951 to 145 (IPI, monetary aggregates, interest rates, financial variables, demand indicators, surveys, trade variables and labor market variables). The input variables have not been smoothed by statistical means, which eliminates some side effects that might have introduced a bias into the old EuroCoin.

For France, Doz and Lenglart (1999) have developed a summary indicator for the industry tendency survey conducted by INSEE by means of six monthly balances of opinion. This approach is now commonly used by INSEE to calculate composite indicators based on data from short-term tendency surveys in the various sectors. Clavel and Minodier (2009) extend this business activity indicator relating to industry by proposing a business climate indicator for the economy as a whole, which incorporates balances of opinion derived from the INSEE tendency surveys conducted in the various sectors such as services, construction, and wholesale and retail sales.

Many economic activity indicators have been developed on the basis of regime-switching factor models. Kim and Nelson (1998) propose an application for the four main U.S. economic series (growth rates of IPI, employment, income, and retail sales), also considered by the

National Bureau of Economic Research (NBER) for the dating of U.S. recessions. These same series are considered by Diebold and Rudebusch (1996), Chauvet (1998) and Chauvet and Piger (2008), who simultaneously estimate a similar model to construct a real-time business cycle indicator. In France, Nguiffo-Boyom (2006) estimates the Kim and Nelson (1998) model simultaneously on four series derived from the INSEE tendency surveys to reproduce the growth cycle, which measures the long-term deviation from trend. For the euro zone, Darné and Ferrara (2011) propose a two-step dynamic factor model. They, first, estimate a factor which, in turn, follows a regime-switching model, based on a set of six business confidence index series for the six main eurozone countries. The authors, thus, develop an indicator to detect the acceleration cycle in the monetary zone in real time.

As for the mixed frequency approach, Mariano and Murasawa (2003) have applied their model to estimate an economic activity indicator in the U.S. Mariano and Murasawa (2010) also use a version of their model to calculate a monthly U.S. GDP series by estimating a mixed frequency VAR that includes the quarterly GDP growth series and the four monthly series traditionally used by NBER to assess the U.S. business cycle (growth rates of employment, income, industrial production, and retail sales). Cornec (2006) also uses this approach, first to provide a monthly dating of the French business cycle using two quarterly series (GDP growth rate and employment) and two monthly series (IPI and household consumption expenditure), and then to estimate a composite activity indicator similar to that of Doz and Lenglart (1999) but which includes the quarterly GDP growth rate series as supplementary information. The empirical results of this application emphasize that the contribution of GDP to the first factor is negligible in comparison with the composite business indicator relating to industry. Again in France, Cornec and Deperraz (2006) use a mixed frequency model to develop an activity indicator in the service sector based on three monthly balances and three quarterly balances from the INSEE tendency survey relating to services. This indicator can usefully supplement the indicator on the business climate in manufacturing for short-term forecasting. Clavel and Minodier (2009) also develop a mixed frequency approach to incorporate the various short-term tendency surveys conducted by INSEE, which are sampled monthly, bimonthly and quarterly, into their business climate indicator. More recently, Camacho and Perez-Quiros (2010, 2011) develop two short-term growth indicators for Spain and the eurozone using a mixed frequency factor model that also includes Markov regime-switching to take account of the business cycle.Frale et al. (2010) develop a small-dimension mixed frequency factor model to provide a measure of monthly GDP in the eurozone. Finally, Frale et al. (2011) propose a mixed frequency model to estimate a monthly GDP indicator in the eurozone called EuroMInd, based on a disaggregation between supply and demand. This indicator is based on the official database for the eurozone developed by Eurostat (Euro-IND).

6.2 Macroeconomic forecasting

Dynamic factor models are widely used, particularly by central banks, as a tool for forecasting various macroeconomic variables, such as the GDP growth rate or inflation (see, for example, a survey in Stock and Watson, 2006, or Eickmeier and Ziegler, 2008). When the forecasting horizon covers the current period, such forecasting is termed “nowcasting” (on this point see Giannone, Reichlin and Small, 2008). The factors are estimated from monthly data used to track countries' economic situation, such as household and business survey data (*soft data*), variables for the real economy (*hard data*), including indices of industrial production, household consumption, retail sales or new vehicle registrations and, finally, financial variables (stock prices, oil prices, interest rates, etc.). For a given country, such a database can include several hundred variables. It is useful, therefore, to be able to synthesize this large data set into a small dimension vector to be included in standard models.

Based on asymptotic theoretical results on the convergence of estimators in this type of model, early work used the largest possible number of variables available. More recent work addresses the question whether including the largest number of variables is appropriate or not to improve the accuracy of forecasts. For example, Barhoumi, Darné and Ferrara (2010) show empirically in the case of France that increasing the data set by disaggregation does not result in significant improvements in the accuracy of short-term GDP forecasts. Boivin and Ng (2006) identify the conditions under which expanding the database could result in less accurate factor estimates and provide empirical rules for eliminating redundant variables. These authors show that expanding the data set is not preferable if the new series add too much idiosyncratic noise and/or increase the cross correlation between idiosyncratic errors too much. Bai and Ng (2008a) use LARS-type statistical methods (least angle regressions), which are weighted regressions, to identify optimal subsets of predictors (targeted predictors) from a large dataset. Schumacher (2010) underscores the effectiveness of this approach in using an international database to predict German growth, which is very sensitive to fluctuations in the international environment. Charpin (2009)'s application of this approach to French data also seems to provide encouraging results.

Once the factors have been estimated, the forecasting of the variable of interest Y_t over a horizon h is derived from either an ARDL-type univariate regression equation (autoregressive distributed lags, see equation (44) below) or a VAR-type multivariate process. When the aim is to forecast over a horizon h exceeding one step, two approaches co-exist: the recursive approach, which uses, for a given step, the forecasts made for the previous steps, and the direct approach, which seeks to predict the value over the horizon h directly, *i.e.* without trying to forecast the variable of interest in the previous steps. In a general context, direct forecasting of the variable over the horizon h makes it possible to reduce the forecast bias resulting from the estimation of the parameters that may appear in the case of a multistep recursive forecast (see, for example, Chevillon, 2007). In the particular context of factor models, according to simulations done by Boivin and Ng (2005), there does not seem to be a significant difference between a direct forecast and a recursive forecast if estimated factors are used. However, the direct approach is preferred in many applications.

Thus, the univariate equation of direct forecasting to a horizon h is written:

$$\hat{Y}_{t+h|t} = \hat{\alpha}_h + \sum_{j=1}^{m} \hat{\beta}'_{hj} \hat{F}_{t-j+1} + \sum_{j=1}^{p} \hat{\phi}_{hj} Y_{t-j+1}, \tag{44}$$

Where $\hat{F}_t$ is the vector of dimension r of the estimated factors, m and p are the autoregressive orders, and $\hat{\beta}_{hj}$ is a vector of estimated coefficients of dimension r. The parameters α_h, β_{hj} and ϕ_{hj} depend on the horizon h since, in the context of a direct forecast, they vary on the basis of the horizon considered. The $mr + p + 1$ parameters of the model are estimated using ordinary least squares. In equation (44), the number of factors r can be specified by one of the tests presented above. However, $r = 3$ is often used in practice, since three factors are often sufficient to explain a significant portion of the data variance. Three variants of the model given in equation (44) are generally used (see Stock and Watson, 2002, Boivin and Ng, 2005). The first, denoted *DI* (diffusion index) is obtained with $m = 1$ and the terms depending on Y_{t-j+1}, j=1, …, p, being suppressed in (44) and, thus, includes only contemporaneous information $\hat{F}_t$. The second, *DI-AR* authorizes a dynamics on the series Y_t and corresponds to $m = 1$ and $1 \leq p \leq 6$ in (44). The optimal autoregressive order p is, then, obtained by minimization of an AIC or BIC-type information criterion. Finally, the specification *DI-AR,Lag* of (44) corresponds to $1 \leq m \leq 3$ and $1 \leq p \leq 3$, thus allowing lags on the factors and on the variable Y_t. Once again, the optimal parameters m and p are obtained by minimization of an information criterion. We should point out that the specification *DI-AR,Lag* is not used for a dynamic factor since such a factor is assumed to already include a time dynamics.

Boivin and Ng (2005) use a simulation to show that the differences between using static or dynamic factors are negligible in forecasting. Barhoumi, Darné and Ferrara (2010) find the same result empirically on French data. Barhoumi, Darné and Ferrara (2010) also show that the specification of the model used to make the forecasts has only a marginal impact on the quality of the forecast, particularly when the number of observations is high.

One of the major problems that appears when these models are used for real-time forecasting results from the fact that data arrive in a staggered fashion, leading to missing values at the end of samples (this is the *ragged-edge data* problem, which is well-known from forecasters). Several solutions have been proposed in the empirical literature, such as the projection of missing data, either using an autoregressive-type parametric model or using moving averages, or the realignment of the data base on the last points available, if the number of variables is high. In the context of factor models, two-step estimation using a Kalman filter (Doz, Giannone and Reichlin, 2011) solves this problem elegantly (Giannone, Reichlin and Small, 2008, or Angelini et al., 2011).

Among the many applications in forecasting the GDP growth rate, we can cite, for example, the articles of Stock and Watson (2002) or Banerjee and Marcellino (2006) for the U.S., those of Barhoumi, Darné and Ferrara (2010, 2013) and Bessec and Doz (2013) for France, Forni, Hallin, Lippi and Reichlin (2000, 2003), Camba-Mendez and Kapetanios (2005), Marcellino, Stock and Watson (2003), Banerjee, Marcellino and Masten (2005), Ruenstler et al. (2009) or Angelini et al. (2011) for the eurozone, Schumacher (2007, 2010), Schumacher and Breitung (2008) and

Eickmeier and Ziegler (2008) or Marcellino and Schumacher (2010) for Germany, Artis, Banerjee and Marcellino (2005) for the United Kingdom, and Van Nieuwenhuyze (2006) for Belgium. Matheson (2011) also develops GDP forecasts for a large number of advanced and emerging countries.

It should be noted that applications to inflation forecasting are much rarer; see, for example, Forni et al. (2003) or Camba-Mendez and Kapetanios (2005) for the eurozone or de Bandt et al. (2007) for France. Boivin and Ng (2005) also consider series of U.S. prices from a forecasting point of view. It appears to be difficult to improve the accuracy of inflation forecasts by using a large number of variables in comparison with an approach based on a precise selection of variables of interest. In contrast, measurements of underlying inflation have been conducted using this type of approach; we refer interested readers to the articles of Cristadoro et al. (2005) for the eurozone or Kapetanios (2004) for the United Kingdom.

Various works have endeavored to identify the contribution of financial variables to macroeconomic forecasting using factor models applied to a database on financial markets activity. For example, Forni et al. (2003) show that, in the eurozone, financial variables help to forecast inflation but cannot be used to accurately forecast industrial output. Bellégo and Ferrara (2009, 2012) use a factor model to assess the likelihood of a recession in the eurozone based on a large set of monthly variables (factor-probit-type model).[17] Specifically, Bellégo and Ferrara (2009) show that, by using only financial variables, this approach would have made it possible to anticipate the 2008-09 recession in the eurozone in real time as early as late 2007.

A review of the literature on the results of factor models in forecasting is found in the article by Eickmeier and Ziegler (2008), who conduct a meta-analysis of the performance of models to forecast GDP and inflation. They conclude that factor models generally improve smaller-scale econometric models, but that methods of combining forecasts[18] constitute a competing alternative approach.

[17] A factor-probit model is obtained by, first, estimating the factors from a data set, and then, incorporating them in a standard probit-type model.

[18] Forecast combinations generally use a weighted average of forecasts of a single target variable, based on a large number of different models. See Timmerman (2006).

6.3 Applications in monetary policy and the international economy

An abundant literature analyzes the impact of monetary policy shocks on the macroeconomy and how the transmission mechanism of these shocks has evolved, particularly for the U.S. Traditionally, the impact of monetary policy shocks is often measured using small dimension VAR models (fewer than 6 variables). Typically, trivariate SVAR models containing the interest rate, output and inflation are used. Beyond this small number of variables, it is difficult to estimate this type of model using standard methods (in a Bayesian context, see, for example, de Mol, Giannone and Reichlin, 2008). Now, as noted in Bernanke and Boivin (2003), monetary policy is conducted by central banks in a data-rich environment. In this context, a category of papers analyzing monetary policy initiated by the article of Bernanke, Boivin and Eliasz (2005) uses FAVAR models. This type of modeling is chosen to remedy the problem of missing variables generally encountered in traditional VAR modeling. Bernanke, Boivin and Eliasz (2005), Stock and Watson (2005a) and Favero, Marcellino and Neglia (2005) use FAVAR models to analyze monetary policy in the U.S. and in some eurozone countries. They all conclude that adding estimated factors from factor models to VAR models allows for a finer-grained analysis of the phenomena in question, particularly in terms of structural shocks. For example, Del Negro and Otrok (2007) estimate a factor common to changes in residential housing prices in various U.S. states, then introduce it into a FAVAR to assess the extent to which monetary easing helped create a housing bubble (the study data stop in 2005). They show that the impact of monetary policy shocks is weak in comparison with the scope of the price fluctuations observed through to the end of their sample.

Another portion of the literature addresses the question of whether the mechanism for the transmission of shocks has changed over time and, if so, how. Time-varying FAVAR (TV-FAVAR) models provide an extremely flexible modeling that can shed light on this question. The literature appears to be in agreement on the fact that this mechanism has changed, although there is no consensus on how. For example, based on a set of 803 quarterly variables from 1972 to 2007, Eickmeier et al. (2011) show that the volatility of monetary shocks in the U.S. declined significantly from the early 1980s to the eve of the subprime crisis and that the negative impact of a shock on U.S. activity and prices declined over this period. The authors also emphasize that the negative impact of a monetary policy shock on inflation expectations and long-term interest rates weakened over time. The reasons given by the authors for these evolutions are the changes in monetary policy and the globalization of trade and finance. Finally, these authors indicate that the transmission mechanism appear to be the same in periods of expansion and periods of contraction. Baumeister, Liu and Mumtaz (2010) also show that, for the U.S. economy, the reaction of GDP, consumption and investment to a monetary shock declined over the period 1960-2008. We should point out, however, that most of these studies do not include the period 2008-2009, during which the industrialized countries suffered the worst recession since the 1920s.

FAVAR models have also been used to analyze changes in the synchronization of global business cycles, making it possible to discriminate between different types of shocks. For example, Stock and Watson (2005b) estimate a FAVAR on the GDP of the G7 countries, enabling them to identify common international shocks and domestic effects due to an international shock and those due to an idiosyncratic shock. They conclude that the reduction in the volatility of cycles in the G7, with the exception of Japan, observed between the mid-1980s and the mid-2000s, is primarily due to a reduction in the amplitude of common international shocks (the so-called "Great Moderation"). Kose, Otrok and Whiteman (2003) consider a similar model to show the existence of a global cycle based on a set of 60 countries. They also show that factors specific to the region play only a minor role in the explanation of macroeconomic fluctuations. Bordo and Helbling (2010) explore the historical angle by using annual GDP data from 1880 to 2008 for 16 industrialized countries and show a trend toward increasing synchronization among those countries. The authors show the role of common shocks in this change using a restricted FAVAR model estimated on this database.

There are also several applications of factor models to the measurement of international cycles and their transmission between countries. Mansour (2003) and Helbling and Bayoumi (2003) estimate a global business cycle for the world and the G7 countries and analyze the contribution of this common cycle to economic changes in each country. Kose, Otrok and Whiteman (2008) use a factor model in a Bayesian context to estimate the common and idosyncratic components of the G7 countries for a set of economic aggregates. They show that the factor common to these countries explains a larger portion of the variation of these aggregates over the period 1986-2003 than in previous periods, thus demonstrating an increased synchronization of the business cycles in the G7. Eickmeier (2007) analyzes the transmission of U.S. structural shocks to Germany using the approach put forward by Forni et al. (2004). After analyzing common economic movements in the eurozone, Marcellino, Stock and Watson (2003) and Eickmeier (2005) try to give economic interpretations to the common factors by linking them to the various countries in the zone and/or certain variables, based on correlation measures.

7 Conclusion

In this article, we have reviewed the recent literature on dynamic factor models. There has been recently an increasing interest in these models on the part of researchers since they can adequately respond to certain problems encountered in practice, particularly inflation of the number of available data. We have presented the models and their most interesting extensions, main estimation methods, and tests of the number of factors. In the final section, we have presented a few recent examples of the application of dynamic factor models to macroeconomic forecasting, the construction of short-term economic indicators and the analysis of monetary policy and the international economy. The success of dynamic factor models means that this review of the literature cannot claim to be exhaustive. Extensions have been developed very recently, particularly to add flexibility by means of non-linearities or mixed frequencies, and numerous applications continue to be published. Moreover, these models are now increasingly weighed against other econometric methods that also make it possible to reduce the scale of the problem. It appears to us, therefore, that the research into dynamic factor models will continue to thrive.

References

Alessi L., Barigozzi M. and Capasso M. (2010). Improved penalization for determining the number of factors in approximate factor models. *Statistics and Probability Letters*, 80, 1806-1813.

Altissimo F., Bassanetti A., Cristadoro R., Forni M., Hallin M., Lippi M. and Reichlin L. (2001). EuroCOIN: A real time coincident indicator of the euro area business cycle. Discussion Paper No. 3108, CEPR.

Altissimo F., Cristadoro R., Forni M., Lippi M. and Veronese G. (2010). New EuroCOIN: Tracking economic growth in real time. *The Review of Economics and Statistics*, 92, 4, 1024-1034.

Amengual D. and Watson M.W. (2007). Consistent estimation of the number of dynamic factors in a large N and T panel. *Journal of Business and Economic Statistics*, 25, 91-96.

Anderson T.W. (1984). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

Angelini E., Camba-Méndez G., Giannone D., Runstler G. and Reichlin L. (2011). Short-term forecasts of Euro area GDP growth. *Econometrics Journal*, 14, 1, C25-C44.

Artis M., Banerjee A. and Marcellino M. (2005). Factor forecasts for the UK. *Journal of Forecasting*, 24, 279-298.

Aruoba S.B., Diebold F.X. and Scotti C. (2009). Real-time measurement of business conditions. *Journal of Business and Economic Statistics*, 27, 4, 417-27.

Bai J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71, 1, 135-171.

Bai J. and Ng S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70, 191-221.

Bai J. and Ng S. (2007). Determining the number of primitive shocks in factor models. *Journal of Business and Economic Statistics*, 25, 52-60.

Bai J. and Ng S. (2008a). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146, 304-317.

Bai J. and Ng S. (2008b). Large dimensional factor analysis, *Foundations and trends in econometrics*, 3, 89-168.

de Bandt, O., Michaux E., Bruneau C.and Flageollet A. (2007). Forecasting inflation using economic indicators: The case of France, *Journal of Forecasting*, 26, 1, 1-22.

Banerjee A. and Marcellino M. (2006). Are there any reliable leading indicators for US inflation and GDP growth?, *International Journal of Forecasting*, 22, 137-151.

Banerjee A. and Marcellino M. (2009). Factor-augmented error correction models. InCastle, J. and Shepard, N. (Eds.), *The Methodology and Practice of Econometrics - A Festschrift for David Hendry,* Oxford: Oxford University Press.

Banerjee A., Marcellino M. and Masten I. (2005). Leading indicators for Euro area inflation and GDP growth. *Oxford Bulletin of Economics and Statistics*, 67, 785-813.

Barhoumi K., Darné, O. and Ferrara, L. (2010). Are disaggregate data useful for forecasting French GDP with dynamic factor models? *Journal of Forecasting*, 29, 1-2, 132-144.

Barhoumi K., Darné, O. and Ferrara, L. (2013). Testing the number of factors: An empirical assessment for forecasting purpose. *Oxford Bulletin of Economics and Statistics*, 75, 1, 64-79.

Baumeister C., Liu, P. and Mumtaz, H. (2010), Changes in the transmission of monetary policy: Evidence from a time-varying factor-augmented VAR. Working Paper No. 401, Bank of England.

Bellégo C. and Ferrara, L. (2009). Forecasting euro area recessions using time-varying binary response models for financial variables. Working Paper No. 259, Banque de France.

Bellégo C. and Ferrara, L. (2012). Macro-financial linkages and business cycles: A factor-augmented Probit approach. *Economic Modelling*, 29, 1793-1797.

Bernanke B.S. and Boivin J. (2003). Monetary policy in a data-rich environment. Journal of Monetary Economics, 50, 525.

Bernanke B. S., Boivin J. and Eliasz P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics*, 120, 387-422.

Bessec, M. and Doz C. (2013). Short-term forecasting of French GDP growth using dynamic factor models, *Journal of Business Cycle Measurement and Analysis*, this issue (published in French in *Économie et Prévision*, under the title: Prévision de court terme de la croissance du PIB français à l'aide de modèles à facteurs dynamiques).

Boivin J. and Ng S. (2005). Understanding and comparing factor-based forecasts. *International Journal of Central Banking*, 1, 117-151.

Boivin J. and Ng S. (2006). Are more data always better for factor analysis? *Journal of Econometrics*, 132, 169-194.

Boivin J., Giannoni M.P. and Mihov I. (2009). Sticky prices and monetary policy: Evidence from disaggregated U.S. data. *American Economic Review*, 99, 350-384.

Bordo M. and Helbling, T. (2010). International business cycle synchronization in historical perspective. Working Paper 16103, NBER.

Breitung J. and Pigorsch U. (2013). A canonical correlation approach for selecting the number of dynamic factors. *Oxford Bulletin of Economics and Statistics*, 75, 1, 23-36.

Brillinger D.R. (1981). *Time Series: Data Analysis and Theory.* Holden-Day, San Francisco.

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.

Camacho M. and Perez-Quiros G. (2011). Spain-STING: Spain short-term indicator of growth. The Manchester School, 79, 594-616.

Camba-Méndez G. and Kapetanios G. (2005). Forecasting euro area inflation using dynamic factor measures of underlying inflation. *Journal of Forecasting*, 25, 491-503.

Campbell J.Y., Lo A.W. and MacKinlay A.C. (1997). *The Econometrics of Financial Markets.* Princeton University Press, Princeton.

Chamberlain G. and Rothschild M. (1983). Arbitrage, factor structure and mean-variance analysis in large asset markets. *Econometrica*, 51, 1305-1324.

Charpin, F. (2009). Estimation précoce de la croissance. De la régression LARS au modèle à facteurs. *Revue de l'OFCE*, 2009/1, 108, 31-48.

Chauvet M. (1998). An econometric characterization of business cycle dynamics with factor structure and regime switching. *International Economic Review*, 39, 969-996.

Chauvet, M. and Piger, J. (2008), A comparison of the real-time performance of business cycle dating methods. *Journal of Business Economics and Statistics*, 26, 1, 42-49.

Chevillon G. (2007). Direct multi-step estimation and forecasting. *Journal of Economic Surveys*, 21, 746-785.

Clavel L. and Minodier C. (2009). A monthly indicator for the French business climate. Document de Travail No. G2009/02, Insee.

Clements, M. and Galvao, A.-B. (2008), Macroeconomic forecasting with mixed-frequency data: forecasting output growth in the United States. *Journal of Business and Economic Statistics*, 26, 546-554.

Connor G. and Korajczyk, R.A. (1986). Performance measurement with the arbitrage pricing theory: A new framework for analysis. *Journal of Financial Economics*, 15, 373-394.

Connor G. and Korajczyk, R.A. (1988). Risk and return in an equilibrium APT: Application of a new test methodology. *Journal of Financial Economics*, 21, 255-289.

Connor G. and Korajczyk, R.A. (1993). A test for the number of factors in an approximate factor model. *Journal of Finance*, 48, 1263-91.

Cornec M. and Deperraz, T. (2006). Un nouvel indicateur synthétique mensuel résumant le climat des affaires dans les services en France. *Économie et Statistique*, 395-396, 13-38.

Cornec M. (2006). Analyse factorielle dynamique multifréquence appliquée à la datation de la conjoncture française. *Économie et Prévision*, 2006/1, 172, 29-43.

Cristadoro R., Forni M., Reichlin L. and Veronese G. (2005). A core inflation index for the Euro area. *Journal of Money, Credit and Banking*, 37, 3, 539-560.

D'Agostino, A. and Giannone D. (2012). Comparing alternative predictors based on large-panel factor models. *Oxford Bulletin of Economics and Statistics*, 74, 2, 306-326.

Darné O. and Ferrara L. (2011). Identification of slowdowns and accelerations for the euro area economy. *Oxford Bulletin of Economics and Statistics*, 73, 3, 335-364.

Del Negro M. and Otrok C. (2007). 99 Luftballons: Monetary policy and the house price boom across US states. *Journal of Monetary Economics*, 54, 1962-1985.

Del Negro M. and Otrok C. (2008). Dynamic factor models with time-varying parameters: Measuring changes in international business cycles. Technical report, Federal Reserve Bank of New York.

De Mol C., Giannone D. and Reichlin L. (2008). Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components?, *Journal of Econometrics*, 146, 318-328.

Dempster A., Laird N. and Rubin D. (1977). Maximum likelihood estimation from incomplete data. *Journal of the Royal Statistical Society*, 14, 1-38.

Diebold F.X. and Rudebusch G.D. (1996). Measuring business cycles: A modern perspective. *Review of Economics and Statistics*, 78, 67-77.

Ding A.A. and Hwang J.T.G. (1999). Prediction intervals, factor analysis models, and high-dimensional empirical linear prediction. *Journal of the American Statistical Association*, 94, 446-455.

Doz C. and Lenglart F. (1999). Analyse factorielle dynamique : test du nombre de facteurs, estimation et application à l'enquête de conjoncture dans l'industrie. *Annales d'Économie et Statistique*, 54, 91-127.

Doz C., Giannone D. and Reichlin L. (2011). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics*, 164, 188-205.

Doz C., Giannone D. and Reichlin L. (2012). A quasi maximum likelihood approach for large approximate dynamic factor models. *Review of Economics and Statistics*, 94, 4, 1014-1024.

Eickmeier S. (2005). Common stationary and non-stationary factors in the euro area analyzed in a large-scale factor model. Discussion Paper No. 2/2005, Bundesbank.

Eickmeier, S. (2007). Business cycle transmission from the US to Germany - a structural factor approach. *European Economic Review*, 51, 521-551.

Eickmeier S., Lemke, W. and Marcellino, M. (2011). Classical time-varying FAVAR models - Estimation, forecasting and structural analysis. Discussion Paper No. 8321, CEPR.

Eickmeier S. and Ziegler C. (2008). How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach. *Journal of Forecasting*, 27, 3, 237-265.

Eklund J. and Kapetanios G. (2008). A review of forecasting techniques for large data sets. Working Paper No 625, Department of Economics, *Queen Mary, University* of London.

Favero C., Marcellino M. and Neglia F. (2005). Principal components at work: The empirical analysis of monetary policy with large datasets. *Journal of Applied Econometrics*, 20, 603-620.

Ferrara, L. and Marsilli, C. (2013). Financial variables as leading indicators of GDP growth: Evidence from a MIDAS approach during the Great Recession. *Applied Economics Letters*, 20, 233-237.

Forni M. and Lippi M. (1997). *Aggregation and the Microfoundations of Dynamic Macroeconomics.* Oxford University Press.

Forni M. and Lippi M. (2011). The general dynamic factor model: One-sided representation results. *Journal of Econometrics*, 163, 23-28.

Forni M. and Reichlin L. (1998). Let's get real: A factor-analytic approach to disaggregated business cycle dynamics. *Review of Economic Studies*, 65, 453-473.

Forni M., Hallin M., Lippi M. and Reichlin L. (2000). The generalized dynamic-factor model: Identification and estimation. *Review of Economics and Statistics*, 82, 540-554.

Forni M., Hallin M., Lippi M. and Reichlin L. (2003). Do financial variables help forecasting inflation and real activity in the euro area? *Journal of Monetary Economics*, 50, 1243-1255.

Forni M., Hallin M., Lippi M. and Reichlin L. (2004). The generalized factor model: Consistency and rates. *Journal of Econometrics*, 119, 231-255.

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, 830-840.

Forni M., Giannone D., Lippi M. and Reichlin L. (2009). Opening the black box: Structural factor models with large cross-sections. *Econometric Theory*, 25, 1319-1347.

Frale, C. Marcellino, M., Mazzi G.L. and Proietti T. (2010). Survey data as coincident or leading indicators. *Journal of Forecasting*, 29, 109-131.

Frale, C. Marcellino, M., Mazzi G.L. and Proietti T. (2011). EUROMIND: A monthly indicator of the euro area economic conditions. *Journal of the Royal Statistical Society*: Series A, 174, 2, 439-470.

Geweke J. (1977). The dynamic factor analysis of economic time series. In *Latent Variables in Socio Economic Models,* Aigner D.J. and Goldberger A.S. (eds). North Holland, Amsterdam.

Ghysels, E., Sinko and Valkanov (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews*, 26, 53-90.

Giannone D., Reichlin L. and Sala L. (2004). Monetary policy in real time. In Gertler M. and Rogoff K. (eds.) *NBER Macroeconomics Annual*, MIT Press.

Giannone D., Reichlin L. and Sala L. (2006). VARs factor models and the empirical validation of equilibrium of business cycle models. *Journal of Econometrics*, 132, 257-279.

Giannone D., Reichlin L. and Small D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55, 665-676.

Gorman W. M. (1981). Some Engel curves. In Deaton A. (ed.) *Essays in the Theory and Measurement of Consumer Behavior in Honor of Sir Richard Stone.* Cambridge University Press, New York.

Hallin M. and Liska R. (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association*, 102, 603-617.

Hallin M. and Liska R. (2011). Dynamic factors in the presence of blocks. *Journal of Econometrics*, 163, 29-41.

Hamilton J.D. (1989), A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, 57, 2, 357-384.

Hamilton J.D. (1994). *Time Series Analysis,* Princeton University Press, Princeton.

Harding M.C. (2009). Structural estimation of high-dimensional factor models. Mimeo, Department of Economics, Stanford University.

Helbling, T. and Bayoumi, T. (2003). Are they all in the same boat? The 2000-2001 growth slowdown and the G7-business cycle linkages. Working Paper, WP/03/46, IMF.

Jacobs J.P.A.M. and Otter P.W. (2008). Determining the number of factors and lag order in dynamic factor models: A minimum entropy approach. *Econometric Reviews*, 27, 385-397.

Jones C.S. (2001). Extracting factors from heteroskedastic asset returns. *Journal of Financial Economics*, 62, 293-325.

Jungbacker B. and Koopman S.J. (2008). Likelihood-based analysis for dynamic factor models. Discussion Paper NO 2008-0007/4, Tinbergen Institute, Amsterdam University.

Jungbacker B., Koopman S.J. 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.

Kapetanios G. (2004). A note on modelling core inflation for the UK using a new dynamic factor estimation method and a large disaggregated price index dataset. *Economics Letters*, 85, 63-69.

Kapetanios G. (2010). An alternative method for determining the number of factors in factor models with large datasets. *Journal of Business and Economic Statistics*, 28, 397-409.

Kapetanios G. and Marcellino M. (2004). A comparison of estimation methods for dynamic factor models of large dimension. Queen Mary Working Paper No. 489, London.

Kim C-J. and Nelson C.R. (1998). Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching. *Review of Economics and Statistics*, 80, 188-201.

Kim C-J. and Yoo, J.S. (1995). New index of coincident indicators: A multivariate Markov Switching factor model approach. *Journal of Monetary Economics*, 36, 607-630.

Kose M. A., Otrok C. and Whiteman C. (2003). International business cycles: World, region and country-specific factors. *American Economic Review*, 93, 4, 1216-1239.

Kose M. A., Otrok C. and Whiteman C. (2008). Understanding the evolution of world business cycles. *Journal of International Economics*, 75, 110-130.

Krolzig H.M. and Hendry D. (2001). Computer automation of general-to-specific model selection procedures. *Journal of Economic Dynamics and Control*, 25, 831-866.

Lawley D.N. and Maxwell A.E. (1971). *Factor Analysis in a Statistical Method*, Butterworth, London.

Lewbel A. (1991). The rank of demand systems: Theory and nonparametric estimation. *Econometrica*, 59, 711-730.

Lopes H.F. and West, M. (2004). Bayesian model assessment in factor analysis. *Statistica Sinica*, 14, 41-67.

Mansour M.J. (2003). Do national business cycles have an international origin? *Empirical Economics*, 28, 223-247.

Marcellino M., Stock, J. and Watson M. (2003). Macroeconomic forecasting in the euro area: Country specific versus euro wide information. *European Economic Review*, 47, 1-18.

Marcellino M. and Schumacher, C. (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.

Mariano R. and Murasawa Y. (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*. 18, 427-443.

Mariano R. and Murasawa Y. (2010). A coincident index, common factors and monthly real GDP. *Oxford Bulletin of Economics and Statistics*. 72, 1, 27-46.

Matheson, T. (2011). New indicators for tracking growth in real time. Working Paper, No. WP/11/43, International Monetary Fund.

Motta G., Hafner C. and von Sachs R. (2011). Locally stationary factor models: Identification and nonparametric estimation. *Econometric Theory*, 27, 6, 1279-1319.

Mumtaz, H. and Surico, P. (2009).The transmission of international shocks: A factor-augmented VAR approach. *Journal of Money, Credit and Banking*, 41, 71-100.

Newbold P. and Harvey, A. (2002), Forecast combination and encompassing, in *A Companion to Economic Forecasting,* Chapter 12, Clements M. and Hendry D. (eds), Basil Blackwell, Oxford.

Nguiffo-Boyom M. (2006). Un indicateur de retournement conjoncturel pour la France: une application du modèle à facteur avec changements de régimes. *Économie et Prévision*, 172, 101-114.

Onatski A. (2009). Testing hypotheses about the number of factors in large factor models. *Econometrica*, 77, 1447-1479.

Onatski A. (2010). Determining the number of factors from empirical distribution of eigenvalues. *The Review of Economics and Statistics*, 92, 4, 1004-1016.

Peña D. and Poncela, P. (2006a). Forecasting with non-stationary dynamic factor models. *Journal of Econometrics*, 119, 291-321.

Peña D. and Poncela, P. (2006b). Non-stationary dynamic factor analysis. *Journal of Statistical Planning and Inference*, 136, 1237-1257.

Ross S. (1976). The arbitrage theory of asset pricing. *Journal of Economic Theory*, 13, 341-360.

Ruenstler G., Barhoumi K., Benk S., Cristadoro R., Den Reijer A., Jakaitiene A., Jelonek P., Rua A., Ruth K., and Van Nieuwenhuyze C. (2009). Short-term forecasting of GDP using large monthly data sets: A pseudo real-time forecast evaluation exercise. *Journal of Forecasting*, 28, 7, 595-611.

Sargent T. and Sims C. (1977). Business cycle modelling without pretending to have too much a priori economic theory. In *New Methods in Business Cycle Research,* Sims C. (ed), Federal Reserve Bank of Minneapolis, Minneapolis.

Schumacher C. (2007). Forecasting German GDP using alternative factor models based on large datasets. *Journal of Forecasting*, 26, 271-302.

Schumacher C. (2010). Factor forecasting using international targeted predictors: The case of German GDP. *Economics Letters*, 107, 2, 95-98.

Schumacher C. and 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, 368-398.

Shumway R.H. and Stoffer D.S. (1982). An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis*, 3, 253-264.

Steiger J.H. (1979). Factor indeterminacy in the 1930s and the 1970s: Some interesting parallels. *Psychometrika*, 40, 157-167.

Stock J. and Watson M. (1989). New Indexes of Coincident and Leading Economic Indicators. In NBER Macroeconomics Annual, NBER.

Stock J. and Watson M. (1998). Diffusion indexes. Working Paper No 6702, NBER.

Stock J. and Watson M. (1999). Forecasting inflation. *Journal of Monetary Economics*, 44, 293-335.

Stock J. and Watson M. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20, 147-162.

Stock J. and Watson M. (2005a). Implications of dynamic factor models for VAR analysis. Working Paper No 11467, NBER.

Stock J. and Watson M. (2005b). Understanding changes in international business cycle dynamics. *Journal of the European Economic Association*, 3, 5, 968-1006.

Stock J. and Watson M. (2006). Forecasting with many predictors. In Elliott G., Granger C.W.J. and Timmermann A. (eds.), *Handbook of Economic Forecasting,* Vol. 1, North-Holland, Amsterdam.

Timmermann, A. (2006), Forecast combinations. In Elliott G., Granger C.W.J. and Timmermann A. (eds.), *Handbook of Economic Forecasting,* Vol. 1, 135-196, North-Holland, Amsterdam.

Van Nieuwenhuyze C. (2006). A generalized dynamic model for the Belgian Economy - useful business cycle indicators and GDP growth forecasts. Working Paper No 80, National Bank of Belgium.

Watson M.W. and Engle R.F. (1983). Alternative algorithms for estimation of dynamic factor, MIMIC, and time varying coefficient regression models. *Journal of Econometrics*, 23, 385-400.

White H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, 50, 1-25.

Yao Q. and Pan J. (2008). Modelling multiple time series via common factors. *Biometrika*, 95, 365-379.