Banco de Portugal

EUROSISTEMA

Estudos e Documentos de Trabalho

Working Papers

9 | 2008

DETERMINING THE NUMBER OF FACTORS IN APPROXIMATE FACTOR MODELS WITH GLOBAL AND GROUP-SPECIFIC FACTORS

Francisco Dias

Maximiano Pinheiro

António Rua

June 2008

Please address correspondence to

Francisco Dias

fadias@bportugal.pt

Banco de Portugal, Av. Almirante Reis no. 71, 1150-012 Lisboa, Portugal;

Tel.: 351 213 128 327

BANCO DE PORTUGAL

Economics and Research Department

Av. Almirante Reis, 71-6th floor

1150-012 Lisboa

www.bportugal.pt

Printed and distributed by

Administrative Services Department

Av. Almirante Reis, 71-2nd floor

1150-012 Lisboa

Number of copies printed

170 issues

Legal Deposit no. 3664/83

ISSN 0870-0117

ISBN 978-989-8061-40-9

Determining the number of factors in approximate factor models with global and group-specific factors

Francisco Dias[*] Maximiano Pinheiro[*,†] António Rua[*]

June 2008

Abstract

For an approximate factor model, in a static representation, with a common component comprising global factors and factors specific to groups of variables, the consistency of the principal components estimator is discussed. An extension of the well known Bai and Ng criteria is proposed for determining the number of global and group-specific factors. The consistency of the suggested criteria is established and the small sample properties are assessed through Monte Carlo simulations. As an empirical illustration, the proposed criteria is applied to estimate the number of global and country-specific macroeconomic factors for the major euro area countries.

Keywords: factor models, global factors, group-specific factors, model selection, criteria.

JEL classification: C32, C52.

[*]Banco de Portugal

[†]ISEG, Technical University of Lisbon

1 Introduction

Factor models are often used in macroeconomics and finance. There has been an ample supporting evidence, based on both single country and multi-country datasets, of the existence of generalized co-movements of macroeconomic series. This evidence has led to the widespread use of factor models to study business cycles and to improve the macroeconomic forecasting performance. The literature was initiated by Geweke (1977) and Sims and Sargent (1977), but until the mid 90's was limited to panels with a small number of variables[1]. Stock and Watson (1998, 1999, 2002a, 2002b), Forni and Reichlin (1998) and Forni *et al.* (2000, 2001, 2004, 2005) made contributions that effectively removed the restriction on the number of variables in macroeconomic factor models. In particular, the seminal work of Stock and Watson, which advocates the use of the principal components method to estimate common factors from a large number of series, became very popular among forecasters. A related field of research integrates factor models into structural vector autegression analysis (*inter alia* Bernanke and Boivin (2003), Bernanke *et al.* (2005), Favero *et al.* (2005) and Giannone *et al.* (2006)). In finance, factor models have also been used since the mid 70's to study asset pricing. Main contributions include Ross (1976), Roll and Ross (1980), Chamberlain and Rothschild (1983), Connor and Korajczyk (1986,1988) and Geweke and Zhou (1996).

In the conventional factor model representation, each variable is assumed to be the sum of two components, a component associated with factors common to all series and an idiosyncratic term. However, in some factor models, the set of variables is partitioned into several subsets and the common component is broken down into two or more levels. There are global common factors, shared by the data generating process of all variables, and group-specific common factors, each shared only by a group of variables. For example, this multi-level common component feature appears often in factor models of the international business cycle (Norrbin and Schlagenhauf (1996), Gregory *et al.* (1997), Gregory and Head (1999), Kose *et al.* (2003) and Canova *et al.* (2007), among others). Typically, in these studies the number of variables is small and the number of global and country-specific factors is imposed *a priori* to be one, rather than being data determined.

Other researchers, but working with datasets with large number of variables, have also admitted partitions of their sets of variables. For instance, Boivin and Ng (2006) use group-specific factors for the US (namely, real, nominal and other) and Marcellino *et al.* (2003) consider both euro area as a whole and country-specific factors for individual euro area countries. Again, the number of global and group specific-factors is fixed *ad-hoc*.

Hence, the determination of the number of global and group-specific factors on data dependent method remains an open issue. The consideration of one

[1] Although Quah and Sargent (1993) already considered a moderate size panel with 60 variables.

global factor and one group-specific factor for each group of variables, as assumed in many empirical applications, may be too restrictive. This is a model selection problem for which no criteria exists in the literature. In this paper, we address the issue of determining the number of global and group-specific common factors when the number of variables in each group is at least moderately large and the partition of the set of variables into groups is exogenously determined.

The framework herein considered is a static approximate factor model allowing for limited dependence in the idiosyncratic terms, in both the time and cross-section dimensions, as well as some contemporaneous correlation between the factors and the idiosyncratic component. In the common component, we distinguish two levels of common factors, global and group-specific. The global factors are allowed to appear in the data generating process of all variables whereas each group-specific factor is restricted *a priori* to the corresponding group of variables.

Connor and Korajczyk (1986,1988) discussed the method of principal components for the conventional static representation, with only global factors, and studied its asymptotic properties for fixed number of time observations. Stock and Watson (1998, 2002a), Bai and Ng (2002) and Amengual and Watson (2005, 2007) considered the asymptotic properties when both panel dimensions grow to infinity, differing slightly on the set of assumptions regarding the data generating processes of the factors and the idiosyncratic errors. In this paper, we begin by extending Bai and Ng (2002) consistency result to our framework, for a given number of global and group-specific factors.

Additionally, Bai and Ng (2002) presented criteria, taking into account the usual trade-off between fit and parsimony, to determine the number of factors in a model with only global factors. In this paper, we also extend the Bai and Ng criteria to the case of factor models with both global and group-specific factors. The asymptotic properties of the proposed criteria are established and a Monte Carlo simulation exercise is conducted to assess the small sample behavior of the criteria.

We illustrate our approach by estimating the number of global and country-specific factors for the major euro area countries (Germany, France, Italy and Spain). The dataset encompasses almost three hundred series over the last two decades. Interestingly, we find two global factors, one related to real activity developments and the other refering to inflation evolution in the euro area.

This paper is organised as follows. In section 2, the static approximate factor model with global and group-specific factors is presented. We discuss the consistency of the principal components estimator for a given number of global and group-specific factors in section 3. In section 4, criteria for determining the number of factors are suggested and the corresponding asymptotic results are established. In section 5, the finite sample properties of the criteria are assessed through a Monte Carlo simulation exercise. In section 6, the empirical illustration is presented. Finally, the last section concludes.

2 The model

Let X_t be a N-dimensional stationary time series observed for $t = 1, \cdots, T$. X_t is partitioned in G sub-vectors (groups) of variables,

$$X_t = \left[X'_{1,t} \cdots X'_{g,t} \cdots X'_{G,t}\right]'$$

where

$$X_{g,t} = \left[X_{g,1t} \cdots X_{g,nt} \cdots X_{g,N_g t}\right]'$$

is the column-vector of observations at time t of the $N_g$variables belonging to group g ($g = 1, \cdots, G$), with $\sum_{g=1}^{G} N_g = N$. We are interested in panels with at least moderately large T and large N. Moreover, we will focus on partitions such that the numbers of variables in each group, N_g, are also at least moderately large.

We suppose that $X_{g,nt}$ admits a static factor representation with r_0 global common factors $F^0_{0,t}$ and r_g factors specific to the g-th group of variables $F^0_{g,t}$,

$$X_{g,nt} = \lambda^{0\prime}_{g0,n} F^0_{0,t} + \lambda^{0\prime}_{gg,n} F^0_{g,t} + e_{g,nt}$$
$$(t = 1, \cdots, T; n = 1, \cdots, N_g; g = 1, \cdots, G) \qquad (1)$$

The vectors of global and group-specific factors, the associated factor loadings $\lambda^0_{g0,n}$ ($r_0 \times 1$) and $\lambda^0_{gg,n}$ ($r_g \times 1$) and the idiosyncratic errors $e_{g,nt}$ are all not observable (the superscript 0 denotes the true factors and loadings). In this representation, the group-specific factors $F^0_{g,t}$ are restricted to the common component of the variables belonging to group g, whilst the global factors may appear in the equations of all variables. We will assume that the partition of variables is given, implying that we know which variables belong to each group.

When $r_g = 0$ for all $g = 1, \cdots, G$ (i.e. when there are no group-specific factors), the equation above reduces to the conventional static representation of the factor model, without any specific contribution of groups of variables to the common component. On the other hand, when $r_0 = 0$ (i.e. there are no global factors), model (1) consists of G conventional static representations assembled together, one for each group of variables.

Let Λ^0_{g0} ($N_g \times r_0$) and Λ^0_{gg} ($N_g \times r_g$) be the matrices of true loadings of the variables of group g ($g = 1, \cdots, G$). Denoting by $e_{g,t}$ the $N_g \times 1$ vector of idiosyncratic errors for the variables $X_{g,t}$ at time t, we have an equivalent expression for model (1):

$$X_{g,t} = \Lambda^0_{g0} F^0_{0,t} + \Lambda^0_{gg} F^0_{g,t} + e_{g,t} \qquad (t = 1, \cdots, T; g = 1, \cdots, G) \qquad (2)$$

or, transposing the equation and stacking the time observations,

$$\underline{X}_g = F^0_0 \Lambda^{0\prime}_{g0} + F^0_g \Lambda^{0\prime}_{gg} + \underline{e}_g \qquad (g = 1, \cdots, G) \qquad (3)$$

where

$$\underline{X}_g = [X_{g,1} \cdots X_{g,t} \cdots X_{g,T}]' \qquad (T \times N_g)$$

$$F_0^0 = \left[F_{0,1}^0 \cdots F_{0,t}^0 \cdots F_{0,T}^0\right]' \qquad (T \times r_0)$$
$$F_g^0 = \left[F_{g,1}^0 \cdots F_{g,t}^0 \cdots F_{g,T}^0\right]' \qquad (T \times r_g)$$
$$\underline{e}_g = [e_{g,1} \cdots e_{g,t} \cdots e_{g,T}]' \qquad (T \times N_g)$$

Instead, if we stack all the variables for time period t, the model may be represented by the set of equations

$$X_t = \Lambda_0^0 F_{0,t}^0 + \Lambda_S^0 F_{S,t}^0 + e_t \qquad (t = 1, \cdots, T)$$

where

$$F_{S,t}^0 = \left[F_{1,t}^{0\prime} \cdots F_{g,t}^{0\prime} \cdots F_{G,t}^{0\prime}\right]' \qquad (\ddot{r}_S \times 1)$$
$$\Lambda_0^0 = \left[\Lambda_{10}^{0\prime} \cdots \Lambda_{g0}^{0\prime} \cdots \Lambda_{G0}^{0\prime}\right]' \qquad (N \times r_0)$$
$$\Lambda_S^0 = diag\left(\Lambda_{11}^0, \cdots, \Lambda_{gg}^0, \cdots, \Lambda_{GG}^0\right) \qquad (N \times \ddot{r}_S)$$

and

$$e_t = \left[e_{1,t}' \cdots e_{g,t}' \cdots e_{G,t}'\right]' \qquad (N \times 1)$$

with $\ddot{r}_S = \sum_{g=1}^{G} r_g$. When considering simultaneously all time observations and all variables, we will adopt the following compact notation

$$X = F^0 \Lambda^{0\prime} + e$$

where

$$X = [X_1 \cdots X_t \cdots X_T]' = \left[\underline{X}_1 \cdots \underline{X}_g \cdots \underline{X}_G\right] \qquad (T \times N)$$
$$F^0 = \left[F_0^0 \ F_S^0\right] \qquad (T \times \ddot{r})$$
$$F_S^0 = \left[F_1^0 \cdots F_g^0 \cdots F_G^0\right] \qquad (T \times \ddot{r}_S)$$
$$\Lambda^0 = \left[\Lambda_0^0 \ \Lambda_S^0\right] = \begin{bmatrix} \Lambda_{10}^0 & \Lambda_{11}^0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ \Lambda_{g0}^0 & 0 & \cdots & \Lambda_{gg}^0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ \Lambda_{G0}^0 & 0 & \cdots & 0 & \cdots & \Lambda_{GG}^0 \end{bmatrix} \qquad (N \times \ddot{r})$$

and

$$e = [e_1 \cdots e_t \cdots e_T]' = \left[\underline{e}_1 \cdots \underline{e}_g \cdots \underline{e}_G\right] \qquad (T \times N)$$

with $\ddot{r} = r_0 + \ddot{r}_S = \sum_{g=0}^{G} r_g$.

For $g = 1, \cdots, G$, let

$$\mathcal{F}_g^0 = \left[F_0^0 \ F_g^0\right] \quad (T \times (r_0 + r_g))$$

and

$$\mathcal{L}_g^0 = \left[\Lambda_{g0}^0 \ \Lambda_{gg}^0\right] \quad (N_g \times (r_0 + r_g))$$

The t-th row of $\mathcal{F}_g^0$ and the n-th row of $\mathcal{L}_g^0$ will be denoted by $\mathcal{F}_{g,t}^{0\prime} = \left[F_{0,t}^{0\prime}\; F_{g,t}^{0\prime}\right]$ and $\mathcal{L}_{g,n}^{0\prime} = \left[\lambda_{g0,n}^{0\prime}\; \lambda_{gg,n}^{0\prime}\right]$, respectively. Using this alternative notation, (2) and (3) may be written as,

$$X_{g,t} = \mathcal{L}_g^0 \mathcal{F}_{g,t}^0 + e_{g,t} \qquad (t = 1, \cdots, T; g = 1, \cdots, G)$$

and

$$\underline{X}_g = \mathcal{F}_g^0 \mathcal{L}_g^{0\prime} + \underline{e}_g \qquad (g = 1, \cdots, G) \tag{4}$$

To develop asymptotic results when both $T \to \infty$ and $N_g \to \infty$ ($g = 1, \cdots, G$), we need to complement the model specification with a set of assumptions, in particular regarding the data generating processes of the factors and of the idiosyncratic errors. Hereafter, the Frobenius norm of a matrix Y will be denoted by $||Y|| = [tr\,(Y'Y)]^{\frac{1}{2}}$, where $tr(Y)$ is the trace of Y. Also let $\gamma_g(s,t) = E\left(N_g^{-1} e_{g,s}' e_{g,t}\right) = E\left(N_g^{-1} \sum_{n=1}^{N_g} e_{g,ns} e_{g,nt}\right)$.

Assumption A (Factors): *For every* $g = 1, \cdots, G$,
A.1. $||T^{-1} \mathcal{F}_g^{0\prime} \mathcal{F}_g^0 - I|| \to 0$ *as* $T \to \infty$*;*
A.2. $E\left(||\mathcal{F}_{g,t}^0||^4\right) < \infty$ *for all* $t = 1, \cdots, T$*;*
A.3. $r_0 + r_g > 0$*;*

Assumption B (Factor loadings): *For every* $g = 1, \cdots, G$,
B.1. $||\mathcal{L}_{g,n}^0|| \leq \ell$ *for some finite positive constant* ℓ *and for all* $n = 1, \cdots, N_g$*;*
B.2. $||N_g^{-1} \mathcal{L}_g^{0\prime} \mathcal{L}_g^0 - \Omega_g|| \to 0$ as $N_g \to \infty$ *for some positive definite matrix* Ω_g $((r_0 + r_g) \times (r_0 + r_g))$.

Assumption C (Idiosyncratic components): *For every* $g = 1, \cdots, G$, *there exists a positive finite constant* M *such that for all* N_g *and* T,
C.1. $E\left(e_{g,nt}^4\right) \leq M$ *for* all $n = 1, \cdots, N_g$ *and* $t = 1, \cdots, T$*;*
C.2. $|\gamma_g\,(s,s)\,| \leq M$ *for all* $s = 1, \cdots, T$ *and* $T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} |\gamma_g\,(s,t)\,| \leq M$*;*
C.3. $|E\,(e_{g,nt} e_{g,mt})\,| \leq \tau_{g,nm}$ $(t = 1, \cdots, T)$ *for some finite* $\tau_{g,nm} > 0$ *such that* $N_g^{-1} \sum_{n=1}^{N_g} \sum_{m=1}^{N_g} \tau_{g,nm} \leq M$*;*
C.4. $[\min\,(N_g, T)]^{j-1}\, E\left(tr\left\{\left[(N_g T)^{-1}\, \underline{e}_g \underline{e}_g'\right]^j\right\}\right) \leq M$ *for all* $j \geq 1$*;*
C.5. $E\left\{N_g^{-\frac{1}{2}} \sum_{n=1}^{N_g} \left[e_{g,nt} e_{g,ns} - E\,(e_{g,nt} e_{g,ms})\right]\right\}^4 \leq M$ *for all* $t = 1, \cdots, T$ *and* $s = 1, \cdots, T$.

Assumption D (Dependence between factors and idiosyncratic components): *For every* $g = 1, \cdots, G$, *there is a positive finite constant* M *such that for all* N_g *and* T,

$$E\left\{\left[\frac{1}{N_g} \sum_{n=1}^{N_g} ||T^{-\frac{1}{2}} \sum_{t=1}^{T} \mathcal{F}_{g,t}^0 e_{g,nt}||^2\right]\right\} \leq M$$

Assumption E (Relative size of the groups of variables): *For every* $g = 1, \cdots, G$, $|N_g/N - \pi_g| \to 0$ for some $\pi_g > 0$ *as* $N \to \infty$.

Assumption F (Irreducibility): *The model is in its irreducible representation, i.e. there is no equivalent representation of the model with the same idiosyncratic errors and a smaller total number of factors.*

Assumptions A to D are mostly an adaptation to our model of the assumptions considered by Bai and Ng (2002). As of Assumption A, for the conventional factor model without group-specific factors, besides restricting the factors to be finite, Bai and Ng admit that the sample uncentered second moments matrix of the true factors converge to a positive definite matrix as $T \to \infty$. Without any loss of generality and to simplify the algebra, for each group g of variables we impose that the limit of the sample uncentered second moments matrix is the identity matrix. Note that Assumption A does not constrain the correlations between the group-specific factors of any two different groups of variables. Indeed, besides ensuring that the model includes at least one factor (global or local) in every group g and that the 4th moments of all factors are bounded, Assumption A only imposes that, as $T \to \infty$,

$$||\frac{1}{T}F_g^{0\prime}F_g^0 - I_{r_g}|| \to 0 \qquad and \qquad ||\frac{1}{T}F_0^{0\prime}F_g^0|| \to 0 \qquad (g = 1, \cdots, G)$$

Nothing is assumed concerning the moments $T^{-1}F_s^{0\prime}F_g^0$ $(s, g > 0)$ but Assumption F, which imposes a weak constraint on the group-specific factors. Assumption F rules out situations where some group-specific factors are common to all groups. Obviously, in this case, the common group-specific factors should have been classified as global factors in order to get an irreducible representation with fewer factors.

Assumption B imposes that the size of the true loadings is bounded and that for every group of variables (and not only for the whole set of variables) $N_g^{-1}\mathcal{L}_g^{0\prime}\mathcal{L}_g^0$ converges to a positive definite matrix as $N_g \to \infty$. In particular, Assumption B.2 requires that for large N_g, (i) each global factor is present (i.e. has a non-null loading) in at least one equation of that group and (ii) each group g specific factor appears in at least one equation of that group. Moreover, note that, together with Assumption E, Assumption B.2 implies that

$$||\frac{1}{N}\Lambda^{0\prime}\Lambda^0 - \Omega|| \to 0$$

as $N \to \infty$ for some definite positive matrix Ω $(\ddot{r} \times \ddot{r})$ (see Appendix I).

For every group of variables, Assumptions C allows for some heteroskedasticity and limited dependence of the idiosyncratic errors in both the time and the cross-section dimensions. Assumption C.1 is a weaker version of the corresponding assumption used by Bai and Ng. To establish the asymptotic results that are presented below, the assumption of zero mean of the idiosyncratic component is not required and we only need to admit that their fourth moments are finite, instead of the 8th moments. Assumptions C.2, C.3 and C.5 are direct adaptations of the corresponding technical assumptions considered by Bai and Ng. Assumption C.4 of Bai and Ng was dropped as it was redundant and replaced by a version of Assumption (A.6) considered by Amengual and Watson (2005, 2007). Bai and Ng (2005) acknowledged that the assumptions admitted in their 2002 paper were not sufficient to prove their results [2] and that some stronger additional assumptions on the dependence of the idiosyncratic errors had to be imposed. They admit two alternatives to complete the set of assumptions, one being our Assumption C.4. The other alternative would consist of admitting that (for our model and in our notation)

$$\underline{e}_g = \Sigma_g^{1/2} \xi_g R_g^{1/2} \qquad (g = 1, \cdots, G)$$

where $\xi_g = \left[\xi_{g,nt}\right]$ are $T \times N$ matrices of independent variables $\xi_{g,nt}$ with zero mean and uniformly bounded 7th moments and where Σ_g $(N_g \times N_g)$ and R_g $(T \times T)$ are arbitrary (possibly random) positive definite matrices with bounded eigenvalues.

Assumption D allows for some contemporaneous correlation between the factors and the idiosyncratic components. Since

$$||\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathcal{F}_{g,t}^{0} e_{g,nt}||^2 = ||\frac{1}{\sqrt{T}} \sum_{t=1}^{T} F_{0,t}^{0} e_{g,nt}||^2 + ||\frac{1}{\sqrt{T}} \sum_{t=1}^{T} F_{g,t}^{0} e_{g,nt}||^2$$

Assumption D imposes that there exists a positive constant $M < \infty$ that satisfies both

$$E\left\{\left[\frac{1}{N_g} \sum_{n=1}^{N_g} ||\frac{1}{\sqrt{T}} \sum_{t=1}^{T} F_{0,t}^{0} e_{g,nt}||^2\right]\right\} \leq M$$

and

$$E\left\{\left[\frac{1}{N_g} \sum_{n=1}^{N_g} ||\frac{1}{\sqrt{T}} \sum_{t=1}^{T} F_{g,t}^{0} e_{g,nt}||^2\right]\right\} \leq M.$$

Equipped with Assumptions C and D, model (1) is an "approximate factor model", in the sense of the terminology introduced by Chamberlain and Rothschild (1983).

Finally, Assumption E simply rules out the possibility that the relative size of any group of variables becomes insignificant when N, the overall number of variables in the model, tends to infinity. In other words, all N_g are admitted to grow to infinity at the same rate, mantaining the number of groups stable throughout the process.

---

[2] More precisely, they were not sufficient to prove their Lemma 4, which was used to prove their Theorem 2. On this issue, see also Amengual and Watson (2007).

3 Model estimation for a given vector of numbers of factors

In this section we adapt the method of asymptotic principal components to the static approximate model with both global and group-specific factors presented in the previous section. Connor and Korajczyk (1986,1988) discussed the method of principal components for the conventional static representation and studied its asymptotic properties for fixed T. Stock and Watson (1998, 2002a), Bai and Ng (2002) and Amengual and Watson (2005, 2007) considered the asymptotic properties when both $N, T \to \infty$, differing slightly on the set of assumptions regarding the data generating processes of the factors and the idiosyncratic errors.

3.1 Overall and group-by-group mean squared idiosyncratic error

We do not know $r = [r_0\ r_1 \cdots r_g \cdots r_G]'$, the true number of factors. When estimating the model, we will have to allow for some vector of numbers of factors $k = [k_0\ k_1 \cdots k_g \cdots k_G]'$. As usual in this litterature, it will be assumed that it is possible to set a ceiling for each element of r, such that there is a known vector of upper bounds to the numbers of factors $k^{\max} = \left[k_0^{\max}\ k_1^{\max} \cdots k_g^{\max} \cdots k_G^{\max}\right]'$ with $r_g \leq k_g^{\max}$ $(g = 0, 1, \cdots, G)$. Hereafter, the superscript (k) will denote the allowance for k factors when estimating the model.

Given k $(\leq k^{\max})$, estimates of $F_g^{(k)}$ $(T \times k_g)$, $\Lambda_{g0}^{(k)}$ $(N_g \times k_0)$ and $\Lambda_{gg}^{(k)}$ $(N_g \times k_g)$ $(g = 0, 1, \cdots, G)$ are obtained minimizing the overall mean squared idiosyncratic error (MSIE) subject to a set of identifying restrictions. The overall MSIE is a function of the factors and loadings:

$$\upsilon\left(\Lambda^{(k)}, F^{(k)}\right) = \frac{1}{NT} \sum\nolimits_{g=1}^{G} \sum\nolimits_{n=1}^{N_g} \sum\nolimits_{t=1}^{T} \left(X_{g,nt} - \lambda_{g0,n}^{(k)\prime} F_{0,t}^{(k)} - \lambda_{gg,n}^{(k)\prime} F_{g,t}^{(k)}\right)^2$$

In a somewhat more compact notation (with $\mathcal{F}_g^{(k)} = \left[F_0^{(k)}\ F_g^{(k)}\right]$ $(T \times (k_0 + k_g))$ and $\mathcal{L}_g^{(k)} = \left[\Lambda_{g0}^{(k)}\ \Lambda_{gg}^{(k)}\right]$ $(N_g \times (k_0 + k_g))$ for $g = 1, \cdots, G$) the MSIE can be rewritten as

$$\upsilon\left(\Lambda^{(k)}, F^{(k)}\right) = \frac{1}{NT} \sum\nolimits_{g=1}^{G} tr\left[\left(\underline{X}_g - \mathcal{F}_g^{(k)} \mathcal{L}_g^{(k)\prime}\right)' \left(\underline{X}_g - \mathcal{F}_g^{(k)} \mathcal{L}_g^{(k)\prime}\right)\right] \quad (5)$$

or, in a fully compact notation,

$$\upsilon\left(\Lambda^{(k)}, F^{(k)}\right) = \frac{1}{NT} tr\left[\left(X - F^{(k)} \Lambda^{(k)\prime}\right)' \left(X - F^{(k)} \Lambda^{(k)\prime}\right)\right]$$

The overall MSIE can be expressed as a weighted average of the MSIE corresponding to each group of variables:

$$\upsilon(\Lambda^{(k)}, F^{(k)}) = \sum_{g=1}^{G} \frac{N_g}{N} \upsilon_g \left(\mathcal{F}_g^{(k)}, \mathcal{L}_g^{(k)} \right)$$

where $\upsilon_g \left(\mathcal{F}_g^{(k)}, \mathcal{L}_g^{(k)} \right)$ refers to the MSIE of the variables of group g:

$$\upsilon_g \left(\mathcal{F}_g^{(k)}, \mathcal{L}_g^{(k)} \right) = \frac{1}{N_g T} tr \left[\left(\underline{X}_g - \mathcal{F}_g^{(k)} \mathcal{L}_g^{(k)\prime} \right)' \left(\underline{X}_g - \mathcal{F}_g^{(k)} \mathcal{L}_g^{(k)\prime} \right) \right] \qquad (6)$$

At the group level and conditional on a given (full rank) solution for $\mathcal{F}_g^{(k)}$, in the absence of restrictions on $\mathcal{L}_g^{(k)}$, the corresponding estimator for this matrix of loadings is the ordinary least squares estimator

$$\mathcal{L}_g^{(k)} = \underline{X}_g' \mathcal{F}_g^{(k)} \left(\mathcal{F}_g^{(k)\prime} \mathcal{F}_g^{(k)} \right)^{-1}$$

Therefore, without incurring in any loss of information, we may simply denote $\upsilon_g \left(\mathcal{F}_g^{(k)} \right)$ instead of $\upsilon_g \left(\mathcal{F}_g^{(k)}, \mathcal{L}_g^{(k)} \right)$ and express (6) equivalently as

$$\upsilon_g \left(\mathcal{F}_g^{(k)} \right) = \frac{1}{N_g T} tr \left[\underline{X}_g' \left(I - P_g^{(k)} \right) \underline{X}_g \right]$$

where $P_g^{(k)} = \mathcal{F}_g^{(k)} \left(\mathcal{F}_g^{(k)\prime} \mathcal{F}_g^{(k)} \right)^{-1} \mathcal{F}_g^{(k)\prime}$ is the matrix of orthogonal projection into the space generated by the columns of $\mathcal{F}_g^{(k)}$. In the same vein, we will denote the overall MSIE simply by $\upsilon \left(F^{(k)} \right)$ instead of $\upsilon \left(\Lambda^{(k)}, F^{(k)} \right)$.

3.2 The principal components estimator

For a given vector k, with $\ddot{k} = \sum_{g=0}^{G} k_g$, and taking into account the zero restrictions imposed on $\Lambda^{(k)}$, a set of $\left(\ddot{k}_S k_0 + \sum_{g=0}^{G} k_g^2 \right)$ identifying restrictions is required in order to obtain a single optimal solution to the minimization problem of $\upsilon \left(\Lambda^{(k)}, F^{(k)} \right)$. As in the conventional model with only global factors, we will consider a partial identification by explicitly imposing the following set of $\left(\ddot{k}_S k_0 + \sum_{g=0}^{G} k_g \left(k_g + 1 \right) / 2 \right)$ restrictions[3]:

$$\frac{1}{T} \mathcal{F}_g^{(k)\prime} \mathcal{F}_g^{(k)} = I \qquad (g = 1, \cdots, G) \qquad (7)$$

[3] In the conventional model with N variables, T time periods and k estimated factors, to achieve exact identification a set of k^2 restrictions is required. In practice, typically only $k(k+1)/2$ restrictions are explicitly considered. That means that the principal component estimator $\tilde{F}^{(k)}$ is only defined up to an orthogonal transformation of matrix $Q^{(k)}$. In other words, any estimator $\mathring{F}^{(k)} = \tilde{F}^{(k)} Q^{(k)}$, with $Q^{(k)}$ orthogonal, also has the same optimal MSIE.

or, equivalently,

$$\frac{1}{T}F_g^{(k)\prime}F_g^{(k)} = I \qquad (g = 0, 1, \cdots, G) \tag{8}$$

and

$$\frac{1}{T}F_0^{(k)\prime}F_g^{(k)} = 0 \qquad (g = 1, \cdots, G) \tag{9}$$

Consistently with Assumption A.1, these identification restrictions do not rule out non-zero correlation between group-specific factors associated with different groups of variables.

We will denote the estimators of factors and loadings obtained when using the set of identifying restrictions (7) (or (8)-(9)) by $\tilde{F}^{(k)}$and $\tilde{\Lambda}^{(k)}$ (and $\tilde{F}_0^{(k)}$, $\tilde{\Lambda}_0^{(k)}$, $\tilde{\Lambda}_{g0}^{(k)}$, $\tilde{\Lambda}_{gg}^{(k)}$, $\tilde{\mathcal{F}}_g^{(k)}$, $\tilde{\mathcal{L}}_g^{(k)}$, etc.). It is straightforward to derive the first order conditions for the problem of minimizing the overall MSIE (5) subject to restrictions (8)-(9):

$$\tilde{\Lambda}_0^{(k)} = \frac{1}{T}X'\tilde{F}_0^{(k)} \qquad \text{and} \qquad \tilde{\Lambda}_{gg}^{(k)} = \frac{1}{T}\underline{X}_g'\tilde{F}_g^{(k)} \qquad (g = 1, \cdots, G) \tag{10}$$

$$\begin{aligned} &\left[\frac{1}{NT}\left(X - \tilde{F}_S^{(k)}\tilde{\Lambda}_S^{(k)\prime}\right)\left(X - \tilde{F}_S^{(k)}\tilde{\Lambda}_S^{(k)\prime}\right)'\right]\frac{1}{\sqrt{T}}\tilde{F}_0^{(k)} = \\ &= \frac{1}{\sqrt{T}}\tilde{F}_0^{(k)}\left\{\frac{1}{\sqrt{T}}\tilde{F}_0^{(k)\prime}\left[\frac{1}{NT}\left(X - \tilde{F}_S^{(k)}\tilde{\Lambda}_S^{(k)\prime}\right)\left(X - \tilde{F}_S^{(k)}\tilde{\Lambda}_S^{(k)\prime}\right)'\right]\frac{1}{\sqrt{T}}\tilde{F}_0^{(k)}\right\} \end{aligned} \tag{11}$$

and, for $g = 1, \cdots, G$,

$$\begin{aligned} &\left[\frac{1}{N_gT}\left(\underline{X}_g - \tilde{F}_0^{(k)}\tilde{\Lambda}_{g0}^{(k)\prime}\right)\left(\underline{X}_g - \tilde{F}_0^{(k)}\tilde{\Lambda}_{g0}^{(k)\prime}\right)'\right]\frac{1}{\sqrt{T}}\tilde{F}_g^{(k)} = \\ &= \frac{1}{\sqrt{T}}\tilde{F}_g^{(k)}\left\{\frac{1}{\sqrt{T}}\tilde{F}_g^{(k)\prime}\left[\frac{1}{N_gT}\left(\underline{X}_g - \tilde{F}_{g0}^{(k)}\tilde{\Lambda}_{g0}^{(k)\prime}\right)\left(\underline{X}_g - \tilde{F}_{g0}^{(k)}\tilde{\Lambda}_{g0}^{(k)\prime}\right)'\right]\frac{1}{\sqrt{T}}\tilde{F}_g^{(k)}\right\} \end{aligned} \tag{12}$$

The equations (10)-(11)-(12), together with the identifying restrictions (8)-(9) tell us that the columns of $T^{-1/2}\tilde{F}_0^{(k)}$ and of $T^{-1/2}\tilde{F}_g^{(k)}$ $(g = 1, \cdots, G)$ are (orthogonal and normalized) eigenvectors, respectively of

$$\frac{1}{NT}\left(X - \tilde{F}_S^{(k)}\tilde{\Lambda}_S^{(k)\prime}\right)\left(X - \tilde{F}_S^{(k)}\tilde{\Lambda}_S^{(k)\prime}\right)'$$

and

$$\frac{1}{N_gT}\left(\underline{X}_g - \tilde{F}_{g0}^{(k)}\tilde{\Lambda}_{g0}^{(k)\prime}\right)\left(\underline{X}_g - \tilde{F}_{g0}^{(k)}\tilde{\Lambda}_{g0}^{(k)\prime}\right)' \qquad (g = 1, \cdots, G)$$

associated with the largest eigenvalues of these two matrices. In Appendix III, we suggest an algorithm to compute the estimates.

Because the model is only partially identified, alternative solutions $\breve{F}^{(k)}$ associated with the same overall MSIE minimum can be generated by block diagonal orthogonal transformations of $\tilde{F}^{(k)}$:

$$\breve{F}^{(k)} = \tilde{F}^{(k)}Q^{(k)}$$

where

$$Q^{(k)} = diag\left(Q_0^{(k)}, Q_1^{(k)}, \cdots, Q_g^{(k)}, \cdots, Q_G^{(k)}\right)$$

such that $Q_g^{(k)}$ are orthogonal matrices $(k_g \times k_g)$ $(g = 0, 1, \cdots, G)$.

3.3 A convenient auxiliary estimator

In close relationship with $\tilde{\mathcal{F}}_g^{(k)}$, we may define the rescaled group-by-group auxiliary estimators

$$\hat{\mathcal{F}}_g^{(k)} = \frac{1}{N_g T}\underline{X}_g \underline{X}_g' \tilde{\mathcal{F}}_g^{(k)} \qquad (g = 1, \cdots, G)$$

Equivalently, $\hat{\mathcal{F}}_g^{(k)} = \left[\hat{\mathcal{F}}_{g0}^{(k)} \ \hat{F}_g^{(k)}\right]$ $(g = 1, \cdots, G)$ with

$$\hat{\mathcal{F}}_{g0}^{(k)} = \frac{1}{N_g T}\underline{X}_g \underline{X}_g' \tilde{F}_0^{(k)}$$

and

$$\hat{F}_g^{(k)} = \frac{1}{N_g T}\underline{X}_g \underline{X}_g' \tilde{F}_g^{(k)}$$

Let us also define the "mean auxiliary estimator" of the global factors

$$\hat{F}_0^{(k)} = \sum\nolimits_{g=1}^{G} \frac{N_g}{N}\hat{\mathcal{F}}_{g0}^{(k)} = \frac{1}{NT} X X' \tilde{F}_0^{(k)}$$

and the "overall auxiliary estimator"

$$\hat{F}^{(k)} = \left[\hat{F}_0^{(k)\prime} \ \hat{F}_1^{(k)\prime} \cdots \hat{F}_g^{(k)\prime} \cdots \hat{F}_G^{(k)\prime}\right]'$$

Note that the matrices $\hat{\mathcal{F}}_{g0}^{(k)}$, are not identical across the groups of variables because the group-by-group rescaling transformations depend on the matrices of observations $\underline{X}_g$ $(g = 1, \cdots, G)$. Additionally, the matrices $T^{-1}\hat{\mathcal{F}}_g^{(k)\prime}\hat{\mathcal{F}}_g^{(k)}$ $(g = 1, \cdots, G)$ may not be full rank $k_0 + k_g$, although usually in practice they are. However, to take into account the possibility of some $T^{-1}\hat{\mathcal{F}}_g^{(k)\prime}\hat{\mathcal{F}}_g^{(k)}$ being singular, the projection matrix associated with the auxiliary estimators has to be defined using the Moore-Penrose pseudoinverse instead of a regular inverse

$$\hat{P}_g^{(k)} = \frac{1}{T}\hat{\mathcal{F}}_g^{(k)}\left(\frac{1}{T}\hat{\mathcal{F}}_g^{(k)\prime}\hat{\mathcal{F}}_g^{(k)}\right)^{+}\hat{\mathcal{F}}_g^{(k)\prime}$$

This matrix is symmetric and idempotent and the estimated MSIE associated with the auxiliary estimator for group g of variables is given by

$$\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) = \frac{1}{N_g T} tr\left[\underline{X}_g'\left(I - \hat{P}_g^{(k)}\right)\underline{X}_g\right]$$

Unlike in the case of the conventional model with only global factors, in general $\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) \neq \upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right)$ and $\upsilon\left(\hat{F}^{(k)}\right) \neq \upsilon\left(\tilde{F}^{(k)}\right)$. In Appendix IV we show that if $T^{-1}\hat{\mathcal{F}}_g^{(k)\prime}\hat{\mathcal{F}}_g^{(k)}$ is full rank $k_0 + k_g$ (the most common case), than $\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) \leq \upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right)$.

3.4 Consistency

Our first main asymptotical result, Theorem 1, states that the rate of convergence to zero of the time average of the squared deviations between the factors as estimated by the auxilary estimator $\hat{\mathcal{F}}_{g,t}$ and $(k_0 + k_g)$ linear combinations of the $(r_0 + r_g)$ true factors $\mathcal{F}_{g,t}^0$ is $[\min(N_g, T)]^{-1}$. Asymptotically, the linear combinations span: (i) the space generated by the columns of $\mathcal{F}_g^0$ when $k_g \geq r_g$ and $k_0 + k_g \geq r_0 + r_g$; or (ii) a subspace of dimension $\min(k_0; r_0) + k_g$ of that space when $k_g < r_g$ or $k_0 + k_g < r_0 + r_g$. Corollary 1.1 translates the result of Theorem 1 to the overall auxiliary estimator $\hat{F}_t^{(k)}$ when the model is not underspecified in any dimension (i.e. when $k_0 \geq r_0$ and $k_g \geq r_g$ for all groups g). Proofs of Theorem 1 and Corollary 1.1 are presented in Appendices VI and VII, respectively.

Theorem 1: *Suppose that Assumptions A to E hold and let* $k = [k_0\ k_1 \cdots k_g \cdots k_G]'$ *with* $k_0 + k_g > 0$ *for all* $g = 1, \cdots, G$. *There is a finite positive constant* L_1 *such that,* for *every* $g = 1, \cdots, G$,

$$\min(N_g, T)\left(\frac{1}{T}\sum\nolimits_{t=1}^{T}||\hat{\mathcal{F}}_{g,t} - H_g^{(k)\prime}\mathcal{F}_{g,t}^0||^2\right) \leq L_1$$

where $\hat{\mathcal{F}}_{g,t}$ and $\mathcal{F}_{g,t}$ *are the transposed* t*-th rows of* $\hat{\mathcal{F}}_g$ *and* $\mathcal{F}_g$, *respectively, and*

$$H_g^{(k)} = \left(\frac{1}{N_g}\mathcal{L}_g^{0\prime}\mathcal{L}_g^0\right)\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}\right)$$

As $T \to \infty$ *and* $N_g \to \infty$, $H_g^{(k)}$ *converges to a matrix of rank*

$$\min(r_0 + r_g; k_0 + k_g; r_0 + k_g)$$

Corollary 1.1: *Suppose that Assumptions A to E hold. For every* $g = 1, \cdots, G$, *let* $k_0 + k_g > 0$ *and* $H_g^{(k)}$ *be partioned into four blocks* $H_{g,00}^{(k)}$ $(r_0 \times k_0)$, $H_{g,0g}^{(k)}$ $(r_0 \times k_g)$, $H_{g,g0}^{(k)}$ $(r_g \times k_0)$ *and* $H_{g,gg}^{(k)}$ $(r_g \times k_g)$:

$$H_g^{(k)} = \begin{bmatrix} H_{g,00}^{(k)} & H_{g,0g}^{(k)} \\ H_{g,g0}^{(k)} & H_{g,gg}^{(k)} \end{bmatrix}$$

Then there is a finite positive constant L_2 such that

$$\min(N,T)\left(\frac{1}{T}\sum_{t=1}^{T}||\hat{F}_t^{(k)} - H^{(k)\prime}F_t^0||^2\right) \leq L_2$$

where $\hat{F}_t^{(k)}$ and F_t^0 *are the transposed* t*-th rows of* $\hat{F}^{(k)}$ *and* F^0, *respectively, and* $H^{(k)}$ *is the* $(\ddot{r} \times \ddot{k})$ matrix

$$H^{(k)} = \begin{bmatrix} \sum_{g=1}^{G}\left(\frac{N_g}{N}H_{g,00}^{(k)}\right) & H_{1,01}^{(k)} & \cdots & H_{g,0g}^{(k)} & \cdots & H_{G,0G}^{(k)} \\ \frac{N_1}{N}H_{1,10}^{(k)} & H_{1,11}^{(k)} & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ \frac{N_g}{N}H_{g,g0}^{(k)} & 0 & \cdots & H_{g,gg}^{(k)} & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ \frac{N_G}{N}H_{G,G0}^{(k)} & 0 & \cdots & 0 & \cdots & H_{G,GG}^{(k)} \end{bmatrix}$$

If $k_g \geq r_g$ $(g = 0, 1, \cdots, G)$, *then* $H^{(k)}$ *converges to a matrix with full rank* $\ddot{r}$ *as* $T \to \infty$ *and* $N_g \to \infty$ $(g = 1, \cdots, G)$.

In the conventional model with only global factors, the MSIE associated with the principal components estimator $\tilde{F}^{(k)}$ and the MSIE associated with the rescaled estimator $\hat{F}^{(k)}$ are identical. As mentioned in the previous subsection, that does not apply to our model with both global and group-specific factors, at least in the finite sample. The following theorem complements the previous propositions by stating that if $k_g \geq r_g$ and $k_0 + k_g \geq r_0 + r_g$ (i.e. if the both the number of group-specific factors and the total number of factors for group g of variables are not underspecified) then the principal components estimator and the auxiliary estimator both attain the same group MSIE asymptotically and therefore span the same factor space. The proof of Theorem 2 is presented in Appendix VIII.

Theorem 2: *Suppose that Assumptions A to E hold and let* $k = [k_0\ k_1 \cdots k_g \cdots k_G]'$ *with* $k_0 + k_g > 0$ *for all* $g = 1, \cdots, G$. *If* $k_g \geq r_g$ *and* $k_0 + k_g \geq r_0 + r_g$, *then*

$$\left|\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right)\right| \to 0$$

as $N_g \to \infty$ and $T \to \infty$.

4 Estimation of the number of global and group-specific factors

4.1 Two classes of procedures to estimate r

Let

$$GPC_g(k) = \upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) + k_0\psi_0\left(N_g, N, T\right) + k_g\psi_g\left(N_g, N, T\right)$$

We can consistently estimate r, the $(G+1)$ vector of the number of the true global and group specific factors, by minimizing a criterion such as

$$GPC(k) = \sum\nolimits_{g=1}^{G} \frac{N_g}{N} GPC_g(k) =$$

$$= \upsilon\left(\hat{F}^{(k)}\right) + \sum\nolimits_{g=1}^{G} \frac{N_g}{N}\left[k_0\psi_0\left(N_g, N, T\right) + k_g\psi_g\left(N_g, N, T\right)\right]$$

where $\psi_g\left(N_g, N, T\right)$ $(g = 0, 1, \cdots, G)$ are positive penalty functions and where GPC stands for "*Group PC*" criterion, because it is a generalization of the PC criterion suggested by Bai and Ng (2002). Theorem 3 establishes conditions for the penalty functions that ensure the consistency of the estimator of r.

Theorem 3: *Suppose that the Assumptions A to F hold and that* $k^{\max} \geq r$. *Let*

$$\hat{k} = \arg \min_{\substack{0 \leq k \leq k^{\max} \\ k_0 + k_g > 0 \ (g=1,\cdots,G)}} GPC(k)$$

If $\psi_g\left(N_g, N, T\right)$ are *positive functions for all* g, N_g, N *and* T *and if*

(i) $\psi_g\left(N_g, N, T\right) \to 0 \quad (g = 0, 1, \cdots, G)$

(ii) $\min\left(N, T\right)\psi_g\left(N_g, N, T\right) \to \infty \quad (g = 0, 1, \cdots, G)$

(iii) $\min\left(N, T\right)\left[\psi_g\left(N_g, N, T\right) - \psi_0\left(N_g, N, T\right)\right] \to \infty \quad (g = 0, 1, \cdots, G)$

as $N \to \infty$ *and* $T \to \infty$, *then* $\text{Prob}\left(\hat{k} = r\right) \to 1$.

When there are only global factors, as in the conventional model, the criterion above reduces to the PC criterion suggested by Bai and Ng (2002). In this particular case, as mentioned above, the estimated MSIE associated with the auxiliary estimator is always identical to the estimated MSIE associated with the principal components estimator. Thus, if there are only global factors, it is indifferent to use either estimator when computing the criterion. However, in our model with more than one group of variables, the two estimated MSIE in general are not identical and the auxiliary estimator is the only one that ensures consistency of the criterion.

As regards the conditions on the penalty functions, conditions (i) and (ii) are identical to those presented by Bai and Ng (2002) in their Theorem 2. The additional condition (iii) is needed to ensure that any solution with $k_0 < r_0$, $k_g > r_g$ and $k_0 + k_g = r_0 + r_g$ for some group or groups of variables g will be discarded asymptotically. In particular, condition (iii) rules out the choice of reducible representations, at least when N and T are sufficiently large. To illustrate this issue, let $k_{(I)}$ and $k_{(II)}$ be two $(G+1)$ vectors of possible number of factors such that

$$k_{(II)} - k_{(I)} = \left[(-1)\ 1\ \cdots\ 1\right]'$$

$k_{(II)}$ corresponds to a representation with one less global factor and one extra group-specific factor than $k_{(I)}$ for every group of variables. To simplify, let us also assume that the size of the panel is large in all dimensions and that the overall and group-by-group estimated MSIEs are the same for both representations. Thus, we are admitting that the loss of one global factor is exactly compensated by an additional group-specific factor in every group. By Assumption F, the true model is in its irreducible representation and the criterion should penalize the representation associated with $k_{(II)}$ more than that associated with $k_{(I)}$ because the former has a larger total number of factors for the same descriptive power. In our illustration, the more parcimonious representation $k_{(I)}$ will be chosen if

$$GPC\left(k_{(II)}\right) - GPC\left(k_{(I)}\right) = -\psi_0\left(N_g, N, T\right) + \sum\nolimits_{g=1}^{G} \frac{N_g}{N} \psi_g\left(N_g, N, T\right) =$$

$$= \sum\nolimits_{g=1}^{G} \frac{N_g}{N} \left[\psi_g\left(N_g, N, T\right) - \psi_0\left(N_g, N, T\right)\right] > 0$$

Asymptotically, condition (iii) in Theorem 3 is sufficient to rule out any reducible representation of the true model. In finite samples, as our illustration just pointed out, it will be convenient that the penalty function verifies

$$\psi_g\left(N_g, N, T\right) - \psi_0\left(N_g, N, T\right) > 0$$

for all N_g, $N\,(> N_g)$ and T.

Corollary 3.1 extends Theorem 3 to a class of criteria which uses a logarithmic transformation of the MSIE, adapting the IC criterion of Bai and Ng:

$$GIC(k) = \sum\nolimits_{g=1}^{G} \frac{N_g}{N} GIC_g(k)$$

where,

$$GIC_g(k) = \ln\left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right)\right] + k_0\psi_0\left(N_g, N, T\right) + k_g\psi_g\left(N_g, N, T\right) \qquad (g = 1, \cdots, G)$$

Corollary 3.1: *Under the Assumptions A to F, the result in Theorem 3 remains valid when criteria $GIC(k)$ and $GIC_g(k)$ are substituted for $GPC(k)$ and $GPC_g(k)$, respectively.*

Proofs of Theorem 3 and Corollary 3.1 are provided in Appendices X and XI, respectively. They are generalizations of the proofs of Theorem 2 and Corollary 1 presented by Bai and Ng (2002).

4.2 The choice of the penalty function

For the conventional model with only global factors, Bai and Ng (2002) suggested the penalty functions $\tilde{\sigma}^2\varphi(N,T)$ and $\varphi(N,T)$ for the PC and IC criteria, respectively, where $\varphi(N,T)$ is one of following three alternatives:

$$\varphi_{(1)}(N,T) = \frac{N+T}{NT}\ln\left(\frac{NT}{N+T}\right) \tag{13}$$

$$\varphi_{(2)}(N,T) = \frac{N+T}{NT}\ln\left[\min\left(N;T\right)\right] \tag{14}$$

$$\varphi_{(3)}(N,T) = \frac{\ln\left[\min\left(N;T\right)\right]}{\min\left(N;T\right)} \tag{15}$$

and where $\tilde{\sigma}^2$ stands for a consistent estimator of the overall MSIE. In applications, Bai and Ng (2002) proposed to use the estimated overall MSIE for $k^{\max}$ as the scaling factor $\tilde{\sigma}^2$. However, in finite samples quite often the estimated number of factors depend on the choice of $k^{\max}$, which is difficult to rationalize. The scaling of the penalty term in IC criteria is much less critical and that explains why the latter class has been preferred by most practicioners. As regards the choice of the alternative specifications for $\varphi(N,T)$, simulations carried out by Bai and Ng (2002) for the conventional model show that $\varphi_{(3)}(N,T)$ behaves worse than the other two when N and/or T are small.

For the model with both global and group specific factors, the choice of the scaling factors is more critical than in the conventional model with only global factors. The estimated group-specific MSIEs depend on all the elements of vector k (and not only on k_0 and k_g). Unless G is very small, the estimated MSIEs associated with $k^{\max}$ may prove to be a very poor choice of the scaling factors due to the large number of factors typically considered in $k^{\max}$. The GIC criteria have the important advantage of being much less sensitive to the small sample scaling problems of the penalty term and hereafter we will restrict our analysis to this class of criteria.

Given our aim to extend Bai and Ng's criteria to a model with both global and group specific factors, we want to keep the penalty functions specification as close as possible to (13), (14) and (15). The simplest specification could have been

$$\psi_{g(i)}(N_g,N,T) = \varphi_{(i)}(N_g,T) \qquad (i=1,2,3; g=0,1,\cdots,G)$$

with $\varphi_{(i)}(\cdot)$ specified as in (13), (14) or (15), but then condition (iii) of Theorem 3 is not satisfied. However, condition (iii) is fulfilled if the penalty associated with the group-specific factors is kept as suggested

$$\psi_{g(i)}(N_g,N,T) = \varphi_{(i)}(N_g,T) \qquad (i=1,2,3;\ g=1,\cdots,G)$$

and the specification for the global factors is sligthly amended by multiplying it by $(1-c)$ where c is a small positive constant:

$$\psi_{0(i)}(N_g,N,T) = (1-c)\,\varphi_{(i)}(N_g,T) \qquad (i=1,2,3)$$

The multiplicative element $(1 - c)$ ensures that, for each group of variables, the penalty associated with an extra group specific factor is slightly larger than N_g/N times the penalty associated with a global factor, thus fulfilling condition (iii) of Theorem 3.

5 Monte Carlo simulation results

To assess the finite sample performance of the criteria suggested in the previous section, we carried out a Monte Carlo exercise. The large number of dimensions encompassed in our framework makes it extremely demanding to cover all possibilities. Hence, our strategy was to focus on a relatively contained number of cases, which nevertheless provide informative insights regarding the finite sample performance of the criteria (as in Bai and Ng (2002)).

Owing to the difficulty in finding adequate scaling factors for *GPC* criteria, as mentioned above, we confined the simulation to the *GIC* criteria. Moreover, we do not present results for the criteria with penalty function based on $\varphi_{(3)}(.)$. Bai and Ng (2002) found that this specification performs poorly when the number of series and/or time observations is small. This also applies to our model.

As to the number of groups, we investigated two cases, $G = 2, 5$. Regarding the true number of factors, we considered two global factors and two specific factors for every group of variables. Although relatively parsimonious, this seems to be sufficient to assess underestimation or overestimation of the number of factors by the criteria.

The data generating processes of all factors are mutually independent first order autoregressive processes with a 0.5 coefficient and unit variance Gaussian innovation. For the idiosyncratic term, four cases were assessed: *i*) as the base case, a Gaussian white noise with the same variance as the common component; *ii*) a Gaussian first order autoregressive process with a 0.5 coefficient and the same variance as the common component; *iii*) a Gaussian white noise with the variance of the idiosyncratic part twice as large as the variance of the common component; *iv*) a Gaussian first order autoregressive process with a 0.5 coefficient and the variance of the idiosyncratic part twice as large as the variance of the common component.

For the panel size, we considered the various combinations of $T = 60$, 100, 200 and $N_g = 60$, 100. Finally, the sensitivity of the criteria performance to the choice of c was assessed considering several values within a reasonable range, in particular, $c = 0.05$, 0.1, 0.2. Overall, for each G, the criterion was applied to 72 different configurations.

For each configuration, one thousand simulations were performed. We report the proportion of times that the criterion gets it right (that is, zero deviation from the true number of factors), underestimates and overestimates by one the true number of factors[4]. To simplify the comparison across cases with a different number of groups, we present the results for the global factors and for the average of the group-specific factors.

We first discuss the configurations with two groups of variables ($G = 2$). In the base case (see Table 1), the results are quite impressive for both $GIC_{(1)}$ and $GIC_{(2)}$. For $T = 100$ and 200, both criteria estimate correctly the true number of global and group-specific factors in almost 100 per cent of the simulations. For $T = 60$ and $N_g = 100$, although both criteria get it right almost always in terms of the number of global factors, $GIC_{(1)}$ performs slightly better than $GIC_{(2)}$, for all values of c. In the most difficult case, with $T = 60$ and $N_g = 60$, $GIC_{(1)}$ clearly outperforms $GIC_{(2)}$, estimating correctly the true number of global factors by near 100 per cent of the simulations and around 96 per cent in the case of the group-specific factors[5]. In fact, $GIC_{(2)}$ seems to bias downwards, that is, it tends to underestimate the true number of factors. Moreover, the results for $GIC_{(1)}$ are robust to the choice of c. Hence, the analysis of the base case seems to favour the use of $GIC_{(1)}$ instead of $GIC_{(2)}$. This conclusion is reinforced by the analysis of the remaining configurations.

In Tables 2, 3 and 4 we present the results for the cases (*ii*), (*iii*) and (*iv*), mentioned above, respectively. Concerning case (*ii*), the conclusions are qualitatively the same for $T = 100$ and 200. For $T = 60$, $GIC_{(2)}$ performs slightly better than $GIC_{(1)}$, although differences are small (for $c = 0.05$ and 0.10). The main difference against the base case refers to the role of c. While in the base case, the results for $GIC_{(1)}$ remain almost unchanged for different values of c, in this case, the performance of the criterion clearly deteriorates when $c = 0.2$ with $T = 60$. This evidence would suggest avoiding the use of $c = 0.2$ with T small. Regarding cases (*iii*) and (*iv*), there is a deterioration in the behavior of the criteria, namely in case (*iii*). However, both these cases are somehow extreme in the sense that they imply a lot of noise in the data set. In such demanding framework, it is natural that the criteria performance is poorer. Note that $GIC_{(2)}$ is much more affected than $GIC_{(1)}$. Regarding the choice of c for $GIC_{(1)}$, from the analysis of case (*iv*), evidence clearly points to discard $c = 0.2$, favouring somehow $c = 0.1$ over $c = 0.05$. Although the results regarding the choice of c for $GIC_{(1)}$ are more mixed, in case (*iii*) those for $c = 0.1$ are slightly better than the ones for $c = 0.05$.

In order to assess the robustness of the conclusions to a different number of groups, we now discuss the results for $G = 5$ (Table 5). To narrow a bit the number of possible combinations and to ease the overall reading of the results, we discarded some of the combinations. Since the results for $c = 0.05$ and $c = 0.1$ are relatively similar, we only report those for $c = 0.1$ and $c = 0.2$. Moreover, since the results for $T = 200$ do not add much to the conclusions, they were also discarded. The summary results for the five groups configurations

[4] The other situations were discarded to ease the presentation of the results

[5] We also considered another variant where the number of series of one of the groups was set to 30 and the other one to 90 (so that the total number of series remains unchanged at 120). As expected, there is a significant deterioration of the performance of the criteria for the group less represented in the data set with the remaining results almost unchanged.

are presented in Table 5. In qualitative terms, the results do not differ much from those obtained with only two groups. Again, in the base case, $GIC_{(1)}$ does always equal or better than $GIC_{(2)}$, with the percentages of success similar to the ones obtained in the two groups case. In the case of variant (*ii*), considering $c = 0.2$ reveals once again to be a poorer choice for $GIC_{(1)}$ than $c = 0.1$, in particular, when $T = 60$. Concerning cases (*iii*) and (*iv*), there is again a tendency to underestimate the true number of factors, particularly strong in case (*iii*), which is much more severe for $GIC_{(2)}$ than for $GIC_{(1)}$. Overall, the main conclusions remain valid when one increases the number of groups from two to five.

6 Empirical results for the euro area

In this section, we provide an empirical application for the $GIC_{(1)}$ criterion for illustrative purposes[6]. A large set of macroeconomic series were collected for Germany, France, Italy and Spain, which together represent around 80 per cent of euro area GDP. Resorting to the Thomson Financial Datastream database, the panel includes a wide range of variables, namely industrial production and sales, labour market variables, price series, monetary aggregates, business and consumer surveys, etc.[7]. The sample covers the period from January 1991, that is after German reunification, up to December 2006. Hence, we have 192 time observations and a total of 295 series (77 series for Germany, 82 for France, 64 for Italy and 72 for Spain). As usual, data are seasonally adjusted and transformed by taking logs and/or differences when necessary. An outlier adjustment procedure was also performed[8].

In this example, the variables for each country constitute a natural group. Each variable is assumed to be driven by factors common to all countries (which therefore reflect overall euro area developments), the global factors, and by country-specific factors. Following the discussion of the previous section, we computed $GIC_{(1)}$ with $c = 0.1$. According to this criterion (Table 6), two global factors were found. The first seems to reflect euro area real activity behavior. This interpretation is apparent from Figure 1 which displays the quarterly euro area GDP growth and the quarterly average of the first global factor. On the other hand, the second global factor captures euro area consumer prices evolution. This can be seen in Figure 2, which displays the year-on-year euro area inflation along with the twelve-months moving average of the second global factor.

Regarding country-specific factors, Germany presents clearly the large number (six), followed by Italy (four), while both France and Spain have two. There-

[6] All Matlab codes are available from the authors upon request.
[7] See the Appendix XII for the list of series.
[8] The outlier adjustment corresponds to replacing observations of the transfomed series with absolute deviations larger than six times the interquartile range by the median value of the preceding five observations (see, for example Stock and Watson (2005)).

fore, the total number of factors that we end up with for Germany is eight (i.e., two global plus six specific factors), six for Italy and four for France and Spain. Interestingly, these figures are almost the same as the number of factors that one obtains when the $IC_{(1)}$ criterion of Bai and Ng (2002) is applied to each country separately (see Table 6).

7 Conclusions

In this paper, within the framework of an approximate factor model, we focused on the issue of determining the number of global and group-specific factors when the number of variables in each group is at least moderately large and the partition of variables into groups is exogenously set. The consistency of the principal components estimator was discussed for given numbers of global and group-specific factors. An extension of the well known Bai and Ng (2002) criteria was proposed and a proof of the consistency was provided. Furthermore, the corresponding finite sample behaviour was investigated through a Monte Carlo simulation exercise. From the set of results, we found that the criterion $GIC_{(1)}$ performs better in finite samples, under most possible conditions. However, the performance of the criteria, including $GIC_{(1)}$, may present a significant negative bias when the idiosyncratic component dominates the data generating process of the variables and the number of time observations is not large.

References

[1] Amengual, D. and M. Watson (2005): "Consistent Estimation of the Number of Dynamic Factors in a Large N and T Panel - Detailed Appendix", mimeo.

[2] Amengual, D. and M. Watson (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.

[3] Bai, J. and S. Ng (2002): "Determining the Number of Factors in Approximate Factor Models", *Econometrica*, 70, 191-221.

[4] Bai, J. and S. Ng (2005): "Determining the Number of Factors in Approximate Factor Models, Errata", mimeo.

[5] Bernanke, B. and J. Boivin (2003) "Monetary policy in a data-rich environment", *Journal of Monetary Economics*, 50, 525-546.

[6] Bernanke, B., J. Boivin and P. Eliasz (2005) "Measuring the effects of monetary policy:a factor-augmented vector autoregressive (FAVAR) approach", *Quarterly Journal of Economics*, 120, no.1, 387-422.

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

[8] Canova, F., M. Ciccarelli and E. Ortega (2007) "Similarities and convergence in G-7 cycles", *Journal of Monetary Economics*, 54, 850-878.

[9] Chamberlain, G. and M. Rothschild (1983): "Arbitrage, Factor Structure and Mean Variance Analysis in Large Asset Markets", *Econometrica*, 51, 1305-1324.

[10] Connor, G. and R. Korajzcyk (1986): "Performance Measurement with the Arbitrage Pricing Theory: A New Framework for Analysis", *Journal of Financial Economics*, 15, 373-394.

[11] Connor, G. and R. Korajzcyk (1988): "Risk and Return in an Equilibrium APT: Application to a New Test Methodology", *Journal of Financial Economics*, 21, 255-289.

[12] Favero, C., M. Marcellino and F. Neglia (2005) "Principal components at work: the empirical analysis of monetary policy with large data sets", *Journal of Applied Econometrics*, 20, no.5, 603-620.

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

[14] Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2001) "Coincident and leading indicators for the euro area", *The Economic Journal*, 111, 62-85.

[15] Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2004) "The generalized dynamic factor model: consistency and rates", *Journal of Econometrics*, 119, 231-255.

[16] 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, 830-840.

[17] Forni, M. and L. Reichlin (1998), "Let's get real: a dynamic factor analytical approach to disaggregated business cycle", *Review of Economic Studies*, 65, 453-474.

[18] Giannone, D., L. Reichlin and L. Sala (2006) "VARs, common factors and the empirical validation of equillibrium business cycle models", *Journal of Econometrics*, 127, no. 1, 257-279.

[19] 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*, North-Holland, Amesterdam, Ch. 19.

[20] Geweke, J. and G. Zhou (1996) "Measuring the price of the Arbitrage Price Theory", *Review of Financial Studies*, 9, no. 2, 557-587.

[21] Gregory, A. W. and A. C. Head (1999) "Common and country-specific fluctuations in productivity, investment, and the current account", *Journal of Monetary Economics*, 44, no. 3, 423-451.

[22] Gregory, A. W., A. C. Head and J. Raynauld (1997) "Measuring world business cycles", *International Economic Review*, vol. 38, no 3, 677-701.

[23] Kose, M. A., C. Otrok and C. H. Whiteman (2003) "International business cycles: world, region, and country-specific factors", *American Economic Review*, vol. 93, no. 4, 1216-1239.

[24] Magnus, J. and H. Neudecker (1988): *Matrix Differential Calculus With Applications in Statistics and Econometrics*, Chichester: John Wiley & Sons.

[25] Marcellino, M., J. Stock and M. Watson (2003) "Macroeconomic forecasting in the euro area: country specific versus area-wide information", *European Economic Review*, vol. 47, no. 1, 1-18.

[26] Norrbin, S. C. and D. E. Schlagenhauf (1996) "The role of international factors in the business cycle: a multicountry study", *Journal of International Economics*, 40, 85-104.

[27] Quah, D. and T. J. Sargent (1993) "A dynamic index model for large cross-sections" in J. Stock and M. Watson (eds.), *Business cycles, indicators and forecasting*, University of Chicago Press, Ch. 7.

[28] Roll, R. and S. Ross (1980) "An empirical investigation of the arbitrage pricing theory", *Journal of Finance*, 35, 1073-1103.

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

[30] Sargent, T. J. and C. A. Sims (1977) "Business cycle modelling without pretending to have too much a-priori economic theory" in C. Sims *et al.* (eds.), *New methods in business cycle research*, Federal Reserve Bank of Minneapolis

[31] Stock, J. and M. Watson (1998): "Diffusion Indexes", NBER Working Paper 6702.

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

[33] Stock, J. and M. Watson (2002a): "Forecasting Using Principal Components From a Large Number of Predictors", *Journal of the American Statistical Association*, 97, 1167-1179.

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

[35] Stock, J. and M. Watson (2005) "Implications of dynamic factor models for VAR analysis", mimeo.

[36] Stock, J. and M. Watson (2007) "Forecasting in dynamic factor models subject to structural stability", mimeo.

Appendix

I - Proof that Ω is positive definite under Assumptions B.2 and E

By Assumptions B.2 and E, as $N_g, T \to \infty$

$$\frac{1}{N}\Lambda^{0\prime}\Lambda^{0} = \begin{bmatrix} \sum_{g=1}^{G}\left(\frac{1}{N}\Lambda_{g0}^{0\prime}\Lambda_{g0}^{0}\right) & \frac{1}{N}\Lambda_{10}^{0\prime}\Lambda_{11}^{0} & \cdots & \frac{1}{N}\Lambda_{g0}^{0\prime}\Lambda_{gg}^{0} & \cdots & \frac{1}{N}\Lambda_{G0}^{0\prime}\Lambda_{GG}^{0} \\ \frac{1}{N}\Lambda_{11}^{0\prime}\Lambda_{10}^{0} & \frac{1}{N}\Lambda_{11}^{0\prime}\Lambda_{11}^{0} & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ \frac{1}{N}\Lambda_{gg}^{0\prime}\Lambda_{g0}^{0} & 0 & \cdots & \frac{1}{N}\Lambda_{gg}^{0\prime}\Lambda_{gg}^{0} & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ \frac{1}{N}\Lambda_{GG}^{0\prime}\Lambda_{G0}^{0} & 0 & \cdots & 0 & \cdots & \frac{1}{N}\Lambda_{GG}^{0\prime}\Lambda_{GG}^{0} \end{bmatrix}$$

$$\to \begin{bmatrix} \sum_{g=1}^{G}\left(\pi_g\Omega_{g,00}\right) & \pi_1\Omega_{1,01} & \cdots & \pi_g\Omega_{g,0g} & \cdots & \pi_G\Omega_{G,0G} \\ \pi_1\Omega_{1,01}' & \pi_1\Omega_{1,11} & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ \pi_g\Omega_{g,0g}' & 0 & \cdots & \pi_g\Omega_{g,gg} & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ \pi_G\Omega_{G,0G}' & 0 & \cdots & 0 & \cdots & \pi_G\Omega_{G,GG} \end{bmatrix} = \Omega$$

where the matrices $\Omega_{g,00}$, $\Omega_{g,0g}$ and $\Omega_{g,gg}$ have dimensions $(r_0 \times r_0)$, $(r_0 \times r_g)$ and $(r_g \times r_g)$, respectively, and are such that

$$\Omega_g = \begin{bmatrix} \Omega_{g,00} & \Omega_{g,0g} \\ \Omega_{g,0g}' & \Omega_{g,gg} \end{bmatrix} \qquad (g = 1, \cdots, G)$$

Being positive semi-definite by construction, Ω will be positive definite if and only if it is non-singular. By Assumption E $\pi_g > 0$ $(g = 1, \cdots, G)$ and, by Assumption B.2, the matrices $\Omega_{g,00}$, $\Omega_{g,gg}$ and $\Omega_{g,00} - \Omega_{g,0g}\Omega_{g,gg}^{-1}\Omega_{g,0g}'$ are all non-singular $(g = 1, \cdots, G)$.

Let

$$\Omega_{SS} = \begin{bmatrix} \pi_1 \Omega_{1,11} & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & \pi_g \Omega_{g,gg} & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & \pi_G \Omega_{G,GG} \end{bmatrix}$$

and

$$\Omega_{0S} = \begin{bmatrix} \pi_1 \Omega_{1,01} & \cdots & \pi_g \Omega_{g,0g} & \cdots & \pi_G \Omega_{G,0G} \end{bmatrix}$$

Note that Ω_{SS} is positive definite.

We may write

$$\Omega^{-1} = \begin{bmatrix} \Omega^{00} & \Omega^{0S} \\ \Omega^{0S\prime} & \Omega^{SS} \end{bmatrix}$$

where

$$\Omega^{0S} = -\Omega^{00}\Omega_{0S}\Omega_{SS}^{-1}$$
$$\Omega^{SS} = \Omega_{SS}^{-1} + \Omega_{SS}^{-1}\Omega_{0S}'\Omega^{00}\Omega_{0S}\Omega_{SS}^{-1}$$

and

$$\Omega^{00} = \left[\sum\nolimits_{g=1}^{G} \left(\pi_g \Omega_{g,00}\right) - \sum\nolimits_{g=1}^{G} \left(\pi_g \Omega_{g,0g}\Omega_{g,gg}^{-1}\Omega_{g,0g}'\right)\right]^{-1} =$$
$$= \left[\sum\nolimits_{g=1}^{G} \pi_g \left(\Omega_{g,00} - \Omega_{g,0g}\Omega_{g,gg}^{-1}\Omega_{g,0g}'\right)\right]^{-1}$$

The latter inverse exists because $\sum_{g=1}^{G} \pi_g \left(\Omega_{g,00} - \Omega_{g,0g}\Omega_{g,gg}^{-1}\Omega_{g,0g}'\right)$, being the sum of positive definite matrices, is also definite positive. □

II - Solution of the optimization problem

$$\max_{Z} tr\left(Z'AZ\right) \qquad \text{subject to} \qquad Z'Z = I \qquad \text{and} \qquad B'Z = 0$$

where Z, A and B are $(m \times n)$, $(m \times m)$ and $(m \times q)$ matrices, respectively, with $n < m$, A symmetric semi-definite positive and B such that $B'B = I$.

Let $B_\perp$ be the $(m \times (m-q))$ orthogonal complement of B, with $B_\perp' B = 0$ and $B_\perp' B_\perp = I$. The restriction $B'Z = 0$ implies that the columns of the optimal solution Z^* are linear combinations of the columns of $B_\perp$ and therefore the solution of the above optimization problem may be expressed as $Z^* = B_\perp W^*$, where W^* $((m-q) \times n)$ is the solution of

$$\max_{W} tr\left(W'B_\perp' A B_\perp W\right) \qquad \text{subject to} \qquad W'W = I$$

By Theorem 11.6 in Magnus and Neudecker (1988, p.205), the columns of W^* are the normalized and orthogonal eigenvectors of $B_\perp' A B_\perp$ associated with the n largest eigenvalues of the latter matrix. □

III - An algorithm to compute the principal components estimates (given k)[9]

(a) First, consider only global factors (i.e., $k_1 = \cdots = k_G = 0$) and estimate the conventional model by computing the eigenvalues and eigenvectors of $(NT)^{-1} XX'$. The resulting estimated global factors and global loadings will be denoted by $\tilde{F}_{g0}^{(k)}(0)$ and $\tilde{\Lambda}_{g0}^{(k)}(0)$. In general, the number within brackets will represent the iteration number.

(b) In iteration i (≥ 1), use the estimated global factors and loadings of the previous iteration in equations (12) and (8)-(9). For each g, the latter equations, after substituting $\tilde{F}_{g0}^{(k)}(i-1)$ for $\tilde{F}_{g0}^{(k)}$ and $\tilde{\Lambda}_{g0}^{(k)}(i-1)$ for $\tilde{\Lambda}_{g0}^{(k)}$, correspond to the first order conditions for the maximization of

$$tr\left[F_g^{(k)\prime}\left(\underline{X}_g - \tilde{F}_0^{(k)}(i-1)\tilde{\Lambda}_{g0}^{(k)}(i-1)'\right)\left(\underline{X}_g - \tilde{F}_0^{(k)}(i-1)\tilde{\Lambda}_{g0}^{(k)}(i-1)'\right)' F_g^{(k)}\right]$$

with respect to $F_g^{(k)}$ and subject to

$$\frac{1}{T}F_g^{(k)\prime}F_g^{(k)} = I \qquad \text{and} \qquad \tilde{F}_0^{(k)}(i-1)'F_g^{(k)} = 0$$

Let $\tilde{F}_0^{(k)}(i-1)_\perp$ be a $T \times (T-k_0)$ matrix such that

$$\frac{1}{T}\tilde{F}_0^{(k)}(i-1)'_\perp \tilde{F}_0^{(k)}(i-1) = 0$$

and

$$\frac{1}{T}\tilde{F}_0^{(k)}(i-1)'_\perp \tilde{F}_0^{(k)}(i-1)_\perp = I$$

The solution of the above maximization problem is (see Appendix II)

$$\tilde{F}_g^{(k)}(i) = \tilde{F}_0^{(k)}(i-1)_\perp \tilde{W}^{(k)}(i-1)$$

where $\tilde{W}^{(k)}(i-1)$ is a $(T-k_0) \times k_g$ matrix with columns that are the orthogonal and normalized eigenvectors of the $(T-k_0) \times (T-k_0)$ matrix

$$\tilde{F}_0^{(k)}(i-1)'_\perp \underline{X}_g \underline{X}_g' \tilde{F}_0^{(k)}(i-1)_\perp$$

[9] A similar algorithm can be envisaged if the following alternative (partial) identification restrictions (on the loadings, instead of on the factors) are considered:

$$\frac{1}{N}\Lambda_0^{(k)\prime}\Lambda_0^{(k)} = I$$

$$\frac{1}{N_g}\Lambda_{gg}^{(k)\prime}\Lambda_{gg}^{(k)} = I \qquad (g = 1, \cdots, G)$$

$$\frac{1}{N_g}\Lambda_{g0}^{(k)\prime}\Lambda_{gg}^{(k)} = 0 \qquad (g = 1, \cdots, G)$$

associated with its largest k_g eigenvalues. From (10), given $F_g^{(k)}(i)$ the corresponding estimate $\tilde{\Lambda}_{gg}^{(k)}(i)$ is simply:

$$\tilde{\Lambda}_{gg}^{(k)}(i) = \frac{1}{T}\underline{X}'_g \tilde{F}_g^{(k)}(i)$$

(c) Compute $\left(X - \tilde{F}_S^{(k)}(i)\tilde{\Lambda}_S^{(k)\prime}(i)\right)\left(X - \tilde{F}_S^{(k)}(i)\tilde{\Lambda}_S^{(k)\prime}(i)\right)'$. From equation (11), the columns of $T^{-1/2}\tilde{F}_0^{(k)}(i)$ are the k_0 orthogonal and normalized eigenvectors of the latter matrix associated with its largest k_0 eigenvalues and

$$\tilde{\Lambda}_0^{(k)}(i) = \frac{1}{T}X'\tilde{F}_0^{(k)}(i)$$

(d) Steps (b) and (c) must be repeated until convergence is achieved.□

IV - Proof that if $T^{-1}\hat{\mathcal{F}}_g^{(k)\prime}\hat{\mathcal{F}}_g^{(k)}$ is full rank, then $\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) \leq \upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right)$

If $T^{-1}\hat{\mathcal{F}}_g^{(k)\prime}\hat{\mathcal{F}}_g^{(k)}$ is positive definite, it is straightforward to show that

$$\tilde{P}_g^{(k)}\left(\frac{1}{N_g T}\underline{X}_g\underline{X}'_g\right)\hat{P}_g^{(k)} = \tilde{P}_g^{(k)}\left(\frac{1}{N_g T}\underline{X}_g\underline{X}'_g\right)$$

Using the latter equality,

$$\upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) = tr\left[\frac{1}{N_g T}\underline{X}'_g\left(\hat{P}_g^{(k)} - \tilde{P}_g^{(k)}\right)\underline{X}_g\right] =$$

$$= tr\left[\left(\hat{P}_g^{(k)} - \tilde{P}_g^{(k)}\right)\left(\frac{1}{N_g T}\underline{X}_g\underline{X}'_g\right)\right] =$$

$$= tr\left[\left(\hat{P}_g^{(k)} - \tilde{P}_g^{(k)}\right)\left(\frac{1}{N_g T}\underline{X}_g\underline{X}'_g\right)\left(\hat{P}_g^{(k)} - \tilde{P}_g^{(k)}\right)\right] \geq 0$$

because the trace of a semi-definite positive matrix is non-negative. □

V - Lemmas

In order to prove the consistency of the principal components estimator, first it is convenient to present and prove several lemmas. Hereafter, rk(A) denotes the rank of matrix A.

<u>Lemma A.1:</u> *Under Assumptions A, B and C, for every $g = 1, \cdots, G$ there exists some $M_1 < \infty$ such that for all N_g and T,*

(i) $T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}\left(\gamma_g(s,t)\right)^2 \leq M_1$;

(ii) $E\left[T^{-1}\sum_{t=1}^{T}||N_g^{-1/2}e'_{g,t}\mathcal{L}_g^0||^2\right] =$

$$= E\left[T^{-1}\sum_{t=1}^{T}||N_g^{-1/2}\sum_{n=1}^{N_g} e_{g,nt}\mathcal{L}_{g,n}^{0}||^2\right] \leq M_1;$$

(iii) $$E\left[T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}\left(N_g^{-1}\sum_{n=1}^{N_g} X_{g,ns}X_{g,nt}\right)^2\right] \leq M_1.$$

Lemma A.2: *Under Assumption C.4, for every* $g = 1, \cdots, G$ *and for any idempotent matrix* P_g $(T \times T)$ *of rank* $k_0 + k_g > 0$ *there exists some* $M_2 < \infty$ *such that for all* N_g *and* T,

$$\frac{\min(N_g, T)}{N_g T} tr\left(\underline{e}_g' P_g \underline{e}_g\right) \leq M_2$$

Lemma A.3: *Let* $\mathcal{F}_g^{(k)}$ *be any matrix* $(T \times m)$ *such that* $T^{-1}\mathcal{F}_g^{(k)\prime}\mathcal{F}_g^{(k)} = I$ *and let*

$$A_{N_g,T} =$$

$$= \left(\frac{1}{\sqrt{T}}\mathcal{F}_g^{(k)}\right)'\left(\frac{1}{N_g T}\underline{X}_g\underline{X}_g'\right)\left(\frac{1}{\sqrt{T}}\mathcal{F}_g^{(k)}\right) - \left(\frac{1}{T}\mathcal{F}_g^{0\prime}\mathcal{F}_g^{(k)}\right)'\Omega_g\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\mathcal{F}_g^{(k)}\right)$$

Under Assumptions A to D, $||A_{N_g,T}|| \to 0$ *and* $\left|tr\left(A_{N_g,T}\right)\right| \to 0$ *as* $N_g \to \infty$ *and* $T \to \infty$.

Lemma A.4: *Consider the* $(G+1) \times 1$ *vectors of non-negative integers*

$$k = [k_0\ k_1 \cdots k_g \cdots k_G]' \text{ and } r = [r_0\ r_1 \cdots r_g \cdots r_G]'$$

For every $g = 1, \cdots, G$, *let*[10]

$$\underset{((r_0 \times r_g)\times(k_0 \times k_g))}{J_g^{(k)}} = \begin{bmatrix} \underset{(r_0 \times k_0)}{J_{00}^{(k)}} & \underset{(r_0 \times k_g)}{J_{g,0g}^{(k)}} \\ \underset{(r_g \times k_0)}{J_{g,g0}^{(k)}} & \underset{(r_g \times k_g)}{J_{g,gg}^{(k)}} \end{bmatrix}$$

and

$$\underset{((T-r_0-r_g)\times(k_0+k_g))}{R_g^{(k)}} = \begin{bmatrix} \underset{((T-r_0-r_g)\times k_0)}{R_{g,0}^{(k)}} & \underset{((T-r_0-r_g)\times k_g)}{R_{g,g}^{(k)}} \end{bmatrix}$$

Also let

$$\chi_g\left(J_g^{(k)}\right) = tr\left(J_g^{(k)\prime}\Omega_g J_g^{(k)}\right)$$

and

$$\chi\left(J_1^{(k)}, \cdots, J_g^{(k)}, \cdots, J_G^{(k)}\right) = \sum_{g=1}^{G}\pi_g\chi_g\left(J_g\right)$$

where Ω_g *are* $(r_0 + r_g) \times (r_0 + r_g)$ *matrices and* $0 < \pi_g \leq 1$ with $\sum_{g=1}^{G}\pi_g = 1$. *Denote by*

$$\left\{\tilde{J}_1^{(k)}, \cdots, \tilde{J}_g^{(k)}, \cdots, \tilde{J}_G^{(k)}; \tilde{R}_1^{(k)}, \cdots, \tilde{R}_g^{(k)}, \cdots, \tilde{R}_G^{(k)}\right\}$$

[10] Note that the block $J_{00}^{(k)}$ does not depend on g.

the optimal solution of problem

$$\max_{\left\{J_g^{(k)},R_g^{(k)}\right\}_g} \chi\left(J_1^{(k)},\cdots,J_g^{(k)},\cdots,J_G^{(k)}\right)$$

subject to:

$$J_g^{(k)\prime}J_g^{(k)}+R_g^{(k)\prime}R_g^{(k)}=I \qquad (g=1,\cdots,G)$$

Under the assumption that Ω_g *is symmetric positive definite* $(g=1,\cdots,G)$*,*

(i) for every g*,* $\operatorname{rk}\left(\tilde{J}_g^{(k)}\right)=\min\left(r_0+r_g;k_0+k_g;r_0+k_g\right)$*;*

(ii) $\chi_g\left(J_g^{(k)}\right)\le tr\left(\Omega_g\right)$ *for all feasible solutions* $J_g^{(k)}$ *and*

if $k_0+k_g\ge r_0+r_g$ *and* $k_g\ge r_g$*, then* $\chi_g\left(\tilde{J}_g^{(k)}\right)=tr\left(\Omega_g\right)$*;*

(iii) if $k_0\ge r_0$ *and* $k_g\ge r_g$*, then* $\tilde{J}_{g,0g}^{(k)}=0$ *and* $\tilde{J}_{g,g0}^{(k)}=0$.

Lemma A.5: *Suppose that Assumptions A to E hold. Let* $k=[k_0\ k_1\cdots k_g\cdots k_G]'$ *and* $k^{\triangleright}=\left[k_0^{\triangleright}\ k_1^{\triangleright}\cdots k_g^{\triangleright}\cdots k_G^{\triangleright}\right]'$ *be* $((G+1)\times 1)$ *vectors of non-negative integers. As* $N_g\to\infty$ $(g=1,\cdots,G)$ *and* $T\to\infty$*, for each* g*:*

(i) There exists a $((r_0+r_g)\times(k_0+k_g))$ *matrix*

$$\underset{((r_0\times r_g)\times(k_0\times k_g))}{\mathring{J}_g^{(k)}}=\begin{bmatrix}\underset{(r_0\times k_0)}{\mathring{J}_{00}^{(k)}} & \underset{(r_0\times k_g)}{\mathring{J}_{g,0g}^{(k)}}\\ \underset{(r_g\times k_0)}{\mathring{J}_{g,g0}^{(k)}} & \underset{(r_g\times k_g)}{\mathring{J}_{g,gg}^{(k)}}\end{bmatrix}$$

with rank $\min\left(r_0+r_g;k_0+k_g;r_0+k_g\right)$ *such that*

$$\left\|\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}-\mathring{J}_g^{(k)}\right\|\to 0$$

(ii) for any $\mathcal{F}_g^{(k)}$ $(T\times(k_0+k_g))$ *such that* $T^{-1}\mathcal{F}_g^{(k)\prime}\mathcal{F}_g^{(k)}=I$*,*

$$\operatorname{plim}\left[\upsilon_g\left(\mathcal{F}_g^{(k)}\right)-\upsilon_g\left(\tilde{\mathcal{F}}_g^{(r)}\right)\right]\ge 0$$

(iii) If $k_0+k_g\ge k_0^{\triangleright}+k_g^{\triangleright}\ge r_0+r_g$ *and* $k_g\ge k_g^{\triangleright}\ge r_g$*, then*

$$\left|\upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right)-\upsilon_g\left(\tilde{\mathcal{F}}_g^{(r)}\right)\right|\to 0$$

and

$$\left|\upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right)-\upsilon_g\left(\tilde{\mathcal{F}}_g^{(k^{\triangleright})}\right)\right|\to 0$$

(iv) If $k_0\ge r_0$ *and* $k_g\ge r_g$*, then* $\mathring{J}_{g,0g}^{(k)}=0$ *and* $\mathring{J}_{g,g0}^{(k)}=0$*;*

Lemma A.1 is a direct adaptation of Lemma 1 in Bai and Ng (2002). Only notation changes are required to adapt their proof. Lemma A.2 is a special case of Result 6 in Amengual and Watson (2005)[11].

[11] As P_g is idempotent of rank k_0+k_g, there exists $\mathcal{F}_g$ $(T\times(k_0+k_g))$ such that $T^{-1}\mathcal{F}_g'\mathcal{F}_g=I$ and $P_g=T^{-1}\mathcal{F}_g\mathcal{F}_g'$. Thus, Lemma 2 is proved by applying Result 6 of Amengual and Watson when their $m\to\infty$. Note that Assumption C.4 corresponds to their Assumption (A.6) with $m\to\infty$.

Proof of Lemma A.3

Taking into account (4), we have

$$||A_{N_g,T}|| = ||(a) + (b) + (c) + (c)'|| \leq ||(a)|| + ||(b)|| + 2||(c)||$$

and

$$\left|tr(A_{N_g,T})\right| = |tr(a) + tr(b) + 2 \times tr(c)| \leq |tr(a)| + tr(b) + 2\,|tr(c)|$$

where

$$(a) = \left(\frac{1}{T}\mathcal{F}_g^{0\prime}\mathcal{F}_g^{(k)}\right)'\left(\frac{1}{N_g}\mathcal{L}_g^{0\prime}\mathcal{L}_g^{0}\right)\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\mathcal{F}_g^{(k)}\right) - \left(\frac{1}{T}\mathcal{F}_g^{0\prime}\mathcal{F}_g^{(k)}\right)'\Omega_g\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\mathcal{F}_g^{(k)}\right)$$

$$(b) = \frac{1}{T^2N_g}\mathcal{F}_g^{(k)\prime}\underline{e}_g\underline{e}_g'\mathcal{F}_g^{(k)}$$

$$(c) = \left(\frac{1}{T\sqrt{N_g}}\mathcal{F}_g^{(k)\prime}\underline{e}_g\right)\left(\frac{1}{\sqrt{N_g}}\mathcal{L}_g^{0}\right)\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\mathcal{F}_g^{(k)}\right)$$

$||(a)|| \to 0$ follows directly from Assumption B.2. Using the fact that $|tr\,(A)\,| \leq m||A||$ for any $m \times m$ matrix A,

$$|tr(a)| \leq (k_0 + k_g)\,||(a)|| \to 0$$

For (b), we have

$$||(b)|| = ||\left(\frac{1}{TN_g^{1/2}}\mathcal{F}_g^{(k)\prime}\underline{e}_g\right)\left(\frac{1}{TN_g^{1/2}}\mathcal{F}_g^{(k)\prime}\underline{e}_g\right)'|| \leq ||\left(\frac{1}{TN_g^{1/2}}\mathcal{F}_g^{(k)\prime}\underline{e}_g\right)||^2 =$$

$$= tr(b) = \frac{1}{N_gT}tr\left(\underline{e}_g'P_g^{(k)}\underline{e}_g\right) \to 0 \qquad (16)$$

by Lemma A.2, , with $P_g^{(k)} = T^{-1}\mathcal{F}_g^{(k)}\mathcal{F}_g^{(k)\prime}$. Now for (c),

$$||(c)||^2 \leq ||N_g^{-1/2}\mathcal{L}_g^0||^2.||T^{-1/2}\mathcal{F}_g^0||^2.||T^{-1/2}\mathcal{F}_g^{(k)}||.tr(b) =$$

$$= O_p\,(1)\,O_p\,(1)\,(k_0 + k_g)\,O_p\left([\min{(N_g, T)}]^{-1}\right) \to 0$$

by (16), Assumptions A.1 and B.2 and because $T^{-1}\mathcal{F}_g^{(k)\prime}\mathcal{F}_g^{(k)} = I$. Thus $||(c)|| \to 0$ and

$$|tr(c)| \leq (k_0 + k_g)\,||(c)|| \to 0$$

Proof of Lemma A.4

First note that any optimal solution is not unique because if $\left\{\tilde{J}_g^{(k)}, \tilde{R}_g^{(k)}\right\}_{g=1,\cdots,G}$ is optimal, then $\left\{\tilde{J}_g^{*(k)}, \tilde{R}_g^{*(k)}\right\}_{g=1,\cdots,G}$ is also optimal with

$$\tilde{J}_g^{*(k)} = \tilde{J}_g^{(k)} \begin{bmatrix} Q_{00} & 0 \\ 0 & Q_{gg} \end{bmatrix} \tag{17}$$

and

$$\tilde{R}_g^{*(k)} = \tilde{R}_g^{(k)} \begin{bmatrix} Q_{00} & 0 \\ 0 & Q_{gg} \end{bmatrix}$$

for any set of (orthogonal) matrices $\{Q_{00}, Q_{11}, \cdots, Q_{GG}\}$, with Q_{gg} $(r_g \times r_g)$, such that $Q_{gg}' Q_{gg} = I$ $(g = 1, \cdots, G)$. In any case, $\tilde{J}_g^{(k)}$ and $\tilde{J}_g^{*(k)}$ share the same rank. Also note that, for any feasible solution, $J_g^{(k)\prime} J_g^{(k)}$ and $J_g^{(k)} J_g^{(k)\prime}$ will have all eigenvalues in the range $[0;1]$. Thus,

$$\chi_g\left(J_g^{(k)}\right) = tr\left(\Omega_g J_g^{(k)} J_g^{(k)\prime}\right) \leq \sum\nolimits_{i=1}^{\min(k_0+k_g; r_0+r_g)} \mu_{g,i} \leq tr\left(\Omega_g\right)$$

where $\mu_{g,1} \geq \mu_{g,2} \geq \cdots \geq \mu_{g,r_0+r_g}$ are the $r_0 + r_g$ (positive) eigenvalues of Ω_g.

We will prove the Lemma by considering in turn all the possible cases:

(1) $k_0 \geq r_0$ and, for every g, $k_g \geq r_g$

$\left\{\tilde{J}_g^{(k)}, \tilde{R}_g^{(k)}\right\}_{g=1,\cdots,G}$ with

$$\tilde{J}_g^{(k)} = \begin{bmatrix} \tilde{J}_{00}^{(k)} & \tilde{J}_{g,0g}^{(k)} \\ \tilde{J}_{g,g0}^{(k)} & \tilde{J}_{g,gg}^{(k)} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} I_{r_0} & 0_{r_0 \times (k_0 - r_0)} \end{bmatrix} & 0 \\ 0 & \begin{bmatrix} I_{r_g} & 0_{r_g \times (k_g - r_g)} \end{bmatrix} \end{bmatrix} \tag{18}$$

(block diagonal and full rank $r_0 + r_g$) and

$$\tilde{R}_g^{(k)} = \begin{bmatrix} \tilde{R}_{g,0}^{(k)} & \tilde{R}_{g,g}^{(k)} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 0 \; W_{g,0}^{(k)} \end{bmatrix} & \begin{bmatrix} 0 \; W_{g,g}^{(k)} \end{bmatrix} \end{bmatrix}$$

$(g = 1, \cdots, G)$ is an optimal solution, for any matrices $W_{g,0}^{(k)}$ $((T - r_0 - r_g) \times (k_0 - r_0))$ and $W_{g,g}^{(k)}$ $((T - r_0 - r_g) \times (k_g - r_g))$ such that $W_{g,0}^{(k)\prime} W_{g,0}^{(k)} = I$, $W_{g,g}^{(k)\prime} W_{g,g}^{(k)} = I$ and $W_{g,0}^{(k)\prime} W_{g,g}^{(k)} = 0$. Indeed, it is an feasible solution and for every g

$$\chi_g\left(\tilde{J}_g^{(k)}\right) = tr\left(\Omega_g\right)$$

(2) $k_0 < r_0$ and, for every g, $k_0 + k_g \geq r_0 + r_g$

$\left\{\tilde{J}_g^{(k)}, \tilde{R}_g^{(k)}\right\}_{g=1,\cdots,G}$ with

$$\tilde{J}_g^{(k)} = \begin{bmatrix} \underset{(r_0 \times k_0)}{J_{00}^{(k)}} & \underset{(r_0 \times k_g)}{J_{g,0g}^{(k)}} \\ \underset{(r_g \times k_0)}{J_{g,g0}^{(k)}} & \underset{(r_g \times k_g)}{J_{g,gg}^{(k)}} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} I_{k_0} \\ 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \\ I_{r_0 - k_0} & 0 & 0 \end{bmatrix} \\ 0 & \begin{bmatrix} 0 & I_{r_g} & 0 \end{bmatrix} \end{bmatrix}$$

(full rank $r_0 + r_g$) and

$$\tilde{R}_g^{(k)} = \begin{bmatrix} 0 & \begin{bmatrix} 0 \; 0 \; W_{g,g}^{(k)} \end{bmatrix} \end{bmatrix}$$

is an optimal solution, for any matrix $W_{g,g}^{(k)}$ $((T - r_0 - r_g) \times (k_0 + k_g - r_0 - r_g))$ such that $W_{g,g}^{(k)\prime} W_{g,g}^{(k)} = I$. As in the previous case,

$$\chi_g \left(\tilde{J}_g^{(k)} \right) = tr \left(\Omega_g \right)$$

for every g.

(3) $k_0 \geq r_0$ and $k_g < r_g$ for some g

Without any loss of generality, let us suppose that $k_g < r_g$ for $g = 1, \cdots, \bar{G}$ and that $k_g \geq r_g$ for $g = \bar{G} + 1, \cdots, G$. We will determine the rank of the optimal solution in this case by comparing it with the optimal solution in the relevant "benchmark case" for which

$$\bar{k} = \begin{bmatrix} k_0 \; r_1 \cdots r_{\bar{G}} \; k_{\bar{G}+1} \cdots k_G \end{bmatrix}'$$

is substituted for $k = \begin{bmatrix} k_0 \; k_1 \cdots k_{\bar{G}} \; k_{\bar{G}+1} \cdots k_G \end{bmatrix}'$. For this benchmark case, the optimal solution $\left\{ \tilde{J}_g^{(\bar{k})}, \tilde{R}_g^{(\bar{k})} \right\}_{g=1,\cdots,G}$ is determined as in case (1). Thus, for every g, $\tilde{J}_g^{(\bar{k})}$ is full rank $r_0 + r_g$. Also, for every g, $\chi_g \left(\tilde{J}_g^{(\bar{k})} \right)$ attains its upper bound $tr \left(\Omega_g \right)$.

By construction, we have that

$$\chi_g \left(\tilde{J}_g^{(\bar{k})} \right) - \chi_g \left(\tilde{J}_g^{(k)} \right) \geq 0 \qquad (g = 1, \cdots, G)$$

and

$$\chi \left(\tilde{J}_1^{(\bar{k})}, \cdots, \tilde{J}_G^{(\bar{k})} \right) - \chi \left(\tilde{J}_1^{(k)}, \cdots, \tilde{J}_G^{(k)} \right) =$$
$$= \sum_{g=1}^{\bar{G}} \frac{N_g}{N} \left[\chi_g \left(\tilde{J}_g^{(\bar{k})} \right) - \chi_g \left(\tilde{J}_g^{(k)} \right) \right] \geq 0$$

An optimal solution for k corresponds to any choice $\left\{ \tilde{J}_g^{(k)}, \tilde{R}_g^{(k)} \right\}_{g=1,\cdots,G}$ that minimizes the latter difference while still complying with the problem constraints

$$J_g^{(k)\prime} J_g^{(k)} + R_g^{(k)\prime} R_g^{(k)} = I \qquad (g = 1, \cdots, G)$$

Given an optimal solution for $\bar{k}$, a candidate for the optimal solution for k is the following: (i) for $g > \bar{G}$, $\tilde{J}_g^{(k)} = \tilde{J}_g^{(\bar{k})}$ and $\tilde{R}_g^{(k)} = \tilde{R}_g^{(\bar{k})}$, ensuring that for these groups of variables $\chi_g \left(\tilde{J}_g^{(k)} \right) = \chi_g \left(\tilde{J}_g^{(\bar{k})} \right) = tr \left(\Omega_g \right)$; (ii) for $g \leq \bar{G}$, delete the

$r_g - k_g$ columns of $\left[\tilde{J}^{(\bar{k})\prime}_{g,0g} \; \tilde{J}^{(\bar{k})\prime}_{g,gg}\right]'$ that have the smaller effect on $\chi_g\left(\tilde{J}^{(\bar{k})}_g\right)$, i.e. make[12]

$$\tilde{J}^{(k)}_g = \begin{bmatrix} \tilde{J}^{(\bar{k})}_{00} & 0 \\ 0 & \tilde{J}^{(\bar{k})}_{g,gg} \end{bmatrix} \begin{bmatrix} I_{k_0} & 0 \\ 0 & S^{(k)}_g \end{bmatrix} = \begin{bmatrix} \tilde{J}^{(\bar{k})}_{00} & 0 \\ 0 & \left(\tilde{J}^{(\bar{k})}_{g,gg} \bar{S}^{(k)}_g\right) \end{bmatrix}$$

and, correspondingly,

$$\tilde{R}^{(k)}_g = \left[\tilde{R}^{(\bar{k})}_{g,0} \;\; 0\right] \begin{bmatrix} I_{k_0} & 0 \\ 0 & \bar{S}^{(k)}_g \end{bmatrix} = \left[\tilde{R}^{(\bar{k})}_{g,0} \;\; 0\right]$$

where $\bar{S}^{(k)}_g$ is a $(r_g \times k_g)$ matrix whose columns are columns of the identity matrix of order r_g with $\bar{S}^{(k)\prime}_g \bar{S}^{(k)}_g = I$. Note that any solution that affects $\tilde{J}^{(\bar{k})}_{00}$ cannot be better than the candidate solution because it affects $\chi_s\left(\tilde{J}^{(\bar{k})}_s\right)$ for all $s \neq g$, thereby (in general) worsening the optimal value of the objective function without addressing the necessity to decrease the number of columns of $\tilde{J}^{(\bar{k})}_{g,gg}$ from r_g to k_g. The same argument rules out changing $\tilde{J}^{(\bar{k})}_g$ or $\tilde{R}^{(\bar{k})}_g$ for any $g > \bar{G}$. Hence, the candidate solution is optimal and

$$\text{(i) for } g \leq \bar{G}, \;\text{rk}\left(\tilde{J}^{(k)}_g\right) = \text{rk}\left(\tilde{J}^{(\bar{k})}_{00}\right) + \text{rk}\left(\tilde{J}^{(\bar{k})}_{g,gg} \bar{S}^{(k)}_g\right) = r_0 + k_g$$

$$\text{(ii) for } g > \bar{G}, \;\text{rk}\left(\tilde{J}^{(k)}_g\right) = \text{rk}\left(\tilde{J}^{(\bar{k})}_{00}\right) + \text{rk}\left(\tilde{J}^{(\bar{k})}_{g,gg}\right) = r_0 + r_g$$

(4) $k_0 < r_0$ and $k_0 + k_g < r_0 + r_g$ for some g

Again without any loss of generality, let us suppose that $k_0 + k_g < r_0 + r_g$ for $g = 1, \cdots, \bar{G}$ and that $k_0 + k_g \geq r_0 + r_g$ for $g = \bar{G} + 1, \cdots, G$. The relevant "benchmark case" is now

$$\mathring{k} = \left[k_0 \;\; (r_1 + r_0 - k_0) \cdots (r_{\bar{G}} + r_0 - k_0) \;\; k_{\bar{G}+1} \cdots k_G\right]'$$

and the associated optimal solution $\left\{\tilde{J}^{(\mathring{k})}_g, \tilde{R}^{(\mathring{k})}_g\right\}_{g=1,\cdots,G}$ is determined as in case (2).

For $g > \bar{G}$, $\tilde{J}^{(k)}_g = \tilde{J}^{(\mathring{k})}_g$ and $\tilde{R}^{(k)}_g = \tilde{R}^{(\mathring{k})}_g$ and, consequently, $\text{rk}\left(\tilde{J}^{(k)}_g\right) = \text{rk}\left(\tilde{J}^{(\mathring{k})}_g\right) = r_0 + r_g$. As for $g \leq \bar{G}$, adapting the argument presented for case (3), let $\mathring{S}^{(k)}_g$ be a $(r_g + r_0 - k_0) \times k_g$ matrix whose columns are k_g columns of the identity matrix of order $(r_g + r_0 - k_0)$, with $\mathring{S}^{(k)\prime}_g \mathring{S}^{(k)}_g = I$. Chose those columns so that the difference

$$\left|\chi_g\left(\tilde{J}^{(\mathring{k})}_g\right) - \chi_g\left(\tilde{J}^{(k)}_g\right)\right|$$

[12] Remark that, by (17) and (18), $\tilde{J}^{(\bar{k})}_{g,g0} = 0$ and $\tilde{J}^{(\bar{k})}_{g,0g} = 0$.

is as small as possible, when setting

$$\tilde{J}_g^{(k)} = \tilde{J}_g^{(\mathring{k})} \begin{bmatrix} I_{k_0} & 0 \\ 0 & \mathring{S}_g^{(k)} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} I_{k_0} \\ 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ I_{r_0-k_0} & 0 \end{bmatrix} \\ 0 & \begin{bmatrix} 0 & I_{r_g} \end{bmatrix} \end{bmatrix} \begin{bmatrix} I_{k_0} & 0 \\ 0 & S_g^{(k)} \end{bmatrix} = \begin{bmatrix} I_{k_0} & 0 \\ 0 & S_g^{(k)} \end{bmatrix}$$

Therefore, for $g \leq \bar{G}$:

$$\text{rk}\left(\tilde{J}_g^{(k)}\right) = \text{rk}\left(I_{k_0}\right) + \text{rk}\left(\mathring{S}_g^{(k)}\right) = k_0 + k_g$$

Proof of Lemma A.5

The principal components estimator was defined as the optimal solution of the problem of minimization of the overall MSIE (5) subject to the restrictions (8)-(9). Let $T > r_0 + r_g$ for every g. Also let $\mathcal{F}_g^{(k)} = \left[F_0^{(k)}\ F_g^{(k)}\right]$ be a $T \times (k_0 + k_g)$ matrix with $T^{-1}\mathcal{F}_g^{(k)\prime}\mathcal{F}_g^{(k)} = I$,

$$\eta_g\left(\mathcal{F}_g^{(k)}\right) = tr\left(\frac{1}{N_g T}\underline{X}_g\underline{X}_g'\right) - \upsilon_g\left(\mathcal{F}_g^{(k)}\right) =$$

$$= tr\left[\left(\frac{1}{\sqrt{T}}\mathcal{F}_g^{(k)}\right)'\left(\frac{1}{N_g T}\underline{X}_g\underline{X}_g'\right)\left(\frac{1}{\sqrt{T}}\mathcal{F}_g^{(k)}\right)\right]$$

and

$$\eta\left(\mathcal{F}_1^{(k)}, \cdots, \mathcal{F}_G^{(k)}\right) = \sum\nolimits_{g=1}^{G} \frac{N_g}{N}\eta_g\left(\mathcal{F}_g^{(k)}\right) = \sum\nolimits_{g=1}^{G} \frac{N_g}{N} tr\left(\frac{1}{N_g T}\underline{X}_g\underline{X}_g'\right) - \upsilon\left(F^{(k)}\right)$$

Hence, the principal components estimator also maximizes $\eta\left(\mathcal{F}_1^{(k)}, \cdots, \mathcal{F}_G^{(k)}\right)$ subject to: (i) $T^{-1}\mathcal{F}_g^{(k)\prime}\mathcal{F}_g^{(k)} = I$ $(g = 1, \cdots, G)$; (ii) identical blocks $F_0^{(k)}$ for all g. Now define $\mathcal{F}_{g\perp}^0$ $(T \times (T - r_0 - r_g))$ such that $T^{-1}\mathcal{F}_g^{0\prime}\mathcal{F}_{g\perp}^0 = 0$ and $||T^{-1}\mathcal{F}_{g\perp}^{0\prime}\mathcal{F}_{g\perp}^0 - I|| \to 0$ as $T \to \infty$. Given Assumption A.1, at least for sufficiently large T the columns of $\left[\mathcal{F}_g^0\ \mathcal{F}_{g\perp}^0\right]$ span the T-dimensional space. Hence, $\mathcal{F}_g^{(k)}$ can be expressed as

$$\mathcal{F}_g^{(k)} = \mathcal{F}_g^0 J_g^{(k)} + \mathcal{F}_{g\perp}^0 R_g^{(k)} \tag{19}$$

for some matrices $J_g^{(k)}$ $((r_0 + r_g) \times (k_0 + k_g))$ and $R_g^{(k)}$ $((T - r_0 - r_g) \times (k_0 + k_g))$. In particular, for large T, there are two matrices $\tilde{J}_g^{(k)}$ and $\tilde{R}_g^{(k)}$ such that

$$\tilde{\mathcal{F}}_g^{(k)} = \mathcal{F}_g^0 \tilde{J}_g^{(k)} + \mathcal{F}_{g\perp}^0 \tilde{R}_g^{(k)} \tag{20}$$

In general for any $\mathcal{F}_g^{(k)}$, from (19) we get

$$\frac{1}{T}\mathcal{F}_g^{(k)\prime}\mathcal{F}_g^{(k)} = I \Longleftrightarrow J_g^{(k)\prime}\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\mathcal{F}_g^0\right)J_g^{(k)} + R_g^{(k)\prime}\left(\frac{1}{T}\mathcal{F}_{g\perp}^{0\prime}\mathcal{F}_{g\perp}^0\right)R_g^{(k)} = I \quad (21)$$

and

$$\frac{1}{T}\mathcal{F}_g^{0\prime}\mathcal{F}_g^{(k)} = \left(\frac{1}{T}\mathcal{F}_g^{0\prime}\mathcal{F}_g^0\right)J_g^{(k)}$$

Thus, by Assumption A.1,

$$||\frac{1}{T}\mathcal{F}_g^{0\prime}\mathcal{F}_g^{(k)} - J_g^{(k)}|| \to 0 \quad (22)$$

$$||J_g^{(k)\prime}\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\mathcal{F}_g^0\right)J_g^{(k)} - J_g^{(k)\prime}J_g^{(k)}|| \to 0 \quad (23)$$

First, note that if we partition $J_g^{(k)}$ into four blocks,

$$\underset{((r_0\times r_g)\times(k_0\times k_g))}{J_g^{(k)}} = \begin{bmatrix} \underset{(r_0\times k_0)}{J_{g,00}^{(k)}} & \underset{(r_0\times k_g)}{J_{g,0g}^{(k)}} \\ \underset{(r_g\times k_0)}{J_{g,g0}^{(k)}} & \underset{(r_g\times k_g)}{J_{g,gg}^{(k)}} \end{bmatrix}$$

by (22) and taking into account that $\mathcal{F}_g^0 = \left[F_0^0\ F_g^0\right]$, we get

$$||\frac{1}{T}F_0^{0\prime}F_0^{(k)} - J_{g,00}^{(k)}|| \to 0$$

As $T^{-1}F_0^{0\prime}F_0^{(k)}$ does not depend on g, $J_{g,00}^{(k)}$ converges to a matrix $J_{00}^{(k)}$ identical for all g. Also note that (4), Assumption E, Lemma A.2 and (22) imply that

$$||\frac{N_g}{N}\eta_g\left(\mathcal{F}_g^{(k)}\right) - \pi_g\chi_g\left(J_g^{(k)}\right)|| \to 0$$

where

$$\chi_g\left(J_g^{(k)}\right) = tr\left(J_g^{(k)\prime}\Omega_g J_g^{(k)}\right)$$

Moreover, for $\chi\left(J_1^{(k)},\cdots,J_G^{(k)}\right) = \sum_{g=1}^G \pi_g\chi_g\left(J_g\right)$, we also have

$$||\eta(\mathcal{F}_1^{(k)},\cdots,\mathcal{F}_G^{(k)}) - \chi\left(J_1^{(k)},\cdots,J_G^{(k)}\right)|| \to 0 \quad (24)$$

Now, let $\left\{\tilde{J}_1^{(k)},\cdots,\tilde{J}_g^{(k)},\cdots,\tilde{J}_G^{(k)};\tilde{R}_1^{(k)},\cdots,\tilde{R}_g^{(k)},\cdots,\tilde{R}_G^{(k)}\right\}$ be an optimal solution of maximizing $\chi\left(J_1^{(k)},\cdots,J_G^{(k)}\right)$ subject to:

$$J_g^{(k)\prime}J_g^{(k)} + R_g^{(k)\prime}R_g^{(k)} = I \qquad (g = 1,\cdots,G)$$

with blocks $J_{00}^{(k)}$ identical for all g. The optimal solution is not unique, because any block diagonal orthogonal transformation of $\tilde{J}_g^{(k)}$ is feasible and attains the same optimal value of the objective function (see first part of the proof of Lemma A.4). Taking this into account, by (24), (21), (22) and (23), for every g there exists a matrix $Q_g = diag\,(Q_{00}, Q_{gg})$ with $Q_{00}\ (k_0 \times k_0)\,,\ Q_{gg}\ (k_g \times k_g)\,,$ $Q'_{00}Q_{00} = I$ and $Q'_{gg}Q_{gg} = I$ such that

$$\mathring{J}_g^{(k)} = \tilde{J}_g^{(k)} Q_g$$

and

$$||\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)} - \mathring{J}_g^{(k)}|| \to 0 \quad (g = 1, \cdots, G)$$

as $N_g \to \infty$ and $T \to \infty$. By Lemma A.4(i),

$$\text{rk}\left(\mathring{J}_g^{(k)}\right) = \min\left(r_0 + r_g; k_0 + k_g; r_0 + k_g\right)$$

Moreover, if $k_g \geq r_g$ and $k_0 + k_g \geq r_0 + r_g$, by Lemma A.4(ii)

$$\chi_g\left(\mathring{J}_g^{(k)}\right) = \chi_g\left(\mathring{J}_g^{(r)}\right) = tr\left(\Omega_g\right) \tag{25}$$

and

$$\text{plim}\,\eta_g\left(\mathcal{F}_g^{(k)}\right) \leq \chi_g\left(\mathring{J}_g^{(r)}\right)$$

for any $\mathcal{F}_g^{(k)}$ such that $T^{-1}\mathcal{F}_g^{(k)\prime}\mathcal{F}_g^{(k)} = I$.

From (25), we get that $\left|\upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\tilde{\mathcal{F}}_g^{(r)}\right)\right| \to 0$. In addition, because (25) is valid for any k such that $k_0 + k_g \geq r_0 + r_g$, it is also verified by any $k^{\triangleright}$ such that

$$k_0 + k_g \geq k_0^{\triangleright} + k_g^{\triangleright} \geq r_0 + r_g$$

implying that $\left|\upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\tilde{\mathcal{F}}_g^{(k^{\triangleright})}\right)\right| \to 0.$

Finally, by Lemma A.4(iii), if $k_0 \geq r_0$ and $k_g \geq r_g$, then $\tilde{J}_{g,0g}^{(k)} = 0$ and $\tilde{J}_{g,g0}^{(k)} = 0$. $\square$

VI - Proof of Theorem 1

The first part of Theorem 1 can be proved following step by step the proof of Bai and Ng's Theorem 1 in, with the necessary adaptations of notation, and therefore the proof will not be repeated here. As regards the asymptotic rank of $H_g^{(k)}$, let $\mathring{H}_g^{(k)}$ be such that

$$||H_g^{(k)} - \mathring{H}_g^{(k)}|| \to 0$$

as $N_g \to \infty$ and $T \to \infty$. From the definition of $H_g^{(k)}$, Assumption B.2 and Lemma A.5(i), $\mathring{H}_g^{(k)} = \Omega_g \mathring{J}_g^{(k)}$ and

$$\text{rk}\left(\mathring{H}_g^{(k)}\right) = \text{rk}\left(\Omega_g \mathring{J}_g^{(k)}\right) = \text{rk}\left(\mathring{J}_g^{(k)}\right) = \min\left(r_0 + r_g; k_0 + k_g; r_0 + k_g\right)$$

□

VII - Proof of Corollary 1.1

By Theorem 1,

$$\min(N_g,T)\left(\frac{1}{T}\sum_{t=1}^{T}||\hat{\mathcal{F}}_{g,t}-H_g^{(k)\prime}\mathcal{F}_{g,t}^{0}||^2\right)=O_p(1)\qquad(g=1,\cdots,G)\quad(26)$$

Multiplying by $\min(N,T)/\min(N_g,T)$ and taking into account Assumption E, we have, for every g,

$$\min(N,T)\left(\frac{1}{T}\sum_{t=1}^{T}||\hat{\mathcal{F}}_{g,t}-H_g^{(k)\prime}\mathcal{F}_{g,t}^{0}||^2\right)=O_p(1)\Longleftrightarrow$$

$$\Longleftrightarrow\min(N,T)\left(\frac{1}{T}\sum_{t=1}^{T}||\hat{\mathcal{F}}_{g0,t}-H_{g,00}^{(k)\prime}F_{0,t}^{0}-H_{g,g0}^{(k)\prime}F_{g,t}^{0}||^2\right)+$$

$$+\min(N,T)\left(\frac{1}{T}\sum_{t=1}^{T}||\hat{F}_{g,t}-H_{g,0g}^{(k)\prime}F_{0,t}^{0}-H_{g,gg}^{(k)\prime}F_{g,t}^{0}||^2\right)=O_p(1)$$

Summing up the latter expressions for $g=1,\cdots,G$, we obtain

$$\min(N,T)\left(\frac{1}{T}\sum_{t=1}^{T}\sum_{g=1}^{G}||\hat{\mathcal{F}}_{g0,t}-H_{g,00}^{(k)\prime}F_{0,t}^{0}-H_{g,g0}^{(k)\prime}F_{g,t}^{0}||^2\right)+$$

$$+\min(N,T)\sum_{g=1}^{G}\left(\frac{1}{T}\sum_{t=1}^{T}||\hat{F}_{g,t}-H_{g,0g}^{(k)\prime}F_{0,t}^{0}-H_{g,gg}^{(k)\prime}F_{g,t}^{0}||^2\right)=O_p(1)$$

Because $N_g/N\leq 1$ $(g=1,\cdots,G)$, the latter expression that

$$\min(N,T)\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{g=1}^{G}||\frac{N_g}{N}\hat{\mathcal{F}}_{g0,t}-\left(\frac{N_g}{N}H_{g,00}^{(k)}\right)'F_{0,t}^{0}-\left(\frac{N_g}{N}H_{g,g0}^{(k)}\right)'F_{g,t}^{0}||^2\right]+$$

$$+\min(N,T)\sum_{g=1}^{G}\left(\frac{1}{T}\sum_{t=1}^{T}||\hat{F}_{g,t}-H_{g,0g}^{(k)\prime}F_{0,t}^{0}-H_{g,gg}^{(k)\prime}F_{g,t}^{0}||^2\right)=O_p(1)\Longrightarrow$$

(because $\hat{F}_0^{(k)}=\sum_{g=1}^{G}(N_g/N)\,\hat{\mathcal{F}}_{g0}^{(k)}$)

$$\min(N,T)\left[\frac{1}{T}\sum_{t=1}^{T}||\hat{F}_{0,t}^{(k)}-\left(\sum_{g=1}^{G}\frac{N_g}{N}H_{g,00}^{(k)}\right)'F_{0,t}^{0}-\sum_{g=1}^{G}\left(\frac{N_g}{N}H_{g,g0}^{(k)\prime}F_{g,t}^{0}\right)||^2\right]+$$

$$+\min(N,T)\sum_{g=1}^{G}\left(\frac{1}{T}\sum_{t=1}^{T}||\hat{F}_{g,t}-H_{g,0g}^{(k)\prime}F_{0,t}^{0}-H_{g,gg}^{(k)\prime}F_{g,t}^{0}||^2\right)=O_p(1)\Longleftrightarrow$$

$$\Longleftrightarrow\min(N,T)\left(\frac{1}{T}\sum_{t=1}^{T}||\hat{F}_t-H^{(k)\prime}F_t^{0}||^2\right)=O_p(1)$$

Now consider $k_0 \geq r_0$ and $k_g \geq r_g$ for every g. Let us denote the limit of $H^{(k)}$ by $\mathring{H}^{(k)}$. By Assumptions B.2 and E and by Lemma A.5(i,iv),

$$\mathring{H}^{(k)} = \begin{bmatrix} \left(\sum_{g=1}^{G} \pi_g \Omega_{g,00} \mathring{J}_{00}^{(k)}\right) & \Omega_{1,01} \mathring{J}_{1,11}^{(k)} & \cdots & \Omega_{g,0g} \mathring{J}_{g,gg}^{(k)} & \cdots & \Omega_{G,0G} \mathring{J}_{G,GG}^{(k)} \\ \pi_1 \Omega_{1,01}' \mathring{J}_{00}^{(k)} & \Omega_{1,11} \mathring{J}_{1,11}^{(k)} & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \pi_g \Omega_{g,0g}' \mathring{J}_{00}^{(k)} & 0 & \cdots & \Omega_{g,gg} \mathring{J}_{g,gg}^{(k)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \pi_G \Omega_{G,0G}' \mathring{J}_{00}^{(k)} & 0 & \cdots & 0 & \cdots & \Omega_{G,GG} \mathring{J}_{G,GG}^{(k)} \end{bmatrix}$$

$$= \Omega \Pi^{-1} \mathring{J}^{(k)}$$

where $\Omega_{g,00}$, $\Omega_{g,0g}$, $\Omega_{g,gg}$ and Ω are as defined in Appendix I, $\mathring{J}_{00}^{(k)}$ and $\mathring{J}_{g,gg}^{(k)}$ are as defined in Lemma A.5,

$$\Pi = diag\left(I_{r_0}, \pi_1 I_{r_1}, \cdots, \pi_g I_{r_g}, \cdots, \pi_G I_{r_G}\right)$$

and

$$\mathring{J}^{(k)} = diag\left(\mathring{J}_{00}^{(k)}, \mathring{J}_{1,11}^{(k)}, \cdots, \mathring{J}_{g,gg}^{(k)}, \cdots, \mathring{J}_{G,GG}^{(k)}\right)$$

$\mathring{H}^{(k)}$ has full rank $\ddot{r} = \sum_{g=0}^{G} r_g$ because Π and Ω are non-singular matrices $\ddot{r} \times \ddot{r}$ (for the latter matrix, see Appendix I) and, by Lemma A.5(i,iv), $\mathring{J}^{(k)}$ has full rank $\ddot{r}$. $\square$

VIII - Proof of Theorem 2

(A) Let $\rho_g^{(k)} = \text{rk}\left(T^{-1} \hat{\mathcal{F}}_g^{(k)\prime} \hat{\mathcal{F}}_g^{(k)}\right)$. The rank $\rho_g^{(k)}$ depends on N_g and T, but we will make this dependence implicit to simplify the notation. First we will show that for sufficiently large N_g and T

$$\min\left(r_0 + r_g; k_0 + k_g; r_0 + k_g\right) \leq \rho_g^{(k)} \leq k_0 + k_g \tag{27}$$

The upper bound results directly from the number of columns of $\hat{\mathcal{F}}_g^{(k)}$ being $k_0 + k_g$. As regards the lower bound, note that

$$\frac{1}{T} \tilde{\mathcal{F}}_g^{(k)\prime} \hat{\mathcal{F}}_g^{(k)} = \left(\frac{1}{\sqrt{T}} \tilde{\mathcal{F}}_g^{(k)}\right)' \left(\frac{1}{N_g T} \underline{X}_g \underline{X}_g'\right) \left(\frac{1}{\sqrt{T}} \tilde{\mathcal{F}}_g^{(k)}\right)$$

For the matrix $\mathring{J}_g^{(k)}$ defined in Lemma A.5(i),

$$||\frac{1}{T} \tilde{\mathcal{F}}_g^{(k)\prime} \hat{\mathcal{F}}_g^{(k)} - \mathring{J}_g^{(k)\prime} \Omega_g \mathring{J}_g^{(k)}|| \leq ||\frac{1}{T} \tilde{\mathcal{F}}_g^{(k)\prime} \hat{\mathcal{F}}_g^{(k)} - \left(\frac{1}{T} \mathcal{F}_g^{0\prime} \tilde{\mathcal{F}}_g^{(k)}\right)' \Omega_g \left(\frac{1}{T} \mathcal{F}_g^{0\prime} \tilde{\mathcal{F}}_g^{(k)}\right) ||+$$

$$+|| \left(\frac{1}{T} \mathcal{F}_g^{0\prime} \tilde{\mathcal{F}}_g^{(k)}\right)' \Omega_g \left(\frac{1}{T} \mathcal{F}_g^{0\prime} \tilde{\mathcal{F}}_g^{(k)}\right) - \mathring{J}_g^{(k)\prime} \Omega_g \mathring{J}_g^{(k)}||$$

The first term on the right hand side converges to zero by Lemma A.3. As for the second term, note that

$$||\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}\right)'\Omega_g\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}\right)-\mathring{J}_g^{(k)\prime}\Omega_g\mathring{J}_g^{(k)}||=$$

$$=||\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}-\mathring{J}_g^{(k)}\right)'\Omega_g\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}-\mathring{J}_g^{(k)}\right)+$$

$$+\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}\right)'\Omega_g\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}-\mathring{J}_g^{(k)}\right)+$$

$$+\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}-\mathring{J}_g^{(k)}\right)'\Omega_g\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}\right)||\leq$$

$$\leq||\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}-\mathring{J}_g^{(k)}\right)'\Omega_g\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}-\mathring{J}_g^{(k)}\right)||+$$

$$+2||\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}\right)'\Omega_g\left(\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}-\mathring{J}_g^{(k)}\right)||\leq$$

$$\leq||\Omega_g||.||\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}-\mathring{J}_g^{(k)}||^2+2.||\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}||.||\Omega_g||.||\frac{1}{T}\mathcal{F}_g^{0\prime}\tilde{\mathcal{F}}_g^{(k)}-\mathring{J}_g^{(k)}||$$

which goes to zero by Assumption B.2 and Lemma A.5(i). Thus

$$||\frac{1}{T}\tilde{\mathcal{F}}_g^{(k)\prime}\hat{\mathcal{F}}_g^{(k)}-\mathring{J}_g^{(k)\prime}\Omega_g\mathring{J}_g^{(k)}||\to 0$$

Therefore, for sufficiently large N_g and T, the rank of $T^{-1}\tilde{\mathcal{F}}_g^{(k)\prime}\hat{\mathcal{F}}_g^{(k)}$ is not smaller than the rank of $\mathring{J}_g^{(k)\prime}\Omega_g\mathring{J}_g^{(k)}$. We use the fact that for any sequence of positive semi-definite matrices $\{A_n\}$ such that $||A_n-B||\to 0$, there exists $\bar{n}$ such that $\mathrm{rk}\,(A_n)\geq \mathrm{rk}\,(B)$ for all $n>\bar{n}$. But

$$\mathrm{rk}\left(\mathring{J}_g^{(k)\prime}\Omega_g\mathring{J}_g^{(k)}\right)=\min\left(r_0+r_g;k_0+k_g;r_0+k_g\right)$$

by Lemma A.5(i) and by Assumption B.2. The lower bound for $\rho_g^{(k)}$ in (27) follows directly from the lower bound on $\mathrm{rk}\left(T^{-1}\tilde{\mathcal{F}}_g^{(k)\prime}\hat{\mathcal{F}}_g^{(k)}\right)$ and the fact that, by construction, $T^{-1}\tilde{\mathcal{F}}_g^{(k)\prime}\tilde{\mathcal{F}}_g^{(k)}$ is the identity matrix of order k_0+k_g.

(B) For suficiently large N_g and T, by (27), if $k_g\geq r_g$ and $k_0+k_g=r_0+r_g$, matrix $T^{-1}\hat{\mathcal{F}}_g^{(k)\prime}\hat{\mathcal{F}}_g^{(k)}$ will be positive definite. From Appendix IV, we know that in this case

$$\upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right)-\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right)=\upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right)-\upsilon_g\left(\check{\mathcal{F}}_g^{(k)}\right)\geq 0$$

(C) Now let $k_g \geq r_g$ and $k_0 + k_g > r_0 + r_g$. By (27), for sufficiently large N_g and T, matrix $T^{-1}\hat{\mathcal{F}}_g^{(k)\prime}\hat{\mathcal{F}}_g^{(k)}$ may be singular. Let $S_g^{(k)}$ be a $(k_0 + k_g) \times \rho_g^{(k)}$ matrix that selects $\rho_g^{(k)}$ linear independent columns of $\hat{\mathcal{F}}_g^{(k)}$. That is, the columns of $S_g^{(k)}$ are $\rho_g^{(k)}$ columns of the identity matrix of order $k_0 + k_g$, implying that $S_g^{(k)\prime}S_g^{(k)}$ is the identity matrix of order $\rho_g^{(k)}$. Let

$$\hat{\mathcal{F}}_{(S)g}^{(k)} = \hat{\mathcal{F}}_g^{(k)}S_g^{(k)} \qquad \text{and} \qquad \tilde{\mathcal{F}}_{(S)g}^{(k)} = \tilde{\mathcal{F}}_g^{(k)}S_g^{(k)}$$

Note that

$$\upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) = \left[\upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\tilde{\mathcal{F}}_{(S)g}^{(k)}\right)\right] +$$

$$+\left[\upsilon_g\left(\tilde{\mathcal{F}}_{(S)g}^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_{(S)g}^{(k)}\right)\right] + \left[\upsilon_g\left(\hat{\mathcal{F}}_{(S)g}^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right)\right]$$

The third term of the right hand side is zero because, by construction, $\upsilon_g\left(\hat{\mathcal{F}}_{(S)g}^{(k)}\right) = \upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right)$. By Lemma A.4(ii) and by Lemma A.5(iii), we also know that the first term converges to zero when $k_g \geq r_g$ and $k_0 + k_g \geq r_0 + r_g$. As regards the second term, for sufficiently large N_g and T, it is non-negative following the same argument as above for the case $k_0 + k_g = r_0 + r_g$, but with $\hat{\mathcal{F}}_{(S)g}^{(k)}$ and $\tilde{\mathcal{F}}_{(S)g}^{(k)}$ instead of $\hat{\mathcal{F}}_g^{(k)}$ and $\tilde{\mathcal{F}}_g^{(k)}$, respectively. All in all, in this Part 2, we conclude that if $k_g \geq r_g$ and $k_0 + k_g \geq r_0 + r_g$, then

$$\text{plim}\left[\upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right)\right] \geq 0$$

(D) Continue to admit that $k_g \geq r_g$ and $k_0 + k_g > r_0 + r_g$ and let $\check{\mathcal{F}}_g^{(k)} = \hat{\mathcal{F}}_{(S)g}^{(k)}\left(T^{-1}\hat{\mathcal{F}}_{(S)g}^{(k)\prime}\hat{\mathcal{F}}_{(S)g}^{(k)}\right)^{-1/2}$. Given that $\upsilon_g\left(\check{\mathcal{F}}_g^{(k)}\right) = \upsilon_g\left(\hat{\mathcal{F}}_{(S)g}^{(k)}\right) = \upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right)$, to complete the proof of Theorem 2 we will now show that

$$\text{plim}\,\upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right) \leq \text{plim}\,\upsilon_g\left(\check{\mathcal{F}}_g^{(k)}\right) \tag{28}$$

Let $\left[\vec{\mathcal{F}}_g\ \vec{W}\right]$ be an optimal solution of the problem

$$\max_{\{\mathcal{F}_g, W\}} \eta_g\left(\mathcal{F}_g\right) = tr\left[\frac{1}{T}\mathcal{F}_g'\left(\frac{1}{N_gT}\underline{X}_g\underline{X}_g'\right)\mathcal{F}_g\right]$$

subject to

$$\begin{bmatrix}\mathcal{F}_g' \\ W'\end{bmatrix}\left[\mathcal{F}_g \quad W\right] = I_{k_0+k_g}$$

where $\mathcal{F}_g$ and W_g are matrices $T \times (r_0 + r_g)$ and $T \times (k_0 + k_g - r_0 - r_g)$, respectively. Let $V_g^{(k)}$ be any $T \times (k_0 + k_g - r_0 - r_g)$ matrix such that $T^{-1}V_g^{(k)\prime}\check{\mathcal{F}}_g^{(k)} =$

0 and $T^{-1}V_g^{(k)\prime}V_g^{(k)} = I$. $\left[\check{\mathcal{F}}_g^{(k)}\ V_g^{(k)}\right]$ is a feasible solution of the above maximization problem. Thus, for all N_g and T, $\eta_g\left(\vec{\mathcal{F}}_g\right) \geq \eta_g\left(\check{\mathcal{F}}_g^{(k)}\right)$, implying that $\upsilon_g\left(\vec{\mathcal{F}}_g\right) \leq \upsilon_g\left(\check{\mathcal{F}}_g^{(k)}\right)$ and

$$\text{plim}\,\upsilon_g\left(\vec{\mathcal{F}}_g\right) \leq \text{plim}\,\upsilon_g\left(\check{\mathcal{F}}_g^{(k)}\right)$$

But by Lemma A.5(ii),

$$\text{plim}\,\upsilon_g\left(\tilde{\mathcal{F}}_g^{(r)}\right) \leq \text{plim}\,\upsilon_g\left(\vec{\mathcal{F}}_g^{(k_0,k_g)}\right)$$

and, by Lemma A.5(iii),

$$\text{plim}\,\upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right) = \text{plim}\,\upsilon_g\left(\tilde{\mathcal{F}}_g^{(r)}\right)$$

Therefore,

$$\text{plim}\,\upsilon_g\left(\tilde{\mathcal{F}}_g^{(k)}\right) \leq \text{plim}\,\upsilon_g\left(\check{\mathcal{F}}_g^{(k)}\right)$$

□

IX - More lemmas

In addition to the lemmas presented in Appendix V, to prove Theorem 3 and Corollary 3.1 we need the following three lemmas.

Lemma A.6: *Suppose that the Assumptions A to E hold and let $k = [k_0\ k_1 \cdots k_g \cdots k_G]'$ be a $(G+1)\times 1$ vector of non-negative integers. If $1 \leq k_0 + k_g \leq r_0 + r_g$, then there exists $M_4 < \infty$ such that for all N_g and T*

$$\min(N_g, T)\left|\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\mathcal{F}_g^0 H_g^{(k)}\right)\right| \leq M_4$$

where $H_g^{(k)}$ is the matrix defined in Theorem 1.

Lemma A.7: *Suppose that the Assumptions A to E hold and let $k = [k_0\ k_1 \cdots k_g \cdots k_G]'$ be a $(G+1)\times 1$ vector of non-negative integers. If $k_0 + k_g < r_0 + r_g$, then there exists $\tau_{g,k} > 0$ such that for all N_g and T*

$$\text{plim}\left\{\inf\left[\upsilon_g\left(\mathcal{F}_g^0 H_g^{(k)}\right) - \upsilon_g\left(\mathcal{F}_g^0\right)\right]\right\} = \tau_{g,k}$$

where $H_g^{(k)}$ is the matrix defined in Theorem 1.

Lemma A.8: *Suppose that the Assumptions A to E hold. Let $k = [k_0\ k_1 \cdots k_G]' \leq k^{\max}$ and $k^{\triangledown} = \left[k_0^{\triangledown}\ k_1^{\triangledown} \cdots k_G^{\triangledown}\right]' \leq k^{\max}$ be $(G+1)\times 1$ vectors of non-negative* *integers. If $k_g \geq r_g$, $k_g^{\bigtriangledown} \geq r_g$, $k_0 + k_g \geq r_0 + r_g$ and $k_0^{\bigtriangledown} + k_g^{\bigtriangledown} \geq r_0 + r_g$, then there exists $M_5 < \infty$ such that for all N_g and T*

$$\min(N_g, T) \left| \upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(k^{\bigtriangledown})}\right) \right| \leq M_5$$

Lemmas A.6, A.7 and A8 are direct adaptations of Lemmas 2, 3 and 4 in Bai and Ng (2002), respectively. The proofs of the latter propositions can easily be adapted step by step to prove Lemmas A.6, A.7 and A.8, with the necessary notation changes. However, two remarks are needed in relation to the proof of Lemma A.8. The first remark regards the adaptation of the first expression in Bai and Ng's proof of their Lemma 4 (page 217):

$$\left| \upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(k^{\bigtriangledown})}\right) \right| \leq \left| \upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\mathcal{F}_g^{0}\right) \right| + \left| \upsilon_g\left(\hat{\mathcal{F}}_g^{(k^{\bigtriangledown})}\right) - \upsilon_g\left(\mathcal{F}_g^{0}\right) \right|$$

$$\leq 2. \max_{\substack{k \leq k^{\max} \\ k_0 + k_g \geq r_0 + r_g \\ k_g \geq r_g}} \left| \upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\mathcal{F}_g^{0}\right) \right|$$

In our model, for the group g of variables, the maximum refers to the maximum for all $k \leq k^{\max}$ such that $k_g \geq r_g$ and $k_0 + k_g \geq r_0 + r_g$. This modification does not change the remaining steps of the proof, because in that case $H_g^{(k)}$ has rank $r_0 + r_g$. The second remark refers to the last part of the proof of Bai and Ng's Lemma 4 (bottom of page 218 and top of page 219), which is not correct. More precisely, as acknowledged by Bai and Ng (2005), the proof that (for our model and in our notation)

$$\frac{1}{N_g T} tr\left(\underline{e}_g' P_g \underline{e}_g\right) = O_p\left(\frac{1}{\min(N_g, T)}\right)$$

is invalid. Under our Assumption C.4, we can use Lemma A.2 (Appendix V) to complete the proof. $\square$

X - Proof of Theorem 3

We need to prove that for all k such that $k \neq r$, $0 \leq k \leq k^{\max}$ and $k_0 + k_g > 0$ $(g = 1, \cdots, G)$

$$\text{Prob}\left[GPC(k) - GPC(r) \leq 0\right] = \text{Prob}\left\{\sum_{g=1}^{G} \frac{N_g}{N}\left[GPC_g(k) - GPC_g(r)\right] \leq 0\right\} =$$

$$= \text{Prob}\{\sum_{g=1}^{G} \frac{N_g}{N}\left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right)\right] \leq$$

$$\leq \sum_{g=1}^{G} \frac{N_g}{N}\left[(r_0 + r_g - k_0 - k_g)\,\psi_0(N_g, N, T) + \right.$$

$$+ (r_g - k_g)\left(\psi_g(N_g,N,T) - \psi_0(N_g,N,T)\right)]\} \to 0 \qquad (29)$$

as $N \to \infty$ and $T \to \infty$. Given k, for every variables group g we will have one of the following (mutually exclusive) cases:

(I) $k_0 = r_0$ and $k_g = r_g$
(II) $k_0 + k_g < r_0 + r_g$
(III) $k_g > r_g$ and $k_0 + k_g \geq r_0 + r_g$
(IV) $k_0 > r_0$ and $k_g = r_g$
(V) $k_0 > r_0$, $k_g < r_g$ and $k_0 + k_g \geq r_0 + r_g$

Let us first admit that for all g, the comparison between (k_0, k_g) falls into one (and a single one) of the previous cases. We will consider in turn those cases:

Case (I): If $k \neq r$, this case can not happen.

Case (II): The left hand side of the inequality in (29) may be rewritten as

$$\sum\nolimits_{g=1}^{G} \frac{N_g}{N}\left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\mathcal{F}_g^0 H_g^{(k)}\right)\right] +$$
$$+ \sum\nolimits_{g=1}^{G} \frac{N_g}{N}\left[\upsilon_g\left(\mathcal{F}_g^0 H_g^{(k)}\right) - \upsilon_g\left(\mathcal{F}_g^0 H_g^{(r)}\right)\right] +$$
$$+ \sum\nolimits_{g=1}^{G} \frac{N_g}{N}\left[\upsilon_g\left(\mathcal{F}_g^0 H_g^{(r)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right)\right]$$

Lemma A.6 implies that the first and third terms converge to zero. As regards the second term, note that $\upsilon_g\left(\mathcal{F}_g^0 H_g^{(r)}\right) = \upsilon_g\left(\mathcal{F}_g^0\right)$ because $\mathcal{F}_g^0 H_g^{(r)}$ and $\mathcal{F}_g^0$ asymptotically span the same space. Thus the second term asymptotically is identical to

$$\sum\nolimits_{g=1}^{G} \frac{N_g}{N}\left[\upsilon_g\left(\mathcal{F}_g^0 H_g^{(k)}\right) - \upsilon_g\left(\mathcal{F}_g^0\right)\right]$$

which has a positive limit by Lemma A.7. Hence, the left hand side of the inequality in (29) has a positive limit and $\psi_0(N_g,N,T) \to 0$ $(g = 0, 1, \cdots, G)$ are sufficient conditions to ensure that the probability converges to zero.

Case (III): Multiplying both sides of the inequality in (29) by $\min(N,T)$ we get

$$\text{Prob}\{\sum\nolimits_{g=1}^{G} \frac{N_g}{N}\min(N,T)\left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right)\right] \leq$$
$$\leq \sum\nolimits_{g=1}^{G} \frac{N_g}{N}[(r_0 + r_g - k_0 - k_g)\min(N,T)\psi_0(N_g,N,T) +$$
$$+ (r_g - k_g)\min(N,T)\left(\psi_g(N_g,N,T) - \psi_0(N_g,N,T)\right)]\}$$

By Lemma A.8, the left hand side of the inequality is bounded. The probability goes to zero because conditions (ii) and (iii) ensure that the right hand side of the inequality diverges to $-\infty$. Note that condition (iii) is required whenever $k_0 + k_g = r_0 + r_g$.

Case (IV): Similar to case (III).

Case (V): The left hand side of the inequality in (29) may be rewritten as

$$\sum_{g=1}^{G} \frac{N_g}{N} \left\{ \left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(\bar{k})}\right) \right] + \left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(\bar{k})}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right) \right] \right\}$$

where $\bar{k} = [r_0\ k_1 \cdots k_{g-1}\ k_g\ k_{g+1} \cdots k_G]'$. Thus, it is sufficient for (29) to prove that both

$$\text{Prob}\{\sum_{g=1}^{G} \frac{N_g}{N} \left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(\bar{k})}\right) \right] \leq$$
$$\leq \sum_{g=1}^{G} \frac{N_g}{N} [(r_0 + r_g - k_0 - k_g)\, \psi_0\,(N_g, N, T) +$$
$$+ (r_g - k_g)\left(\psi_g\,(N_g, N, T) - \psi_0\,(N_g, N, T)\right)]\} \to 0 \tag{30}$$

and

$$\text{Prob}\{\sum_{g=1}^{G} \frac{N_g}{N} \left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(\bar{k})}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right) \right] \leq$$
$$\leq \sum_{g=1}^{G} \frac{N_g}{N} [(r_0 + r_g - k_0 - k_g)\, \psi_0\,(N_g, N, T) +$$
$$+ (r_g - k_g)\left(\psi_g\,(N_g, N, T) - \psi_0\,(N_g, N, T)\right)]\} \to 0 \tag{31}$$

The proofs of (30) and of (31) are similar to those presented for cases (IV) and (II), respectively.

Turning now to "mixed cases", given k, let $\delta_{(j)}$ (≥ 0) be the number of groups of variables g that fall into case (j) ($j =$ I,II,III,IV,V). We have $\delta_{(\mathrm{I})} < G$ (because, by construction, $k \neq r$) and $\sum_j \delta_{(j)} = G$. Without loss of generality, admit that groups with $g \leq \delta_{(\mathrm{I})}$fall into case (I) and that groups with $\sum_{i \leq j-1} \delta_{(i)} < g \leq \sum_{i \leq j} \delta_{(i)}$ fall into case (j).

If $\delta_{(\mathrm{I})} = 0$, to prove (29) then it is sufficient to show that for $j =$II,III,IV,V

$$\text{Prob}\left\{ \sum_{g=\delta_{(\mathrm{II})}+\cdots+\delta_{(j-1)}+1}^{g=\delta_{(\mathrm{II})}+\cdots+\delta_{(j-1)}+\delta_{(j)}} \frac{N_g}{N} \left[GPC_g(k) - GPC_g(r)\right] \leq 0 \right\} \to 0 \tag{32}$$

The proofs are similar to those for the corresponding "pure cases" presented in Part 1.

Finally, when $0 < \delta_{(\mathrm{I})} < G$, note that for any $1 \leq g \leq \delta_{(\mathrm{I})}$, by Lemma A.8 we have

$$\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right) = O_p\left(\frac{1}{\min(N,T)}\right) \tag{33}$$

Consider any (II) $\leq (j) \leq$ (V) for which $\delta_j > 0$ (it exists because $k \neq r$). Substitute

$$\text{Prob}\{\sum_{g=1}^{g=\delta_{(\mathrm{I})}} \left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right) \right] +$$

$$\sum_{g=\delta_{(\mathrm{II})}+\cdots+\delta_{(j-1)}+1}^{g=\delta_{(\mathrm{II})}+\cdots+\delta_{(j-1)}+\delta_{(j)}} \frac{N_g}{N} \left[GPC_g(k) - GPC_g(r)\right] \leq 0\} \to 0$$

for (32). Owing to (33), the additional terms do not substantially change the proof and the same arguments apply. □

XI - Proof of Corollary 3.1

For $k_0 + k_g < r_0 + r_g$, Lemmas A.1(iii), A.6 and A.7 imply that[13]

$$\operatorname{plim} \frac{\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right)}{\upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right)} = 1 + \frac{\operatorname{plim}\left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right)\right]}{\operatorname{plim}\left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right)\right]} = 1+$$

$$= 1 + \frac{\operatorname{plim}\left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\mathcal{F}_g^0 H_g^{(k)}\right)\right]}{\operatorname{plim}\left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right)\right]} + \frac{\operatorname{plim}\left[\upsilon_g\left(\mathcal{F}_g^0 H_g^{(k)}\right) - \upsilon_g\left(\mathcal{F}_g^0 H_g^{(r)}\right)\right]}{\operatorname{plim}\left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right)\right]} +$$

$$+ \frac{\operatorname{plim}\left[\upsilon_g\left(\mathcal{F}_g^0 H_g^{(r)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right)\right]}{\operatorname{plim}\left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right)\right]} = 1 + \frac{\operatorname{plim}\left[\upsilon_g\left(\mathcal{F}_g^0 H_g^{(k)}\right) - \upsilon_g\left(\mathcal{F}_g^0\right)\right]}{\operatorname{plim}\left[\upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right)\right]} > 1 + \varepsilon_{g,k}$$

for some $\varepsilon_{g,k} > 0$. Thus

$$\operatorname{plim}\left\{\ln\left[\frac{\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right)}{\upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right)}\right]\right\} > \delta_{g,k}$$

for some $\delta_{g,k} > 0$.

Now, when $k_0 + k_g \geq r_0 + r_g$, let $k^{\bigtriangledown} = \left[k_0^{\bigtriangledown}\ k_1 \cdots k_{g-1}\ k_g^{\bigtriangledown}\ k_{g+1} \cdots k_G\right]' \leq k^{\max}$ be any $(G+1) \times 1$ vector of non-negative integers. If $k_g \geq k_g^{\bigtriangledown} \geq r_g$ and $k_0 + k_g \geq k_0^{\bigtriangledown} + k_g^{\bigtriangledown} \geq r_0 + r_g$, Lemmas A.1(iii) and A.8 imply that

$$\frac{\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right)}{\upsilon_g\left(\hat{\mathcal{F}}_g^{()}\right)} = 1 + \frac{\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(k^{\bigtriangledown})}\right)}{\upsilon_g\left(\hat{\mathcal{F}}_g^{(\bar{k})}\right)} = 1 + O_p\left(\frac{1}{\min\left(N_g, T\right)}\right)$$

and thus that

$$\ln\left[\frac{\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right)}{\upsilon_g\left(\hat{\mathcal{F}}_g^{(k^{\bigtriangledown})}\right)}\right] = O_p\left(\frac{1}{\min\left(N_g, T\right)}\right)$$

[13] The argument is similar to the one presented for case (II) in Part 1 of the proof of Theorem 3.

Therefore, the proofs in Appendix X remain valid after substituting

$$\ln\left[\frac{\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right)}{\upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right)}\right]$$

for $\upsilon_g\left(\hat{\mathcal{F}}_g^{(k)}\right) - \upsilon_g\left(\hat{\mathcal{F}}_g^{(r)}\right)$. □

XII - Data set

SERIES	Thomson Financial Datastream code
GERMANY	
PRODUCTION OF TOTAL INDUSTRY (EXCLUDING CONSTRUCTION) VOLA	BDOPRI35G
PRODUCTION IN TOTAL MANUFACTURING VOLA	BDOPRI38G
PRODUCTION OF TOTAL CONSTRUCTION VOLA	BDOPRI30G
PRODUCTION OF TOTAL MANUFACTURED INTERMEDIATE GOODS VOLA	BDOPRI61G
PRODUCTION OF TOTAL MANUFACTURED INVESTMENT GOODS VOLA	BDOPRI70G
ORDERS FOR TOTAL MANUFACTURED GOODS (VOLUME) VOLA	BDOODI45G
ORDERS FOR EXPORTED MANUFACTURED GOODS (VOLUME) VOLA	BDOODI54G
ORDERS FOR MANUFACTURED GOODS FROM DOM. MARKET (VOLUME) VOLA	BDOODI53G
ORDERS FOR MANUFACTURED INTERMEDIATE GOODS (VOLUME) VOLA	BDOODI51G
ORDERS FOR MANUFACTURED INVESTMENT GOODS (VOLUME) VOLA	BDOODI52G
SALES OF TOTAL MANUFACTURED GOODS (VOLUME) VOLN	BDOSLI69H
SALES OF MANUFACTURED INTERMEDIATE GOODS (VOLUME) VOLN	BDOSLI26H
SALES OF MANUFACTURED INVESTMENT GOODS (VOLUME) VOLN	BDOSLI27H
TOTAL WHOLESALE TRADE (VOLUME) VOLN	BDOSLI22H
TOTAL RETAIL TRADE (VOLUME) VOLA	BDOSLI15G
TOTAL CAR REGISTRATIONS VOLA	BDOSLI05O
PASSENGER CAR REGISTRATIONS SADJ	BDOSLI12E
PERMITS ISSUED FOR DWELLINGS VOLA	BDOODI15O
IMPORTS CIF CURA	BDOXT009B
EXPORTS FOB CURA	BDOXT003B
UNEMPLOYMENT: % CIVILIAN LABOUR(% DEPENDENT LABOUR TO DEC 196	BDUN%TOTQ
PERSONS IN EMPLOYMENT - MINING AND MANUFACTURINGVOLN	BDUUOA01P
UNFILLED VACANCIES VOLA	BDOOL015O
PPI - ALL ITEMS NADJ	BDOPP019F
PPI - MANUFACTURING INDUSTRY NADJ	BDOPP017F
PPI - FOOD, BEVERAGES & TOBACCO NADJ	BDOPP013F
PPI - INVESTMENT GOODS NADJ	BDOPP068F
PPI - INTERMEDIATE GOODS NADJ	BDOPP064F
WPI NADJ	BDOWP005F
CPI -HOUSING RENTAL SERVICES NADJ	BDOCP053F
CPI - ENERGY (EXCL. GASOLINE BEFORE 1991) NADJ	BDOCP041F
CPI - EXCLUDING FOOD & ENERGY NADJ	BDOCP042F
CPI - FOOD AND ALCOHOL-FREE DRINKS (EXCL. REST)NADJ	BDOCP019F
CPI NADJ	BDOCP009F
EXPORT PRICE INDEX SADJ	BDEXPPRCE
IMPORT PRICE INDEX SADJ	BDIMPPRCE
MONEY SUPPLY-GERMAN CONTRIBUTION TO EURO M1(PAN M0690)	BDM1....A
MONEY SUPPLY - M2 (CONTINUOUS SERIES) CURA	BDM2C...B
MONEY SUPPLY - M3 (CONTINUOUS SERIES) CURA	BDM3C...B
FIBOR - 3 MONTH (MTH.AVG.)	BDINTER3
YIELD 10-YEAR GOVT.BONDS(PROXY- 9-10+ YEAR FEDERAL SECUR NADJ	BDOIR080R
SHARE PRICES - CDAX NADJ	BDOSP001F
GERMAN MARKS TO US$ (MTH.AVG.)	BDXRUSD.
UK MARKET PRICE - UK BRENT CURN	UKI76AAZA
ECONOMIC SENTIMENT INDICATOR - GERMANY SADJ	BDEUSESIG
CONSTRUCTION CONFIDENCE INDICATOR - GERMANY SADJ	BDEUSBCIQ
CONSTRUCTION SURVEY: ACT.COMPARED TO LAST MONTH-GERMANY SADJ	BDEUSBACQ
CONSTRUCTION SURVEY: EMPLOYMENT EXPECTATIONS - GERMANY SADJ	BDEUSBEMQ
CONSTRUCTION SURVEY: ORDER BOOK POSITION - GERMANY SADJ	BDEUSBOBQ
CONSTRUCTION SURVEY: PRICE EXPECTATIONS - GERMANY SADJ	BDEUSBPRQ
CONSUMER CONFIDENCE INDICATOR - GERMANY SADJ	BDEUSCCIQ
CONSUMER SURVEY: ECONOMIC SITUATION LAST 12 MTH-GERMANY SADJ	BDEUSCECQ
CONSUMER SURVEY: ECONOMIC SITUATION NEXT 12 MTH-GERMANY SADJ	BDEUSCEYQ
CONSUMER SURVEY: FINANCIAL SITUATION LAST 12 MTH-GERMANY SADJ	BDEUSCFNQ
CONSUMER SURVEY: FINANCIAL SITUATION NEXT 12 MTH-GERMANY SADJ	BDEUSCFYQ
CONSUMER SURVEY: MAJOR PURCH.OVER NEXT 12 MONTHS-GERMANY SADJ	BDEUSCPCQ
CONSUMER SURVEY: MAJOR PURCHASES AT PRESENT - GERMANY SADJ	BDEUSCMPQ
CONSUMER SURVEY: PRICES LAST 12 MONTHS - GERMANYSADJ	BDEUSCPRQ
CONSUMER SURVEY: PRICES NEXT 12 MONTHS - GERMANYSADJ	BDEUSCPYQ
CONSUMER SURVEY: SAVINGS AT PRESENT - GERMANY SADJ	BDEUSCSAQ
CONSUMER SURVEY: SAVINGS OVER NEXT 12 MONTHS - GERMANY SADJ	BDEUSCSYQ
CONSUMER SURVEY: STATEMENT ON FIN.SITUATION OF HOUSEHOLD SADJ	BDEUSCFHQ
CONSUMER SURVEY: UNEMPLOYMENT NEXT 12 MONTHS - GERMANY SADJ	BDEUSCUNQ
INDUSTRIAL CONFIDENCE INDICATOR - GERMANY SADJ	BDEUSICIQ
INDUSTRY SURVEY: EMP.EXPECTATIONS FOR MO.AHEAD -GERMANY SADJ	BDEUSIEMQ
INDUSTRY SURVEY: EXPORT ORDER BOOK POSITION - GERMANY SADJ	BDEUSIEBQ
INDUSTRY SURVEY: ORDER BOOK POSITION - GERMANY SADJ	BDEUSIOBQ
INDUSTRY SURVEY: PROD.EXPECTATION FOR MTH.AHEAD-GERMANY SADJ	BDEUSIPAQ
INDUSTRY SURVEY: PRODN. TRENDS IN RECENT MTH. - GERMANY SADJ	BDEUSIPRQ
INDUSTRY SURVEY: SELLING PRC.EXPECT.MTH. AHEAD -GERMANY SADJ	BDEUSISPQ
INDUSTRY SURVEY: STOCKS OF FINISHED GOODS - GERMANY SADJ	BDEUSIFPQ
RETAIL CONFIDENCE INDICATOR - GERMANY SADJ	BDEUSRCIQ
RETAIL SURVEY: CURRENT BUSINESS SITUATION - GERMANY SADJ	BDEUSRPBQ

RETAIL SURVEY: EMPLOYMENT - GERMANY SADJ	BDEUSREMQ
RETAIL SURVEY: FUTURE BUSINESS SITUATION - GERMANY SADJ	BDEUSREBQ
RETAIL SURVEY: ORDERS PLACED WITH SUPPLIERS - GERMANY SADJ	BDEUSROSQ
RETAIL SURVEY: STOCKS - GERMANY SADJ	BDEUSRSTQ

FRANCE

PRODUCTION OF TOTAL INDUSTRY (EXCLUDING CONSTRUCTION) VOLA	FROPRI35G
PRODUCTION IN TOTAL MANUFACTURING VOLA	FROPRI38G
PRODUCTION OF TOTAL MANUFACTURED CONSUMER GOODS VOLA	FROPRI49G
PRODUCTION OF TOTAL MANUFACTURED INTERMEDIATE GOODS VOLA	FROPRI61G
PRODUCTION OF TOTAL MANUFACTURED INVESTMENT GOODS VOLA	FROPRI70G
PRODUCTION OF TOTAL ENERGY VOLA	FROPRI44G
PRODUCTION IN TOTAL AGRICULTURE VOLA	FROPRI47G
PRODUCTION OF TOTAL CONSTRUCTION VOLA	FROPRI30G
PRODUCTION OF TOTAL VEHICLES VOLA	FROPRI58G
PERMITS ISSUED FOR DWELLINGS VOLA	FROODI15O
WORK STARTED FOR DWELLINGS VOLA	FROWSI41O
TOTAL RETAIL TRADE (VOLUME) VOLA	FROSLI15G
HOUSEHOLD CONSUMPTION - MANUFACTURED GOODS CONA	FRHCONMGD
HOUSEHOLD CONSUMPTION - MANUFACTURED GOODS, RETAIL GOODS CONA	FRHCONMCD
HOUSEHOLD CONSUMPTION - AUTOMOBILES CONA	FRHCONAUD
HOUSEHOLD CONSUMPTION - DURABLE GOODS CONA	FRHCONDGD
HOUSEHOLD CONSUMPTION - TEXTILES & LEATHER CONA	FRHCONTLD
HOUSEHOLD CONSUMPTION - OTHER MANUFACTURED GOODSCONA	FRHCONOTD
HOUSEHOLD CONSUMPTION - FURNITURE CONA	FRHCONFND
HOUSEHOLD CONSUMPTION - HOUSEHOLD APPLIANCES CONA	FRHCONHAD
HOUSEHOLD CONSUMPTION - ELECTRICAL GOODS CONA	FRHCONELD
PASSENGER CAR REGISTRATIONS SADJ	FROSLI12E
TOTAL CAR REGISTRATIONS VOLA	FROSLI05O
IMPORTS FOB CURA	FROXT009B
EXPORTS FOB CURA	FROXT003B
UNEMPLOYMENT VOLA	FROUN010O
NEW UNEMPLOYMENT CLAIMS SADJ	FROUN007G
UNEMPLOYMENT RATE (% OF TOTAL LABOUR FORCE) SADJ	FROUN015Q
NEW JOB VACANCIES FULL & PART-TIME REGISTERED DURING MONTH	FRVACTOTO
PPI - AGRICULTURAL GOODS NADJ	FROPP004F
PPI - INTERMEDIATE GOODS EXCLUDING ENERGY NADJ	FROPP065F
PPI - CHEMICALS NADJ	FROPP054F
PPI - METAL PRODUCTS NADJ	FROPP037F
PPI - PETROLEUM PRODUCTS NADJ	FROPP057F
PPI - MANUFACTURED PRODUCTS NADJ	FROPP017F
CPI NADJ	FROCP009F
CPI - FOOD NADJ	FROCP019F
CPI - ENERGY NADJ	FROCP041F
CPI - EXCLUDING FOOD & ENERGY NADJ	FROCP042F
CPI - RENT NADJ	FROCP054F
CPI - SERVICES EXCLUDING RENT NADJ	FROCP064F
MONEY SUPPLY - M1 (NATIONAL CONTRIBUTION TO M1) CURN	FRM1....A
MONEY SUPPLY - M2 (NATIONAL CONTRIBUTION TO M2) CURN	FRM2....A
MONEY SUPPLY - M3 (NATIONAL CONTRIBUTION TO M3) CURN	FRM3....A
PIBOR / EURIBOR - 3-MONTH (MTH.AVG.)	FRINTER3
YIELD 10-YEAR GOVERNMENT BENCHMARK BONDS NADJ	FROIR080R
SHARE PRICES - SBF 250 NADJ	FROSP001F
FRENCH FRANC TO US $	FRXRUSD.
UK MARKET PRICE - UK BRENT CURN	UKI76AAZA
ECONOMIC SENTIMENT INDICATOR - FRANCE SADJ	FREUSESIG
CONSTRUCTION CONFIDENCE INDICATOR - FRANCE SADJ	FREUSBCIQ
CONSTRUCTION SURVEY: ACT.COMPARED TO LAST MONTH - FRANCE SADJ	FREUSBACQ
CONSTRUCTION SURVEY: EMPLOYMENT EXPECTATIONS - FRANCE SADJ	FREUSBEMQ
CONSTRUCTION SURVEY: ORDER BOOK POSITION - FRANCE SADJ	FREUSBOBQ
CONSTRUCTION SURVEY: PRICE EXPECTATIONS - FRANCESADJ	FREUSBPRQ
CONSUMER CONFIDENCE INDICATOR - FRANCE SADJ	FREUSCCIQ
CONSUMER SURVEY: ECONOMIC SITUATION LAST 12 MTH.- FRANCE SADJ	FREUSCECQ
CONSUMER SURVEY: ECONOMIC SITUATION NEXT 12 MTH.- FRANCE SADJ	FREUSCEYQ
CONSUMER SURVEY: FINANCIAL SITUATION LAST 12 MTH- FRANCE SADJ	FREUSCFNQ
CONSUMER SURVEY: FINANCIAL SITUATION NEXT 12 MTH- FRANCE SADJ	FREUSCFYQ
CONSUMER SURVEY: MAJOR PURCH.OVER NEXT 12 MONTHS- FRANCE SADJ	FREUSCPCQ
CONSUMER SURVEY: MAJOR PURCHASES AT PRESENT - FRANCE SADJ	FREUSCMPQ
CONSUMER SURVEY: PRICES LAST 12 MONTHS - FRANCE SADJ	FREUSCPRQ
CONSUMER SURVEY: PRICES NEXT 12 MONTHS - FRANCE SADJ	FREUSCPYQ
CONSUMER SURVEY: SAVINGS AT PRESENT - FRANCE SADJ	FREUSCSAQ
CONSUMER SURVEY: SAVINGS OVER NEXT 12 MONTHS - FRANCE SADJ	FREUSCSYQ
CONSUMER SURVEY: STATEMENT ON FIN.SITUATION OF HOUSEHOLD SADJ	FREUSCFHQ
CONSUMER SURVEY: UNEMPLOYMENT NEXT 12 MONTHS - FRANCE SADJ	FREUSCUNQ
INDUSTRIAL CONFIDENCE INDICATOR - FRANCE SADJ	FREUSICIQ
INDUSTRY SURVEY: EMP.EXPECTATIONS FOR MO. AHEAD - FRANCE SADJ	FREUSIEMQ
INDUSTRY SURVEY: EXPORT ORDER BOOK POSITION - FRANCE SADJ	FREUSIEBQ
INDUSTRY SURVEY: ORDER BOOK POSITION - FRANCE SADJ	FREUSIOBQ
INDUSTRY SURVEY: PROD.EXPECTATION FOR MTH.AHEAD - FRANCE SADJ	FREUSIPAQ
INDUSTRY SURVEY: PRODN. TRENDS IN RECENT MTH. - FRANCE SADJ	FREUSIPRQ

INDUSTRY SURVEY: SELLING PRC.EXPECT. MTH. AHEAD - FRANCE SADJ	FREUSISPQ
INDUSTRY SURVEY: STOCKS OF FINISHED GOODS - FRANCE SADJ	FREUSIFPQ
RETAIL CONFIDENCE INDICATOR - FRANCE SADJ	FREUSRCIQ
RETAIL SURVEY: CURRENT BUSINESS SITUATION - FRANCE SADJ	FREUSRPBQ
RETAIL SURVEY: EMPLOYMENT - FRANCE SADJ	FREUSREMQ
RETAIL SURVEY: FUTURE BUSINESS SITUATION - FRANCE SADJ	FREUSREBQ
RETAIL SURVEY: ORDERS PLACED WITH SUPPLIERS - FRANCE SADJ	FREUSROSQ
RETAIL SURVEY: STOCKS - FRANCE SADJ	FREUSRSTQ

ITALY

PRODUCTION OF TOTAL INDUSTRY (EXCLUDING CONSTRUCTION) VOLA	ITOPRI35G
PRODUCTION OF TOTAL MANUFACTURED CONSUMER GOODS VOLA	ITOPRI49G
PRODUCTION OF TOTAL MANUFACTURED INTERMEDIATE GOODS VOLA	ITOPRI61G
PRODUCTION OF TOTAL MANUFACTURED INVESTMENT GOODS VOLA	ITOPRI70G
SALES OF TOTAL MANUFACTURED GOODS (VALUE) NADJ	ITOSLI09F
SALES OF TOTAL MANUFACTURED CONSUMER GOODS (VALUE) NADJ	ITOSLI61F
SALES OF MANUFACTURED INTERMEDIATE GOODS (VALUE)NADJ	ITOSLI64F
SALES OF MANUFACTURED INVESTMENT GOODS (VALUE) NADJ	ITOSLI65F
ORDERS FOR TOTAL MANUFACTURED GOODS (VALUE) SADJ	ITOODI32E
TOTAL RETAIL TRADE (VOLUME) VOLA	ITOSLI15G
TOTAL CAR REGISTRATIONS VOLA	ITOSLI05O
PASSENGER CAR REGISTRATIONS SADJ	ITOSLI12E
IMPORTS CIF CURA	ITOXT009B
EXPORTS FOB CURA	ITOXT003B
STANDARDIZED UNEMPLOYMENT RATE SADJ	ITOUN014Q
PPI NADJ	ITOPP019F
CPI NADJ	ITOCP009F
CPI - FOOD NADJ	ITOCP019F
CPI - ENERGY NADJ	ITOCP041F
CPI - EXCLUDING FOOD & ENERGY NADJ	ITOCP042F
CPI - SERVICES LESS HOUSING NADJ	ITOCP064F
CPI - HOUSING NADJ	ITOCP057F
EXPORT UNVALUE INDEX NADJ	ITEXPPRCF
IMPORT UNVALUE INDEX NADJ	ITIMPPRCF
MONEY SUPPLY: M1 - ITALIAN CONTRIBUTION TO THE EURO AREA CURN	ITM1....A
MONEY SUPPLY: M2 - ITALIAN CONTRIBUTION TO THE EURO AREA CURN	ITM2....A
MONEY SUPPLY: M3 - ITALIAN CONTRIBUTION TO THE EURO AREA CURN	ITM3....A
TREASURY BOND NET YIELD -SECONDARY MKT. (EP)	ITGBOND.
SHARE PRICES - ISE MIB STORICO NADJ	ITOSP001F
ITALIAN LIRE TO US $ (MTH.AVG.)	ITXRUSD.
UK MARKET PRICE - UK BRENT CURN	UKI76AAZA
ECONOMIC SENTIMENT INDICATOR - ITALY SADJ	ITEUSESIG
CONSTRUCTION CONFIDENCE INDICATOR - ITALY SADJ	ITEUSBCIQ
CONSTRUCTION SURVEY: ACT.COMPARED TO LAST MONTH - ITALY SADJ	ITEUSBACQ
CONSTRUCTION SURVEY: EMPLOYMENT EXPECTATIONS - ITALY SADJ	ITEUSBEMQ
CONSTRUCTION SURVEY: ORDER BOOK POSITION - ITALYSADJ	ITEUSBOBQ
CONSTRUCTION SURVEY: PRICE EXPECTATIONS - ITALY SADJ	ITEUSBPRQ
CONSUMER CONFIDENCE INDICATOR - ITALY SADJ	ITEUSCCIQ
CONSUMER SURVEY: ECONOMIC SITUATION LAST 12 MTH.- ITALY SADJ	ITEUSCECQ
CONSUMER SURVEY: ECONOMIC SITUATION NEXT 12 MTH.- ITALY SADJ	ITEUSCEYQ
CONSUMER SURVEY: FINANCIAL SITUATION LAST 12 MTH.- ITALY SADJ	ITEUSCFNQ
CONSUMER SURVEY: FINANCIAL SITUATION NEXT 12 MTH.- ITALY SADJ	ITEUSCFYQ
CONSUMER SURVEY: MAJOR PURCH.OVER NEXT 12 MONTHS- ITALY SADJ	ITEUSCPCQ
CONSUMER SURVEY: MAJOR PURCHASES AT PRESENT - ITALY SADJ	ITEUSCMPQ
CONSUMER SURVEY: PRICES LAST 12 MONTHS - ITALY SADJ	ITEUSCPRQ
CONSUMER SURVEY: PRICES NEXT 12 MONTHS - ITALY SADJ	ITEUSCPYQ
CONSUMER SURVEY: SAVINGS AT PRESENT - ITALY SADJ	ITEUSCSAQ
CONSUMER SURVEY: SAVINGS OVER NEXT 12 MONTHS - ITALY SADJ	ITEUSCSYQ
CONSUMER SURVEY: STATEMENT ON FIN.SITUATION OF HOUSEHOLD SADJ	ITEUSCFHQ
CONSUMER SURVEY: UNEMPLOYMENT NEXT 12 MONTHS - ITALY SADJ	ITEUSCUNQ
INDUSTRIAL CONFIDENCE INDICATOR - ITALY SADJ	ITEUSICIQ
INDUSTRY SURVEY: EMP. EXPECTATIONS FOR MO. AHEAD- ITALY SADJ	ITEUSIEMQ
INDUSTRY SURVEY: EXPORT ORDER BOOK POSITION - ITALY SADJ	ITEUSIEBQ
INDUSTRY SURVEY: ORDER BOOK POSITION - ITALY SADJ	ITEUSIOBQ
INDUSTRY SURVEY: PROD.EXPECTATION FOR MTH. AHEAD- ITALY SADJ	ITEUSIPAQ
INDUSTRY SURVEY: PRODN. TRENDS IN RECENT MTH. - ITALY SADJ	ITEUSIPRQ
INDUSTRY SURVEY: SELLING PRC. EXPECT. MTH. AHEAD- ITALY SADJ	ITEUSISPQ
INDUSTRY SURVEY: STOCKS OF FINISHED GOODS - ITALY SADJ	ITEUSIFPQ
RETAIL CONFIDENCE INDICATOR - ITALY SADJ	ITEUSRCIQ
RETAIL SURVEY: CURRENT BUSINESS SITUATION - ITALY SADJ	ITEUSRPBQ
RETAIL SURVEY: EMPLOYMENT - ITALY SADJ	ITEUSREMQ
RETAIL SURVEY: FUTURE BUSINESS SITUATION - ITALYSADJ	ITEUSREBQ
RETAIL SURVEY: ORDERS PLACED WITH SUPPLIERS - ITALY SADJ	ITEUSROSQ
RETAIL SURVEY: STOCKS - ITALY SADJ	ITEUSRSTQ

SPAIN

PRODUCTION OF TOTAL INDUSTRY (EXCLUDING CONSTRUCTION) VOLA	ESOPRI35G
PRODUCTION IN TOTAL MANUFACTURING VOLA	ESOPRI38G

PRODUCTION IN TOTAL MINING VOLN	ESOPRI36H
PRODUCTION OF TOTAL MANUFACTURED CONSUMER GOODS VOLN	ESOPRI49H
PRODUCTION OF TOTAL MANUFACTURED INTERMEDIATE GOODS VOLN	ESOPRI61H
PRODUCTION OF TOTAL MANUFACTURED INVESTMENT GOODS VOLN	ESOPRI70H
PRODUCTION OF CEMENT VOLA	ESOPRI01O
PRODUCTION OF ACCOMMODATION: NIGHTS IN HOTEL VOLA	ESOPRI21O
PASSENGER CAR REGISTRATIONS VOLA	ESOSLI12O
CONSUMPTION: PETROL - CARS (VOLA) VOLA	ESPCA313O
CONSUMPTION: DIESEL OIL (VOLA) VOLA	ESOIL562O
ELECTRICITY CONSUMPTION (VOLA) VOLA	ESECO312O
ELECTRICITY CONSUMPTION - INDUSTRIAL SECTOR (VOLA) VOLA	ESELE629G
CONSUMPTION: VISIBLE - CEMENT (VOLA) VOLA	ESCEM301O
IMPORTS CIF CURA	ESOXT009B
EXPORTS FOB CURA	ESOXT003B
STANDARDIZED UNEMPLOYMENT RATE SADJ	ESOUN014Q
PPI NADJ	ESOPP019F
PPI - AGRICULTURAL PRODUCTS NADJ	ESOPP004F
PPI - MANUFACTURING ALL ITEMS NADJ	ESOPP017F
PPI - INTERMEDIATE GOODS NADJ	ESOPP064F
PPI - CONSUMER GOODS NADJ	ESOPP062F
PPI - INVESTMENT GOODS NADJ	ESOPP068F
PPI - ENERGY NADJ	ESOPP022F
CPI NADJ	ESOCP009F
CPI - ENERGY NADJ	ESOCP041F
CPI - EXCLUDING FOOD & ENERGY NADJ	ESOCP042F
CPI - SERVICEXCLUDING RENT NADJ	ESOCP064F
CPI - RENT NADJ	ESOCP057F
CONSTRUCTION COST INDEX NADJ	ESOOP005F
EXPORT UNIT VALUE INDEX NADJ	ESEXPPRCF
IMPORT UNIT VALUE INDEX NADJ	ESIMPPRCF
MONEY SUPPLY: M2 - SPANISH CONTRIBUTION TO EURO M2 CURN	ESM2....A
MONEY SUPPLY: M3 - SPANISH CONTRIBUTION TO EURO M3 CURN	ESM3....A
INTERBANK RATE - 3 MONTH (WEIGHTED AVERAGE, EP)	ESINTER3
YIELD 10-YEAR GOVERNMENT BONDS NADJ	ESOIR080R
SHARE PRIC- MSE GENERAL INDEX NADJ	ESOSP001F
SPANISH PESETAS TO US $ (MTH.AVG.)	ESXRUSD.
UK MARKET PRICE - UK BRENT CURN	UKI76AAZA
ECONOMIC SENTIMENT INDICATOR - SPAIN SADJ	ESEUSESIG
CONSTRUCTION CONFIDENCE INDICATOR - SPAIN SADJ	ESEUSBCIQ
CONSTRUCTION SURVEY: ACT.COMPARED TO LAST MONTH - SPAIN SADJ	ESEUSBACQ
CONSTRUCTION SURVEY: EMPLOYMENT EXPECTATIONS - SPAIN SADJ	ESEUSBEMQ
CONSTRUCTION SURVEY: ORDER BOOK POSITION - SPAINSADJ	ESEUSBOBQ
CONSTRUCTION SURVEY: PRICE EXPECTATIONS - SPAIN SADJ	ESEUSBPRQ
CONSUMER CONFIDENCE INDICATOR - SPAIN SADJ	ESEUSCCIQ
CONSUMER SURVEY: ECONOMIC SITUATION LAST 12 MTH.- SPAIN SADJ	ESEUSCECQ
CONSUMER SURVEY: ECONOMIC SITUATION NEXT 12 MTH.- SPAIN SADJ	ESEUSCEYQ
CONSUMER SURVEY: FINANCIAL SITUATION LAST 12 MTH.- SPAIN SADJ	ESEUSCFNQ
CONSUMER SURVEY: FINANCIAL SITUATION NEXT 12 MTH.- SPAIN SADJ	ESEUSCFYQ
CONSUMER SURVEY: MAJOR PURCH.OVER NEXT 12 MONTHS- SPAIN SADJ	ESEUSCPCQ
CONSUMER SURVEY: MAJOR PURCHASAT PRESENT - SPAIN SADJ	ESEUSCMPQ
CONSUMER SURVEY: PRICLAST 12 MONTHS - SPAIN SADJ	ESEUSCPRQ
CONSUMER SURVEY: PRICNEXT 12 MONTHS - SPAIN SADJ	ESEUSCPYQ
CONSUMER SURVEY: SAVINGS AT PRESENT - SPAIN SADJ	ESEUSCSAQ
CONSUMER SURVEY: SAVINGS OVER NEXT 12 MONTHS - SPAIN SADJ	ESEUSCSYQ
CONSUMER SURVEY: STATEMENT ON FIN.SITUATION OF HOUSEHOLD SADJ	ESEUSCFHQ
CONSUMER SURVEY: UNEMPLOYMENT NEXT 12 MONTHS - SPAIN SADJ	ESEUSCUNQ
INDUSTRIAL CONFIDENCE INDICATOR - SPAIN SADJ	ESEUSICIQ
INDUSTRY SURVEY: EMP. EXPECTATIONS FOR MO. AHEAD- SPAIN SADJ	ESEUSIEMQ
INDUSTRY SURVEY: EXPORT ORDER BOOK POSITION - SPAIN SADJ	ESEUSIEBQ
INDUSTRY SURVEY: ORDER BOOK POSITION - SPAIN SADJ	ESEUSIOBQ
INDUSTRY SURVEY: PROD.EXPECTATION FOR MTH. AHEAD- SPAIN SADJ	ESEUSIPAQ
INDUSTRY SURVEY: PRODN. TRENDS IN RECENT MTH. - SPAIN SADJ	ESEUSIPRQ
INDUSTRY SURVEY: SELLING PRC. EXPECT. MTH. AHEAD- SPAIN SADJ	ESEUSISPQ
INDUSTRY SURVEY: STOCKS OF FINISHED GOODS - SPAIN SADJ	ESEUSIFPQ
RETAIL CONFIDENCE INDICATOR - SPAIN SADJ	ESEUSRCIQ
RETAIL SURVEY: CURRENT BUSINESS SITUATION - SPAIN SADJ	ESEUSRPBQ
RETAIL SURVEY: EMPLOYMENT - SPAIN SADJ	ESEUSREMQ
RETAIL SURVEY: FUTURE BUSINESS SITUATION - SPAINSADJ	ESEUSREBQ
RETAIL SURVEY: ORDERS PLACED WITH SUPPLIERS - SPAIN SADJ	ESEUSROSQ
RETAIL SURVEY: STOCKS - SPAIN SADJ	ESEUSRSTQ

Table 1 - Simulation results for GIC $_{(1)}$ and GIC$_{(2)}$: base case and G=2

		Deviation from the true number of factors	GIC$_{(1)}$						GIC$_{(2)}$					
			T=60		T=100		T=200		T=60		T=100		T=200	
			Global	Group	Global	Group	Global	Group	Global	Group	Global	Group	Global	Group
Ng=60	c=0.05	-1	0.02	0.04	0.00	0.00	0.00	0.00	0.23	0.15	0.00	0.00	0.00	0.00
		0	0.98	0.96	1.00	1.00	1.00	1.00	0.76	0.81	1.00	1.00	1.00	1.00
		1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	c=0.10	-1	0.01	0.04	0.00	0.00	0.00	0.00	0.11	0.20	0.00	0.01	0.00	0.00
		0	0.99	0.96	1.00	1.00	1.00	1.00	0.88	0.78	1.00	1.00	1.00	1.00
		1	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00
	c=0.20	-1	0.00	0.05	0.00	0.00	0.00	0.00	0.02	0.21	0.00	0.01	0.00	0.00
		0	1.00	0.95	1.00	1.00	1.00	1.00	0.93	0.73	1.00	1.00	1.00	1.00
		1	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00
Ng=100	c=0.05	-1	0.00	0.02	0.00	0.00	0.00	0.00	0.01	0.06	0.00	0.00	0.00	0.00
		0	1.00	0.98	1.00	1.00	1.00	1.00	0.99	0.94	1.00	1.00	1.00	1.00
		1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	c=0.10	-1	0.00	0.02	0.00	0.00	0.00	0.00	0.01	0.06	0.00	0.00	0.00	0.00
		0	1.00	0.98	1.00	1.00	1.00	1.00	0.99	0.94	1.00	1.00	1.00	1.00
		1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	c=0.20	-1	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.00
		0	1.00	0.98	1.00	1.00	1.00	1.00	1.00	0.93	1.00	1.00	1.00	1.00
		1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 2 - Simulation results for GIC $_{(1)}$ and GIC$_{(2)}$: case (ii) and G=2

		Deviation from the true number of factors	$GIC_{(1)}$						$GIC_{(2)}$					
			T=60		T=100		T=200		T=60		T=100		T=200	
			Global	Group	Global	Group	Global	Group	Global	Group	Global	Group	Global	Group
Ng=60	c=0.05	-1	0.00	0.01	0.00	0.00	0.00	0.00	0.05	0.09	0.00	0.00	0.00	0.00
		0	0.97	0.92	1.00	1.00	1.00	1.00	0.95	0.90	1.00	1.00	1.00	1.00
		1	0.03	0.06	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	c=0.10	-1	0.00	0.05	0.00	0.00	0.00	0.00	0.02	0.10	0.00	0.00	0.00	0.00
		0	0.87	0.91	1.00	1.00	1.00	1.00	0.97	0.90	1.00	1.00	1.00	1.00
		1	0.13	0.04	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	c=0.20	-1	0.00	0.40	0.00	0.03	0.00	0.00	0.00	0.12	0.00	0.00	0.00	0.00
		0	0.25	0.44	0.95	0.97	1.00	1.00	0.92	0.86	1.00	1.00	1.00	1.00
		1	0.36	0.01	0.05	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.00
Ng=100	c=0.05	-1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00
		0	0.99	0.97	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	1.00	1.00
		1	0.01	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	c=0.10	-1	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00
		0	0.92	0.96	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	1.00	1.00
		1	0.08	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	c=0.20	-1	0.00	0.41	0.00	0.01	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.00
		0	0.22	0.43	0.98	0.99	1.00	1.00	0.92	0.93	1.00	1.00	1.00	1.00
		1	0.34	0.01	0.02	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00

Table 3 - Simulation results for GIC $_{(1)}$ and GIC$_{(2)}$: case (iii) and G=2

		Deviation from the true number of factors	GIC$_{(1)}$						GIC$_{(2)}$					
			T=60		T=100		T=200		T=60		T=100		T=200	
			Global	Group	Global	Group	Global	Group	Global	Group	Global	Group	Global	Group
Ng=60	c=0.05	-1	0.77	0.47	0.51	0.24	0.02	0.04	0.13	0.28	0.72	0.48	0.17	0.09
		0	0.07	0.37	0.48	0.70	0.98	0.95	0.00	0.07	0.06	0.34	0.84	0.89
		1	0.00	0.04	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.01
	c=0.10	-1	0.66	0.45	0.27	0.27	0.00	0.05	0.32	0.30	0.67	0.47	0.04	0.13
		0	0.27	0.35	0.70	0.66	1.00	0.95	0.02	0.04	0.23	0.32	0.95	0.87
		1	0.02	0.01	0.03	0.01	0.00	0.00	0.00	0.00	0.04	0.01	0.01	0.00
	c=0.20	-1	0.15	0.34	0.03	0.31	0.00	0.05	0.39	0.22	0.13	0.34	0.00	0.14
		0	0.48	0.22	0.81	0.54	1.00	0.95	0.30	0.02	0.47	0.22	0.98	0.85
		1	0.29	0.00	0.12	0.00	0.00	0.00	0.15	0.00	0.31	0.00	0.02	0.00
Ng=100	c=0.05	-1	0.77	0.33	0.12	0.13	0.00	0.00	0.75	0.56	0.71	0.30	0.00	0.03
		0	0.21	0.58	0.88	0.85	1.00	1.00	0.04	0.27	0.26	0.62	1.00	0.97
		1	0.00	0.05	0.00	0.02	0.00	0.00	0.00	0.01	0.00	0.05	0.00	0.00
	c=0.10	-1	0.50	0.36	0.03	0.18	0.00	0.00	0.72	0.50	0.42	0.33	0.00	0.03
		0	0.42	0.51	0.96	0.81	1.00	1.00	0.17	0.25	0.51	0.56	1.00	0.97
		1	0.07	0.01	0.01	0.00	0.00	0.00	0.04	0.01	0.06	0.01	0.00	0.00
	c=0.20	-1	0.07	0.39	0.00	0.19	0.00	0.00	0.21	0.32	0.05	0.36	0.00	0.03
		0	0.59	0.30	0.97	0.79	1.00	1.00	0.42	0.15	0.66	0.38	1.00	0.97
		1	0.26	0.00	0.03	0.00	0.00	0.00	0.30	0.00	0.23	0.00	0.00	0.00

Table 4 - Simulation results for GIC $_{(1)}$ and GIC$_{(2)}$: case (iv) and G=2

		Deviation from the true number of factors	GIC$_{(1)}$						GIC$_{(2)}$					
			T=60		T=100		T=200		T=60		T=100		T=200	
			Global	Group	Global	Group	Global	Group	Global	Group	Global	Group	Global	Group
Ng=60	c=0.05	-1	0.45	0.25	0.08	0.18	0.00	0.03	0.56	0.48	0.38	0.32	0.05	0.07
		0	0.54	0.67	0.92	0.80	1.00	0.97	0.03	0.22	0.58	0.57	0.95	0.93
		1	0.01	0.06	0.01	0.00	0.00	0.00	0.00	0.03	0.03	0.01	0.00	0.01
	c=0.10	-1	0.20	0.32	0.08	0.18	0.00	0.03	0.64	0.46	0.38	0.32	0.05	0.07
		0	0.72	0.61	0.92	0.80	1.00	0.97	0.19	0.21	0.58	0.57	0.95	0.93
		1	0.08	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.03	0.01	0.00	0.01
	c=0.20	-1	0.00	0.37	0.00	0.21	0.00	0.03	0.19	0.32	0.05	0.35	0.00	0.08
		0	0.24	0.24	0.85	0.72	1.00	0.97	0.43	0.12	0.71	0.43	1.00	0.92
		1	0.28	0.00	0.12	0.00	0.00	0.00	0.29	0.00	0.19	0.00	0.01	0.00
Ng=100	c=0.05	-1	0.16	0.17	0.01	0.05	0.00	0.00	0.66	0.31	0.32	0.20	0.00	0.01
		0	0.84	0.79	0.99	0.94	1.00	1.00	0.33	0.60	0.68	0.76	1.00	0.99
		1	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.03	0.00	0.00
	c=0.10	-1	0.05	0.24	0.00	0.06	0.00	0.00	0.36	0.36	0.14	0.26	0.00	0.01
		0	0.87	0.72	1.00	0.94	1.00	1.00	0.59	0.54	0.85	0.71	1.00	0.99
		1	0.07	0.01	0.00	0.00	0.00	0.00	0.05	0.02	0.02	0.00	0.00	0.00
	c=0.20	-1	0.00	0.37	0.00	0.06	0.00	0.00	0.02	0.39	0.00	0.31	0.00	0.01
		0	0.24	0.28	0.97	0.93	1.00	1.00	0.51	0.30	0.90	0.62	1.00	0.99
		1	0.25	0.00	0.03	0.00	0.00	0.00	0.31	0.00	0.09	0.00	0.00	0.00

Table 5 - Simulation results for GIC $_{(1)}$ and GIC $_{(2)}$: G=5

			Deviation from the true number of factors	$GIC_{(1)}$				$GIC_{(2)}$			
				T=60		T=100		T=60		T=100	
				Global	Group	Global	Group	Global	Group	Global	Group
Base case	Ng=60	c=0.10	-1	0.00	0.04	0.00	0.00	0.04	0.22	0.00	0.01
			0	1.00	0.96	1.00	1.00	0.96	0.77	1.00	0.99
			1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		c=0.20	-1	0.00	0.04	0.00	0.00	0.00	0.23	0.00	0.01
			0	1.00	0.96	1.00	1.00	1.00	0.77	1.00	0.99
			1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Ng=100	c=0.10	-1	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00
			0	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00
			1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		c=0.20	-1	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00
			0	1.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00
			1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Case ii)	Ng=60	c=0.10	-1	0.00	0.01	0.00	0.00	0.01	0.10	0.00	0.01
			0	1.00	0.92	1.00	1.00	0.99	0.90	1.00	0.99
			1	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00
		c=0.20	-1	0.00	0.10	0.00	0.00	0.00	0.10	0.00	0.01
			0	0.69	0.84	1.00	1.00	1.00	0.90	1.00	0.99
			1	0.26	0.05	0.00	0.00	0.00	0.00	0.00	0.00
	Ng=100	c=0.10	-1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
			0	1.00	0.96	1.00	1.00	1.00	1.00	1.00	1.00
			1	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00
		c=0.20	-1	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00
			0	0.71	0.89	1.00	1.00	1.00	1.00	1.00	1.00
			1	0.24	0.03	0.00	0.00	0.00	0.00	0.00	0.00
Case iii)	Ng=60	c=0.10	-1	0.79	0.49	0.20	0.37	0.32	0.48	0.79	0.51
			0	0.18	0.36	0.80	0.59	0.01	0.09	0.19	0.34
			1	0.00	0.05	0.00	0.02	0.00	0.01	0.00	0.05
		c=0.20	-1	0.14	0.56	0.00	0.42	0.62	0.43	0.11	0.59
			0	0.85	0.27	0.99	0.56	0.26	0.04	0.86	0.25
			1	0.02	0.00	0.00	0.00	0.01	0.00	0.03	0.00
	Ng=100	c=0.10	-1	0.32	0.30	0.00	0.04	0.80	0.45	0.21	0.25
			0	0.68	0.65	1.00	0.96	0.15	0.43	0.80	0.72
			1	0.00	0.02	0.00	0.00	0.00	0.03	0.00	0.02
		c=0.20	-1	0.05	0.36	0.00	0.04	0.24	0.52	0.01	0.28
			0	0.95	0.60	1.00	0.96	0.75	0.33	0.99	0.71
			1	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Case iv)	Ng=60	c=0.10	-1	0.13	0.31	0.02	0.21	0.81	0.49	0.35	0.40
			0	0.87	0.62	0.98	0.78	0.07	0.26	0.65	0.53
			1	0.00	0.04	0.00	0.00	0.00	0.03	0.00	0.03
		c=0.20	-1	0.00	0.37	0.00	0.22	0.25	0.52	0.02	0.47
			0	0.80	0.55	1.00	0.78	0.73	0.19	0.98	0.48
			1	0.16	0.02	0.00	0.00	0.02	0.00	0.00	0.00
	Ng=100	c=0.10	-1	0.03	0.11	0.00	0.01	0.23	0.28	0.00	0.01
			0	0.97	0.87	1.00	0.99	0.77	0.68	1.00	0.99
			1	0.00	0.02	0.00	0.00	0.00	0.02	0.00	0.00
		c=0.20	-1	0.00	0.16	0.00	0.01	0.03	0.32	0.00	0.12
			0	0.82	0.82	1.00	0.99	0.97	0.65	1.00	0.88
			1	0.15	0.01	0.00	0.00	0.00	0.00	0.00	0.00

Table 6 - Number of global and group specific factors based on GIC(1)

		no. series	no. factors	no. factors based on Bai-Ng IC1 criterion
Global		295	2	-
Groups	Germany	77	6	8
	France	82	2	4
	Italy	64	4	6
	Spain	72	2	3

Figure 1 - The first global factor and euro area GDP growth rate

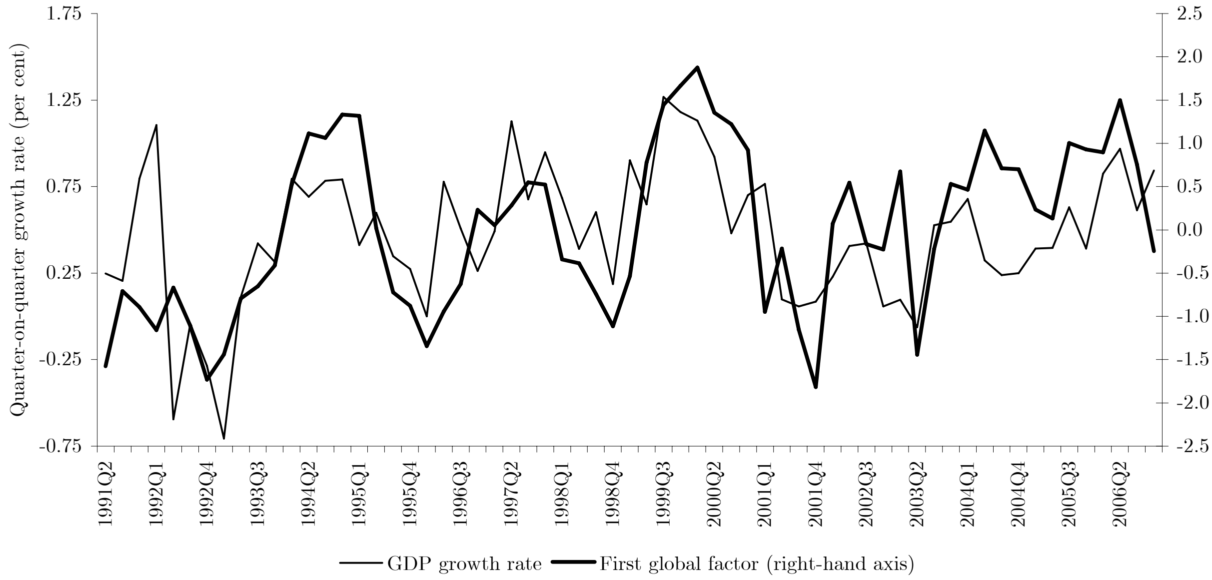
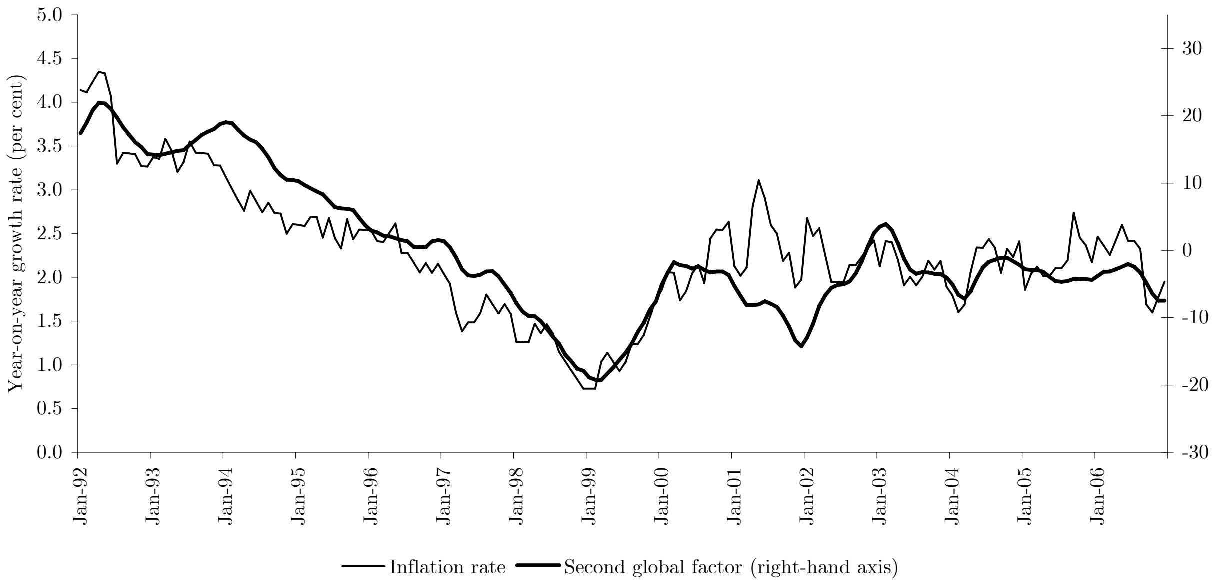


Figure 2 - The second global factor and euro area consumer inflation rate



WORKING PAPERS

2000

1/00 UNEMPLOYMENT DURATION: COMPETING AND DEFECTIVE RISKS
— *John T. Addison, Pedro Portugal*

2/00 THE ESTIMATION OF RISK PREMIUM IMPLICIT IN OIL PRICES
— *Jorge Barros Luís*

3/00 EVALUATING CORE INFLATION INDICATORS
— *Carlos Robalo Marques, Pedro Duarte Neves, Luís Morais Sarmento*

4/00 LABOR MARKETS AND KALEIDOSCOPIC COMPARATIVE ADVANTAGE
— *Daniel A. Traça*

5/00 WHY SHOULD CENTRAL BANKS AVOID THE USE OF THE UNDERLYING INFLATION INDICATOR?
— *Carlos Robalo Marques, Pedro Duarte Neves, Afonso Gonçalves da Silva*

6/00 USING THE ASYMMETRIC TRIMMED MEAN AS A CORE INFLATION INDICATOR
— *Carlos Robalo Marques, João Machado Mota*

2001

1/01 THE SURVIVAL OF NEW DOMESTIC AND FOREIGN OWNED FIRMS
— *José Mata, Pedro Portugal*

2/01 GAPS AND TRIANGLES
— *Bernardino Adão, Isabel Correia, Pedro Teles*

3/01 A NEW REPRESENTATION FOR THE FOREIGN CURRENCY RISK PREMIUM
— *Bernardino Adão, Fátima Silva*

4/01 ENTRY MISTAKES WITH STRATEGIC PRICING
— *Bernardino Adão*

5/01 FINANCING IN THE EUROSYSTEM: FIXED VERSUS VARIABLE RATE TENDERS
— *Margarida Catalão-Lopes*

6/01 AGGREGATION, PERSISTENCE AND VOLATILITY IN A MACROMODEL
— *Karim Abadir, Gabriel Talmain*

7/01 SOME FACTS ABOUT THE CYCLICAL CONVERGENCE IN THE EURO ZONE
— *Frederico Belo*

8/01 TENURE, BUSINESS CYCLE AND THE WAGE-SETTING PROCESS
— *Leandro Arozamena, Mário Centeno*

9/01 USING THE FIRST PRINCIPAL COMPONENT AS A CORE INFLATION INDICATOR
— *José Ferreira Machado, Carlos Robalo Marques, Pedro Duarte Neves, Afonso Gonçalves da Silva*

10/01 IDENTIFICATION WITH AVERAGED DATA AND IMPLICATIONS FOR HEDONIC REGRESSION STUDIES
— *José A.F. Machado, João M.C. Santos Silva*

2002

1/02 QUANTILE REGRESSION ANALYSIS OF TRANSITION DATA
— *José A.F. Machado, Pedro Portugal*

2/02 SHOULD WE DISTINGUISH BETWEEN STATIC AND DYNAMIC LONG RUN EQUILIBRIUM IN ERROR CORRECTION MODELS?
— *Susana Botas, Carlos Robalo Marques*

3/02 MODELLING TAYLOR RULE UNCERTAINTY
— *Fernando Martins, José A. F. Machado, Paulo Soares Esteves*

4/02 PATTERNS OF ENTRY, POST-ENTRY GROWTH AND SURVIVAL: A COMPARISON BETWEEN DOMESTIC AND FOREIGN OWNED FIRMS
— *José Mata, Pedro Portugal*

5/02 BUSINESS CYCLES: CYCLICAL COMOVEMENT WITHIN THE EUROPEAN UNION IN THE PERIOD 1960-1999. A FREQUENCY DOMAIN APPROACH
— *João Valle e Azevedo*

6/02 AN “ART”, NOT A “SCIENCE”? CENTRAL BANK MANAGEMENT IN PORTUGAL UNDER THE GOLD STANDARD, 1854 -1891
— *Jaime Reis*

7/02 MERGE OR CONCENTRATE? SOME INSIGHTS FOR ANTITRUST POLICY
— *Margarida Catalão-Lopes*

8/02 DISENTANGLING THE MINIMUM WAGE PUZZLE: ANALYSIS OF WORKER ACCESSIONS AND SEPARATIONS FROM A LONGITUDINAL MATCHED EMPLOYER-EMPLOYEE DATA SET
— *Pedro Portugal, Ana Rute Cardoso*

9/02 THE MATCH QUALITY GAINS FROM UNEMPLOYMENT INSURANCE
— *Mário Centeno*

10/02 HEDONIC PRICES INDEXES FOR NEW PASSENGER CARS IN PORTUGAL (1997-2001)
— *Hugo J. Reis, J.M.C. Santos Silva*

11/02 THE ANALYSIS OF SEASONAL RETURN ANOMALIES IN THE PORTUGUESE STOCK MARKET
— *Miguel Balbina, Nuno C. Martins*

12/02 DOES MONEY GRANGER CAUSE INFLATION IN THE EURO AREA?
— *Carlos Robalo Marques, Joaquim Pina*

13/02 INSTITUTIONS AND ECONOMIC DEVELOPMENT: HOW STRONG IS THE RELATION?
— *Tiago V.de V. Cavalcanti, Álvaro A. Novo*

2003

1/03 FOUNDING CONDITIONS AND THE SURVIVAL OF NEW FIRMS
— *P.A. Geroski, José Mata, Pedro Portugal*

2/03 THE TIMING AND PROBABILITY OF FDI: AN APPLICATION TO THE UNITED STATES MULTINATIONAL ENTERPRISES
— *José Brandão de Brito, Felipa de Mello Sampayo*

3/03 OPTIMAL FISCAL AND MONETARY POLICY: EQUIVALENCE RESULTS
— *Isabel Correia, Juan Pablo Nicolini, Pedro Teles*

4/03 FORECASTING EURO AREA AGGREGATES WITH BAYESIAN VAR AND VECM MODELS
— *Ricardo Mourinho Félix, Luís C. Nunes*

5/03 CONTAGIOUS CURRENCY CRISES: A SPATIAL PROBIT APPROACH
— *Álvaro Novo*

6/03 THE DISTRIBUTION OF LIQUIDITY IN A MONETARY UNION WITH DIFFERENT PORTFOLIO RIGIDITIES
— *Nuno Alves*

7/03 COINCIDENT AND LEADING INDICATORS FOR THE EURO AREA: A FREQUENCY BAND APPROACH
— *António Rua, Luís C. Nunes*

8/03 WHY DO FIRMS USE FIXED-TERM CONTRACTS?
— *José Varejão, Pedro Portugal*

9/03 NONLINEARITIES OVER THE BUSINESS CYCLE: AN APPLICATION OF THE SMOOTH TRANSITION AUTOREGRESSIVE MODEL TO CHARACTERIZE GDP DYNAMICS FOR THE EURO-AREA AND PORTUGAL
— *Francisco Craveiro Dias*

10/03 WAGES AND THE RISK OF DISPLACEMENT
— *Anabela Carneiro, Pedro Portugal*

11/03 SIX WAYS TO LEAVE UNEMPLOYMENT
— *Pedro Portugal, John T. Addison*

12/03 EMPLOYMENT DYNAMICS AND THE STRUCTURE OF LABOR ADJUSTMENT COSTS
— *José Varejão, Pedro Portugal*

13/03 THE MONETARY TRANSMISSION MECHANISM: IS IT RELEVANT FOR POLICY?
— *Bernardino Adão, Isabel Correia, Pedro Teles*

14/03 THE IMPACT OF INTEREST-RATE SUBSIDIES ON LONG-TERM HOUSEHOLD DEBT: EVIDENCE FROM A LARGE PROGRAM
— *Nuno C. Martins, Ernesto Villanueva*

15/03 THE CAREERS OF TOP MANAGERS AND FIRM OPENNESS: INTERNAL VERSUS EXTERNAL LABOUR MARKETS
— *Francisco Lima, Mário Centeno*

16/03 TRACKING GROWTH AND THE BUSINESS CYCLE: A STOCHASTIC COMMON CYCLE MODEL FOR THE EURO AREA
— *João Valle e Azevedo, Siem Jan Koopman, António Rua*

17/03 CORRUPTION, CREDIT MARKET IMPERFECTIONS, AND ECONOMIC DEVELOPMENT
— *António R. Antunes, Tiago V. Cavalcanti*

18/03 BARGAINED WAGES, WAGE DRIFT AND THE DESIGN OF THE WAGE SETTING SYSTEM
— *Ana Rute Cardoso, Pedro Portugal*

19/03 UNCERTAINTY AND RISK ANALYSIS OF MACROECONOMIC FORECASTS: FAN CHARTS REVISITED
— *Álvaro Novo, Maximiano Pinheiro*

2004

1/04 HOW DOES THE UNEMPLOYMENT INSURANCE SYSTEM SHAPE THE TIME PROFILE OF JOBLESS DURATION?
— *John T. Addison, Pedro Portugal*

2/04 REAL EXCHANGE RATE AND HUMAN CAPITAL IN THE EMPIRICS OF ECONOMIC GROWTH
— *Delfim Gomes Neto*

3/04 ON THE USE OF THE FIRST PRINCIPAL COMPONENT AS A CORE INFLATION INDICATOR
— *José Ramos Maria*

4/04 OIL PRICES ASSUMPTIONS IN MACROECONOMIC FORECASTS: SHOULD WE FOLLOW FUTURES MARKET EXPECTATIONS?
— *Carlos Coimbra, Paulo Soares Esteves*

5/04 STYLISED FEATURES OF PRICE SETTING BEHAVIOUR IN PORTUGAL: 1992-2001
— *Mónica Dias, Daniel Dias, Pedro D. Neves*

6/04 A FLEXIBLE VIEW ON PRICES
— *Nuno Alves*

7/04 ON THE FISHER-KONIECZNY INDEX OF PRICE CHANGES SYNCHRONIZATION
— *D.A. Dias, C. Robalo Marques, P.D. Neves, J.M.C. Santos Silva*

8/04 INFLATION PERSISTENCE: FACTS OR ARTEFACTS?
— *Carlos Robalo Marques*

9/04 WORKERS' FLOWS AND REAL WAGE CYCLICALITY
— *Anabela Carneiro, Pedro Portugal*

10/04 MATCHING WORKERS TO JOBS IN THE FAST LANE: THE OPERATION OF FIXED-TERM CONTRACTS
— *José Varejão, Pedro Portugal*

11/04 THE LOCATIONAL DETERMINANTS OF THE U.S. MULTINATIONALS ACTIVITIES
— *José Brandão de Brito, Felipa Mello Sampayo*

12/04 KEY ELASTICITIES IN JOB SEARCH THEORY: INTERNATIONAL EVIDENCE
— *John T. Addison, Mário Centeno, Pedro Portugal*

13/04 RESERVATION WAGES, SEARCH DURATION AND ACCEPTED WAGES IN EUROPE
— *John T. Addison, Mário Centeno, Pedro Portugal*

14/04 THE MONETARY TRANSMISSION N THE US AND THE EURO AREA: COMMON FEATURES AND COMMON FRICTIONS
— *Nuno Alves*

15/04 NOMINAL WAGE INERTIA IN GENERAL EQUILIBRIUM MODELS
— *Nuno Alves*

16/04 MONETARY POLICY IN A CURRENCY UNION WITH NATIONAL PRICE ASYMMETRIES
— *Sandra Gomes*

17/04 NEOCLASSICAL INVESTMENT WITH MORAL HAZARD
— *João Ejarque*

18/04 MONETARY POLICY WITH STATE CONTINGENT INTEREST RATES
— *Bernardino Adão, Isabel Correia, Pedro Teles*

19/04 MONETARY POLICY WITH SINGLE INSTRUMENT FEEDBACK RULES
— *Bernardino Adão, Isabel Correia, Pedro Teles*

20/04 ACOUNTING FOR THE HIDDEN ECONOMY: BARRIERS TO LAGALITY AND LEGAL FAILURES
— *António R. Antunes, Tiago V. Cavalcanti*

2005

1/05 SEAM: A SMALL-SCALE EURO AREA MODEL WITH FORWARD-LOOKING ELEMENTS
— *José Brandão de Brito, Rita Duarte*

2/05 FORECASTING INFLATION THROUGH A BOTTOM-UP APPROACH: THE PORTUGUESE CASE
— *Cláudia Duarte, António Rua*

3/05 USING MEAN REVERSION AS A MEASURE OF PERSISTENCE
— *Daniel Dias, Carlos Robalo Marques*

4/05 HOUSEHOLD WEALTH IN PORTUGAL: 1980-2004
— *Fátima Cardoso, Vanda Geraldes da Cunha*

5/05 ANALYSIS OF DELINQUENT FIRMS USING MULTI-STATE TRANSITIONS
— *António Antunes*

6/05 PRICE SETTING IN THE AREA: SOME STYLIZED FACTS FROM INDIVIDUAL CONSUMER PRICE DATA
— *Emmanuel Dhyne, Luis J. Álvarez, Hervé Le Bihan, Giovanni Veronese, Daniel Dias, Johannes Hoffmann, Nicole Jonker, Patrick Lünnemann, Fabio Rumler, Jouko Vilmunen*

7/05 INTERMEDIATION COSTS, INVESTOR PROTECTION AND ECONOMIC DEVELOPMENT
— *António Antunes, Tiago Cavalcanti, Anne Villamil*

8/05 TIME OR STATE DEPENDENT PRICE SETTING RULES? EVIDENCE FROM PORTUGUESE MICRO DATA
— *Daniel Dias, Carlos Robalo Marques, João Santos Silva*

9/05 BUSINESS CYCLE AT A SECTORAL LEVEL: THE PORTUGUESE CASE
— *Hugo Reis*

10/05 THE PRICING BEHAVIOUR OF FIRMS IN THE EURO AREA: NEW SURVEY EVIDENCE
— *S. Fabiani, M. Druant, I. Hernando, C. Kwapil, B. Landau, C. Loupias, F. Martins, T. Mathä, R. Sabbatini, H. Stahl, A. Stokman*

11/05 CONSUMPTION TAXES AND REDISTRIBUTION
— *Isabel Correia*

12/05 UNIQUE EQUILIBRIUM WITH SINGLE MONETARY INSTRUMENT RULES
— *Bernardino Adão, Isabel Correia, Pedro Teles*

13/05 A MACROECONOMIC STRUCTURAL MODEL FOR THE PORTUGUESE ECONOMY
— *Ricardo Mourinho Félix*

14/05 THE EFFECTS OF A GOVERNMENT EXPENDITURES SHOCK
— *Bernardino Adão, José Brandão de Brito*

15/05 MARKET INTEGRATION IN THE GOLDEN PERIPHERY – THE LISBON/LONDON EXCHANGE, 1854-1891
— *Rui Pedro Esteves, Jaime Reis, Fabiano Ferramosca*

2006

1/06 THE EFFECTS OF A TECHNOLOGY SHOCK IN THE EURO AREA
— *Nuno Alves , José Brandão de Brito , Sandra Gomes, João Sousa*

2/02 THE TRANSMISSION OF MONETARY AND TECHNOLOGY SHOCKS IN THE EURO AREA
— *Nuno Alves, José Brandão de Brito, Sandra Gomes, João Sousa*

3/06 MEASURING THE IMPORTANCE OF THE UNIFORM NONSYNCHRONIZATION HYPOTHESIS
— *Daniel Dias, Carlos Robalo Marques, João Santos Silva*

4/06 THE PRICE SETTING BEHAVIOUR OF PORTUGUESE FIRMS EVIDENCE FROM SURVEY DATA
— *Fernando Martins*

5/06 STICKY PRICES IN THE EURO AREA: A SUMMARY OF NEW MICRO EVIDENCE
— *L. J. Álvarez, E. Dhyne, M. Hoeberichts, C. Kwapil, H. Le Bihan, P. Lünnemann, F. Martins, R. Sabbatini, H. Stahl, P. Vermeulen and J. Vilmunen*

6/06 NOMINAL DEBT AS A BURDEN ON MONETARY POLICY
— *Javier Díaz-Giménez, Giorgia Giovannetti , Ramon Marimon, Pedro Teles*

7/06 A DISAGGREGATED FRAMEWORK FOR THE ANALYSIS OF STRUCTURAL DEVELOPMENTS IN PUBLIC FINANCES
— *Jana Kremer, Cláudia Rodrigues Braz, Teunis Brosens, Geert Langenus, Sandro Momigliano, Mikko Spolander*

8/06 IDENTIFYING ASSET PRICE BOOMS AND BUSTS WITH QUANTILE REGRESSIONS
— *José A. F. Machado, João Sousa*

9/06 EXCESS BURDEN AND THE COST OF INEFFICIENCY IN PUBLIC SERVICES PROVISION
— *António Afonso, Vítor Gaspar*

10/06 MARKET POWER, DISMISSAL THREAT AND RENT SHARING: THE ROLE OF INSIDER AND OUTSIDER FORCES IN WAGE BARGAINING
— *Anabela Carneiro, Pedro Portugal*

11/06 MEASURING EXPORT COMPETITIVENESS: REVISITING THE EFFECTIVE EXCHANGE RATE WEIGHTS FOR THE EURO AREA COUNTRIES
— *Paulo Soares Esteves, Carolina Reis*

12/06 THE IMPACT OF UNEMPLOYMENT INSURANCE GENEROSITY
ON MATCH QUALITY DISTRIBUTION
— *Mário Centeno, Alvaro A. Novo*

13/06 U.S. UNEMPLOYMENT DURATION: HAS LONG BECOME LONGER OR SHORT BECOME SHORTER?
— *José A.F. Machado, Pedro Portugal e Juliana Guimarães*

14/06 EARNINGS LOSSES OF DISPLACED WORKERS: EVIDENCE FROM A MATCHED EMPLOYER-EMPLOYEE DATA SET
— *Anabela Carneiro, Pedro Portugal*

15/06 COMPUTING GENERAL EQUILIBRIUM MODELS WITH OCCUPATIONAL CHOICE AND FINANCIAL FRICTIONS
— *António Antunes, Tiago Cavalcanti, Anne Villamil*

16/06 ON THE RELEVANCE OF EXCHANGE RATE REGIMES FOR STABILIZATION POLICY
— *Bernardino Adao, Isabel Correia, Pedro Teles*

17/06 AN INPUT-OUTPUT ANALYSIS: LINKAGES VS LEAKAGES
— *Hugo Reis, António Rua*

2007

1/07 RELATIVE EXPORT STRUCTURES AND VERTICAL SPECIALIZATION: A SIMPLE CROSS-COUNTRY INDEX
— *João Amador, Sónia Cabral, José Ramos Maria*

2/07	THE FORWARD PREMIUM OF EURO INTEREST RATES — *Sónia Costa, Ana Beatriz Galvão*
3/07	ADJUSTING TO THE EURO — *Gabriel Fagan, Vítor Gaspar*
4/07	SPATIAL AND TEMPORAL AGGREGATION IN THE ESTIMATION OF LABOR DEMAND FUNCTIONS — *José Varejão, Pedro Portugal*
5/07	PRICE SETTING IN THE EURO AREA: SOME STYLISED FACTS FROM INDIVIDUAL PRODUCER PRICE DATA — *Philip Vermeulen, Daniel Dias, Maarten Dossche, Erwan Gautier, Ignacio Hernando, Roberto Sabbatini, Harald Stahl*
6/07	A STOCHASTIC FRONTIER ANALYSIS OF SECONDARY EDUCATION OUTPUT IN PORTUGAL — *Manuel Coutinho Pereira, Sara Moreira*
7/07	CREDIT RISK DRIVERS: EVALUATING THE CONTRIBUTION OF FIRM LEVEL INFORMATION AND OF MACROECONOMIC DYNAMICS — *Diana Bonfim*
8/07	CHARACTERISTICS OF THE PORTUGUESE ECONOMIC GROWTH: WHAT HAS BEEN MISSING? — *João Amador, Carlos Coimbra*
9/07	TOTAL FACTOR PRODUCTIVITY GROWTH IN THE G7 COUNTRIES: DIFFERENT OR ALIKE? — *João Amador, Carlos Coimbra*
10/07	IDENTIFYING UNEMPLOYMENT INSURANCE INCOME EFFECTS WITH A QUASI-NATURAL EXPERIMENT — *Mário Centeno, Alvaro A. Novo*
11/07	HOW DO DIFFERENT ENTITLEMENTS TO UNEMPLOYMENT BENEFITS AFFECT THE TRANSITIONS FROM UNEMPLOYMENT INTO EMPLOYMENT — *John T. Addison, Pedro Portugal*
12/07	INTERPRETATION OF THE EFFECTS OF FILTERING INTEGRATED TIME SERIES — *João Valle e Azevedo*
13/07	EXACT LIMIT OF THE EXPECTED PERIODOGRAM IN THE UNIT-ROOT CASE — *João Valle e Azevedo*
14/07	INTERNATIONAL TRADE PATTERNS OVER THE LAST FOUR DECADES: HOW DOES PORTUGAL COMPARE WITH OTHER COHESION COUNTRIES? — *João Amador, Sónia Cabral, José Ramos Maria*
15/07	INFLATION (MIS)PERCEPTIONS IN THE EURO AREA — *Francisco Dias, Cláudia Duarte, António Rua*
16/07	LABOR ADJUSTMENT COSTS IN A PANEL OF ESTABLISHMENTS: A STRUCTURAL APPROACH — *João Miguel Ejarque, Pedro Portugal*
17/07	A MULTIVARIATE BAND-PASS FILTER — *João Valle e Azevedo*
18/07	AN OPEN ECONOMY MODEL OF THE EURO AREA AND THE US — *Nuno Alves, Sandra Gomes, João Sousa*
19/07	IS TIME RIPE FOR PRICE LEVEL PATH STABILITY? — *Vitor Gaspar, Frank Smets , David Vestin*

20/07 IS THE EURO AREA M3 ABANDONING US?
— *Nuno Alves, Carlos Robalo Marques, João Sousa*

21/07 DO LABOR MARKET POLICIES AFFECT EMPLOYMENT COMPOSITION? LESSONS FROM EUROPEAN COUNTRIES
— *António Antunes, Mário Centeno*

2008

1/08 THE DETERMINANTS OF PORTUGUESE BANKS' CAPITAL BUFFERS
— *Miguel Boucinha*

2/08 DO RESERVATION WAGES REALLY DECLINE? SOME INTERNATIONAL EVIDENCE ON THE DETERMINANTS OF RESERVATION WAGES
— *John T. Addison, Mário Centeno, Pedro Portugal*

3/08 UNEMPLOYMENT BENEFITS AND RESERVATION WAGES: KEY ELASTICITIES FROM A STRIPPED-DOWN JOB SEARCH APPROACH
— *John T. Addison, Mário Centeno, Pedro Portugal*

4/08 THE EFFECTS OF LOW-COST COUNTRIES ON PORTUGUESE MANUFACTURING IMPORT PRICES
— *Fátima Cardoso, Paulo Soares Esteves*

5/08 WHAT IS BEHIND THE RECENT EVOLUTION OF PORTUGUESE TERMS OF TRADE?
— *Fátima Cardoso, Paulo Soares Esteves*

6/08 EVALUATING JOB SEARCH PROGRAMS FOR OLD AND YOUNG INDIVIDUALS: HETEROGENEOUS IMPACT ON UNEMPLOYMENT DURATION
— *Luis Centeno, Mário Centeno, Álvaro A. Novo*

7/08 FORECASTING USING TARGETED DIFFUSION INDEXES
— *Francisco Dias, Maximiano Pinheiro, António Rua*

8/08 STATISTICAL ARBITRAGE WITH DEFAULT AND COLLATERAL
— *José Fajardo, Ana Lacerda*

9/08 DETERMINING THE NUMBER OF FACTORS IN APPROXIMATE FACTOR MODELS WITH GLOBAL AND GROUP-SPECIFIC FACTORS
— *Francisco Dias, Maximiano Pinheiro, António Rua*