

The generalized dynamic factor model: an application to forecasting

Il modello dinamico a fattori generalizzato: una applicazione alla previsione

Marco Lippi

Dipartimento di Scienze Economiche, Università di Roma “La Sapienza”

Riassunto: Modelli a fattori con un grande numero di variabili sono stati recentemente applicati intensivamente in macroeconomia. In particolare, è stato mostrato che i modelli a fattori danno buoni risultati in previsione se confrontati con modelli univariati o con modelli multivariati di piccole dimensioni. Questa presentazione si basa su Forni, Hallin, Lippi e Reichlin (2005), in cui si mostra come il modello dinamico rappresenti una buona alternativa rispetto ai modelli statici studiati in Stock e Watson (2002).

Keywords: Factor models, Forecasting.

1. Factor model in the analysis of (large) panels of time series data: a brief review.

1.1. Panels of time series data.

Denote by $\mathbf{X}_n^T = (x_{it}, i = 1, 2, \dots, n, t = 1, 2, \dots, T)$ an $n \times T$ rectangular array of observations. Throughout, we assume that

(F0a) $\mathbf{X}_n^T$ is a finite realization of a real-valued stochastic process $\mathbf{X} = \{x_{it}, i \in \mathbb{N}, t \in \mathbb{Z}\}$ indexed by $\mathbb{N} \times \mathbb{Z}$, where the n -dimensional vector process $\{\mathbf{x}_{nt} = (x_{1t} \ x_{2t} \ \dots \ x_{nt})', t \in \mathbb{Z}\}$ is stationary with finite first and second moments for any n .

The spectral techniques to be used in the sequel furthermore require the following technical assumption

(F0b) For all $n \in \mathbb{N}$, the vector $\mathbf{x}_{nt}$ admits a Wold representation $\mathbf{x}_{nt} = \sum_{k=0}^{\infty} \mathbf{C}_k^n \mathbf{w}_{nt-k}$, where the full-rank innovations $\mathbf{w}_{nt}$ have finite moments of order four, and the matrices $\mathbf{C}_k^n = (C_{ij,k}^n)$ satisfy $\sum_{k=0}^{\infty} |C_{ij,k}^n| k^{1/2} < \infty$ for all $n, i, j \in \mathbb{N}$.

Assumptions (F0a) and (F0b) jointly will be referred to as (F0).

It is important to point out that the roles played by the cross-sectional and the time dimensions are strongly asymmetric. A strong order in the *time* index $t \in \mathbb{Z}$ has been implicitly imposed, allowing for meaningful measures of serial dependence at given lags, the stationarity assumption (F0), spectral densities, etc. By contrast, we are not going to assume *cross-sectional stationarity*: for example, given t , there is no reason for the correlation between x_{1t} and x_{2t} being the same as the correlation between x_{3t} and x_{4t} . In the same way, the observation of $\mathbf{X}_n^T$ alone cannot help predicting x_{n+1t} , for any t . Indeed, we will need a regularity assumption in the cross-sectional dimension: this is a crucial requirement on the eigenvalues of the spectral density matrix of the vector $\mathbf{x}_t^n$ (see Section

1.3 below). However, such assumption, as the reader can easily check, remains invariant under permutation of the n cross-sectional items in the panel. As a rule, in the various factor models we are now describing, both assumptions and results are independent of the particular cross-sectional indexing.

1.2. Static and dynamic factor models.

The usual vector time-series models deal with situations in which T is large as compared to n , so that the relevant asymptotic theory takes n as given with T tending to infinity. The factor models to which we refer in Forni et al. (2000) and in the present paper are a useful alternative tool when n is too large with respect to T for vector time-series methods being applicable. In that case the relevant asymptotic theory must produce results for both T and n tending to infinity.

Standard factor models decompose a given finite number n of observed processes x_{it} into two non-observable components, a *common* component χ_{it} driven by q common factors F_{jt} , where q is very small as compared to n , and an *idiosyncratic* component ξ_{it} . In the simplest case the factors are *loaded* only contemporaneously:

$$x_{it} = \chi_{it} + \xi_{it} = a_{i1}F_{1t} + a_{i2}F_{2t} + \cdots + a_{iq}F_{qt} + \xi_{it}, \quad (1.1)$$

the usual assumptions being:

- (F1a) The processes $\{F_{jt}, j = 1, 2, \dots, q, t \in \mathbb{Z}\}$ and $\{\xi_{it}, i = 1, 2, \dots, n, t \in \mathbb{Z}\}$ are orthogonal (so that the processes $\{\chi_{it}, i = 1, 2, \dots, n, t \in \mathbb{Z}\}$ and $\{\xi_{it}, i = 1, 2, \dots, n, t \in \mathbb{Z}\}$ are orthogonal).
- (F1b) the idiosyncratic processes $\{\xi_{i't}\}$ and $\{\xi_{i''t}\}$ are orthogonal for all $i' \neq i''$.

Under Assumptions (F1a) and (F1b), model (1.1) can be identified (up to rotations of the factors) and estimated. The model has been extended by Sargent and Sims (1977) and Geweke (1977) to allow for a *dynamic* loading of the factors:

- (F2) There exists a q -dimensional (q very small with respect to n) non-observed orthonormal white noise process $\{\mathbf{u}_t = (u_{1t} \ u_{2t} \ \dots \ u_{qt})'\}$, the vector of *common shocks*, such that for $i = 1, 2, \dots, n$,

$$x_{it} = b_{i1}(L)u_{1t} + b_{i2}(L)u_{2t} + \cdots + b_{iq}(L)u_{qt} + \xi_{it}, \quad (1.2)$$

where $b_{ij}(L)$ is a square-summable filter, each idiosyncratic variable ξ_{it} being orthogonal to $\xi_{i't}$ at any lead and lag for $i \neq i'$, and to the shocks u_{jt} at any lead and lag for $j = 1, 2, \dots, q$.

1.3. Infinite cross-section: the generalized dynamic factor model

In most applications, the assumption that the idiosyncratic processes associated with distinct cross-sectional items are mutually orthogonal (see (F1b) and (F2)) is much too restrictive. On the other hand, dropping orthogonality of distinct idiosyncratic terms would destroy identifiability of model (1.1) or (1.2) for a given finite n . Chamberlain and Rothschild (1983) have introduced a more general definition of idiosyncratic component under the assumption of an infinite cross-sectional dimension for a static model like (1.1). Forni et al. (2000) have provided the following generalization to the dynamic

case. Considering the spectral density matrices $\Sigma_n^x(\theta)$ and $\Sigma_n^\xi(\theta)$, $\theta \in [-\pi, \pi]$, associated with $\{\mathbf{x}_{nt} = (\chi_{1t} \chi_{2t} \dots \chi_{nt})'\}$ and $\{\boldsymbol{\xi}_{nt} = (\xi_{1t} \xi_{2t} \dots \xi_{nt})'\}$, respectively, denote by λ_{nk}^x and λ_{nk}^ξ the corresponding *dynamic eigenvalues*, namely, the mappings $\theta \mapsto \lambda_{nk}^x(\theta)$ and $\theta \mapsto \lambda_{nk}^\xi(\theta)$, where $\lambda_{nk}^x(\theta)$ and $\lambda_{nk}^\xi(\theta)$ stand for $\Sigma_n^x(\theta)$ and $\Sigma_n^\xi(\theta)$'s respective k -th eigenvalues, in decreasing order, $k = 1, 2, \dots, n$. The *generalized dynamic factor model* analysed in Forni et al. (2000) and Forni and Lippi (2001) is model (F2), with infinite cross-sectional dimension, thus $i \in \mathbb{N}$, and the additional assumption:

(F3a) $\lambda_{nq}^x(\theta) \rightarrow \infty$ θ -a.e. in $[-\pi, \pi]$, and there exists a real Λ such that $\lambda_{n1}^\xi(\theta) \leq \Lambda$ for any $\theta \in [-\pi, \pi]$ and any $n \in \mathbb{N}$.

Assumption (F3a) has been discussed at length in Forni et al. (2000). Here we only recall that the condition on the asymptotic behaviour of $\lambda_{nk}^\xi(\theta)$ includes as a special case the usual mutually orthogonal idiosyncratic components, with the additional condition of bounded variance (unnecessary when n is finite), but allows the idiosyncratic components to be correlated in finite groups or even in more general ways. Moreover, as Forni and Lippi (2001) show, Assumption (F3a) is equivalent to the following condition on the eigenvalues $\lambda_{nk}(\theta)$ of the “observable” spectral density matrix $\Sigma_n(\theta) = \Sigma_n^x(\theta) + \Sigma_n^\xi(\theta)$:

(F3b) $\lambda_{nq}(\theta) \rightarrow \infty$ θ -a.e. in $[-\pi, \pi]$ as $n \rightarrow \infty$, and there exists a real Λ such that $\lambda_{nq+1}(\theta) \leq \Lambda$ for any $\theta \in [-\pi, \pi]$ and any $n \in \mathbb{N}$.

Forni et al. (2000) prove that under Assumptions (F0), (F2) and (F3a), or equivalently (F3b), Model (1.2) is identified, namely the common and the idiosyncratic components are unique (not the vector of common shocks $\mathbf{u}_t$). Moreover, Forni et al. (2000) construct a consistent estimator of the components χ_{it} and ξ_{it} . The definition of the estimator will be necessary in the sequel and is therefore recalled below.

Denote by $\Sigma_n^T(\theta) = (\sigma_{ij}^T(\theta))$, $\theta \in [-\pi, \pi]$, an arbitrary periodogram-smoothing or lag-window estimator of the spectral density $\Sigma_n(\theta) = (\sigma_{ij}(\theta))$ of $\mathbf{x}_{nt}$. Denote by $\lambda_{nj}^T(\theta)$ the j -th eigenvalue of $\Sigma_n^T(\theta)$ in decreasing order, $j = 1, 2, \dots, n$, and by $\mathbf{p}_{nj}^T(\theta)$ the corresponding row-eigenvector. We indicate by $\underline{\mathbf{p}}_{nj}^T(L)$ the filter obtained as the time-domain counterpart of the Fourier expansion of $\mathbf{p}_{nj}^T(\theta)$. Then consider the *dynamic principal components* $\underline{\mathbf{p}}_{nj}^T(L)\mathbf{x}_{nt}$, $j = 1, 2, \dots, n$. For the projection of x_{it} on the space spanned by the first q dynamic principal components, call it $\chi_{it,n}$, we have

$$\chi_{it,n}^T = \tilde{\underline{\mathbf{p}}}_{n1,i}^T(L)\underline{\mathbf{p}}_{n1}^T(L)\mathbf{x}_{nt} + \tilde{\underline{\mathbf{p}}}_{n2,i}^T(L)\underline{\mathbf{p}}_{n2}^T(L)\mathbf{x}_{nt} + \dots + \tilde{\underline{\mathbf{p}}}_{nq,i}^T(L)\underline{\mathbf{p}}_{nq}^T(L)\mathbf{x}_{nt}, \quad (1.3)$$

where $\tilde{\underline{\mathbf{p}}}_{nj,i}^T(\theta)$ is the i -th component of $\mathbf{p}_{nj}^T(\theta)$. Defining $\boldsymbol{\chi}_{nt}^T = (\chi_{1t,n}^T \chi_{2t,n}^T \dots \chi_{nt,n}^T)'$, we have

$$\boldsymbol{\chi}_{nt}^T = \tilde{\underline{\mathbf{p}}}_{n1}^T(L)\underline{\mathbf{p}}_{n1}^T(L)\mathbf{x}_{nt} + \tilde{\underline{\mathbf{p}}}_{n2}^T(L)\underline{\mathbf{p}}_{n2}^T(L)\mathbf{x}_{nt} + \dots + \tilde{\underline{\mathbf{p}}}_{nq}^T(L)\underline{\mathbf{p}}_{nq}^T(L)\mathbf{x}_{nt}. \quad (1.4)$$

Forni et al. (2000) prove that for T and n tending to infinity the projection $\chi_{it,n}^T$ tend in probability to χ_{it} . It is important to remark that, for fixed n and $T \rightarrow \infty$, the vector $\boldsymbol{\chi}_{nt}^T$ does not converge to $\boldsymbol{\chi}_{nt} = (\chi_{1t} \chi_{2t} \dots \chi_{nt})'$ but to

$$\check{\boldsymbol{\chi}}_{nt} = (\chi_{1t,n} \chi_{2t,n} \dots \chi_{nt,n})' = \tilde{\underline{\mathbf{p}}}_{n1}(L)\underline{\mathbf{p}}_{n1}(L)\mathbf{x}_{nt} + \tilde{\underline{\mathbf{p}}}_{n2}(L)\underline{\mathbf{p}}_{n2}(L)\mathbf{x}_{nt} + \dots + \tilde{\underline{\mathbf{p}}}_{nq}(L)\underline{\mathbf{p}}_{nq}(L)\mathbf{x}_{nt},$$

where $\mathbf{p}_{nj}(\theta)$ is the j -th eigenvector of $\Sigma_n(\theta)$. Convergence to the common components χ_{it} requires that n tends to infinity as well.

The drawback of the filters used in (1.3) is that they are in general two-sided and therefore the performance of the estimator $\chi_{it,n}^T$ deteriorates at the end of the sample and is of little use when forecasting is the main purpose. A strategy still based on the dynamic principal components of $\Sigma_n^T(\theta)$, but using only one-sided filters, thus suitable for end-of-sample estimation and forecasting, is developed in Section 2.

2. A two-step forecasting strategy

Given the decomposition $x_{it} = \chi_{it} + \xi_{it}$ and orthogonality at any lead and lag of χ_{it} and ξ_{it} , forecasting the variables x_{it} can be split into forecasting the χ 's and the ξ 's separately. Since we have assumed orthogonality or weak cross correlation between different idiosyncratic terms, their forecast can be left to univariate or small size vector modeling. Here we concentrate on forecasting the common components χ_{it} by means of the factor structure.

To provide a first idea of our method, let us suppose that we are trying to predict χ_{iT+h} for a given i . Let us denote by $\mathcal{G}(\chi, \tau)$ the space spanned by the variables χ_{jt} , $t \leq \tau$, $j \in \mathbb{N}$. If an estimate of $\mathcal{G}(\chi, T)$ is available, then the optimal linear forecast of χ_{iT+h} in the MMS sense is obtained by linearly projecting χ_{iT+h} onto the estimate of $\mathcal{G}(\chi, T)$. We proceed as follows. Firstly, the dynamic analysis developed in Forni et al. (2000), and summarized here in Section 1.3, is employed to obtain estimates of the covariance matrices $\Gamma_{nk}^\chi = E[\chi_{nt}\chi'_{nt-k}]$ and $\Gamma_{nk}^\xi = E[\xi_{nt}\xi'_{nt-k}]$. This is Step 1. In Step 2 we obtain an estimate of the space $\mathcal{G}(\chi, T)$, and of the projection $\text{proj}(\chi_{iT+h}|\mathcal{G}(\chi, T))$, by means of linear combinations of contemporaneous and past values of the x 's, i.e. by *one-sided* filtering of the x 's, the coefficients of the filters being based on the estimates of Γ_{nk}^χ and Γ_{nk}^ξ .

2.1. Step 1: Estimation of common and idiosyncratic cross-covariances

To estimate Γ_{nk}^χ and Γ_{nk}^ξ we start with the matrix $\Sigma_n^T(\theta)$, defined in Section 1.3 as an arbitrary periodogram-smoothing or lag-window estimator of the spectral density $\Sigma_n(\theta)$. Under (F0), this estimator is such that, for all $n \in \mathbb{N}$,

$$\lim_{T \rightarrow \infty} P \left[\max_{1 \leq i, j \leq n} \sup_{\theta \in [-\pi, \pi]} |\sigma_{ij}^T(\theta) - \sigma_{ij}(\theta)| > \epsilon \right] = 0 \quad (2.1)$$

(see Brockwell and Davis, 1987, p. 433). Put

$$\mathbf{P}_n^T(\theta) = (\mathbf{p}_{n1}^{T'}(\theta) \mathbf{p}_{n2}^{T'}(\theta) \dots \mathbf{p}_{nq}^{T'}(\theta))' \quad \text{and} \quad \mathbf{Q}_n^T(\theta) = (\mathbf{p}_{nq+1}^{T'}(\theta) \mathbf{p}_{nq+2}^{T'}(\theta) \dots \mathbf{p}_{nn}^{T'}(\theta))',$$

and

$$\begin{aligned} \Sigma_n^{\chi T}(\theta) &= \tilde{\mathbf{P}}_n^T(\theta) \text{Diag} \left(\lambda_{n1}^T(\theta) \lambda_{n2}^T(\theta) \dots \lambda_{nq}^T(\theta) \right) \mathbf{P}_n^T(\theta) \\ \Sigma_n^{\xi T}(\theta) &= \tilde{\mathbf{Q}}_n^T(\theta) \text{Diag} \left(\lambda_{nq+1}^T(\theta) \lambda_{nq+2}^T(\theta) \dots \lambda_{nn}^T(\theta) \right) \mathbf{Q}_n^T(\theta) \end{aligned}$$

By (1.4), the matrices $\Sigma_n^{\chi T}(\theta)$ and $\Sigma_n^{\xi T}(\theta)$ are the spectral density matrices of the estimated χ_{nt}^T and $\xi_{nt}^T = \mathbf{x}_{nt} - \chi_{nt}^T$ respectively, and are therefore consistent estimates, for $T \rightarrow \infty$,

of the spectral density matrices $\tilde{\Sigma}_n^\chi(\theta)$ and $\tilde{\Sigma}_n^\xi(\theta)$, corresponding to the vectors $\tilde{\chi}_{nt}$ and $\tilde{\xi}_{nt} = \mathbf{x}_{nt} - \tilde{\chi}_{nt}$ respectively. Note that $\Sigma_n^T(\theta) = \Sigma_n^{\chi T}(\theta) + \Sigma_n^{\xi T}(\theta)$. Note also that $\Sigma_n(\theta) = \tilde{\Sigma}_n^\chi(\theta) + \tilde{\Sigma}_n^\xi(\theta) = \Sigma_n^\chi(\theta) + \Sigma_n^\xi(\theta)$ but that in general $\tilde{\Sigma}_n^\chi(\theta) \neq \Sigma_n^\chi(\theta)$ and $\tilde{\Sigma}_n^\xi(\theta) \neq \Sigma_n^\xi(\theta)$.

Now, continuity of eigenvalues and eigenvectors as functions of the matrix entries and (2.1) imply that

$$\mathbf{\Gamma}_{nk}^{\chi T} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \Sigma_n^{\chi T}(\theta) d\theta \quad \text{and} \quad \mathbf{\Gamma}_{nk}^{\xi T} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \Sigma_n^{\xi T}(\theta) d\theta \quad (2.2)$$

are consistent estimates of

$$\check{\mathbf{\Gamma}}_{nk}^\chi = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \tilde{\Sigma}_{n,n}^\chi(\theta) d\theta \quad \text{and} \quad \check{\mathbf{\Gamma}}_{nk,n}^\xi = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\theta} \tilde{\Sigma}_n^\xi(\theta) d\theta$$

respectively, as T tends to infinity. Again, we must remark that $\check{\mathbf{\Gamma}}_{nk}^\chi \neq \mathbf{\Gamma}_{nk}^\chi$ and $\check{\mathbf{\Gamma}}_{nk,n}^\xi \neq \mathbf{\Gamma}_{nk}^\xi$. (However, since $\chi_{it,n} \rightarrow \chi_{it}$ for any i as $n \rightarrow \infty$, as proved in Forni et al. (2000), then each entry of $\tilde{\Sigma}_n^\chi(\theta)$ converges in first mean to the corresponding entry of $\Sigma_n^\chi(\theta)$. As a consequence, each entry of $\check{\mathbf{\Gamma}}_{nk}^\chi$ converges to the corresponding entry of $\mathbf{\Gamma}_{nk}^\chi$. The same holds substituting ξ for χ .)

Note that the estimates (2.2) of the matrices $\mathbf{\Gamma}_{nk}^{\chi T}$ and $\mathbf{\Gamma}_{nk}^{\xi T}$ are not obtained from estimated covariances of the estimates χ_{it}^T and ξ_{it}^T , as defined by (1.3), but directly from the spectral density matrix $\Sigma_n^T(\theta)$, through the eigenvectors $\mathbf{p}_{nj}^T(\theta)$. A one-sided re-estimation of χ_{it} and ξ_{it} , based on the matrices $\mathbf{\Gamma}_{nk}^{\chi T}$ and $\mathbf{\Gamma}_{nk}^{\xi T}$ will be obtained as a by-product of Step 2.

2.2. Step 2: Forecasting the common component

We want to estimate the projection

$$\text{proj}(\chi_{iT+h} | \mathcal{G}(\chi, T)).$$

In Forni et al. (2000) the variables χ_{iT-k} , for $k \geq 0$, spanning the space $\mathcal{G}(\chi, T)$, are consistently (as n and T tend to infinity) estimated by sequences $\mathbf{a}_n^T(L)\mathbf{x}_{nt}$, where the filters $\mathbf{a}_n^T(L)$ are two-sided (see (1.3)). Now we are constrained to use sequences $\mathbf{b}_n^T(L)\mathbf{x}_{nt}$ where the filters $\mathbf{b}_n^T(L)$ are one-sided.

We need some preliminary definition. Firstly, define the $ng \times 1$ stack

$$\mathbf{x}_{nt,g} = (\mathbf{x}_{nt}' \mathbf{x}_{nt-1}' \cdots \mathbf{x}_{nt-g}')'.$$

In the same way, the stacks $\chi_{nt,g}^T$ and $\xi_{nt,g}^T$ are obtained from the estimates χ_{nt}^T and ξ_{nt}^T respectively. The $ng \times ng$ variance-covariance matrices of the stacks are denoted by $\mathbf{\Upsilon}_{n,g}^T$, $\mathbf{\Upsilon}_{n,g}^{\chi T}$ and $\mathbf{\Upsilon}_{n,g}^{\xi T}$ respectively, and can be obtained in an obvious way from the matrices $\mathbf{\Gamma}_{nk}^T$, $\mathbf{\Gamma}_{nk}^{\chi T}$ and $\mathbf{\Gamma}_{nk}^{\xi T}$. Then define the $1 \times ng$ vector $\mathbf{Z}_{n\ell,g}^T$, for $\ell \leq ng$ in the following way:

$$\begin{aligned} \mathbf{Z}_{n\ell,g}^T = \text{Arg} \quad & \max_{\mathbf{a}} \text{var} \left(\mathbf{a} \chi_{nt,g}^T \right) \\ \text{s.t.} \quad & \text{var} \left(\mathbf{a} \xi_{nt,g}^T \right) = 1, \quad \mathbf{a} \xi_{nt,g}^T \perp \mathbf{Z}_{nm,g}^T \xi_{nt,g}^T, \quad \text{for } m < \ell \end{aligned} \quad (2.3)$$

where $\mathbf{a}$ is $1 \times ng$. Equivalently:

$$\begin{aligned} \mathbf{Z}_{n\ell,g}^T = \text{Arg} \quad & \max_{\mathbf{a}} \mathbf{a} \mathbf{\Upsilon}_{n,g}^{\chi T} \mathbf{a}' \\ \text{s.t.} \quad & \mathbf{a} \mathbf{\Upsilon}_{n,g}^{\xi T} \mathbf{a}' = 1, \quad \mathbf{a} \mathbf{\Upsilon}_{n,g}^{\xi T} \mathbf{Z}_{nm,g}^T = 0, \quad \text{for } m < \ell \end{aligned} \quad (2.4)$$

Lastly, define r sequences $\{W_{nt,g}^{kT}, n \in \mathbb{N}\}$, as follows

$$W_{nt,g}^{kT} = \mathbf{Z}_{nk,g}^T \mathbf{x}_{nt,g}, \quad k = 1, 2, \dots, r.$$

In conclusion, we have constructed r sequences of aggregates of the x 's and their lagged values up to lag g , thus using one-sided filtering of the x 's. The aggregation weights are obtained exploiting the information contained in the matrices $\mathbf{\Upsilon}_{n,g}^{\chi T}$ and $\mathbf{\Upsilon}_{n,g}^{\xi T}$, estimated in Step 1. Precisely, given the estimated variance-covariance matrices of common and idiosyncratic components, in (2.3) the weights are chosen in such a way that the common-idiosyncratic variance ratio is maximized in the aggregates.

In Forni, Hallin, Lippi and Reichlin (2005) we show that, under simple additional assumptions on the process $\mathbf{X} = \{x_{it}, i \in \mathbb{N}, t \in \mathbb{Z}\}$, and a specification of g as $g(n, T)$ and r as $r(n, T)$, the projection of x_{iT+h} onto the space spanned by $W_{nT,g}^{kT}$, $k = 1, 2, \dots, r$, converges in probability to $\text{proj}(\chi_{iT+h} | \mathcal{G}(\chi, T))$. Obviously, the same method will produce the result that the projection of x_{it} , $t \leq T$, onto the space spanned by $W_{nt,g}^{kT}$, $k = 1, 2, \dots, r$, converges in probability to χ_{it} , so that a one-sided consistent re-estimation of χ_{it} , for $t \leq T$, is obtained, solving the end-of-sample estimation problem mentioned in Section 1.3.

A comparison between our weights $\mathbf{Z}_{nh,g}^T$ and those proposed in Stock and Watson (2002) can be useful. Stock and Watson's weights are defined (in our notation) as the (static) principal components of the stacked x 's:

$$\begin{aligned} \mathbf{S}_{n\ell,g}^T = \text{Arg} \quad & \max_{\mathbf{a}} \mathbf{a} \mathbf{\Upsilon}_{n,g}^T \mathbf{a}' \\ \text{s.t.} \quad & \mathbf{a} \mathbf{a}' = 1, \quad \mathbf{a} \mathbf{S}_{nm,g}^T = 0, \quad \text{for } m < \ell \end{aligned} \quad (2.5)$$

This one-step procedure produces a consistent estimation of $\mathcal{G}(\chi, T)$. If the matrices $\mathbf{\Upsilon}_{n,g}^{\chi T}$ and $\mathbf{\Upsilon}_{n,g}^{\xi T}$ provide a good estimate of the matrices $\mathbf{\Upsilon}_{n,g}^{\chi}$ and $\mathbf{\Upsilon}_{n,g}^{\xi}$ respectively, then our $\mathbf{Z}$ -weights, which are tailored with the purpose of minimizing the idiosyncratic with respect to the common term in the aggregate, should perform better than the $\mathbf{S}$ -weights to approximate the space $\mathcal{G}(\chi, T)$. However, goodness of such estimates will depend in a very complicated way on the underlying model, on T and on n , so that an *a priori* assessment of Stock and Watson's and our method seems very difficult. In the present paper we limit ourselves to presenting results obtained with simulated data and an empirical panel.

References

- Brockwell, P.J. and R. A. Davis (1991), *Time Series: Theory and Methods* (2nd Ed.), New York: Springer-Verlag.
- Chamberlain, G. and M. Rothschild (1983). Arbitrage, factor structure and mean-variance analysis in large asset markets. *Econometrica* **51**, 1305-1324.
- Forni, M. and M. Lippi (2001). The generalized dynamic factor model : representation theory. *Econometric Theory* **17**, 1113-1141.

- Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000). The generalized factor model : identification and estimation. *The Review of Economics and Statistics* **82**, 540-554.
- Forni, M., M. Hallin, M. Lippi and L. Reichlin (2005), The generalized dynamic factor model: one-sided estimation and forecasting, *Journal of the American Statistical Association*, **100**, 830-40.
- Sargent, T.J. and C.A. Sims (1977). Business cycle modeling without pretending to have too much *a priori* economic theory. In C.A. Sims, Ed., *New Methods in Business Research*, Federal Reserve Bank of Minneapolis, Minneapolis.
- Geweke, J. (1977). The dynamic factor analysis of economic time series. In D.J. Aigner and A.S. Goldberger, Eds., *Latent Variables in Socio-Economic Models*, Amsterdam: North Holland.
- Stock, J.H. and M.W. Watson (2002). Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business and Economic Statistics* **20**, 147-162.