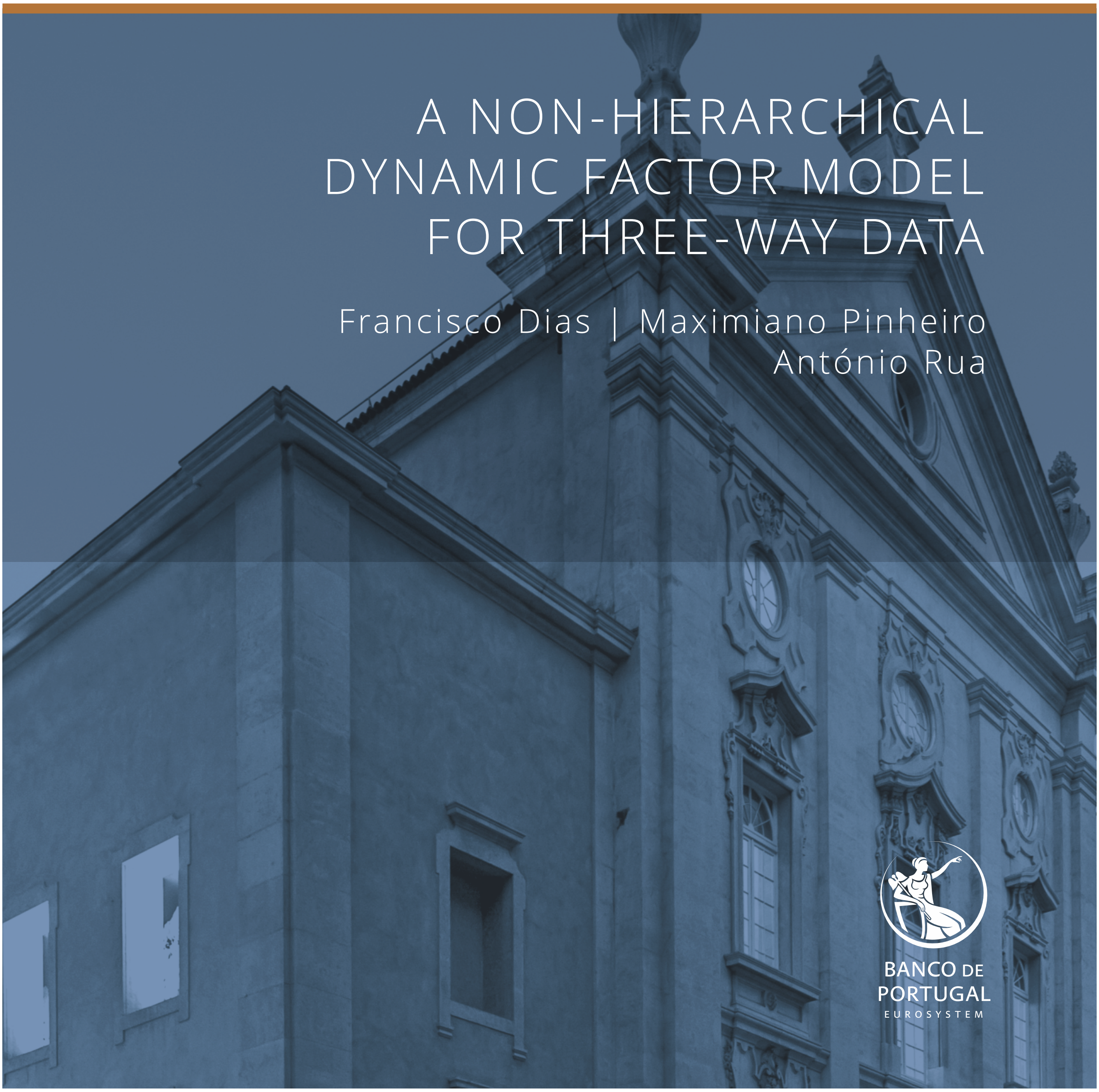
07

WORKING PAPERS 2020



A NON-HIERARCHICAL DYNAMIC FACTOR MODEL FOR THREE-WAY DATA

Francisco Dias | Maximiano Pinheiro
António Rua

BANCO DE PORTUGAL
EUROSYSTEM

07
WORKING
PAPERS 2020

A NON-HIERARCHICAL DYNAMIC FACTOR MODEL FOR THREE-WAY DATA

Francisco Dias | Maximiano Pinheiro
António Rua

MAY 2020

Please address correspondence to
Banco de Portugal, Economics and Research Department
Av. Almirante Reis, 71, 1150-012 Lisboa, Portugal
Tel.: +351 213 130 000, email: estudos@bportugal.pt

BANCO DE PORTUGAL
EUROSYSTEM

Lisboa, 2020 • www.bportugal.pt

Working Papers | Lisbon 2020 • Banco de Portugal Av. Almirante Reis, 71 | 1150-012 Lisboa • www.bportugal.pt •
Edition Economics and Research Department • ISBN (online) 978-989-678-731-8 • ISSN (online) 2182-0422

A non-hierarchical dynamic factor model for three-way data

Francisco Dias
Banco de Portugal

Maximiano Pinheiro
Banco de Portugal

António Rua
Banco de Portugal
Nova SBE

May 2020

Abstract
Along with the advances of statistical data collection worldwide, dynamic factor models have gained prominence in economics and finance when dealing with data rich environments. Although factor models have been typically applied to two-dimensional data, three-way array data sets are becoming increasingly available. Motivated by the tensor decomposition literature, we propose a dynamic factor model for three-way data. We show that this modeling strategy is flexible while remaining quite parsimonious, in sharp contrast with previous approaches. We discuss identification and put forward a set of identifying restrictions that enhance the interpretation of the model. We propose an estimation procedure based on maximum likelihood using the Expectation-Conditional Maximization algorithm and assess the finite sample properties of the estimator through a Monte Carlo study. In the empirical application, we apply the model to inflation data for nineteen euro area countries and fifty-five products covering the last two decades.

JEL: C38, C55, E31

Keywords: Dynamic factor model; Multi-way data; Expectation-Conditional maximization; Inflation.

1. Introduction

Macroeconomic data sets consist typically of variables addressed by one or two-dimensional indices. In the latter case, one of the dimensions refers to time periods and the other dimension represents *inter alia* macroeconomic aggregates, countries, regions, and industries. However, an increasingly number of big data sets is becoming available with variables whose observations need to be addressed by more than a single index besides time.

The emergence of these large data sets with observations referenced to more than two-dimensional indices recommends macroeconometric models developed specifically to address the issues raised by this type of data. Models proposed for 'flat-view' data, even when they appear to allow direct adaptation to multi-dimensional indexed data, lose their parsimony and clearly become over-parameterized when applied to more complex data.

In this respect, there are a few statistical fields which have been faster in the development of models targeted to deal with data organized in multi-way arrays, commonly called tensors, instead of the traditional vectors and matrices. Psychometrics and chemometrics have historically been two areas driving theoretical and algorithmic developments in models for high-order (i.e. multi-mode) tensor data since as early as the 1960s. Signal processing and machine learning followed in the 1990s and in the 2000s, respectively.[1]

In this paper, we propose a dynamic factor model for three-way array data, as this type of data sets is becoming increasingly available in economics and finance. In the empirical application, we estimate such a model using detailed consumer price indices for elementary product items for the nineteen euro area countries over the last two decades. Naturally, the model's empirical relevance is not confined to this particular application and can be easily applied with other multi-way data sets exactly as it is formulated or with minor changes.

In a dynamic factor model, the data generating process of each variable is the sum of a common component and an idiosyncratic component. The commonality is driven by a relatively small number of latent time-factors generated by a finite order vector autoregression. The literature on dynamic factor models in economics and finance traces back to the seminal contributions by Geweke (1977), Sargent and Sims (1977), Geweke and Singleton (1981) and Watson and Engel (1983). In this early literature, the analysis was limited to two-way data arrays with a small number of variables, and the model was estimated by Gaussian maximum likelihood using either frequency or time domain approaches. For fixed cross-sectional size, the consistency of the maximum likelihood estimator was

1. There is a vast literature on these topics. See, for example, Kolda and Balder (2009) for a review of tensor models and available algorithms, Smilde *et al.* (2004) focus on tensor decompositions with applications in chemometrics, Cichocki *et al.* (2015) highlight the applications to signal processing, and Sidiropoulos *et al.* (2017) provide an overview of models and algorithms of tensor rank decomposition and factorization for signal processing and machine learning.

ensured under the assumptions of cross-sectional and serial independence of the idiosyncratic components, as well as of independence between the latter and the time-factors.

The growing availability of large two-way panel data sets steered the development of a non-parametric estimation procedure drawing on least squares. The resulting principal components estimator overcomes the feasibility issues of the maximum likelihood estimator in the context of large cross-sections. The consistency of the principal components estimator has been addressed by Connor and Korajczyk (1986, 1988, 1993) when the number of variables grow to infinity and the time dimension remains fixed. When both the number of variables and the time dimension tend to infinity, Stock and Watson (1998, 2002a), Bai and Ng (2002), and Amengual and Watson (2007) have shown that the first principal components span the time-factor space, even in the presence of heteroskedasticity and limited time and cross-sectional dependence of the idiosyncratic terms, as well as moderate correlation between the latter and the factors. Related work using frequency domain methods includes Forni and Reichlin (1998), Forni and Lippi (2001), and Forni *et al.* (2000, 2004, 2005).

The classical approach based on maximum likelihood has been reconciled with the estimation of models for large cross-sections by Doz, Giannone and Reichlin (2012). In the sense of White (1982), the classical dynamic factor model is treated as a possibly misspecified model which may be used for estimation purposes. Resorting to the Expectation-Maximization (EM) algorithm of Dempster, Laird and Rubin (1977), the estimation by maximum likelihood becomes feasible for large cross-sections, and the factor space is estimated consistently even if the underlying data generating process deviates slightly from the classical assumptions of homoskedasticity and serial independence of idiosyncratic components. Such specification has been enhanced to allow explicitly for serially correlated idiosyncratic components, significantly improving the model fit in many applications (Reis and Watson, 2010; Pinheiro, Rua and Dias, 2013; Jungbacker and Koopman, 2015).

In some empirical applications, the set of variables may be partitioned into blocks which in turn may or may not be further partitioned. In more complex data structures, the variables are organized along more than one partition, for instance, a partition of variables by region and a partition of the same variables by industry. If one estimates an unrestricted dynamic factor model with such data sets, the issue of over-parameterization is very present and becomes overwhelming even with small to moderate sample sizes. Furthermore, in such situations, if we limit the number of factors to a manageable level, we will potentially face strong cross-correlation of estimated idiosyncratic components between variables belonging to the same group or subgroup, suggesting under-specification of the unrestricted model. One way to tackle this problem is to specify models not only with global factors, shared by the data generating process of all variables and thus with unrestricted loadings, but also with block-of-variables-specific factors. By construction, the latter are restricted to appear only in the measurement equation of those variables belonging to the particular block for which they are targeted.

When in a data set there is a single partition of variables by primary blocks, but these may be organized into sub-blocks, the dynamic factor models developed to address these multi-level data structures are usually referred to as hierarchical models. Such models are rather popular when comparing the business cycles of different countries (Norrbin and Schlagenhauf, 1996; Gregory, Head and Raynauld, 1997; Gregory and Head, 1999; Kose, Otrok and Whiteman, 2003, 2008; Marcellino, Stock and Watson, 2003; Crucini, Kose and Otrok, 2011; Mumtaz, Simonelli and Surico, 2011; Kose, Otrok and Prasad, 2012; Dias, Pinheiro and Rua, 2013; among others).

There are also several examples of dynamic factor models which explore multi-level data structures and focus on economic issues other than international business cycles. For instance, Diebold, Li and Yue (2008) propose a factor model to study the evolution of sovereign bond yields in the US, Germany, Japan and UK; Giannone, Reichlin and Small (2008) aim at improving nowcasts and organize the variables according to their timing of release; Stock and Watson (2008) develop a model for the number of building permits in the different US states, with global and state-specific factors, as well as stochastic volatility of both factors and idiosyncratic components; Beck, Hubrich and Marcellino (2009) employ a dynamic factor model to describe regional inflation dynamics in the euro area with area-wide factors and country-specific factors; Foerster, Sarte and Watson (2011) use a hierarchical factor model to decompose US industrial production (117 industries) into components arising from aggregate shocks, sectoral shocks and pure idiosyncratic shocks; Moench, Ng and Potter (2013) propose a general framework for hierarchical dynamic factor models and apply it to estimate a model of real activity in the US, with five first-level and nine second-level blocks of variables.

Besides dynamic factor models developed for hierarchical data structures, there are relatively few papers in the literature with models developed specifically for more than one non-time data modes (i.e. for data addressed by more than one index besides time). For example, Karadimitropoulou and León-Ledesma (2013) provide an assessment of G7 countries' business cycles considering, besides a partition of variables by country, a parallel overlapping partition of variables by industry. They specify 38 factors, one of which global, seven country-specific (one for each country) and 30 sector-specific (as many as the industries in their data set). Beck, Hubrich and Marcellino (2016) investigate the sources of prices changes in the euro area and estimate aggregate, sectoral-specific and country-specific factors using a dataset for 11 products and six countries, further broken down in 61 regions. They specify 18 factors, one global, 11 product-specific and 6 country-specific, with the intra-country heterogeneity conveyed by the idiosyncratic component.

If we postulate one global factor and also one specific factor for each block (and sub-block) of variables, as done in most of the above literature, in terms of empirical feasibility we will face a rather small upper bound to the sample size for its non-time modes. In this paper, we have a partition of variables by 19 countries that coexists at the same level with a partition of the same variables by 55 products. Adopting the modeling approach of one factor by block of variables, plus an global factor, one would reach the excessive number of 75 factors. So, inspired by the tensor decomposition literature, we propose a specification of the measurement equation which, while remaining rather flexible to convey global and relative price movements, it is very parsimonious.

In order to illustrate the proposed modeling approach, let us consider that, besides the time mode, the data structure has (at least) two other modes and the variables are partitioned in blocks defined for each non-time mode so that a given variable belongs to as many blocks as index-dimensions besides time. In our empirical application, each variable is referred to a consumer price item and to a country. Given one of those non-time modes of the data set, for instance our 55 products, if we modeled a specific factor for each item, and the corresponding loadings in the measurement equation were regarded as vectors spanning a given space, likely it would be possible to approximate that space by a subspace of much smaller dimension, say three. So instead of 55 specific factors we would only need three to approximately span the loadings space. Applying the same reasoning to each data mode besides time, one obtains a multiplicative and very parsimonious structure for the loadings.

In the empirical application, we end up with a representation of consumer price co-movements in the euro area with the relatively small number of six factors, including one global factor, two factor specific to consumer price items (expressing the non-idiosyncratic changes of relative prices across products), and one factor specific to countries. Given the common currency in the euro area, cross-country relative price changes correspond to real exchange rates movements. Hence, the number of loading parameters in the suggested specification is around 2% and 6% of the corresponding number when adopting the unrestricted specification or the approach with one factor for each block of variables in the overlapping data partitions, respectively. Note that the suggested approach can make an otherwise unfeasible sample size to became quite manageable. Moreover, in terms of computational cost, the number of variables in each mode is less of a concern and one would have managed with larger numbers of consumer price items and countries in our application.

The identification of the proposed specification is addressed and a set of identifying restrictions which improve the interpretation of model results is discussed.[2] By considering those identifying restrictions, it can be shown that the measurement equation may be averaged out in each and every non-time mode of the data, giving rise to simpler equations with conventional specifications. In this way, it is possible to put forward the above mentioned interpretation that

2. In contrast with simpler dynamic factor models, in this specification factors cannot be made contemporaneously uncorrelated (i.e. orthogonal) in the sample and therefore it becomes less straightforward to obtain variance decompositions. To address this issue, we explore and assess two alternative re-parameterizations and discuss their limitations.

some factors are conveying pure relative price changes across products or across countries. Moreover, exploring the aggregative properties of the model, we compute a global inflation indicator for the euro area which reflects inflation developments in the euro area purged from relative prices changes across products and countries and idiosyncratic terms.

When averaging out the country dimension in the proposed measurement equation, variables become indexed with time and products (for the euro area as a whole) and the suggested specification collapses to the dynamic factor model considered by Reis and Watson (2010) to identify and estimate a 'pure' inflation indicator, which aims at identifying equiproportional changes in all prices.[3]. In particular, they do so by transforming their global factor in a way such that it becomes uncorrelated in the sample with all leads, lags and contemporaneous values of the remaining factors in the model. We address how such an indicator can be computed in the proposed framework and apply it to our data set.

The paper is organized as follows. In section 2, we present a dynamic factor model specified and identified to tackle our three-mode and non-hierarchical data structure of price changes (by products and countries). We also discuss identification alternatives, and how to compute the pure inflation indicator in the context of our model. In section 3, we propose the estimation procedure based on maximum likelihood and the Expectation-Conditional Maximization (ECM) algorithm. In section 4, we conduct a Monte Carlo analysis to assess the small sample performance of the suggested maximum likelihood estimator. In section 5, the empirical application is presented and the main estimation results are discussed. Finally, in section 6 we sum up and make some concluding remarks.

2. Model specification

For the sake of exposition, let us assume that the data is organized in a three-way array $\mathcal{X} \in \mathbb{R}^{T \times I \times J}$ with generic element $x_{t,i,j}$ representing, say, the year-on-year price changes. The first, second and third mode indices t, i and j refer to time periods, to products and to countries, respectively. The time-mode matricization of tensor $\mathcal{X}$ yields the following matrix:

$$\underset{(T \times IJ)}{X} = \begin{bmatrix} \mathcal{X}_{\bullet,\bullet,1} & \cdots & \mathcal{X}_{\bullet,\bullet,j} & \cdots & \mathcal{X}_{\bullet,\bullet,J} \end{bmatrix}$$

3. Reis and Watson (2010) use quarterly data on deflators for 187 detailed categories of personal consumption expenditures in the United States.

with

$$\underset{(T \times I)}{\mathcal{X}_{\bullet,\bullet,j}} = \begin{bmatrix} x_{1,1,j} & \cdots & x_{1,i,j} & \cdots & x_{1,I,j} \\ \vdots & & \vdots & & \vdots \\ x_{t,1,j} & \cdots & x_{t,i,j} & \cdots & x_{t,I,j} \\ \vdots & & \vdots & & \vdots \\ x_{T,1,j} & \cdots & x_{T,i,j} & \cdots & x_{T,I,j} \end{bmatrix}$$

2.1. A three-way factor model

Let us postulate that the data is generated by a factor model with measurement equation:

$$x_{t,i,j} = \kappa_{i,j} + \delta_{1,1} f_{t,1,1} + \sum_{m=2}^{M} \delta_{m,1} \alpha_{i,m} f_{t,m,1} + \sum_{n=2}^{N} \delta_{1,n} \beta_{j,n} f_{t,1,n} +$$

$$+ \sum_{m=2}^{M} \sum_{n=2}^{N} \delta_{m,n} \alpha_{i,m} \beta_{j,n} f_{t,m,n} + u_{t,i,j} \quad (t = 1, \cdots, T; i = 1, \cdots, I; j = 1, \cdots, J) \tag{1}$$

where $\kappa_{i,j}$ are (unknown) time-invariant additive terms, $f_{t,m,n}$ denote latent (i.e. non-observable) time-factors, $\alpha_{i,m}$, $\beta_{j,n}$ and $\delta_{m,n}$ are (unknown) time-invariant multiplicative loading parameters, and $u_{t,i,j}$ stand for residual idiosyncratic components. We assume that the latter may be serial correlated:

$$u_{t,i,j} = \rho_{i,j} u_{t-1,i,j} + \varepsilon_{t,i,j} \quad (t = 1, \cdots, T; i = 1, \cdots, I; j = 1, \cdots, J) \tag{2}$$

where $\rho_{i,j}$ denote (unknown) first order autocorrelation coefficients, with $|\rho_{i,j}| < 1$ for all (i, j), and $\varepsilon_{t,i,j}$ are innovations independent and identically distributed according to

$$\varepsilon_{t,i,j} \sim N(0; \sigma_{i,j}) \quad (t = 1, \cdots, T; i = 1, \cdots, I; j = 1, \cdots, J) \tag{3}$$

with (unknown) variances $\sigma_{i,j}$ strictly positive for all i and j.

The state transition equation of the latent time-factors may be written as:

$$f_{t,m,n} = \sum_{p=1}^{P} \sum_{\tilde{m}=1}^{M} \sum_{\tilde{n}=1}^{N} \gamma_{p,m,n,\tilde{m},\tilde{n}} f_{t-p,\tilde{m},\tilde{n}} + \eta_{t,m,n}$$

$$(t = 1, \cdots, T; m = 1, \cdots, M; n = 1, \cdots, N) \tag{4}$$

where $\gamma_{p,m,n,\tilde{m},\tilde{n}}$ and $\eta_{t,m,n}$ respectively denote (unknown) coefficients and innovations of the vector autoregressive process of order P generating the latent time-factors. We postulate that innovations are independent from the idiosyncratic components, and independent and identically generated over time periods from the following multivariate normal distribution (Ω being a $MN \times MN$ positive definite matrix):

$$\eta_t = [\eta_{t,1,1}\ \eta_{t,2,1} \cdots \eta_{t,m,n} \cdots \eta_{t,M,N}]' \sim N(0;\Omega) \quad (t = 1, \cdots, T)^4 \tag{5}$$

We also assume that time-factors in the initial periods are normal distributed:

$$\mathbf{f}_0 \sim N(\mu; \Psi) \tag{6}$$

where

$$\mathbf{f}_0 = [f_{0,1,1}\ f_{0,2,1} \cdots f_{0,m,n} \cdots f_{0,M,N}\ f_{-1,1,1} \cdots f_{-1,M,N}\ f_{1-P,1,1} \cdots f_{1-P,M,N}]'$$

and μ are column-vectors of dimension MNP, and Ψ is a $(MNP \times MNP)$ positive definite matrix. Both μ and Ψ are unknown.

The total number of time-factors in our model is MN including:

- One global time-factor $\{f_{t,1,1}\}$ with time-invariant loadings $\delta_{1,1}$;
- $(M-1)$ time-factors $\{f_{t,m,1}\}$ $(m = 2, \cdots, M)$, with loadings $\{\delta_{m,1}\alpha_{i,m}\}$ which vary with the different products and are country-invariant;
- $(N-1)$ time-factors $\{f_{t,1,n}\}$ $(n = 2, \cdots, N)$, with loadings $\{\delta_{1,n}\beta_{j,n}\}$ which vary by country and are invariant to products;
- $(M-1)(N-1)$ interaction time-factors $\{f_{t,m,n}\}$ $(m = 2, \cdots, M; n = 2, \cdots, N)$, with loadings $\{\delta_{m,n}\alpha_{i,m}\beta_{j,n}\}$ varying with both product and country.

Measurement equation (1) could have been written as

$$x_{t,i,j} = \kappa_{i,j} + \sum_{m=1}^{M}\sum_{n=1}^{N} \delta_{m,n}\alpha_{i,m}\beta_{j,n}f_{t,m,n} + u_{t,i,j}$$

with

$$\alpha_{i,1} = 1 \quad (i = 1, \cdots, I); \quad \beta_{j,1} = 1 \quad (j = 1, \cdots, J). \tag{7}$$

By taking on board these latter restrictions on the values of $\alpha_{i,1}$ $(i = 1, \cdots, I)$ and $\beta_{j,1}$ $(j = 1, \cdots, J)$, we accept a negative impact on the maximum likelihood when estimating the model, in exchange for much greater parsimony and improved ability to interpret the specification.

As regards parsimony, the smaller M and N are set relative to I and J, respectively, the more parsimonious is the model. Abstracting from identification restrictions to be discussed in the next subsection, the number of parameters making up the factor loadings in measurement equation (1) is $MN + I(M-1) +$

4. Note that η_t is a column-vector in spite of its elements being written with three subindices. Subindex t is shared by all elements and does not impact on the ordering of elements. As for the second and third subindices m and n, they are such that element $\eta_{t,m,n}$ is the $((n-1)M+m)$th element of column-vector η_t (for instance, if $M=4$ and $N=3$, $\eta_{t,4,2}$ denotes the 7th element of η_t). Hereafter,we will follow this notation convention.

$J(N-1)$, whereas the measurement equation of the unrestricted factor model with the same MN time-factors,

$$x_{t,i,j} = \kappa_{i,j} + \sum_{k=1}^{MN} \lambda_{i,j,k} f_{t,k} + u_{t,i,j} \quad (t = 1, \cdots, T; i = 1, \cdots, I; j = 1, \cdots, J)$$

would have $IJMN$ loading parameters $\lambda_{i,j,k}$. In our empirical application we have $I = 55$ and $J = 19$, and we will end up with $M = 3$ and $N = 2$, thus a total of 135 loading parameters, corresponding to just 2.2% of the 6,270 loading parameters in the unrestricted model. The disparity in the numbers increases sharply with I, J, M and N.

Instead of the unrestricted measurement equation, we could have considered the comparison of our specification with that of a dynamic factor model with one global factor, one product-specific factor associated with each product, and one country-specific factor associated with each country. We would have a total of $1 + I + J$ factors (75 in our empirical illustration) and, after taking into account all zero restrictions on the factor loadings, as well as the over-identifying restriction of having the global factor impacting equally on all variables, $1 + 2IJ$ loading parameters (2,091 in our empirical case).

In spite of its parsimony, our measurement equation is rather flexible to capture relative price movements. Indeed, for any (i_1, i_2) or (j_1, j_2), from (1) we get:

$$x_{t,i_1,j} - x_{t,i_2,j} = (\kappa_{i_1,j} - \kappa_{i_2,j}) +$$
$$+ \sum_{m=2}^{M} (\alpha_{i_1,m} - \alpha_{i_2,m}) \left(\delta_{m,1} f_{t,m,1} + \sum_{n=2}^{N} \delta_{m,n} \beta_{j,n} f_{t,m,n} \right) + (u_{t,i_1,j} - u_{t,i_2,j})$$

and

$$x_{t,i,j_1} - x_{t,i,j_2} = (\kappa_{i,j_1} - \kappa_{i,j_2}) +$$
$$+ \sum_{n=2}^{N} (\beta_{j_1,n} - \beta_{j_2,n}) \left(\delta_{1,n} f_{t,1,n} + \sum_{m=2}^{M} \delta_{m,n} \alpha_{i,m} f_{t,m,n} \right) + (u_{t,i,j_1} - u_{t,i,j_2})$$

meaning that inflation differentials between products in the same country, or inflation differentials between different countries for the same product, depend on the corresponding differentials of the specific factor loadings, as well as on the influence of interaction terms (besides being affected by the differences in the respective idiosyncratic components).

In more compact matrix notation, model (1) to (6) may be written as:

$$x_t = \kappa + \Theta(\bar{A}, \bar{B}) \text{diag}(\delta) f_t + u_t \quad (t = 1, \cdots, T) \tag{8}$$

$$f_t = \Gamma \mathbf{f}_{t-1} + \eta_t \quad (t = 1, \cdots, T) \tag{9}$$

$$u_t = \text{diag}(\rho) u_{t-1} + \varepsilon_t \quad (t = 1, \cdots, T) \tag{10}$$

$$\varepsilon_t \sim i.i.d.N(0; \text{diag}(\sigma)) \quad (t = 1, \cdots, T) \tag{11}$$

$$\eta_t \sim i.i.d.N(0;\Omega) \quad (t = 1, \cdots, T) \tag{12}$$

$$E(\varepsilon_t \eta_s') = 0 \quad \forall t, s \tag{13}$$

$$\mathbf{f}_0 \sim N(\mu; \Psi) \tag{14}$$

where $i.i.d.$ stands for independent and identically distributed, and

$$x_t = \begin{bmatrix} x_{t,1,1} & x_{t,2,1} & \cdots & x_{t,i,j} & \cdots & x_{t,I,J} \end{bmatrix}' \quad (IJ \times 1)$$

$$f_t = \begin{bmatrix} f_{t,1,1} & f_{t,2,1} & \cdots & f_{t,m,n} & \cdots & f_{t,M,N} \end{bmatrix}' \quad (MN \times 1)$$

$$\mathbf{f}_t = \begin{bmatrix} f_t' & f_{t-1}' & \cdots & f_{t-P+1}' \end{bmatrix}' \quad (MNP \times 1)$$

$$u_t = \begin{bmatrix} u_{t,1,1} & u_{t,2,1} & \cdots & u_{t,i,j} & \cdots & u_{t,I,J} \end{bmatrix}' \quad (IJ \times 1)$$

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{t,1,1} & \varepsilon_{t,2,1} & \cdots & \varepsilon_{t,i,j} & \cdots & \varepsilon_{t,I,J} \end{bmatrix}' \quad (IJ \times 1)$$

$$\kappa = \begin{bmatrix} \kappa_{1,1} & \kappa_{2,1} & \cdots & \kappa_{i,j} & \cdots & \kappa_{I,J} \end{bmatrix}' \quad (IJ \times 1)$$

$$\Theta(\bar{A}, \bar{B}) = (B \otimes A) \quad (IJ \times MN)^5$$

$$A = \begin{bmatrix} \mathbf{1}_I & \bar{A} \end{bmatrix} \quad (I \times M), \quad B = \begin{bmatrix} \mathbf{1}_J & \bar{B} \end{bmatrix} \quad (J \times N)$$

$$\bar{A} = \begin{bmatrix} \alpha_{1,2} & \cdots & \alpha_{1,m} & \cdots & \alpha_{1,M} \\ \vdots & & \vdots & & \vdots \\ \alpha_{i,2} & \cdots & \alpha_{i,m} & \cdots & \alpha_{i,M} \\ \vdots & & \vdots & & \vdots \\ \alpha_{I,2} & \cdots & \alpha_{Im} & \cdots & \alpha_{I,M} \end{bmatrix} \quad (I \times (M-1))$$

$$\bar{B} = \begin{bmatrix} \beta_{1,2} & \cdots & \beta_{1,n} & \cdots & \beta_{1,N} \\ \vdots & & \vdots & & \vdots \\ \beta_{j,2} & \cdots & \beta_{j,n} & \cdots & \beta_{j,N} \\ \vdots & & \vdots & & \vdots \\ \beta_{J,2} & \cdots & \beta_{J,n} & \cdots & \beta_{J,N} \end{bmatrix} \quad (J \times (N-1))$$

$$\delta = \begin{bmatrix} \delta_{1,1} & \delta_{2,1} & \cdots & \delta_{m,n} & \cdots & \delta_{M,N} \end{bmatrix}' \quad (MN \times 1)$$

$$\rho = \begin{bmatrix} \rho_{1,1} & \rho_{2,1} & \cdots & \rho_{i,j} & \cdots & \rho_{IJ} \end{bmatrix}' \quad (IJ \times 1)$$

$$\sigma = \begin{bmatrix} \sigma_{1,1} & \sigma_{2,1} & \cdots & \sigma_{i,j} & \cdots & \sigma_{I,J} \end{bmatrix}' \quad (IJ \times 1)$$

$$\Gamma = \begin{bmatrix} \Gamma_1 & \cdots & \Gamma_p & \cdots & \Gamma_P \end{bmatrix} \quad (MN \times MNP)$$

5. $(B \otimes A)$ represents the kronecker product of B and A.

$$\underset{(MN \times MN)}{\Gamma_p} = \begin{bmatrix} \gamma_{p,1,1,1,1} & \gamma_{p,1,1,2,1} & \cdots & \gamma_{p,1,1,\tilde{m},\tilde{n}} & \cdots & \gamma_{p,1,1,M,N} \\ \gamma_{p,2,1,1,1} & \gamma_{p,2,1,2,1} & \cdots & \gamma_{p,2,1,\tilde{m},\tilde{n}} & \cdots & \gamma_{p,2,1,M,N} \\ \vdots & \vdots & & \vdots & & \vdots \\ \gamma_{p,m,n,1,1} & \gamma_{p,m,n,2,1} & \cdots & \gamma_{p,m,n,\tilde{m},\tilde{n}} & \cdots & \gamma_{p,m,n,M,N} \\ \vdots & \vdots & & \vdots & & \vdots \\ \gamma_{p,M,N,1,1} & \gamma_{p,M,N,2,1} & \cdots & \gamma_{p,M,N,\tilde{m},\tilde{n}} & \cdots & \gamma_{p,M,N,M,N} \end{bmatrix}$$

and $\mathbf{1}_K$ denotes a K-dimensional column-vector of ones (K being any positive integer).[6]

2.2. Identification

The model specified above is not identified. With the purpose of reinforcing the economic interpretation of loadings and factors, we will consider the following identification restrictions:

- For given weights $w_i^{(\alpha)}$ $(i = 1, \ldots, I)$, such that $w_i^{(\alpha)} \geq 0$ $(i = 1, \cdots, I)$ and $\sum_{i=1}^{I} w_i^{(\alpha)} = 1$, weighted standardization of loading-factors $\alpha_{i,m}$ in the sample, i.e.

$$\sum_{i=1}^{I} w_i^{(\alpha)} \alpha_{i,m} = 0 \quad (m = 2, \cdots, M) \tag{15}$$

  and

$$\sum_{i=1}^{I} w_i^{(\alpha)} \alpha_{i,m}^2 = 1 \quad (m = 2, \cdots, M) \tag{16}$$

- For given weights $w_j^{(\beta)}$ $(j = 1, \ldots, J)$, such that $w_j^{(\beta)} \geq 0$ $(j = 1, \cdots, J)$ and $\sum_{j=1}^{J} w_i^{(\beta)} = 1$, weighted standardization of loading-factors $\beta_{j,m}$ in the sample, i.e.

$$\sum_{j=1}^{J} w_j^{(\beta)} \beta_{j,n} = 0 \quad (n = 2, \cdots, N) \tag{17}$$

  and

$$\sum_{j=1}^{J} w_j^{(\beta)} \beta_{j,n}^2 = 1 \quad (n = 2, \cdots, N) \tag{18}$$

- Normalization of time-factors in the sample, i.e.

$$\frac{1}{T} \sum_{t=1}^{T} f_{t,m,n}^2 = 1 \quad (m = 1, \cdots, M; n = 1, \cdots, N) \tag{19}$$

6. As regards Γ_p, following our notation convention, element $\gamma_{p,m,n,\tilde{m},\tilde{n}}$ belongs to its $((n-1)N + m)$th row and to its $((\tilde{n}-1)N + \tilde{m})$th column.

A simple choice of weights could be $w_i^{(\alpha)} = I^{-1}$ $(i = 1, \cdots, I)$ and $w_j^{(\beta)} = J^{-1}$ $(j = 1, \cdots, J)$. But we will see below that a preferable choice in the case of inflation data consists of setting:

- $w_i^{(\alpha)}$ equal to the weight of the i-th product in the basket for the euro area as a whole; and
- $w_j^{(\beta)}$ equal to the weight of country j in the same basket.

In matrix notation, identifying restrictions (15) to (19) may be expressed equivalently as:

$$\bar{A}' w^{(\alpha)} = \mathbf{0}_{M-1} \tag{20}$$

$$\operatorname{diag}\left(\bar{A}' \operatorname{diag}\left(w^{(\alpha)}\right) \bar{A}\right) = \mathbf{1}_{M-1} \tag{21}$$

$$\bar{B}' w^{(\beta)} = \mathbf{0}_{N-1} \tag{22}$$

$$\operatorname{diag}\left(\bar{B}' \operatorname{diag}\left(w^{(\beta)}\right) \bar{B}\right) = \mathbf{1}_{N-1} \tag{23}$$

$$\operatorname{diag}\left(\frac{1}{T}\sum_{t=1}^{T} f_t f_t'\right) = \mathbf{1}_{MN} \tag{24}$$

where $\mathbf{0}_K$ denotes a K-dimensional column-vector of zeros (K being any positive integer), and

$$w^{(\alpha)} = \left[\begin{array}{ccccccc} w_1^{(\alpha)} & w_2^{(\alpha)} & \cdots & w_i^{(\alpha)} & \cdots & w_I^{(\alpha)} \end{array}\right]' \quad (I \times 1)$$

$$w^{(\beta)} = \left[\begin{array}{ccccccc} w_1^{(\beta))} & w_2^{(\beta)} & \cdots & w_j^{(\beta)} & \cdots & w_J^{(\beta)} \end{array}\right]' \quad (J \times 1)$$

with $w^{(\alpha)} \geq 0$, $w^{(\beta)} \geq 0$, and $\mathbf{1}_I' w^{(\alpha)} = \mathbf{1}_J' w^{(\beta)} = 1$.

By normalizing time-factors and multiplicative loading-factors $\alpha_{i,m}$ and $\beta_{j,n}$, the role and meaning of parameters $\delta_{m,n}$ become more clear. Without affecting the likelihood value, the latter parameters will absorb the global scale effects in the commonalities, freeing time-factors and loading-factors of that role.

For a moment let us consider that the above identifying restrictions were not imposed. Then for each set of time-factors and loading parameters, it is straightforward to confirm that there would be an alternative set of factors and parameters associated with the same likelihood value. Indeed, for each commonality term in the measurement equation, we would have:

$$\delta_{m,n}\alpha_{i,m}\beta_{j,n}f_{t,m,n} = \delta_{m,n}^{*}\alpha_{i,m}^{*}\beta_{j,n}^{*}f_{t,m,n}^{*} \tag{25}$$

for

$$\delta_{m,n}^{*} = \delta_{m,n}\sqrt{\left(\frac{1}{T}\sum_{t=1}^{T} f_{t,m,n}^{2}\right)\left(\sum_{i=1}^{I}\omega_i^{(\alpha)}\alpha_{i,m}^{2}\right)\left(\sum_{j=1}^{J}\omega_j^{(\beta)}\beta_{j,n}^{2}\right)}$$

$$f^*_{t,m,n} = \frac{f_{t,m,n}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T} f^2_{t,m,n}}};\ \alpha^*_{i,m} = \frac{\alpha_{i,m}}{\sqrt{\sum_{i=1}^{I}\omega^{(\alpha)}_i \alpha^2_{i,m}}};\ \beta^*_{j,n} = \frac{\beta_{j,n}}{\sqrt{\sum_{j=1}^{J}\omega^{(\beta)}_j \beta^2_{j,n}}} \tag{26}$$

and the transition equation (4) would also admit the equivalent alternative formulation:

$$f^*_{t,m,n} = \sum_{p=1}^{P}\sum_{\tilde{m}=1}^{M}\sum_{\tilde{n}=1}^{N}\gamma^*_{p,m,n,\tilde{m},\tilde{n}} f^*_{t-p,\tilde{m},\tilde{n}} + \eta^*_{t,m,n}$$

with

$$\gamma^*_{p,m,n,\tilde{m},\tilde{n}} = \gamma_{p,m,n,\tilde{m},\tilde{n}}\sqrt{\frac{\sum_{t=1}^{T} f^2_{t,\tilde{m},\tilde{n}}}{\sum_{t=1}^{T} f^2_{t,m,n}}}$$

$$\Omega^* = \mathrm{Var}\left(\eta^*_t\right) = \Pi^{-1}\Omega\Pi^{-1}$$

and[7]

$$\Pi = \mathrm{diag}\left(\left[\sqrt{\frac{1}{T}\sum_{t=1}^{T} f^2_{t,1,1}}\cdots\sqrt{\frac{1}{T}\sum_{t=1}^{T} f^2_{t,m,n}}\cdots\sqrt{\frac{1}{T}\sum_{t=1}^{T} f^2_{t,M,N}}\right]'\right)$$

Thus, conditions (16), (18) and (19) (or, equivalently, (21), (23) and (24)), by setting particular values for the denominators in (26), resolve the ambiguity in the measurement equation expressed by (25).

There is another source of non-identification in the measurement equation, which restrictions (15) and (17) allow addressing. Let us admit that restriction (15) is not verified (the argument for restriction (17) is similar). Denoting $\sum_{i=1}^{I} w^{(\alpha)}_i \alpha_{i,m}$ by $\tilde{\alpha}_m$, we may express the commonality as:

$$\delta_{1,1}f_{t,1,1} + \sum_{m=2}^{M}\delta_{m,1}\alpha_{i,m}f_{t,m,1} + \sum_{n=2}^{N}\delta_{1,n}\beta_{j,n}f_{t,1,n} + \sum_{m=2}^{M}\sum_{n=2}^{N}\delta_{m,n}\alpha_{i,m}\beta_{j,n}f_{t,m,n}$$

$$= \delta_{1,1}f_{t,1,1} + \sum_{m=2}^{M}\delta_{m,1}\left(\alpha_{i,m} - \tilde{\alpha}_m + \tilde{\alpha}_m\right)f_{t,m,1} + \sum_{n=2}^{N}\delta_{1,n}\beta_{j,n}f_{t,1,n} +$$

$$+ \sum_{m=2}^{M}\sum_{n=2}^{N}\delta_{m,n}\left(\alpha_{i,m} - \tilde{\alpha}_m + \tilde{\alpha}_m\right)\beta_{j,n}f_{t,m,n} =$$

$$= \delta^*_{1,1}f^*_{t,1,1} + \sum_{m=2}^{M}\delta_{m,1}\alpha^*_{i,m}f_{t,m,1} + \sum_{n=2}^{N}\delta^*_{1,n}\beta_{j,n}f^*_{t,1,n} +$$

7. Let C and c be a square matrix $(K \times K)$ and the K-dimensional column-vector obtained stacking all its diagonal elements, respectively. Hereafter $\mathrm{diag}()$ may be used with two different meanings depending on the context: (i) $\mathrm{diag}(C) = c$; or (ii) $\mathrm{diag}(c)$ as the diagonal matrix with diagonal elements taken from c.

$$+\sum_{m=2}^{M}\sum_{n=2}^{N}\delta_{m,n}\alpha^*_{i,m}\beta_{j,n}f_{t,m,n}$$

where

$$\alpha^*_{i,m}=\alpha_{i,m}-\tilde{\alpha}_m \quad (i=1,\cdots,I;m=2,\cdots,M)$$

$$\delta^*_{1,n}f^*_{t,1,n}=\sum_{n=1}^{N}\left[\delta_{1,n}f_{t,1,n}+\sum_{m=2}^{M}\tilde{\alpha}_m\delta_{m,n}f_{t,m,n}\right] \quad (t=1,\cdots,T;n=1,\cdots,N)$$

In order to separate $\delta^*_{1,n}$ from $f^*_{t,1,n}$ in the latter identities, for each n we need to normalize $\{\delta^*_{1,n}f^*_{t,1,n}\}_{t=1,\cdots,T}$. The normalized output is $\{f^*_{t,1,n}\}_{t=1,\cdots,T}$ whereas the normalizing factor becomes $\delta^*_{1,n}$.[8]

Let

$$\mathring{x}_{t,i}=\sum_{j=1}^{J}w_j^{(\beta)}x_{t,i,j}; \quad \tilde{x}_{t,j}=\sum_{i=1}^{I}w_i^{(\alpha)}x_{t,i,j}; \quad \bar{\bar{x}}_t=\sum_{i=1}^{I}\sum_{j=1}^{J}w_i^{(\alpha)}w_j^{(\beta)}x_{t,i,j}$$

$$\mathring{\kappa}_i=\sum_{j=1}^{J}w_j^{(\beta)}\kappa_{i,j}; \quad \tilde{\kappa}_j=\sum_{i=1}^{I}w_i^{(\alpha)}\kappa_{i,j}; \quad \bar{\bar{\kappa}}=\sum_{i=1}^{I}\sum_{j=1}^{J}w_i^{(\alpha)}w_j^{(\beta)}\kappa_{i,j}$$

$$\mathring{u}_{t,i}=\sum_{j=1}^{J}w_j^{(\beta)}u_{t,i,j}; \quad \tilde{u}_{t,j}=\sum_{i=1}^{I}w_i^{(\alpha)}u_{t,i,j}; \quad \bar{\bar{u}}_t=\sum_{i=1}^{I}\sum_{j=1}^{J}w_i^{(\alpha)}w_j^{(\beta)}u_{t,i,j}$$

Over-identifying restrictions (7), together with identifying restrictions (15) to (17), give the model convenient aggregation properties which become apparent when we collapse either the product mode of the data, or the country mode, or both. Indeed, when averaging (using the appropriate weights) over the country mode $(j=1,\cdots,J)$, the product mode $(i=1,\cdots,I)$, and both, we get respectively:

$$\mathring{x}_{t,i}=\mathring{\kappa}_i+\delta_{1,1}f_{t,1,1}+\sum_{m=2}^{M}\delta_{m,1}\alpha_{i,m}f_{t,m,1}+\mathring{u}_{t,i} \tag{27}$$

$$\tilde{x}_{t,j}=\tilde{\kappa}_j+\delta_{1,1}f_{t,1,1}+\sum_{n=2}^{N}\delta_{1,n}\beta_{j,n}f_{t,1,n}+\tilde{u}_{t,j} \tag{28}$$

$$\bar{\bar{x}}_t=\bar{\bar{\kappa}}+\delta_{1,1}f_{t,1,1}+\bar{\bar{u}}_t \tag{29}$$

One direct implication of (29) is that $\{\pi_t^{Global}\}_{t=1,\cdots,T}$, defined as

$$\pi_t^{Global}=\bar{\bar{\kappa}}+\delta_{1,1}f_{t,1,1} \quad (t=1,\cdots,T) \tag{30}$$

8. As time-factors $\{f_{t,1,n}\}_{t=1,\cdots,T}$ for $n=1,\cdots,N$ are transformed into $\{f^*_{t,1,n}\}_{t=1,\cdots,T}$, the re-parametrization also requires that a corresponding transformation be performed to Γ and Ω, so that the likelihood value remains unchanged.

can be interpreted as a global inflation indicator for the euro area as a whole, which abstracts from the inflation differentials associated with relative price changes both across products and across countries, as well as from idiosyncratic price movements.

It should be remarked that the global time-factor $\{f_{t,1,1}\}_{t=1,\cdots,T}$, the product-relative-price time-factors $\{f_{t,m,1}\}_{t=1,\cdots,T}$ $(m = 1,\cdots,M)$ and the country-relative-prices time-factors $\{f_{t,1,n}\}_{t=1,\cdots,T}$ $(n = 1,\cdots,N)$ will not be orthogonal in the sample. Here we are using the term orthogonality in the sense of zero contemporaneous linear correlation. However, with the proposed identification strategy, we argue that the most relevant interpretations remain valid in spite of non-orthogonality. Indeed, when doubly aggregating the measurement equation over both product and country data modes, as in (29), no factors remain to 'explain' the inflation for the euro area $\{\bar{\bar{x}}_t\}_{t=1,\cdots,T}$ but the global time-factor $\{f_{t,1,1}\}_{t=1,\cdots,T}$.[9] The disappearance of all time-factors but $\{f_{t,1,1}\}_{t=1,\cdots,T}$ from (29) when aggregating the measurement equation across both modes justifies the interpretation of $\left\{\pi_t^{Global}\right\}_{t=1,\cdots,T}$ as a global inflation indicator for the euro area as a whole.

2.3. Alternative identifications

It is not possible to impose identifying restrictions (i.e. restrictions which do not negatively impact on the likelihood function) ensuring orthogonality in the sample between all time-factors without forfeiting some of the convenient interpretations and aggregation properties of the identification above proposed. In this subsection, we discuss the pros and cons of two alternative identification strategies. For ease of exposition, we will limit the discussion of re-parameterizations to the partially aggregated measurement equation (27), in which the observables are addressed by only two-dimensional indices (time and products), the country mode having been canceled out by summing with weights w_j^{β} $(j = 1,\cdots,J)$.

Let us rewrite equation (27) in matrix notation, both for the single generic time period t:

$$\mathring{x}_t = \mathring{\kappa} + \delta_{1,1}\mathbf{1}_I f_{t,1,1} + \bar{A}\mathrm{diag}\left(\delta_{\bullet,1}\right) f_{t,\bullet,1} + \mathring{u}_t \quad (t = 1,\cdots,T) \tag{31}$$

and jointly for all $t = 1 \cdots, T$:

$$\mathring{X} = \mathbf{1}_T\mathring{\kappa}' + \delta_{1,1}f_{1,1}\mathbf{1}_I' + F_{\bullet,1}\mathrm{diag}\left(\delta_{\bullet,1}\right)\bar{A}' + \mathring{U} \tag{32}$$

where

$$\mathring{x}_t = \left[\begin{array}{ccccc} \mathring{x}_{t,1} & \cdots & \mathring{x}_{t,i} & \cdots & \mathring{x}_{t,I} \end{array}\right]' \quad (I \times 1)$$

9. Actually, $\{\bar{\bar{x}}_t\}_{t=1,\cdots,T}$ is a proxy to the actual change of the euro area consumer price index because the latter is based on weights for products which differ slightly from country to country while our proxy uses the same product weights $w_i^{(\alpha)}$ $(i = 1,\cdots,I)$ (the product weights for the euro area as whole) for all countries j $(j = 1,\cdots,J)$.

$$\mathring{X} = \begin{bmatrix} \mathring{x}_1' \\ \vdots \\ \mathring{x}_t' \\ \vdots \\ \mathring{x}_T' \end{bmatrix} = \begin{bmatrix} \mathring{x}_{1,1} & \cdots & \mathring{x}_{1,i} & \cdots & \mathring{x}_{1,I} \\ \vdots & & \vdots & & \vdots \\ \mathring{x}_{t,1} & \cdots & \mathring{x}_{t,i} & \cdots & \mathring{x}_{t,I} \\ \vdots & & \vdots & & \vdots \\ \mathring{x}_{T,1} & \cdots & \mathring{x}_{T,i} & \cdots & \mathring{x}_{T,I} \end{bmatrix} \quad (T \times I)$$

$$\mathring{u}_t = \begin{bmatrix} \mathring{u}_{t,1} & \cdots & \mathring{u}_{t,i} & \cdots & \mathring{u}_{t,I} \end{bmatrix}' \quad (I \times 1)$$

$$\mathring{U} = \begin{bmatrix} \mathring{u}_1' \\ \vdots \\ \mathring{u}_t' \\ \vdots \\ \mathring{u}_T' \end{bmatrix} = \begin{bmatrix} \mathring{u}_{1,1} & \cdots & \mathring{u}_{1,i} & \cdots & \mathring{u}_{1,I} \\ \vdots & & \vdots & & \vdots \\ \mathring{u}_{t,1} & \cdots & \mathring{u}_{t,i} & \cdots & \mathring{u}_{t,I} \\ \vdots & & \vdots & & \vdots \\ \mathring{u}_{T,1} & \cdots & \mathring{u}_{T,i} & \cdots & \mathring{u}_{T,I} \end{bmatrix} \quad (T \times I)$$

$$f_{1,1} = \begin{bmatrix} f_{1,1,1} & \cdots & f_{t,1,1} & \cdots & f_{T,1,1} \end{bmatrix}' \quad (T \times 1)$$

$$f_{t,\bullet,1} = \begin{bmatrix} f_{t,2,1} & \cdots & f_{t,m,1} & \cdots & f_{t,M,1} \end{bmatrix}' \quad ((M-1) \times 1)$$

$$F_{\bullet,1} = \begin{bmatrix} f_{1,\bullet,1}' \\ \vdots \\ f_{t,\bullet,1}' \\ \vdots \\ f_{T,\bullet,1}' \end{bmatrix} = \begin{bmatrix} f_{1,2,1} & \cdots & f_{1,m,1} & \cdots & f_{M,1} \\ \vdots & & \vdots & & \vdots \\ f_{t,2,1} & \cdots & f_{t,m,1} & \cdots & f_{M,1} \\ \vdots & & \vdots & & \vdots \\ f_{T,2,1} & \cdots & f_{T,m,1} & \cdots & f_{M,1} \end{bmatrix} \quad (T \times (M-1))$$

$$\mathring{\kappa} = \begin{bmatrix} \mathring{\kappa}_1 & \cdots & \mathring{\kappa}_i & \cdots & \mathring{\kappa}_I \end{bmatrix}' \quad (I \times 1)$$

and

$$\delta_{\bullet,1} = \begin{bmatrix} \delta_{2,1} & \cdots & \delta_{m,1} & \cdots & \delta_{M,1} \end{bmatrix}' \quad ((M-1) \times 1)$$

We obtain our first re-parametrization by substituting $\{f_{1,1}^{\perp}\}_{t=1,\cdots,T}$ for $\{f_{1,1}\}_{t=1,\cdots,T}$, such that $T^{-1}F_{\bullet,1}'f_{1,1}^{\perp} = 0$, and ensuring the necessary adjustment of the loadings associated to factors $F_{\bullet,1}$ so that the commonalities remain unchanged. Denoting $\mathbf{I}_K$ the $(K \times \ K)$ identity matrix, for any positive integer K, we have

$$\mathring{X} = \mathbf{1}_T\mathring{\kappa}' + \delta_{1,1}f_{1,1}^{\perp}\mathbf{1}_I' + F_{\bullet,1}\bar{A}^{\perp'} + \mathring{U}$$

with

$$f_{1,1}^{\perp} = \left[\mathbf{I}_T - F_{\bullet,1}\left(F_{\bullet,1}'F_{\bullet,1}\right)^{-1}F_{\bullet,1}'\right]f_{1,1} \quad (T \times 1)$$

and

$$\bar{A}^{\perp} = \bar{A}\mathrm{diag}\left(\delta_{\bullet,1}\right) + \delta_{1,1}\mathbf{1}_I f_{1,1}'F_{\bullet,1}\left(F_{\bullet,1}'F_{\bullet,1}\right)^{-1} \quad (I \times (M-1))$$

Given $f_{1,1}$, in practical terms $f_{1,1}^{\perp}$ corresponds to the residuals vector of the linear regression, estimated by ordinary least squares,

$$f_{t,1,1} = \theta' f_{t,\bullet,1} + f_{t,1,1}^{\perp} \quad (t = 1, \cdots, T)$$

Compared to the identification laid out in the previous subsection, the advantage of this re-parameterization is that it makes the transformed global factor $\left\{f^{\perp}_{t,1,1}\right\}_{t=1,\cdots,T}$ contemporaneously uncorrelated with the remaining factors. Its drawback is that, unlike in the case of the identification discussed in subsection 2.2, when computing the weighted average of equation (32) over the products mode using weights $w^{(\alpha)}_i$ $(i = 1, \cdots, I)$, the effect of factors $\{F_{\bullet,1}\}_{t=1,\cdots,T}$ does not cancel out anymore because in general $\bar{A}^{\perp\prime} w^{(\alpha)} \neq \mathbf{0}_{M-1}$. This means that the part of the commonality associated with factors $\{f_{t,\bullet,1}\}_{t=1,\cdots,T}$ cannot be interpreted anymore as conveying the effect of relative price movements and, as a consequence, neither should factor $\left\{f^{\perp}_{t,1,1}\right\}_{t=1,\cdots,T}$ be viewed as expressing global price movements.

The second possible re-parameterization is similar to the previous one except that it is based on the orthogonalization of factors $\{f_{t,\bullet,1}\}_{t=1,\cdots,T}$ relative to $\{f_{t,1,1}\}_{t=1,\cdots,T}$ instead of the other way around:[10]

$$F^{\perp}_{\bullet,1} = \left[\mathbf{I}_T - T^{-1} f_{1,1} f'_{1,1}\right] F_{\bullet,1} \quad (T \times (M-1))$$

This re-parameterization leaves unchanged the loadings of transformed factors $\left\{f^{\perp}_{t,\bullet,1}\right\}_{t=1,\cdots,T}$, as well as the global factor $\{f_{t,1,1}\}_{t=1,\cdots,T}$, but modifies the loadings associated with the latter. The re-parameterized measurement equation becomes:

$$\mathring{x}_t = \mathring{\kappa} + \delta^{\perp}_{1,1} f_{t,1,1} + \bar{A} \mathrm{diag}\left(\delta_{\bullet,1}\right) f^{\perp}_{t,\bullet,1} + \mathring{u}_t \quad (t = 1, \cdots, T) \tag{33}$$

where $\delta^{\perp}_{1,1}$ denotes a column-vector instead of a scalar:

$$\delta^{\perp}_{1,1} = \delta_{1,1} \mathbf{1}_I + \bar{A} \mathrm{diag}\left(\delta_{\bullet,1}\right) \left(T^{-1} F'_{\bullet,1} f_{1,1}\right) \quad (I \times 1)$$

As the loadings associated with the transformed factors $\left\{f^{\perp}_{t,\bullet,1}\right\}_{t=1,\cdots,T}$ remain unchanged in relation to our original parameterization, these factors are canceled out when time-observations are aggregated using weights $w^{(\alpha)}_i$ $(i = 1, \cdots, I)$. However, the loadings associated with the global factor $\{f_{t,1,1}\}_{t=1,\cdots,T}$ are no more constant across products, undermining the interpretation of the global factor as price changes with an equiproportional effect on all products. Notwithstanding this non-uniformity of the loadings associated with the global factor after the re-parameterization, it is interesting to remark that, when averaging out the products mode from (33) with weights $w^{(\alpha)}_i$ $(i = 1, \cdots, I)$, we get the same doubly aggregate specification as in (30), i.e. the global inflation indicator $\left\{\pi^{Global}_t\right\}_{t=1,\cdots,T}$ remains unchanged by this re-parameterization due to the fact that $w^{(\alpha)\prime} \delta^{\perp}_{1,1} = \delta_{1,1}$.

10. Continuing to consider the parameterization presented in subsection 2.2 as the base case, in the following expression we make use of restriction (19) (or (24)) that $T^{-1} f'_{1,1} f_{1,1} = 1$.

2.4. Pure inflation

Reis and Watson (2010), after estimating a dynamic factor model for two-dimensional indexed data with measurement equation similar to (31) and (32), transform the global factor in such a way that it becomes orthogonal to the remaining factors not only contemporaneously but also at all their leads and lags:

$$f_{t,1,1}^{(p)} = E\left[f_{t,1,1} \middle| \{f_{\tau,\bullet,1}\}_{\tau=1,\cdots,T} \right] \quad (t = 1, \cdots, T)$$

In the notation of our model, and with $\bar{\bar{\kappa}}$ as in (29), the pure inflation indicator is defined as:

$$\pi_t^{(p)} = \bar{\bar{\kappa}} + \delta_{1,1} f_{t,1,1}^{(p)} \quad (t = 1, \cdots, T) \tag{34}$$

Reis and Watson (2010) named it so because of its lack of correlation at all leads and lags with the remaining factors and because loadings in (34) are all identical.[11]

In order to calculate conditional expectations, a model of the joint dynamics of $\{f_{t,1,1}\}_{t=1,\cdots,T}$ and $\{f_{t,\bullet,1}\}_{t=1,\cdots,T}$ is required. After estimating the dynamic factor model and obtaining the smoothed factors and estimated parameters, they propose the Kalman smoother applied to the state-space system consisting of the same state transition vector autoregression as the estimated dynamic factor model. The measurement equations are set as noiseless identities between the previously obtained smoothed estimates of the factors associated with relative price movements, which become the observables, and the corresponding state variables.

Denoting $f_{t,\bullet,1|\tau} = E(f_{t,\bullet,1}|x_1, \cdots, x_\tau)$ the $M-1$ smoothed factors associated with relative price movements across products obtained from the estimation of our dynamic factor model (8) to (14) with identifying restrictions (20) to (24), the system in state-space representation required to estimate pure inflation as in Reis and Watson (2010) is the following, for $t = 1, \cdots, T$:

$$f_{t,\bullet,1|T} = \begin{bmatrix} \mathbf{0}_{M-1} & \mathbf{I}_{M-1} & \mathbf{0}_{M-1} & \cdots & \mathbf{0}_{M-1} \end{bmatrix} f_t$$

$$f_t = \Gamma \mathbf{f}_{t-1} + \eta_t$$

The orthogonalized factor $f_{t,1,1}^{(p)}$ in (34) is constructed as the smoothed estimate of the global factor $\{f_{t,1,1}\}_{t=1,\cdots,T}$ obtained from our dynamic factor model minus the smoothed estimate of the same vector from the supplementary state-space system.

Note that, in the spirit of the collapsed measurement equation (31) and (32), where the country mode has been averaged out, in the supplementary state space system we only considered as observables the smoothed versions of the $M-1$ factors conveying relative price changes across products. A specification leading

11. Actually, Reis and Watson's dynamic factor model neither does admit a loading parameter associated with the global factor nor imposes the normalization of the latter, but our model does both and thus we had to include $\delta_{1,1}$ in the latter definition.

to an alternative pure inflation indicator may be obtained by including in the supplementary state-space system as observables the smoothed versions of all time-factors but the global factor (because all of them reflect relative price changes, although of different nature), i.e.

$$f_{t,\bullet,\bullet|T} = \left[\begin{array}{ccccc} \mathbf{0}_{MN-1} & \mathbf{I}_{MN-1} & \mathbf{0}_{MN-1} & \cdots & \mathbf{0}_{MN-1} \end{array} \right] f_t$$

$$f_t = \Gamma \mathbf{f}_{t-1} + \eta_t$$

where

$$f_{t,\bullet,\bullet|T} = \left[\begin{array}{cccccc} f_{t,2,1|T} & f_{t,3,1|T} & \cdots & f_{t,m,n|T} & \cdots & f_{t,M,N|T} \end{array} \right]' \; ((MN-1) \times 1)$$

With three-dimensional indexed data, any of the two versions of the supplementary state space system is *a priori* as warranted as the other. However, in practice, they may deliver quite different results for the pure inflation indicator, as it is the case in our empirical application. The indicator tends to become almost flat when conditioning the global factor on all remaining factors instead of only on a subset of relative price factors. One should note that when controlling for all leads and lags of the factors associated with relatives prices, as the dimension of the vector space spanned by all these leads and lags tends to expand very fast when the number of factors increases, it leaves little to be identified as pure inflation.

3. Model estimation

The estimation of the proposed model is performed by maximum likelihood using the ECM algorithm (Meng and Durbin, 1993), which is a generalization of the EM algorithm.

The EM algorithm consists of iterating an expectation step (E-step) and a maximization step (M-step) until convergence, i.e. until the improvement in the log-likelihood function in consecutive steps is smaller than some tolerance level. When applied to a dynamic factor model, given a set of values for the model parameters, the E-step uses the Kalman smoother to compute the first- and second-order moments of the time-factors. Having obtained the estimated factor moments, the M-step corresponds to maximizing with respect to model parameters the expected value of the joint likelihood of observables $\{x_t\}$ and time-factors $\{f_t\}$.

When the set of model parameters may be partitioned into subsets, and it is simpler to maximize the expected log-likelihood conditional on observables with respect to each subset of parameters at a time, with all other parameters held fixed, the ECM algorithm replaces each M-step of the EM algorithm by a sequence of conditional maximization steps (CM-steps) intercalated by E-steps. In other words, given an initial guess for all parameters, the E-step is performed, followed by the CM-step regarding a first subset of the parameters partition. Then another E-step is carried on, followed by a CM-step regarding a second subset of parameters, and so on until all parameter subsets are exhausted. The algorithm continues by repeating the cycle of E-steps followed by CM-steps until convergence is achieved.

As in the case of the EM algorithm, and under essentially the same conditions, each pair (E-step,CM-step) of the ECM algorithm monotonically increases the value of the likelihood function, and the sequence converges to a stationary point of the latter function.

3.1. Joint log likelihood of the complete data and its expectation conditional on observables

In order to implement the ECM algorithm, as mentioned above, first we need to obtain the joint log likelihood function of the complete data (i.e. observables and latent time-factors) and calculate its expectation conditional on the observables.

Let $\varphi(z;\theta)$ denote a generic probability density function with arguments z and parameters θ. The joint likelihood of the complete data for our model is simple to obtain:

$$\mathcal{L}(\mathbf{f}_0, f_1, \cdots, f_t, \cdots, f_T, x_1, \cdots, x_t, \cdots, x_T; \bar{A}, \bar{B}, \delta, \kappa, \rho, \sigma, \Gamma, \Omega, \mu, \Psi) =$$

$$= \ln\varphi(\mathbf{f}_0; \mu, \Psi) + \sum_{t=1}^{T} \ln\varphi(f_t|\mathbf{f}_0, f_1, \cdots, f_{t-1}; \Gamma, \Omega) + \ln\varphi(x_1|f_1; \bar{A}, \bar{B}, \delta, \kappa, \rho, \sigma) +$$

$$+ \sum_{t=2}^{T} \ln\varphi(x_t|x_1, \cdots, x_{t-1}, f_1, f_2, \cdots, f_t; \bar{A}, \bar{B}, \delta, \kappa, \rho, \sigma) \propto$$

$$\propto \frac{1}{2}\ln\left[\det(\Psi^{-1})\right] - \frac{1}{2}(\mathbf{f}_0 - \mu)'\,\Psi^{-1}(\mathbf{f}_0 - \mu) + \frac{T}{2}\ln\left[\det(\Omega^{-1})\right] +$$

$$-\frac{1}{2}\sum_{t=1}^{T}(f_t - \Gamma\mathbf{f}_{t-1})'\Omega^{-1}(f_t - \Gamma\mathbf{f}_{t-1}) + \frac{1}{2}\sum_{i=1}^{I}\sum_{j=1}^{J}\ln(1-\rho_{i,j}^2) + \frac{T}{2}\sum_{i=1}^{I}\sum_{j=1}^{J}\ln(\sigma_{i,j}^{-1}) +$$

$$-\frac{1}{2}\left[x_1 - \kappa - \Theta(\bar{A}, \bar{B})\mathrm{diag}(\delta)f_1\right]'\left[I_{IJ} - \mathrm{diag}(\rho)^2\right]\mathrm{diag}(\sigma)^{-1}.$$

$$.\left[x_1 - \kappa - \Theta(\bar{A}, \bar{B})\mathrm{diag}(\delta)f_1\right] - \frac{1}{2}\sum_{t=2}^{T}\left\{\left[x_t - \kappa - \Theta(\bar{A}, \bar{B})\mathrm{diag}(\delta)f_t\right] +\right.$$

$$\left. -\mathrm{diag}(\rho)\left[x_{t-1} - \kappa - \Theta(\bar{A}, \bar{B})\mathrm{diag}(\delta)f_{t-1}\right]\right\}'\mathrm{diag}(\sigma)^{-1}.$$

$$.\left\{\left[x_t - \kappa - \Theta(\bar{A}, \bar{B})\mathrm{diag}(\delta)f_t\right] - \mathrm{diag}(\rho)\left[x_{t-1} - \kappa - \Theta(\bar{A}, \bar{B})\mathrm{diag}(\delta)f_{t-1}\right]\right\} \tag{35}$$

where $\propto$ means ‘equal up to a constant’.

Taking conditional expectations yields:

$$\ell(x_1, \cdots, x_t, \cdots, x_T; \bar{A}, \bar{B}, \delta, \kappa, \rho, \sigma, \Gamma, \Omega, \mu, \Psi) =$$

$$= E\left[\mathcal{L}(\mathbf{f}_0, f_1, \cdots, f_T, x_1, \cdots, x_T; \bar{A}, \bar{B}, \delta, \kappa, \rho, \sigma, \Gamma, \Omega, \mu, \Psi)|x_1, \cdots, x_T\right] \propto$$

$$
\begin{aligned}
&\propto \frac{1}{2}\ln\left[\det\left(\Psi^{-1}\right)\right]-\frac{1}{2}\mathrm{tr}\left\{\Psi^{-1}\left[\mathbf{V}_{0,0|T}+(\mathbf{f}_{0|T}-\mu)(\mathbf{f}_{0|T}-\mu)'\right]\right\}+\\
&+\frac{T}{2}\ln\left[\det(\Omega^{-1})\right]-\frac{T}{2}\mathrm{tr}\left[\Omega^{-1}\left(\bar{F}_{0,0}+\Gamma\bar{\mathbf{F}}_{\bullet,\bullet}\Gamma'-\bar{\mathbf{F}}_{0,\bullet}\Gamma'-\Gamma\bar{\mathbf{F}}'_{0,\bullet}\right)\right]+\\
&+\frac{1}{2}\sum_{i=1}^{I}\sum_{j=1}^{J}\ln(1-\rho_{i,j}^2)+\frac{T}{2}\sum_{i=1}^{I}\sum_{j=1}^{J}\ln(\sigma_{i,j}^{-1})+\\
&-\frac{T-1}{2}\mathrm{tr}\Bigg\{\mathrm{diag}(\sigma)^{-1}\Bigg[\left[\bar{\mathbb{X}}_{0,0}+\frac{T}{T-1}\left(\bar{x}_{0,0}-\kappa\right)\left(\bar{x}_{0,0}-\kappa\right)'\right]+\mathrm{diag}(\rho)^2\Big[\bar{\mathbb{X}}_{1,1}+\\
&+\frac{T-2}{T-1}\left(\bar{x}_{1,1}-\kappa\right)\left(\bar{x}_{1,1}-\kappa\right)'\Big]-2\mathrm{diag}(\rho)\left[\bar{\mathbb{X}}_{0,1}+\left(\bar{x}_{0,1}-\kappa\right)\left(\bar{x}_{1,0}-\kappa\right)'\right]+\\
&+\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{\mathbb{F}}_{0,0}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+\\
&+\mathrm{diag}(\rho)^2\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{\mathbb{F}}_{1,1}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+\\
&-2\mathrm{diag}(\rho)\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{\mathbb{F}}_{0,1}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+\\
&-2\left[\bar{\mathbb{W}}_{0,0}+\left(\bar{x}_{0,0}-\kappa\right)\bar{f}'_{0,0}\right]\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+\\
&-2\mathrm{diag}(\rho)^2\left[\bar{\mathbb{W}}_{1,1}+\left(\bar{x}_{1,1}-\kappa\right)\bar{f}'_{1,1}\right]\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+\\
&+2\mathrm{diag}(\rho)\left[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+\left(\bar{x}_{0,1}-\kappa\right)\bar{f}'_{1,0}+\left(\bar{x}_{1,0}-\kappa\right)\bar{f}'_{0,1}\right]\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'\Bigg]\Bigg\}
\end{aligned}
\tag{36}
$$

where we are resorting to the following additional notation:

$$f_{t|\tau}=E\left(f_t|x_1,\cdots,x_\tau\right)$$

$$V_{t-q_1,t-q_2|\tau}=E\left[\left(f_{t-q_1}-f_{t-q_1|\tau}\right)\left(f_{t-q_2}-f_{t-q_2|\tau}\right)'|x_1,\cdots,x_\tau\right]$$

$$\bar{F}_{q_1,q_2}=\frac{1}{T}\sum_{t=1}^{T}E\left(f_{t-q_1}f'_{t-q_2}|x_1,\cdots,x_T\right)=\frac{1}{T}\sum_{t=1}^{T}\left(V_{t-q_1,t-q_2|T}+f_{t-q_1|T}f'_{t-q_2|T}\right)$$

$$\mathbf{f}_{t|\tau}=E\left(\mathbf{f}_t|x_1,\cdots,x_\tau\right)=\begin{bmatrix} f'_{t|\tau} & f'_{t-1|\tau} & \cdots & f'_{t-P+1|\tau}\end{bmatrix}'$$

$$\mathbf{V}_{0,0|T}=E\left[\left(\mathbf{f}_0-\mathbf{f}_{0|T}\right)\left(\mathbf{f}_0-\mathbf{f}_{0|T}\right)'|x_1,\cdots,x_T\right]$$

$$\bar{\mathbf{F}}_{\bullet,\bullet}=\begin{bmatrix}\bar{F}_{1,1} & \cdots & \bar{F}_{1,p} & \cdots & \bar{F}_{1,P}\\ \vdots & & \vdots & & \vdots\\ \bar{F}_{p,1} & \cdots & \bar{F}_{p,p} & \cdots & \bar{F}_{p,P}\\ \vdots & & \vdots & & \vdots\\ \bar{F}_{P,1} & \cdots & \bar{F}_{P,p} & \cdots & \bar{F}_{P,P}\end{bmatrix}$$

$$\bar{\mathbf{F}}_{0,\bullet}=\begin{bmatrix}\bar{F}_{0,1} & \cdots & \bar{F}_{0,p} & \cdots & \bar{F}_{0,P}\end{bmatrix}$$

$$\bar{f}_{0,0} = \frac{1}{T-1}\sum_{t=1}^{T} f_{t|T} \quad ; \quad \bar{f}_{1,1} = \frac{1}{T-1}\sum_{t=2}^{T-1} f_{t|T}$$

$$\bar{f}_{0,1} = \frac{1}{T-1}\sum_{t=2}^{T} f_{t|T} \quad ; \quad \bar{f}_{1,0} = \frac{1}{T-1}\sum_{t=1}^{T-1} f_{t|T} = \frac{1}{T-1}\sum_{t=2}^{T} f_{t-1|T}$$

$$\bar{\mathbb{F}}_{0,0} = \frac{1}{T-1}\sum_{t=1}^{T}\left(V_{t,t|T} + f_{t|T}f'_{t|T}\right) \; ; \; \bar{\mathbb{F}}_{1,1} = \frac{1}{T-1}\sum_{t=2}^{T-1}\left(V_{t,t|T} + f_{t|T}f'_{t|T}\right)$$

$$\bar{\mathbb{F}}_{0,1} = \frac{1}{T-1}\sum_{t=2}^{T}\left(V_{t,t-1|T} + f_{t|T}f'_{t-1|T}\right)$$

$$\bar{x}_{0,0} = \frac{1}{T}\sum_{t=1}^{T} x_t \quad ; \quad \bar{x}_{1,1} = \frac{1}{T-1}\sum_{t=2}^{T-1} x_t$$

$$\bar{x}_{0,1} = \frac{1}{T-1}\sum_{t=2}^{T} x_t \quad ; \quad \bar{x}_{1,0} = \frac{1}{T-1}\sum_{t=1}^{T-1} x_t = \frac{1}{T-1}\sum_{t=2}^{T} x_{t-1}$$

$$\bar{\mathbb{X}}_{0,0} = \frac{1}{T-1}\sum_{t=1}^{T}\left(x_t - \bar{x}_{0,0}\right)\left(x_t - \bar{x}_{0,0}\right)'$$

$$\bar{\mathbb{X}}_{1,1} = \frac{1}{T-1}\sum_{t=2}^{T-1}\left(x_t - \bar{x}_{1,1}\right)\left(x_t - \bar{x}_{1,1}\right)'$$

$$\bar{\mathbb{X}}_{0,1} = \frac{1}{T-1}\sum_{t=2}^{T}\left(x_t - \bar{x}_{0,1}\right)\left(x_{t-1} - \bar{x}_{1,0}\right)'$$

$$\bar{\mathbb{W}}_{0,0} = \frac{1}{T-1}\sum_{t=1}^{T}\left(x_t - \bar{x}_{0,0}\right)f'_{t|T} \quad ; \quad \bar{\mathbb{W}}_{1,1} = \frac{1}{T-1}\sum_{t=2}^{T-1}\left(x_t - \bar{x}_{1,1}\right)f'_{t|T}$$

$$\bar{\mathbb{W}}_{0,1} = \frac{1}{T-1}\sum_{t=2}^{T}\left(x_t - \bar{x}_{0,1}\right)f'_{t-1|T} \quad ; \quad \bar{\mathbb{W}}_{1,0} = \frac{1}{T-1}\sum_{t=2}^{T}\left(x_{t-1} - \bar{x}_{1,0}\right)f'_{t|T}$$

3.2. First order conditions

The first order conditions for the maximization of

$$\ell(x_1, \cdots, x_t, \cdots, x_T; \bar{A}, \bar{B}, \delta, \kappa, \rho, \sigma, \Omega, \mu, \Psi)$$

with respect to the parameters $(\bar{A}, \bar{B}, \delta, \kappa, \rho, \sigma, \Omega, \mu, \Psi)$ are the backbone of the CM-steps of the ECM algorithm. The calculations leading to the following conditions are detailed in Annex A.1.[12]

12. $d_\chi\ell()$ denotes the differential of $\ell()$ with respect to the scalar, vector or matrix of parameters χ.

From $d_\mu \ell(x_1, \cdots, x_t, \cdots, x_T; \bar{A}, \bar{B}, \delta, \kappa, \rho, \sigma, \Gamma, \Omega, \mu, \Psi) = 0 \; \forall d\mu$ we get:

$$\mu = \mathbf{f}_{0|T} \tag{37}$$

From $d_{\Psi^{-1}} \ell(x_1, \cdots, x_t, \cdots, x_T; \bar{A}, \bar{B}, \delta, \kappa, \rho, \sigma, \Gamma, \Omega, \mu, \Psi) = 0 \; \forall d(\Psi^{-1})$ we get:

$$\Psi = \mathbf{V}_{0,0|T} \tag{38}$$

From $d_\Gamma \ell(x_1, \cdots, x_t, \cdots, x_T; \bar{A}, \bar{B}, \delta, \kappa, \rho, \sigma, \Gamma, \Omega, \mu, \Psi) = 0 \; \forall d\Gamma$ we get:

$$\Gamma = \bar{\mathbf{F}}_{0,\bullet} \bar{\mathbf{F}}_{\bullet,\bullet}^{-1} \tag{39}$$

From $d_{\Omega^{-1}} \ell(x_1, \cdots, x_t, \cdots, x_T; \bar{A}, \bar{B}, \delta, \kappa, \rho, \sigma, \Gamma, \Omega, \mu, \Psi) = 0 \; \forall d(\Omega^{-1})$ we get:

$$\Omega = \bar{F}_{0,0} - \Gamma \bar{\mathbf{F}}_{\bullet,\bullet} \Gamma' \tag{40}$$

From $d_\kappa \ell(x_1, \cdots, x_t, \cdots, x_T; \bar{A}, \bar{B}, \delta, \kappa, \rho, \sigma, \Gamma, \Omega, \mu, \Psi) = 0 \; \forall d\kappa$ we get:

$$\kappa = \left(\frac{T}{T-1} I_{IJ} + \frac{T-2}{T-1} \mathrm{diag}(\rho)^2 - 2\mathrm{diag}(\rho) \right)^{-1} \cdot$$

$$\cdot \left[\frac{T}{T-1} \bar{x}_{0,0} + \frac{T-2}{T-1} \mathrm{diag}(\rho)^2 \bar{x}_{1,1} - \mathrm{diag}(\rho) \left(\bar{x}_{0,1} + \bar{x}_{1,0} \right) - \Theta(\bar{A}, \bar{B}) \mathrm{diag}(\delta) \bar{f}_{0,0} + \right.$$

$$\left. - \mathrm{diag}(\rho)^2 \Theta(\bar{A}, \bar{B}) \mathrm{diag}(\delta) \bar{f}_{1,1} + \mathrm{diag}(\rho) \Theta(\bar{A}, \bar{B}) \mathrm{diag}(\delta) \left(\bar{f}_{0,1} + \bar{f}_{1,0} \right) \right] \tag{41}$$

From $d_\delta \ell(x_1, \cdots, x_t, \cdots, x_T; \bar{A}, \bar{B}, \delta, \kappa, \rho, \sigma, \Gamma, \Omega, \mu, \Psi) = 0 \; \forall d\delta$ we get:

$$\delta = \mathbf{S}_\delta^{-1} \left(\bar{\mathbb{F}}_{0,0}, \bar{\mathbb{F}}_{1,1}, \bar{\mathbb{F}}_{0,1}; \bar{A}, \bar{B}, \rho, \sigma \right) \cdot$$

$$\cdot \, \mathbf{c}_\delta \left(\bar{\mathbb{W}}_{0,0}, \bar{\mathbb{W}}_{1,1}, \bar{\mathbb{W}}_{0,1}, \bar{\mathbb{W}}_{1,0}, \bar{x}_{0,0}, \bar{x}_{1,1}, \bar{x}_{0,1}, \bar{x}_{1,0}, \bar{f}_{0,0}, \bar{f}_{1,1}, \bar{f}_{0,1}, \bar{f}_{1,0}; \bar{A}, \bar{B}, \kappa, \rho, \sigma \right) \tag{42}$$

where $\mathbf{S}_\delta\,()$ is a $(MN \times MN)$ matrix and and $\mathbf{c}_\delta\,()$ is a $(MN \times 1)$ vector. Adopting the notation $[Z]_{q_1,q_2}$ to represent element (q_1, q_2) of matrix Z, and $[z]_q$ to represent the q-th element of column-vector z, the generic elements of $\mathbf{S}_\delta\,()$ and $\mathbf{c}_\delta\,()$ are ($m, \tilde{m} = 1, \cdots, M$; $n, \tilde{n} = 1, \cdots, N$; $\alpha_{i,1} = \beta_{j,1} = 1$ for all i and j):

$$[\mathbf{S}_\delta]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}} =$$

$$= \left[\bar{\mathbb{F}}_{0,0} \right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}} \left(\sum_{i=1}^{I} \sum_{j=1}^{J} \sigma_{i,j}^{-1} \alpha_{i,m} \alpha_{i,\tilde{m}} \beta_{j,n} \beta_{j,\tilde{n}} \right) +$$

$$+ \left[\bar{\mathbb{F}}_{1,1} \right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}} \left(\sum_{i=1}^{I} \sum_{j=1}^{J} \rho_{i,j}^2 \sigma_{i,j}^{-1} \alpha_{i,m} \alpha_{i,\tilde{m}} \beta_{j,n} \beta_{j,\tilde{n}} \right) +$$

$$- \left[\bar{\mathbb{F}}_{0,1} + \bar{\mathbb{F}}_{0,1}' \right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}} \left(\sum_{i=1}^{I} \sum_{j=1}^{J} \rho_{i,j} \sigma_{i,j}^{-1} \alpha_{i,m} \alpha_{i,\tilde{m}} \beta_{j,n} \beta_{j,\tilde{n}} \right)$$

$$[\mathbf{c}_\delta]_{(n-1)M+m} =$$

$$= \sum_{i=1}^{I}\sum_{j=1}^{J}\Big\{\sigma_{i,j}^{-1}\left[\bar{\mathbb{W}}_{0,0}+(\bar{x}_{0,0}-\kappa)\,\bar{f}'_{0,0}\right]_{(j-1)I+i,(n-1)M+m}\alpha_{i,m}\beta_{j,n}+$$

$$+\sigma_{i,j}^{-1}\left[\bar{\mathbb{W}}_{1,1}+(\bar{x}_{1,1}-\kappa)\,\bar{f}'_{1,1}\right]_{(j-1)I+i,(n-1)M+m}\alpha_{i,m}\beta_{j,n}+-\rho_{i,j}\sigma_{i,j}^{-1}\left[\bar{\mathbb{W}}_{0,1}\right.$$

$$\left.+\bar{\mathbb{W}}_{1,0}+(\bar{x}_{0,1}-\kappa)\,\bar{f}'_{1,0}+(\bar{x}_{1,0}-\kappa)\,\bar{f}'_{0,1}\right]_{(j-1)I+i,(n-1)M+m}\alpha_{i,m}\beta_{j,n}\Big\}$$

From $d_{\bar{A}}\ell(x_1,\cdots,x_t,\cdots,x_T;\bar{A},\bar{B},\delta,\kappa,\rho,\sigma,\Gamma,\Omega,\mu,\Psi)=0 \quad \forall d\bar{A}$ we get, for $i=1,\cdots,I$:

$$\alpha_{[i]} = \mathbf{S}_{\alpha_{[i]}}^{-1}\left(\bar{\mathbb{F}}_{0,0},\bar{\mathbb{F}}_{1,1},\bar{\mathbb{F}}_{0,1};\bar{B},\delta,\rho,\sigma\right).$$

$$\cdot\,\mathbf{c}_{\alpha_{[i]}}\left(\bar{\mathbb{W}}_{0,0},\bar{\mathbb{W}}_{1,1},\bar{\mathbb{W}}_{0,1},\bar{\mathbb{W}}_{1,0},\bar{x}_{0,0},\bar{x}_{1,1},\bar{x}_{0,1},\bar{x}_{1,0},\bar{f}_{0,0},\bar{f}_{1,1},\bar{f}_{0,1},\bar{f}_{1,0};\bar{B},\delta,\kappa,\rho,\sigma\right) \tag{43}$$

where $\alpha_{[i]}=[\alpha_{i,2}\cdots\alpha_{i,m}\cdots\alpha_{i,M}]'$ $(((M-1)\times 1)$ is the transposed i-th row of $\bar{A}$, $\mathbf{S}_{\alpha_{[i]}}()$ is a $((M-1)\times(M-1))$ matrix, and $\mathbf{c}_{\alpha_{[i]}}()$ is a $((M-1)\times 1)$ vector. Their generic elements are (for $m,\tilde{m}=2,\cdots,M$ and with $\beta_{j,1}=1$ for all j):

$$\left[\mathbf{S}_{\alpha_{[i]}}\right]_{m-1,\tilde{m}-1} =$$

$$= \sum_{\tilde{n}=1}^{N}\sum_{n=1}^{N}\sum_{j=1}^{J}\Big\{\sigma_{i,j}^{-1}\delta_{m,n}\delta_{\tilde{m},\tilde{n}}\beta_{j,n}\beta_{j,\tilde{n}}\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{(n-1)M+m,(\tilde{n}-1)N+\tilde{m}}+$$

$$+\rho_{i,j}^2\left[\bar{\mathbb{F}}_{1,1}\right]_{(n-1)M+m,(\tilde{n}-1)N+\tilde{m}}-\rho_{i,j}\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}'_{0,1}\right]_{(n-1)M+m,(\tilde{n}-1)N+\tilde{m}}\Big\}\Big\}$$

$$\left[\mathbf{c}_{\alpha_{[i]}}\right]_{m-1} =$$

$$= \sum_{n=1}^{N}\sum_{j=1}^{J}\Big\{\sigma_{i,j}^{-1}\delta_{m,n}\beta_{j,n}\Big\{\Big\{\left[\bar{\mathbb{W}}_{0,0}+(\bar{x}_{0,0}-\kappa)\,\bar{f}'_{0,0}\right]_{(j-1)I+i,(n-1)N+m}+$$

$$+\rho_{i,j}^2\left[\bar{\mathbb{W}}_{1,1}+(\bar{x}_{1,1}-\kappa)\,\bar{f}'_{1,1}\right]_{(j-1)I+i,(n-1)N+m}+$$

$$-\rho_{i,j}\left[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+(\bar{x}_{0,1}-\kappa)\,\bar{f}'_{1,0}+(\bar{x}_{1,0}-\kappa)\,\bar{f}'_{0,1}\right]_{(j-1)I+i,(n-1)N+m}\Big\}+$$

$$-\sum_{\tilde{n}=1}^{N}\delta_{1,\tilde{n}}\beta_{j,\tilde{n}}\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{(n-1)M+m,(\tilde{n}-1)N+1}+\rho_{i,j}^2\left[\bar{\mathbb{F}}_{1,1}\right]_{(n-1)M+m,(\tilde{n}-1)N+1}+$$

$$-\rho_{i,j}\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}'_{0,1}\right]_{(n-1)M+m,(\tilde{n}-1)N+1}\Big\}\Big\}\Big\}$$

From $d_{\bar{B}}\ell(x_1,\cdots,x_t,\cdots,x_T;\bar{A},\bar{B},\delta,\kappa,\rho,\sigma,\Gamma,\Omega,\mu,\Psi)=0\ \forall d\bar{B}$ we get, for $j=1,\cdots,J$:

$$\beta_{[j]} = \mathbf{S}_{\beta_{[j]}}^{-1}\left(\bar{\mathbb{F}}_{0,0},\bar{\mathbb{F}}_{1,1},\bar{\mathbb{F}}_{0,1};\bar{A},\delta,\rho,\sigma\right).$$

$$\cdot\, \mathbf{c}_{\beta_{[j]}}\left(\bar{\mathbb{W}}_{0,0}, \bar{\mathbb{W}}_{1,1}, \bar{\mathbb{W}}_{0,1}, \bar{\mathbb{W}}_{1,0}, \bar{x}_{0,0}, \bar{x}_{1,1}, \bar{x}_{0,1}, \bar{x}_{1,0}, \bar{f}_{0,0}, \bar{f}_{1,1}, \bar{f}_{0,1}, \bar{f}_{1,0}; \bar{A}, \delta, \kappa, \rho, \sigma\right) \tag{44}$$

where $\beta_{[j]} = [\beta_{j,2} \cdots \beta_{j,n} \cdots \beta_{j,N}]'$ $(((N-1)\times 1)$ is the transposed j-th row of $\bar{B}$, $\mathbf{S}_{\beta_{[j]}}()$ is a $((N-1)\times(N-1))$ matrix, and $\mathbf{c}_{\beta_{[j]}}()$ is a $((N-1)\times 1)$ vector. Let:

$$\delta_{[n]} = [\delta_{1,n} \cdots \delta_{m,n} \cdots \delta_{M,n}]' \quad (M \times 1)$$

$$\rho_{[j]} = [\rho_{1,j} \cdots \rho_{i,j} \cdots \rho_{I,j}]' \quad (I \times 1)$$

$$\sigma_{[j]} = [\sigma_{1,j} \cdots \sigma_{i,j} \cdots \sigma_{I,j}]' \quad (I \times 1)$$

and $[Z]_{[n,\tilde{n}]}$ denote the $(M \times M)$ block of $(MN \times MN)$ matrix Z:

$$\begin{bmatrix} z_{(n-1)M+1,(\tilde{n}-1)M+1} & \cdots & z_{(n-1)M+1,(\tilde{n}-1)M+\tilde{m}} & \cdots & z_{(n-1)M+1,\tilde{n}M} \\ \vdots & & \vdots & & \vdots \\ z_{(n-1)M+m,(\tilde{n}-1)M+1} & \cdots & z_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}} & \cdots & z_{(n-1)M+m,\tilde{n}M} \\ \vdots & & \vdots & & \vdots \\ z_{nM,(\tilde{n}-1)M+1} & \cdots & z_{nM,(\tilde{n}-1)M+\tilde{m}} & \cdots & z_{nM,\tilde{n}M} \end{bmatrix}$$

Similarly, for $(IJ \times MN)$ matrix $\hat{Z}$, $\left[\hat{Z}\right]_{[j,n]}$ will denote the $(I \times M)$ block:

$$\begin{bmatrix} \hat{z}_{(j-1)I+1,(n-1)M+1} & \cdots & \hat{z}_{(j-1)I+1,(n-1)M+m} & \cdots & \hat{z}_{(j-1)I+1,nM} \\ \vdots & & \vdots & & \vdots \\ \hat{z}_{(j-1)I+i,(n-1)M+1} & \cdots & \hat{z}_{(j-1)I+i,(n-1)M+m} & \cdots & \hat{z}_{(j-1)I+i,nM} \\ \vdots & & \vdots & & \vdots \\ \hat{z}_{jI,(n-1)M+1} & \cdots & \hat{z}_{jI,(n-1)M+m} & \cdots & \hat{z}_{jI,nM} \end{bmatrix}$$

Using this block notation, the generic elements of $\mathbf{S}_{\beta_{[j]}}()$ and $\mathbf{c}_{\beta_{[j]}}()$ are the following (for $n, \tilde{n} = 2, \cdots, N$ and with $tr(Q)$ denoting the trace of square matrix Q):

$$\left[\mathbf{S}_{\beta_{[j]}}\right]_{n-1,\tilde{n}-1} = \mathrm{tr}\Bigg\{\Big\{ \left[\bar{\mathbb{F}}_{0,0}\right]_{[n,\tilde{n}]} \mathrm{diag}(\delta_{[\tilde{n}]}) \left[\mathbf{1}_I\ \bar{A}\right]' +$$

$$+ \left[\bar{\mathbb{F}}_{1,1}\right]_{[n,\tilde{n}]} \mathrm{diag}(\delta_{[\tilde{n}]}) \left[\mathbf{1}_I\ \bar{A}\right]' \mathrm{diag}(\rho_{[j]})^2 +$$

$$- \left[\bar{\mathbb{F}}_{0,1} + \bar{\mathbb{F}}'_{0,1}\right]_{[n,\tilde{n}]} \mathrm{diag}(\delta_{[\tilde{n}]}) \left[\mathbf{1}_I\ \bar{A}\right]' \mathrm{diag}(\rho_{[j]}) \Big\} \mathrm{diag}(\sigma_{[j]})^{-1} \left[\mathbf{1}_I\ \bar{A}\right] \mathrm{diag}(\delta_{[n]}) \Bigg\}$$

$$\left[\mathbf{c}_{\beta_{[j]}}\right]_{n-1} = \mathrm{tr}\Bigg\{\Bigg\{\Big\{ \left[\bar{\mathbb{W}}_{0,0} + (\bar{x}_{0,0} - \kappa)\, \bar{f}'_{0,0}\right]'_{[j,n]} +$$

$$+ \left[\bar{\mathbb{W}}_{1,1} + (\bar{x}_{1,1} - \kappa)\, \bar{f}'_{1,1}\right]'_{[j,n]} \mathrm{diag}(\rho_{[j]})^2 +$$

$$- \left[\bar{\mathbb{W}}_{0,1} + \bar{\mathbb{W}}_{1,0} + (\bar{x}_{0,1} - \kappa)\, \bar{f}'_{1,0} + (\bar{x}_{1,0} - \kappa)\, \bar{f}'_{0,1}\right]'_{[jn]} \mathrm{diag}(\rho_{[j]}) \Big\} +$$

$$- \Big\{ \left[\bar{\mathbb{F}}_{0,0}\right]_{[n,1]} \mathrm{diag}(\delta_{[1]}) \left[\mathbf{1}_I\ \bar{A}\right]' + \left[\bar{\mathbb{F}}_{1,1}\right]_{[n,1]} \mathrm{diag}(\delta_{[1]}) \left[\mathbf{1}_I\ \bar{A}\right]' \mathrm{diag}(\rho_{[j]})^2 +$$

$$-\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}_{0,1}'\right]_{[n,1]}\mathrm{diag}(\delta_{[1]})\left[\mathbf{1}_I\ \bar{A}\right]'\mathrm{diag}(\rho_{[j]})\Big\}\Bigg\}\mathrm{diag}(\sigma_{[j]})^{-1}\left[\mathbf{1}_I\ \bar{A}\right]\mathrm{diag}(\delta_{[n]})\Bigg\}$$

From $d_{\sigma_{i,j}^{-1}}\ell(x_1,\cdots,x_t,\cdots,x_T;\bar{A},\bar{B},\delta,\kappa,\rho,\sigma,\Gamma,\Omega,\mu,\Psi)=0\ \forall d\sigma_{i,j}^{-1}$ ($i=1,\cdots,I$ and $j=1,\cdots,J$), we get:

$$\begin{gathered}
\sigma=\frac{T-1}{T}\mathrm{diag}\Bigg\{\left[\bar{\mathbb{X}}_{0,0}+\frac{T}{T-1}\left(\bar{x}_{0,0}-\kappa\right)\left(\bar{x}_{0,0}-\kappa\right)'\right]+\\
+\mathrm{diag}(\rho)^2\left[\bar{\mathbb{X}}_{1,1}+\frac{T-2}{T-1}\left(\bar{x}_{1,1}-\kappa\right)\left(\bar{x}_{1,1}-\kappa\right)'\right]+\\
-2\mathrm{diag}(\rho)\left[\bar{\mathbb{X}}_{0,1}+\left(\bar{x}_{0,1}-\kappa\right)\left(\bar{x}_{1,0}-\kappa\right)'\right]+\\
+\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{\mathbb{F}}_{0,0}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+\\
+\mathrm{diag}(\rho)^2\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{\mathbb{F}}_{1,1}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+\\
-2\mathrm{diag}(\rho)\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{\mathbb{F}}_{0,1}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+\\
-2\left[\bar{\mathbb{W}}_{0,0}+\left(\bar{x}_{0,0}-\kappa\right)\bar{f}_{0,0}'\right]\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+\\
-2\mathrm{diag}(\rho)^2\left[\bar{\mathbb{W}}_{1,1}+\left(\bar{x}_{1,1}-\kappa\right)\bar{f}_{1,1}'\right]\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+\\
+2\mathrm{diag}(\rho)\left[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+\left(\bar{x}_{0,1}-\kappa\right)\bar{f}_{1,0}'+\left(\bar{x}_{1,0}-\kappa\right)\bar{f}_{0,1}'\right]\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'\Bigg\}
\end{gathered}\tag{45}$$

Finally, from $d_{\rho_{i,j}}\ell(x_1,\cdots,x_t,\cdots,x_T;\bar{A},\bar{B},\delta,\kappa,\rho,\sigma,\Gamma,\Omega,\mu,\Psi)=0\ \forall d\rho_{i,j}$ ($i=1,\cdots,I;j=1,\cdots,J$) we get:

$$\begin{gathered}
\rho_{i,j}=\Bigg\{\left[\bar{\mathbb{X}}_{1,1}+\frac{T-2}{T-1}\left(\bar{x}_{1,1}-\kappa\right)\left(\bar{x}_{1,1}-\kappa\right)'\right]_{(j-1)I+i,(j-1)I+i}+\\
+\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{\tilde{m}=1}^{M}\sum_{\tilde{n}=1}^{N}\delta_{m,n}\delta_{\tilde{m},\tilde{n}}\alpha_{i,m}\alpha_{i,\tilde{m}}\beta_{j,n}\beta_{j,\tilde{n}}\left[\bar{\mathbb{F}}_{1,1}\right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}}+\\
-2\sum_{m=1}^{M}\sum_{n=1}^{N}\delta_{m,n}\alpha_{i,m}\beta_{j,n}\left[\bar{\mathbb{W}}_{1,1}+\left(\bar{x}_{1,1}-\kappa\right)\bar{f}_{1,1}'\right]_{(j-1)I+i,(n-1)M+m}+\\
+\frac{\sigma_{ij}}{(T-1)\left(1-\rho_{i,j}^2\right)}\Bigg\}^{-1}\Bigg\{\left[\bar{\mathbb{X}}_{0,1}+\left(\bar{x}_{0,1}-\kappa\right)\left(\bar{x}_{1,0}-\kappa\right)'\right]_{(j-1)I+i,(j-1)I+i}+\\
+\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{\tilde{m}=1}^{M}\sum_{\tilde{n}=1}^{N}\delta_{m,n}\delta_{\tilde{m}\tilde{n}}\alpha_{i,m}\alpha_{i,\tilde{m}}\beta_{j,n}\beta_{j,\tilde{n}}\left[\bar{\mathbb{F}}_{0,1}\right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}}+\\
-\sum_{m=1}^{M}\sum_{n=1}^{N}\delta_{m,n}\alpha_{i,m}\beta_{j,n}\Big[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+
\end{gathered}$$

$$\left. + \left(\bar{x}_{0,1} - \kappa\right)\bar{f}'_{1,0} + \left(\bar{x}_{1,0} - \kappa\right)\bar{f}'_{0,1}\right]_{(j-1)I+i,(n-1)M+m} \Bigg\} \tag{46}$$

Unlike all other first order conditions, note that equation (46) is not fully solved with respect to $\rho_{i,j}$. The right hand side of (46) depends on $\rho_{i,j}$ through the term

$$\frac{\sigma_{i,j}}{(T-1)\left(1-\rho_{i,j}^2\right)} \tag{47}$$

For large T this term will be negligible. For not so large T, it will be straightforward to solve equation (46) by starting to ignore term (47) and then iterating the complete equation until convergence (conditional on the remaining moments and parameters), restricting $|\rho_{i,j}| \leq \rho^{max}$ with ρ^{max} set to a value lower than but very close to one.

3.3. E-step of the ECM algorithm

The ECM algorithm E-step consists of updating the values of sufficient statistics based on the first and second order moments of observables and time-factors. The updating relies on the Kalman smoother applied to the model in state-space representation, for given estimates of the model parameters. In this section, we will use the following slightly more compact formulation of the model than in (8) to (14):

$$x_1 = \kappa + \Xi_0(\bar{A},\bar{B},\delta)\mathbf{f}_1^e + u_1$$

$$x_t = \left(I_{IJ} - \mathrm{diag}(\rho)\right)\kappa + \mathrm{diag}(\rho)x_{t-1} + \Xi(\bar{A},\bar{B},\delta)\mathbf{f}_t^e + \varepsilon_t \quad (t = 2,\cdots,T)$$

$$\mathbf{f}_t^e = \mathbf{\Gamma}^e\mathbf{f}_{t-1}^e + \mathbf{v}_t \quad (t = 1,\cdots,T)$$

$$\varepsilon_t \sim N(0;\mathrm{diag}(\sigma)) \quad (t = 1,\cdots,T)$$

$$\mathbf{v}_t \sim N(0;\mathbf{\Omega}^e) \quad (t = 1,\cdots,T)$$

where

$$\mathbf{f}_t^e = \begin{cases} \left[f_t'\ f_{t-1}'\right]' & (2MN \times 1) \quad if\ P = 1 \\ \mathbf{f}_t & (PMN \times 1) \quad if\ P > 1 \end{cases}$$

$$\Xi_0(\bar{A},\bar{B},\delta) = \left[\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta) \quad \mathbf{0}_{IJ\times MN} \cdots \mathbf{0}_{IJ\times MN}\right] \quad (IJ \times MNP^e)$$

$$\underset{(IJ \times MNP^e)}{\Xi(\bar{A},\bar{B},\delta)} = \left[\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta) \quad -\mathrm{diag}(\rho)\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta) \quad \mathbf{0}_{IJ\times MN} \cdots \mathbf{0}_{IJ\times MN}\right]$$

$$\underset{(MNP^e \times MNP^e)}{\mathbf{\Gamma}^e} = \begin{bmatrix} \Gamma_1 & \Gamma_2 & \cdots & \Gamma_{P^e-1} & \Gamma_{P^e} \\ \mathbf{I}_{MN} & \mathbf{0}_{MN\times MN} & \cdots & \mathbf{0}_{MN\times MN} & \mathbf{0}_{MN\times MN} \\ \mathbf{0}_{MN\times MN} & \mathbf{I}_{MN} & \cdots & \mathbf{0}_{MN\times MN} & \mathbf{0}_{MN\times MN} \\ \vdots & \vdots & & \vdots & \vdots \\ \mathbf{0}_{MN\times MN} & \mathbf{0}_{MN\times MN} & \cdots & \mathbf{I}_{MN} & \mathbf{0}_{MN\times MN} \end{bmatrix}$$

$$\mathbf{v}_t = \left[\eta_t' \quad \mathbf{0}_{1\times MN} \cdots \mathbf{0}_{1\times MN}\right]' \quad (MNP^e \times 1)$$

$$\underset{(MNP^e \times MNP^e)}{\boldsymbol{\Omega}^e} = \begin{bmatrix} \Omega & \mathbf{0}_{MN\times MN} & \cdots & \mathbf{0}_{MN\times MN} \\ \mathbf{0}_{MN\times MN} & \mathbf{0}_{MN\times MN} & \cdots & \mathbf{0}_{MN\times MN} \\ \vdots & \vdots & & \vdots \\ \mathbf{0}_{MN\times MN} & \mathbf{0}_{MN\times MN} & \cdots & \mathbf{0}_{MN\times MN} \end{bmatrix}$$

with

$$P^e = \max\{2; P\}$$

$$\Gamma_p = \mathbf{0}_{MN\times MN} \quad \forall p > P$$

Let us also consider the notation:

$$\mathbf{f}^e_{t|\tau} = E\left(\mathbf{f}^e_t | x_1, \cdots, x_\tau\right)$$

$$\mathbf{V}^e_{t-q_1,t-q_2|\tau} = E\left[\left(\mathbf{f}^e_{t-q_1} - \mathbf{f}^e_{t-q_1|\tau}\right)\left(\mathbf{f}^e_{t-q_2} - \mathbf{f}^e_{t-q_2|\tau}\right)' | x_1, \cdots, x_\tau\right]$$

The **Kalman forward recursions** for $t = 2, \cdots \Theta$ are given by:

$$\mathbf{f}^e_{t|t-1} = \boldsymbol{\Gamma}^e \mathbf{f}^e_{t-1|t-1}$$

$$\mathbf{V}^e_{t,t|t-1} = \boldsymbol{\Gamma}^e \mathbf{V}^e_{t-1,t-1|t-1} (\boldsymbol{\Gamma}^e)' + \boldsymbol{\Omega}^e$$

$$\mathbf{V}^e_{t,t-1|t-1} = \boldsymbol{\Gamma}^e \mathbf{V}^e_{t-1,t-1|t-1}$$

$$\mathbf{K}_t = \mathbf{V}^e_{t,t|t-1} \Xi(\bar{A}, \bar{B}, \delta)' \mathrm{diag}(\sigma)^{-1} \Big\{ \mathbf{I}_{IJ} - \Xi(\bar{A}, \bar{B}, \delta) \left[\mathbf{I}_{MNP^e} + \right.$$

$$\left. + \mathbf{V}^e_{t,t|t-1} \Xi(\bar{A}, \bar{B}, \delta)' \mathrm{diag}(\sigma)^{-1} \Xi(\bar{A}, \bar{B}, \delta)\right]^{-1} \mathbf{V}^e_{t,t|t-1} \Xi(\bar{A}, \bar{B}, \delta)' \mathrm{diag}(\sigma)^{-1} \Big\}$$

$$\mathbf{f}^e_{t|t} = \mathbf{f}^e_{t|t-1} + \mathbf{K}_t \left[x_t - (I_{IJ} - \mathrm{diag}(\rho))\,\kappa - \mathrm{diag}(\rho) x_{t-1} - \Xi(\bar{A}, \bar{B}, \delta) \mathbf{f}^e_{t|t-1}\right]$$

$$\mathbf{V}^e_{t,t|t} = \left[\mathbf{I}_{MNP^e} - \mathbf{K}_t \Xi(\bar{A}, \bar{B}, \delta)\right] \mathbf{V}^e_{t,t|t-1}$$

starting with

$$\mathbf{f}^e_{1|1} = \begin{cases} [\mu' \; \mu']' & if\ P = 1 \\ \mu & if\ P > 1 \end{cases}$$

$$\mathbf{V}^e_{1,1|1} = \begin{cases} \begin{bmatrix} \Psi & \Gamma^e \Psi \\ \Psi (\Gamma^e)' & \Psi \end{bmatrix} & if\ P = 1 \\ \Psi & if\ P > 1 \end{cases}$$

Built upon the results of the Kalman forward recursions, we have the **Kalman smoother backward recursions** (for $t = T-1, T-2, \cdots, 1$):

$$\mathbf{J}_t = \mathbf{V}^e_{t-1,t-1|t-1} (\boldsymbol{\Gamma}^e)' \left(\mathbf{V}^e_{t,t|t-1}\right)^{-1}$$

$$\mathbf{f}^e_{t|T} = \mathbf{f}^e_{t|t} + \mathbf{J}_{t+1} \left(\mathbf{f}^e_{t+1|T} - \mathbf{f}^e_{t+1|t}\right)$$

$$\mathbf{V}^e_{t,t|T} = \mathbf{V}^e_{t,t|t} + \mathbf{J}_{t+1}\left(\mathbf{V}^e_{t+1,t+1|T} - \mathbf{V}^e_{t+1,t+1|t}\right)\mathbf{J}'_{t+1}$$

$$\mathbf{V}^e_{t,t-1|T} = \mathbf{V}^e_{t,t|t}\mathbf{J}'_t + \mathbf{J}_{t+1}\left(\mathbf{V}^e_{t+1,t|T} - \mathbf{V}^e_{t+1,t|t}\right)\mathbf{J}'_t$$

starting with

$$\mathbf{J}_T = \mathbf{V}^e_{T-1,T-1|T-1}(\mathbf{\Gamma}^e)'\left(\mathbf{V}^e_{T,T|T-1}\right)^{-1}$$

$$\mathbf{V}^e_{T,T-1|T} = \left[\mathbf{I}_{MNP^e} - \mathbf{K}_T\Xi(\bar{A},\bar{B},\delta)\right]\mathbf{V}^e_{T,T-1|T-1}$$

3.4. Log-likelihood evaluation

Given estimates of the model parameters, we may use the prediction error decomposition to evaluate the log-likelihood function:

$$\varphi(\bar{A},\bar{B},\delta,\kappa,\rho,\sigma,\Gamma,\Omega,\mu,\Psi;x_1,\cdots,x_T) =$$

$$= \varphi_1(\bar{A},\bar{B},\delta,\kappa,\rho,\sigma,\mu,\Psi;x_1) + \sum_{t=2}^{T}\varphi_t(\bar{A},\bar{B},\delta,\kappa,\rho,\sigma,\Gamma,\Omega,\mu,\Psi;x_t|x_1,\cdots,x_{t-1}) =$$

$$= -\frac{1}{2}\Big\{IJ\ln(2\pi) + \ln\left[\det\left(\widehat{\mathrm{Var}}\left(x_1\right)\right)\right] +$$

$$+\left[x_1 - \widehat{E}\left(x_1\right)\right]'\left[\widehat{\mathrm{Var}}\left(x_1\right)\right]^{-1}\left[x_1 - \widehat{E}\left(x_1\right)\right]\Big\} +$$

$$-\frac{1}{2}\sum_{t=2}^{T}\Big\{IJ\ln(2\pi) + \ln\left[\det\left(\widehat{\mathrm{Var}}_{t-1}\left(x_t\right)\right)\right] +$$

$$+\left[x_t - \widehat{E}_{t-1}\left(x_t\right)\right]'\left[\widehat{\mathrm{Var}}_{t-1}\left(x_t\right)\right]^{-1}\left[x_t - \widehat{E}_{t-1}\left(x_t\right)\right]\Big\}$$

where

$$\widehat{E}\left(x_1\right) = \kappa + \Xi_0(\bar{A},\bar{B},\delta)\mathbf{f}_{1|0}$$

$$\widehat{\mathrm{Var}}\left(x_1\right) = \left[\mathbf{I}_{IJ} - \mathrm{diag}(\rho)^2\right]^{-1}\mathrm{diag}(\sigma) + \Xi_0(\bar{A},\bar{B},\delta)\ \mathbf{V}_{1,1|0}\ \Xi_0(\bar{A},\bar{B},\delta)'$$

and for $t = 2,\cdots,T$:

$$\widehat{E}_{t-1}\left(x_t\right) = \widehat{E}\left(x_t|x_1,\cdots,x_{t-1}\right) = \left(I_{IJ} - \mathrm{diag}(\rho)\right)\kappa + \mathrm{diag}(\rho)x_{t-1} + \Xi(\bar{A},\bar{B},\delta)\mathbf{f}_{t|t-1}$$

$$\widehat{\mathrm{Var}}_{t-1}\left(x_t\right) = \widehat{\mathrm{Var}}\left(x_t|x_1,\cdots,x_{t-1}\right) = \mathrm{diag}(\sigma) + \Xi(\bar{A},\bar{B},\delta)\mathbf{V}_{t,t|t-1}\Xi(\bar{A},\bar{B},\delta)'$$

Instead of directly computing determinants and inverses of $(IJ \times IJ)$ covariance matrices $\widehat{\mathrm{Var}}\left(x_1\right)$ and $\widehat{\mathrm{Var}}\left(x_t|x_1,\cdots,x_{t-1}\right)$, it is more convenient to make use of the so-called matrix determinant and inversion lemmas:[13]

$$\ln\left[\det \widehat{\mathrm{Var}}\,(x_1)\right] =$$

$$= \ln\left\{ \det\left[\mathbf{I}_{MNP^e} + \mathbf{V}_{1,1|0}\Xi_0(\bar{A},\bar{B},\delta)'\Phi\Xi_0(\bar{A},\bar{B},\delta)\right] \right\} + \sum_{i=1}^{I}\sum_{j=1}^{J}\ln\frac{\sigma_{ij}}{1-\rho_{ij}^2}$$

$$\ln\left[\det \widehat{\mathrm{Var}}_{t-1}\,(x_t)\right] =$$

$$= \ln\left\{ \det\left[\mathbf{I}_{MNP^e} + \mathbf{V}_{t,t|t-1}\Xi(\bar{A},\bar{B},\delta)'\mathrm{diag}(\sigma)^{-1}\Xi(\bar{A},\bar{B},\delta)\right] \right\} + \sum_{i=1}^{I}\sum_{j=1}^{J}\ln\left(\sigma_{ij}\right)$$

$$\left[\widehat{\mathrm{Var}}\,(x_1)\right]^{-1} = \Phi +$$

$$-\Phi\Xi_0(\bar{A},\bar{B},\delta)\left[\mathbf{I}_{MNP^e} + \mathbf{V}_{1,1|0}\Xi_0(\bar{A},\bar{B},\delta)'\Phi\Xi_0(\bar{A},\bar{B},\delta)\right]^{-1}\mathbf{V}_{1,1|0}\Xi_0(\bar{A},\bar{B},\delta)'\Phi$$

$$\left[\widehat{\mathrm{Var}}_{t-1}\,(x_t)\right]^{-1} = \mathrm{diag}(\sigma)^{-1} - \mathrm{diag}(\sigma)^{-1}\Xi(\bar{A},\bar{B},\delta).$$

$$.\left[\mathbf{I}_{MNP^e} + \mathbf{V}_{t,t|t-1}\Xi(\bar{A},\bar{B},\delta)'\mathrm{diag}(\sigma)^{-1}\Xi(\bar{A},\bar{B},\delta)\right]^{-1}.$$

$$.\mathbf{V}_{t,t|t-1}\Xi(\bar{A},\bar{B},\delta)'\mathrm{diag}(\sigma)^{-1}$$

where

$$\Phi = \left[\mathbf{I}_{IJ} - \mathrm{diag}(\rho)^2\right]\mathrm{diag}(\sigma)^{-1}$$

These expressions only require the computation of determinants and inverses of order $MNP^e << IJ$ (besides those of diagonal matrices).

4. Monte Carlo analysis

In this section, a simulation study is conducted to assess the small sample performance of the maximum likelihood estimator applied to the proposed model considering that the parameters are unknown but otherwise the equations are correctly specified. First, we generate the data for numbers of variables and factors similar to the ones we deal with in the empirical illustration presented in section 5. Then, we perform a sensitivity analysis to some of the simulation settings.

13. If all matrices are conformable and C is non-singular, one has:

$$\det\left(A + UCV'\right) = \det\left(I + CV'A^{-1}U\right)\det\left(A\right)$$

$$\left(A + UCV'\right)^{-1} = A^{-1} - A^{-1}U\left(I + CV'A^{-1}U\right)^{-1}CV'A^{-1}$$

4.1. The base case

Take model (8)-(14) with identifying restrictions (20)-(24). The results for the base case were obtained from 1,000 simulated sample draws with $T = 200$, $I = 50$, $J = 25$, and $M = N = 3$ (respectively, the number of time periods, the number of variables in the first non-time data mode, the number of variables in the second non-time mode, and the number of time-factors associated with these non-time modes).[14]

Regarding the generation of time-factors, we set the vector autoregression to be of order one (i.e. $P = 1$) with a diagonal coefficient matrix such that, for each simulation of the paths of factors, the coefficients in the diagonal are drawn from a uniform distribution on $[0.3, 0.7]$.[15] In order to obtain each set of simulated factor paths, the innovations of the autoregressive process were generated by drawing standard Gaussian noise and then applying the transformation associated with the lower triangular Cholesky decomposition of previously constructed covariance matrices. The latter were obtained in three steps. Firstly, for each simulated sample, MN eigenvalues were drawn from a uniform distribution on $[0, 1]$ and rescaled in order to average one. A correlation matrix was then generated by resorting to the procedure proposed by Davies and Higham (2000), which is based on subjecting the diagonal matrix of given eigenvalues to a random orthogonal similarity transformation and then to a sequence of Givens rotations. Lastly, the correlation matrix thus generated was rescaled by pre- and post-multiplying it by a diagonal matrix constructed such that the sample variances of all factors became equal to one.

Concerning the measurement equation, for each simulated sample, the elements of matrices $\bar{A}$ and $\bar{B}$ were drawn from a uniform distribution on $[0, 1]$ and adjusted by subtracting the corresponding column sum from each element (in order to comply with identifying restrictions (20) and (22)). All elements of vectors κ and δ were simply set to zero and to one, respectively. As to the idiosyncratic components, their autoregressive coefficients were drawn from a uniform distribution on $[-0.9, 0.9]$, while, as it is usual in the literature, the innovation variances were set indirectly by drawing the ratios between the variance of each idiosyncratic component $u_{t,i,j}$ and the variance of the associated observable $x_{t,i,j}$ from a uniform distribution (in our case on $[0.5, 0.9]$).

To evaluate the estimation performance, we consider the trace R^2, as in Stock and Watson (2002a), which measures the closeness of the space generated by the true factors to that generated by the estimated factors. The results of the simulation

14. For each sample draw, 10,000 burning-in periods were considered.

15. The autoregressive coefficients considered when generating the time-factors also varied from sample draw to sample draw. In the literature, generating dynamic factors from a VAR(1) process with a diagonal coefficient matrix is standard in simulations (see, for example, Stock and Watson, 2002a, and Doz, Giannone and Reichlin, 2012, although in these papers the same autoregressive coefficient is common for all factors).

study are presented in Table 1. For the base case, we obtain a R^2 very close to one which denotes a very good fit of the space spanned by the true factors. To assess the robustness of the findings, we conduct in the next subsection a sensitivity analysis.

	Model specification					
	T	I	J	M	N	R^2
Base case						
	200	50	25	3	3	0.999
Sensitivity analysis						
	100	50	25	3	3	0.999
	50	50	25	3	3	0.999
	200	50	**50**	3	3	0.999
	200	50	25	**2**	**2**	0.999
	200	50	25	**4**	**4**	0.998
	200	50	25	**5**	**5**	0.996
	100	50	**50**	**5**	**5**	0.998

Note: The bold figures denote specification values different from the base case.

Table 1. Simulation results

4.2. Sensitivity analysis

The sensitivity analysis was carried out in the following way. Firstly, we changed one setting of the simulation design in each turn, while maintaining unchanged all the other specifications, to assess the robustness of the results to each particular dimension.

To start with, the number of time periods in the sample was reduced from $T = 200$ to $T = 100$ and also to $T = 50$, in order to assess how the quality of estimation depends on the sample size in the time mode. We find no visible deterioration of the estimation performance.

The data set in the empirical illustration presented in section 5 is characterized by having the size of one non-time mode (the number of countries) much smaller than the size of the other non-time mode (the number of products). The base case for our simulations was designed to approximately convey such disparity of non-time mode sizes. We assessed the sensitivity of the simulation results to more balanced sizes of non-time modes, by making $I = J = 50$ (instead of $I = 50$ and $J = 25$ as in the base case). Again, the R^2 is very close to one in this scenario.

Simulations were also carried out after changing the number of dynamic factors. Instead of $M = N = 3$, we considered a lower figure, $M = N = 2$, as well as higher numbers, $M = N = 4$ and $M = N = 5$. The results are not sensitive either in this respect.

Finally, we considered the case where all the above specifications were changed at the same time. In particular, the number of time periods was decreased from $T = 200$ to $T = 100$, the number of units in the second non-time mode was increased from 25 to 50 (that is, $I = J = 50$) and, regarding the number of dynamic factors, we set $M = N = 5$. The R^2 is also close to one in this case which reinforces the previous simulation results concerning the robustness of the estimation method.

5. Empirical analysis

5.1. Data

For the empirical application, we collected disaggregated data concerning inflation in the euro area, namely by considering its breakdown by country and by product. The data is compiled and provided by the Eurostat and inflation is measured as the year-on-year rate of change of the Harmonized Index of Consumer Prices (HICP) as used by the European Central Bank for the monitoring of inflation in the Economic and Monetary Union.

In particular, we collected data for all countries currently belonging to the euro area. Hence we considered nineteen countries namely Austria, Belgium, Finland, France, Germany, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Greece, Slovenia, Cyprus, Malta, Slovakia, Estonia, Latvia and Lithuania. For each country, we gathered detailed data by product considering the highest level of disaggregation available for all countries. It was possible to consider 55 products.[16]

The resulting monthly data set basically spans the last two decades, running from January 2002 up to October 2019. Hence, in our empirical application we have $T = 214$, $I = 55$ and $J = 19$.

5.2. Results

The estimation of the model involves setting the number of factors, M and N. As discussed previously, the model can be collapsed over the country mode or over the products mode. This feature of the model suggests the following empirical strategy. Given an estimate of the global factor (and $\delta_{1,1}$), one can purge the data set disaggregated only by products using (27) and then apply an established criterion to determine the number of factors of the resulting series. This provides a value for $M - 1$. In a similar fashion, by considering the data set disaggregated by countries, one can take out the effect of the global factor using (28) and then determine a value for $N - 1$.

Regarding the criteria, we use the Eigenvalue Ratio (ER) and Growth Ratio (GR) estimators proposed by Ahn and Horenstein (2013) which determine the number of factors based on the empirical distribution of the eigenvalues. Ahn and Horenstein (2013) show that the ER and GR estimators perform the best among a

16. In the Appendix, we report the list of countries and products along with the corresponding weights for the euro area.

set of alternative criteria suggested in the literature, including the well-known Bai and Ng (2002) and Onatski (2010) criteria.

Therefore, we firstly obtain an estimate for the global factor by considering a model with a sufficiently higher M and N so as not to affect the consistency of the estimator due to under-specification. In particular, we estimate a model with $M = N = 5$.[17] Applying the ER and GR estimators to the data disaggregated only by products, after purging the effect of the estimated global factor, both criteria points to the presence of two factors. When considering the data disaggregated only by countries, we get one factor. Hence, we set $M = 3$ and $N = 2$.[18] Regarding the order of the vector autoregressive process of the latent factors, we set $P = 1$.[19]

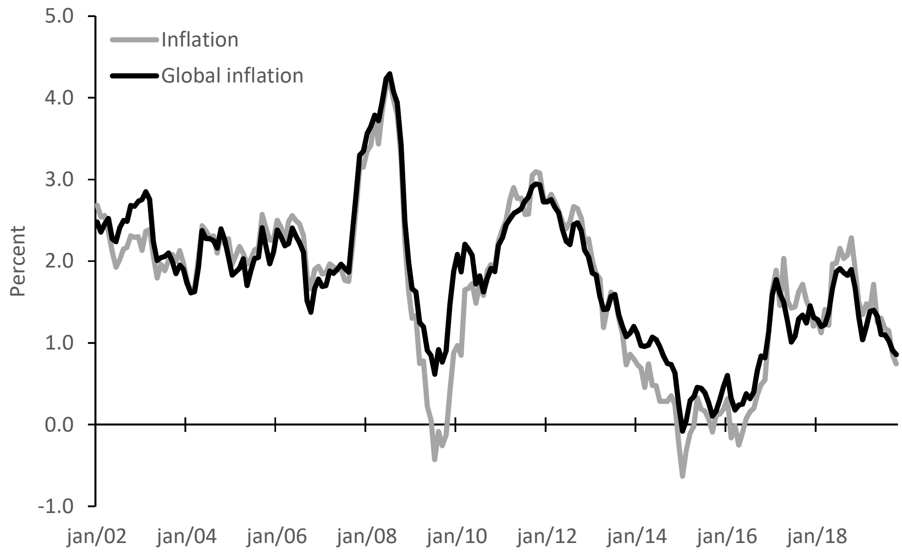


Figure 1: Inflation and the global inflation indicator for the euro area

In Figure 1, we display the actual inflation for the euro area along the resulting global inflation indicator as defined in (30). Despite the strong correlation, the global inflation indicator is smoother than actual inflation for the euro area. In fact, this global inflation indicator reflects the underlying common inflation developments within the euro area discarding idiosyncratic movements and the influence of relative price changes across countries and products. In this sense, such a measure

17. The results are not sensitive to increasing or decreasing both M and N.

18. We have also considered the criteria proposed by Bai and Ng (2002) but the selected number of factors is quite sensitive to the choice regarding the maximum possible number of factors (see also the discussion in Onatski (2010) and Ahn and Horenstein (2013)). Alternatively, if one applies the estimator developed by Onatski (2010), we end up with $M = 3$ and $N = 4$ but the main results presented in this section remain qualitatively unchanged.

19. The results are qualitatively similar for higher values of P.

can be seen as more meaningful indicator of price developments in the euro area than headline inflation.

One should also mention that estimated dynamic factor model despite its parsimony has a relatively high explanatory power. In particular, we assess the fraction of inflation variability explained by the factor model for the euro area as a whole as well as for each individual country and product. Following Stock and Watson (2002b), we compute the R^2 of the regression of observed inflation on the estimated factors. We find that the global factor captures a large share of the variance of euro area inflation, in particular, 90 percent. By country, the average proportion captured by total communality is 64 percent and for most countries (14 out of 19) the proportion is above 50 percent. By product, we find that, on average, 39 percent of the variance is explained by the model.

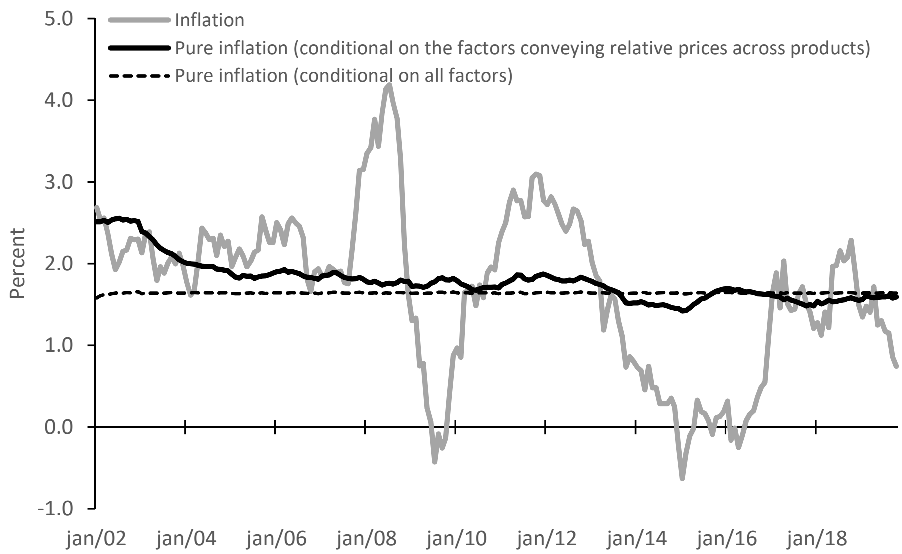


Figure 2: Inflation and pure inflation indicators for the euro area

As discussed previously, a pure inflation indicator can be computed as defined in (34). When we apply the approach of Reis and Watson (2010) to our collapsed model (with only the global factor and two factors conveying relative prices across products), we are able to generate a pure inflation indicator. The resulting pure inflation indicator is displayed in Figure 2 (solid black line) along the observed inflation for the euro area. This pure inflation indicator seems to capture the very slow moving trend developments in inflation. After declining until the end of 2004, it stood slightly below 2 percent before declining again throughout 2013 to a level around 1.6 percent thereafter. Alternatively, one can compute a pure inflation indicator by using all the factors to orthogonalize $f_{t,1,1}^{(p)}$. This indicator is also displayed in Figure 2 (black dashed line). We find that when such an approach is applied to the fully fledged model with three-mode data (and factors specific to real exchange rate movements as well as interactions between relative prices across products and countries), we get an almost flat indicator (around 1.6 percent), i.e. very few price movements are identified as 'pure'. The rationale for this diminished relevance of the pure inflation indicator lies with its sensitivity to the number of factors in the model. In fact, when controlling for all leads and lags of the factors associated with relative prices, because the dimension of the vector space spanned by all these leads and lags tends to expand very quickly when the number of factors increases, there is not much variability to be imputed to the pure inflation indicator.

6. Conclusion

In an increasingly data rich environment, dynamic factor models have become one of the most established methods when dealing with large data sets. Though factor models are usually applied to two-way data, the availability of data sets with observations addressed by three-dimensional indices has been growing in economics and finance. Most of the literature that takes on board three-way data sets within the factor model framework postulates one global factor and one specific factor for each group (and sub-group) of variables. However, as the sample size for non-time modes grow such an approach becomes unfeasible. Hence, inspired by the tensor decomposition literature, we proposed a non-hierarchical dynamic factor model for three-way data which is rather flexible while remaining quite parsimonious.

We discussed the identification of the proposed specification and put forward a set of identifying restrictions which improve the interpretation of the model. Based on these assumptions, we have shown that the model may be averaged out in each and every non-time mode of the data, nesting standard dynamic factor model specifications. We have shown that the model can be conveniently estimated by maximum likelihood using the ECM algorithm, which is a generalization of the well-known EM algorithm. We assessed the finite sample performance of the proposed estimator through a Monte Carlo simulation study and found that it performs remarkably well.

We applied the proposed model to inflation data. In particular, we collected consumer price indices for 55 products for the 19 euro area countries over the last two decades. The estimated model ends up being quite parsimonious while presenting a noteworthy explanatory power across countries and products. Drawing on the estimated model, it was possible to obtain a global inflation indicator that captures inflation developments within the euro area abstracting from inflation differentials associated with relative price changes both across products and across countries as well as from idiosyncratic price movements. Furthermore, we addressed the estimation of pure inflation in the context of the suggested model, with the resulting empirical measures capturing the very slow moving trend developments in inflation.

References

Ahn, S. C. and A. R. Horenstein (2013) “Eigenvalue ratio test for the number of factors”, *Econometrica*, 81(3): 1203–1227.

Amengual, D. and M. W. 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.

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

Beck, G. W., K. Hubrich and M. Marcellino (2009) “Regional inflation dynamics within and across euro area countries and a comparison with the US”, *Economic Policy*, 57: 141-184.

Beck, G. W., K. Hubrich and M. Marcellino (2016) “On the importance of sectoral and regional shocks for price-setting”, *Journal of Applied Econometrics*, 31: 1234-1253.

Chichocki, A., D. P. Mandic, A. H. Phan, C. F. Caiafa, G. Zhou, Q. Zhao and L. D. Lathauwer (2015) “Tensor decompositions for signal processing applications: From two-way to multiway component analysis”, *IEEE Signal Processing Magazine*, March: 145-163.

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

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

Connor, G. and R. Korajczyk (1993) “A test for the number of factors in an approximate factor model”, *Journal of Finance*, 48: 1263-1291.

Crucini, M. J., M. A. Kose and C. Otrok (2011) “What are the driving forces of international business cycles?”, *Review of Economic Dynamics*, 14: 156-175.

Davies, P.I. and N.J. Higham (2000) “Numerically Stable Generation of Correlation Matrices and Their Factors”, *BIT Numerical Mathematics*, 40: 640–651.

Dempster, A. P., N. M. Laird and D. B. Rubin (1977) “Maximum likelihood from incomplete data via the EM algorithm”, *Journal of the Royal Statistical Society*, Series B, 39: 1-22.

Dias, F., M. Pinheiro and A. Rua (2013) “Determining the number of global and country-specific factors in the euro area”, *Studies in Nonlinear Dynamics and Econometrics*, 17: 537-617.

Diebold, F. X., C. Li and V. Z. Yue (2008) “Global yield curve dynamics and interactions: A dynamic Nelson–Siegel approach”, *Journal of Econometrics*, 146: 351-363.

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

Foerster, A. T., P. D. G. Sarte and M. W. Watson (2011) "Sectoral versus aggregate shocks: A structural factor analysis of industrial production", *Journal of Political Economy*, 119: 1-38.

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

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.

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

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.

Forni, M. and M. Lippi (2001) "The generalized factor model: Representation theory", *Econometric Theory*, 17: 1113-1141.

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

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

Geweke, J. and K. Singleton (1981) "Maximum likelihood 'confirmatory' factor analysis of economic time series", *International Economic Review*, 22: 37-54.

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: 423-451.

Gregory, A. W., A. C. Head and J. Raynauld (1997) "Measuring world business cycles", *International Economic Review*, 38: 677-701.

Jungbacker, B. and S. Koopman (2015) "Likelihood-based dynamic factor analysis for measurement and forecasting", *The Econometrics Journal*, 18: C1-C21.

Karadimitropoulou, A. and M. León (2013) "World, country, and sector factors in international business cycles", *Journal of Economic Dynamics & Control*, 37: 2913-2927.

Kolda, T. G. and B. W. Bader (2009) "Tensor decompositions and applications", *SIAM Review*, 51: 455-500.

Kose, M. A., C. Otrok and E. Prasad (2012) "Global business cycles: Convergence or decoupling?", *International Economic Review*, 53: 511-538.

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

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

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

Meng, X.-L. and D. B. Rubin (1993) "Maximum likelihood estimation via ECM algorithm: A general framework", *Biometrika*, 80: 267-278.

Moench, E., S. Ng and S. Potter (2013) "Dynamic hierarchical factor models", *Review of Economics and Statistics*, 95: 1811-1817.

Mumtaz, H., S. Simonnelli and P. Surico (2011) "International comovements, business cycle and inflation: A historical perspective", *Review of Economic Dynamics*, 14: 176-198.

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.

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

Pinheiro, M., A. Rua and F. Dias (2013) "Dynamic factor models with jagged edge panel data: Taking on board the dynamics of the idiosyncratic components", *Oxford Bulletin of Economics and Statistics*, 75: 80-102.

Reis, R. and M. W. Watson (2010) "Relative goods' prices, pure inflation, and the Phillips correlation", *American Economic Journal: Macroeconomics*, 2: 128-157.

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

Sidiropoulos, N. D., L. D. Lathauwer, X. Fu, K. Huang, E. Papalexakis and C. Faloutsos (2017) "Tensor decomposition for signal processing and machine learning", *IEEE Transactions On Signal Processing*, 65: 3551-3582.

Smilde, A., R. Bro and P. Geladi (2004) *Multi-way Analysis With Applications in the Chemical Sciences*, Wiley, Chichester.

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

Stock, J. and M. W. Watson (2002a) "Forecasting using principal components from a large number of predictors", *Journal of the American Statistical Association*, 97: 1167-1179.

Stock, J. and M. W. Watson (2002b) "Macroeconomic Forecasting Using Diffusion Indexes", *Journal of Business and Economic Statistics*, 20(2): 147-162.

Stock, J. and M. W. Watson (2008) "The evolution of national and regional factors in U.S. housing construction"in T. Bollerslev, J. Russell and M. Watson (eds.) *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, Oxford University Press, Oxford.

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

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

Appendix

A.1. First order conditions - proofs

In what follows, we will denote:

- By $e_k^{(K)}$ the k-th column of the identity matrix of order K, i.e. the column-vector of size K with the value one in the k-th position and zeros elsewhere;
- For any real $(K_1 \times K_2)$ matrix C, by $\mathrm{trim}(C)$ the truncated $((K_1-1)\times K_2)$ matrix obtained from C by dropping its first row.

A.1.1. μ.

$$d_\mu \ell() = 0 \quad \forall d\mu \;\Leftrightarrow\; \left(\mathbf{f}_{0|T} - \mu\right)' \Psi^{-1} d\mu = 0 \quad \forall d\mu \;\Leftrightarrow\; \mu = \mathbf{f}_{0|T}$$

A.1.2. Ψ.

$$d_{\Psi^{-1}} \ell() = 0 \; \forall d(\Psi^{-1}) \;\Leftrightarrow\; \text{(using A.1.1)}$$

$$\Leftrightarrow\; \mathrm{tr}\left[\left(\Psi - \mathbf{V}_{0,0|T}\right) d(\Psi^{-1})\right] = 0 \; \forall d(\Psi^{-1}) \;\Leftrightarrow\; \Psi = \mathbf{V}_{0,0|T}$$

A.1.3. Γ.

$$d_\Gamma \ell() = 0 \quad \forall d\Gamma \quad \Leftrightarrow \quad \mathrm{tr}\left[\left(\bar{\mathbf{F}}_{\bullet,\bullet}\Gamma' - \bar{\mathbf{F}}'_{0,\bullet}\right) d\Gamma\right] = 0 \quad \forall d\Gamma \quad \Leftrightarrow \quad \Gamma = \bar{\mathbf{F}}_{0,\bullet}\bar{\mathbf{F}}^{-1}_{\bullet,\bullet}$$

A.1.4. Ω.

$$d_{\Omega^{-1}} \ell() = 0 \quad \forall d(\Omega^{-1}) \;\Leftrightarrow$$

$$\Leftrightarrow\; \mathrm{tr}\left\{\left[\Omega - \bar{F}_{0,0} - \Gamma\bar{\mathbf{F}}_{\bullet,\bullet}\Gamma' + +\bar{\mathbf{F}}_{0,\bullet}\Gamma' + \Gamma\bar{\mathbf{F}}'_{0,\bullet}\right] d(\Omega^{-1})\right\} = 0 \quad \forall d(\Omega^{-1})$$

$$\text{(using A.1.2)} \quad \Leftrightarrow \quad \Omega = \bar{\mathbf{F}}_{0,0} - \Gamma\bar{\mathbf{F}}_{\bullet,\bullet}\Gamma'$$

A.1.5. κ.

$$d_\kappa \ell() = 0 \quad \forall d\kappa \quad \Leftrightarrow \quad \Bigg[\frac{T}{T-1}\left(\bar{x}_{0,0} - \kappa\right) + \frac{T-2}{T-1}\mathrm{diag}(\rho)^2\left(\bar{x}_{1,1} - \kappa\right) +$$

$$-\mathrm{diag}(\rho)\left(\bar{x}_{0,1} + \bar{x}_{1,0} - 2\kappa\right) - \Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{f}_{0,0} - \mathrm{diag}(\rho)^2\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{f}_{1,1} +$$

$$+\mathrm{diag}(\rho)\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\left(\bar{f}_{0,1} + \bar{f}_{1,0}\right)\Bigg]' \mathrm{diag}(\sigma)^{-1} d\kappa = 0 \quad \forall d\kappa \quad \Leftrightarrow$$

$$\Leftrightarrow \kappa = \left(\frac{T}{T-1} I_{IJ} + \frac{T-2}{T-1}\mathrm{diag}(\rho)^2 - 2\mathrm{diag}(\rho)\right)^{-1} \cdot$$

$$\cdot\Bigg[\frac{T}{T-1}\bar{x}_{0,0} + \frac{T-2}{T-1}\mathrm{diag}(\rho)^2\bar{x}_{1,1} - \mathrm{diag}(\rho)\left(\bar{x}_{0,1} + \bar{x}_{1,0}\right) - \Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{f}_{0,0} +$$

$$-\mathrm{diag}(\rho)^2\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{f}_{1,1} + \mathrm{diag}(\rho)\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\left(\bar{f}_{0,1} + \bar{f}_{1,0}\right)\Bigg]$$

A.1.6. δ.

$$d_{\delta_{m,n}}\ell() = 0 \quad \forall d\delta_{mn} \quad \Leftrightarrow$$ [20]

$$\Leftrightarrow \quad e_{(n-1)M+m}^{(MN)\prime}\Theta(\bar{A},\bar{B})'\mathrm{diag}(\sigma)^{-1}\Big\{\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{\mathbb{F}}_{0,0}+$$

$$+\mathrm{diag}(\rho)^2\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{\mathbb{F}}_{1,1} - \mathrm{diag}(\rho)\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\left(\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}_{0,1}'\right)+$$

$$-\left[\bar{\mathbb{W}}_{0,0}+(\bar{x}_{0,0}-\kappa)\bar{f}_{0,0}'\right] - \mathrm{diag}(\rho)^2\left[\bar{\mathbb{W}}_{1,1}+(\bar{x}_{1,1}-\kappa)\bar{f}_{1,1}'\right]+$$

$$+\mathrm{diag}(\rho)\left[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+(\bar{x}_{0,1}-\kappa)\bar{f}_{1,0}'+(\bar{x}_{1,0}-\kappa)\bar{f}_{0,1}'\right]\Big\}e_{(n-1)M+m}^{(MN)} = 0 \quad \Leftrightarrow$$

$$\Leftrightarrow \sum_{\tilde{m}=1}^{M}\sum_{\tilde{n}=1}^{N}\delta_{\tilde{m},\tilde{n}}\Bigg\{\left[\bar{\mathbb{F}}_{0,0}\right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}}\sum_{i=1}^{I}\sum_{j=1}^{J}\sigma_{i,j}^{-1}\alpha_{i,m}\alpha_{i,\tilde{m}}\beta_{j,n}\beta_{j,\tilde{n}}+$$

$$+\left[\bar{\mathbb{F}}_{1,1}\right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}}\sum_{i=1}^{I}\sum_{j=1}^{J}\rho_{i,j}^2\sigma_{i,j}^{-1}\alpha_{i,m}\alpha_{i,\tilde{m}}\beta_{j,n}\beta_{j,\tilde{n}}+$$

$$-\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}_{0,1}'\right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}}\sum_{i=1}^{I}\sum_{j=1}^{J}\rho_{i,j}\sigma_{i,j}^{-1}\alpha_{i,m}\alpha_{i,\tilde{m}}\beta_{j,n}\beta_{j,\tilde{n}}\Bigg\} =$$

$$=\sum_{i=1}^{I}\sum_{j=1}^{J}\sigma_{i,j}^{-1}\left[\bar{\mathbb{W}}_{0,0}+(\bar{x}_{0,0}-\kappa)\bar{f}_{0,0}'\right]_{(j-1)I+i,(n-1)M+m}\alpha_{i,m}\beta_{j,n}+$$

$$+\sum_{i=1}^{I}\sum_{j=1}^{J}\rho_{i,j}^2\sigma_{i,j}^{-1}\left[\bar{\mathbb{W}}_{1,1}+(\bar{x}_{1,1}-\kappa)\bar{f}_{1,1}'\right]_{(j-1)I+i,(n-1)M+m}\alpha_{i,m}\beta_{j,n}+$$

$$-\sum_{i=1}^{I}\sum_{j=1}^{J}\rho_{i,j}\sigma_{i,j}^{-1}\Big[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+$$

$$+(\bar{x}_{0,1}-\kappa)\bar{f}_{1,0}'+(\bar{x}_{1,0}-\kappa)\bar{f}_{0,1}'\Big]_{(j-1)I+i,(n-1)M+m}\alpha_{i,m}\beta_{j,n}$$

Stacking these equations for $m = 1,\cdots,M$ and $n = 1,\cdots,N$, we have a linear system on δ conditional on values for the remaining parameters and the moments of observables and latent variables:

$$\mathbf{S}_\delta\delta = \mathbf{c}_\delta$$

with (for $m,\tilde{m} = 1,\cdots,M$, $n,\tilde{n} = 1,\cdots,N$, and with $\alpha_{i,1} = \beta_{j,1} = 1$ for all i and j):

$$[\mathbf{S}_\delta]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}} =$$

20. Note that $\delta_{m,n}$ is the $((n-1)M+m)$-th element of vector δ and thus it is the element $((n-1)M+m,(n-1)M+m)$ of $\mathrm{diag}(\delta)$, implying that $d_{\delta_{m,n}}\mathrm{diag}(\delta) = e_{(n-1)M+m}^{(MN)}e_{(n-1)M+m}^{(MN)\prime}d\delta_{m,n}$.

$$= \left[\bar{\mathbb{F}}_{0,0}\right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}} \sum_{i=1}^{I}\sum_{j=1}^{J} \sigma_{i,j}^{-1}\alpha_{i,m}\alpha_{i,\tilde{m}}\beta_{j,n}\beta_{j,\tilde{n}}+$$

$$+ \left[\bar{\mathbb{F}}_{1,1}\right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}} \sum_{i=1}^{I}\sum_{j=1}^{J} \rho_{i,j}^{2}\sigma_{i,j}^{-1}\alpha_{i,m}\alpha_{i,\tilde{m}}\beta_{j,n}\beta_{j,\tilde{n}}+$$

$$- \left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}_{0,1}'\right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}} \sum_{i=1}^{I}\sum_{j=1}^{J} \rho_{i,j}\sigma_{i,j}^{-1}\alpha_{i,m}\alpha_{i,\tilde{m}}\beta_{j,n}\beta_{j,\tilde{n}}$$

$$[\mathbf{c}_{\delta}]_{(n-1)M+m} =$$

$$= \sum_{i=1}^{I}\sum_{j=1}^{J} \Big\{ \sigma_{i,j}^{-1}\left[\bar{\mathbb{W}}_{0,0} + (\bar{x}_{0,0}-\kappa)\bar{f}_{0,0}'\right]_{(j-1)I+i,(n-1)M+m}\alpha_{i,m}\beta_{j,n}+$$

$$+\sigma_{i,j}^{-1}\left[\bar{\mathbb{W}}_{1,1} + (\bar{x}_{1,1}-\kappa)\bar{f}_{1,1}'\right]_{(j-1)I+i,(n-1)M+m}\alpha_{i,m}\beta_{j,n} - \rho_{i,j}\sigma_{i,j}^{-1}\cdot$$

$$\cdot\left[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0} + (\bar{x}_{0,1}-\kappa)\bar{f}_{1,0}' + (\bar{x}_{1,0}-\kappa)\bar{f}_{0,1}'\right]_{(j-1)I+i,(n-1)M+m}\alpha_{i,m}\beta_{j,n}\Big\}$$

A.1.7. $\bar{A}$ and $\bar{B}$. $\ell()$ depends on $\bar{A}$ and $\bar{B}$ only through $\Theta(\bar{A},\bar{B})$. Therefore, in order to obtain the first order conditions with respect to $\bar{A}$ and $\bar{B}$, we will first calculate the differential of $\ell()$ with respect to Θ:

$$d_{\Theta}\ell() = -(T-1)\mathrm{tr}\Big\{\mathrm{diag}(\delta)\Big[\bar{\mathbb{F}}_{0,0}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+$$

$$\bar{\mathbb{F}}_{1,1}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'\mathrm{diag}(\rho)^2 - \left(\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}_{0,1}'\right)\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'\mathrm{diag}(\rho)+$$

$$-\Big[\bar{\mathbb{W}}_{0,0} + (\bar{x}_{0,0}-\kappa)\bar{f}_{0,0}'\Big]' - \Big[\bar{\mathbb{W}}_{1,1} + (\bar{x}_{1,1}-\kappa)\bar{f}_{1,1}'\Big]'\mathrm{diag}(\rho)^2+$$

$$+\Big[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0} + (\bar{x}_{0,1}-\kappa)\bar{f}_{1,0}' + (\bar{x}_{1,0}-\kappa)\bar{f}_{0,1}'\Big]'\mathrm{diag}(\rho)\Big]\mathrm{diag}(\sigma)^{-1}d\Theta(\bar{A},\bar{B})\Big\}$$

Thus, taking into account that

$$d_{\bar{A}}\Theta(\bar{A},\bar{B}) = \left(\left[\begin{array}{cc}\mathbf{1}_J & \bar{B}\end{array}\right]\otimes \mathbf{I}_I\right)\left(\mathbf{I}_N\otimes\left[\begin{array}{cc}\mathbf{0}_I & d\bar{A}\end{array}\right]\right)$$

$$d_{\bar{B}}\Theta(\bar{A},\bar{B}) = \left(\mathbf{I}_J\otimes\left[\begin{array}{cc}\mathbf{1}_I & \bar{A}\end{array}\right]\right)\left(\left[\begin{array}{cc}\mathbf{0}_J & d\bar{B}\end{array}\right]\otimes \mathbf{I}_M\right)$$

we have:

(i) $\bar{A}$

$$d_{\bar{A}}\ell() = 0 \quad \forall d\bar{A} \quad \Leftrightarrow$$

$$\Leftrightarrow \quad \mathrm{tr}\Big\{Y\left(x_1,\cdots,x_T;\bar{A},\bar{B},\delta,\kappa,\rho,\sigma\right)\left(\mathbf{I}_N\otimes\left[\mathbf{0}_I\ d\bar{A}\right]\right)\Big\} = 0 \quad \forall d\bar{A}$$

where[21]

$$Y(\bar{A},\bar{B},\delta,\kappa,\rho,\sigma) =$$

$$= \mathrm{diag}(\delta)\Big[\bar{\mathbb{F}}_{0,0}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})' + \bar{\mathbb{F}}_{1,1}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'\mathrm{diag}(\rho)^2 +$$

$$-\left(\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}_{0,1}'\right)\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'\mathrm{diag}(\rho) - \Big[\bar{\mathbb{W}}_{0,0} + (\bar{x}_{0,0}-\kappa)\,\bar{f}_{0,0}'\Big]' +$$

$$-\Big[\bar{\mathbb{W}}_{1,1} + (\bar{x}_{1,1}-\kappa)\,\bar{f}_{1,1}'\Big]'\mathrm{diag}(\rho)^2 +$$

$$+\Big[\bar{\mathbb{W}}_{0,1} + \bar{\mathbb{W}}_{1,0} + (\bar{x}_{0,1}-\kappa)\,\bar{f}_{1,0}' + (\bar{x}_{1,0}-\kappa)\,\bar{f}_{0,1}'\Big]'\mathrm{diag}(\rho)\Big].$$

$$.\mathrm{diag}(\sigma)^{-1}\left(\begin{bmatrix} \mathbf{1}_J & \bar{B} \end{bmatrix}\otimes \mathbf{I}_I\right)$$

Therefore, making use of the equality between traces of matrices presented in Annex A.2.1, the first order conditions of the maximization of $\ell()$ with respect to $\bar{A}$ may be written as:

$$\mathrm{trim}\left(\sum_{n=1}^{N} Y_{[n,n]}(\bar{A},\bar{B},\delta,\kappa,\rho,\sigma)\right) = 0$$

with

$$Y_{[n,n]} = \mathrm{diag}(\delta_{[n]})\Big\{\sum_{\tilde{n}=1}^{N}\sum_{j=1}^{J}\beta_{j,n}\beta_{j,\tilde{n}}\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\mathbf{1}_I\ \bar{A}\right]' +$$

$$+\left[\bar{\mathbb{F}}_{1,1}\right]_{[n\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\mathbf{1}_I\ \bar{A}\right]'\mathrm{diag}(\rho_{[j]})^2 +$$

$$-\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}_{0,1}'\right]_{[n\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\mathbf{1}_I\ \bar{A}\right]'\mathrm{diag}(\rho_{[j]})\Big\}\mathrm{diag}(\sigma_{[j]})^{-1} +$$

$$-\sum_{j=1}^{J}\beta_{jn}\Big\{\left[\bar{\mathbb{W}}_{0,0} + (\bar{x}_{0,0}-\kappa)\,\bar{f}_{0,0}'\right]_{[jn]}' + \left[\bar{\mathbb{W}}_{1,1} + (\bar{x}_{1,1}-\kappa)\,\bar{f}_{1,1}'\right]_{[jn]}'\mathrm{diag}(\rho_{[j]})^2$$

$$-\left[\bar{\mathbb{W}}_{0,1} + \bar{\mathbb{W}}_{1,0} + (\bar{x}_{0,1}-\kappa)\,\bar{f}_{1,0}' + (\bar{x}_{1,0}-\kappa)\,\bar{f}_{0,1}'\right]_{[jn]}'\mathrm{diag}(\rho_{[j]})\Big\}\mathrm{diag}(\sigma_{[j]})^{-1}\Big\}$$

Thus,

$$\mathrm{trim}\left(\sum_{n=1}^{N} Y_{[n,n]}(\bar{A},\bar{B},\delta,\kappa,\rho,\sigma)\right) = 0 \quad \Leftrightarrow$$

$$\Leftrightarrow \quad \sum_{\tilde{n}=1}^{N}\sum_{n=1}^{N}\sum_{j=1}^{J}\beta_{j,n}\beta_{j,\tilde{n}}\mathrm{trim}\left(\mathrm{diag}(\delta_{[n]})\right)\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\mathbf{1}_I\ \bar{A}\right]' +$$

21. In order to simplify notation, we will keep implicit the dependence of $Y()$ on observables and time-factors.

$$+\left[\bar{\mathbb{F}}_{1,1}\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\mathbf{1}_I\ \bar{A}\right]'\mathrm{diag}(\rho_{[j]})^2+$$

$$-\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}_{0,1}'\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\mathbf{1}_I\ \bar{A}\right]'\mathrm{diag}(\rho_{[j]})\Big\}\mathrm{diag}(\sigma_{[j]})^{-1}=$$

$$=\sum_{n=1}^{N}\sum_{j=1}^{J}\beta_{j,n}\mathrm{trim}\left(\mathrm{diag}(\delta_{[n]})\right)\Big\{\left[\bar{\mathbb{W}}_{0,0}+(\bar{x}_{0,0}-\kappa)\bar{f}_{0,0}'\right]_{[j,n]}'+$$

$$+\left[\bar{\mathbb{W}}_{1,1}+(\bar{x}_{1,1}-\kappa)\bar{f}_{1,1}'\right]_{[j,n]}'\mathrm{diag}(\rho_{[j]})^2+$$

$$-\left[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+(\bar{x}_{0,1}-\kappa)\bar{f}_{1,0}'+(\bar{x}_{1,0}-\kappa)\bar{f}_{0,1}'\right]_{[j,n]}'\mathrm{diag}(\rho_{[j]})\Big\}\mathrm{diag}(\sigma_{[j]})^{-1}$$

or equivalently, denoting the transposed i-th row of $\bar{A}$ by $\alpha_{[i]}$, for $i=1,\cdots,I$:[22]

$$\sum_{\tilde{n}=1}^{N}\sum_{n=1}^{N}\sum_{j=1}^{J}\sigma_{i,j}^{-1}\beta_{j,n}\beta_{j,\tilde{n}}\mathrm{trim}\left(\mathrm{diag}(\delta_{[n]})\right)\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{[n,\tilde{n}]}+\rho_{i,j}^2\left[\bar{\mathbb{F}}_{1,1}\right]_{[n,\tilde{n}]}+$$

$$-\rho_{i,j}\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}_{0,1}'\right]_{[n\tilde{n}]}\Big\}\mathrm{diag}(\delta_{[\tilde{n}]})\begin{bmatrix}1\\ \alpha_{[i]}\end{bmatrix}=$$

$$=\sum_{n=1}^{N}\sum_{j=1}^{J}\sigma_{i,j}^{-1}\beta_{j,n}\mathrm{trim}\left(\mathrm{diag}(\delta_{[n]})\right)\Big\{\left[\bar{\mathbb{W}}_{0,0}+(\bar{x}_{0,0}-\kappa)\bar{f}_{0,0}'\right]_{[j,n]}'+$$

$$+\rho_{i,j}^2\left[\bar{\mathbb{W}}_{1,1}+(\bar{x}_{1,1}-\kappa)\bar{f}_{1,1}'\right]_{[j,n]}'+$$

$$-\rho_{i,j}\left[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+(\bar{x}_{0,1}-\kappa)\bar{f}_{1,0}'+(\bar{x}_{1,0}-\kappa)\bar{f}_{0,1}'\right]_{[j,n]}'\Big\}e_i^{(I)}$$

which, for $i=1,\cdots,I$ and $m=2,\cdots,M$, may also be written as:

$$\sum_{\tilde{m}=2}^{N}\alpha_{i,\tilde{m}}\Bigg\{\sum_{\tilde{n}=1}^{N}\sum_{n=1}^{N}\sum_{j=1}^{J}\Big\{\sigma_{i,j}^{-1}\delta_{m,n}\delta_{\tilde{m},\tilde{n}}\beta_{j,n}\beta_{j,\tilde{n}}\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}}+$$

$$+\rho_{i,j}^2\left[\bar{\mathbb{F}}_{1,1}\right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}}-\rho_{i,j}\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}_{0,1}'\right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}}\Big\}\Big\}\Bigg\}=$$

$$=\sum_{n=1}^{N}\sum_{j=1}^{J}\Bigg\{\sigma_{i,j}^{-1}\delta_{m,n}\beta_{j,n}\Big\{\Big\{\left[\bar{\mathbb{W}}_{0,0}+(\bar{x}_{0,0}-\kappa)\bar{f}_{0,0}'\right]_{(j-1)I+i,(n-1)M+m}+$$

$$+\rho_{i,j}^2\left[\bar{\mathbb{W}}_{1,1}+(\bar{x}_{1,1}-\kappa)\bar{f}_{1,1}'\right]_{(j-1)I+i,(n-1)M+m}+$$

$$-\rho_{i,j}\left[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+(\bar{x}_{0,1}-\kappa)\bar{f}_{1,0}'+(\bar{x}_{1,0}-\kappa)\bar{f}_{0,1}'\right]_{(j-1)I+i,(n-1)M+m}\Big\}+$$

$$-\sum_{\tilde{n}=1}^{N}\delta_{1,\tilde{n}}\beta_{j,\tilde{n}}\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{(n-1)M+m,(\tilde{n}-1)M+1}+\rho_{i,j}^2\left[\bar{\mathbb{F}}_{1,1}\right]_{(n-1)M+m,(\tilde{n}-1)M+1}+$$

22. With $\beta_{j,1}=1$ for all j.

$$-\rho_{i,j}\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}'_{0,1}\right]_{(n-1)M+m,(\tilde{n}-1)M+1}\Big\}\Big\}\Bigg\}$$

$$\Leftrightarrow \quad \mathbf{S}_{\alpha_{[i]}}\alpha_{[i]}=\mathbf{c}_{\alpha_{[i]}} \quad (i=1,\cdots,I)$$

with (for $m,\tilde{m}=2,\cdots,M$):

$$\left[\mathbf{S}_{\alpha_{[i]}}\right]_{m-1,\tilde{m}-1}=$$

$$=\sum_{\tilde{n}=1}^{N}\sum_{n=1}^{N}\sum_{j=1}^{J}\Bigg\{\sigma_{i,j}^{-1}\delta_{m,n}\delta_{\tilde{m},\tilde{n}}\beta_{j,n}\beta_{j,\tilde{n}}\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{(n-1)M+m,(\tilde{n}-1)N+\tilde{m}}+$$

$$+\rho_{i,j}^2\left[\bar{\mathbb{F}}_{1,1}\right]_{(n-1)M+m,(\tilde{n}-1)N+\tilde{m}}-\rho_{i,j}\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}'_{0,1}\right]_{(n-1)M+m,(\tilde{n}-1)N+\tilde{m}}\Big\}\Bigg\}$$

and

$$\left[\mathbf{c}_{\alpha_{[i]}}\right]_{m-1}=$$

$$=\sum_{n=1}^{N}\sum_{j=1}^{J}\Bigg\{\sigma_{i,j}^{-1}\delta_{m,n}\beta_{j,n}\Big\{\Big\{\left[\bar{\mathbb{W}}_{0,0}+\left(\bar{x}_{0,0}-\kappa\right)\bar{f}'_{0,0}\right]_{(j-1)I+i,(n-1)N+m}+$$

$$+\rho_{i,j}^2\left[\bar{\mathbb{W}}_{1,1}+\left(\bar{x}_{1,1}-\kappa\right)\bar{f}'_{1,1}\right]_{(j-1)I+i,(n-1)N+m}+$$

$$-\rho_{i,j}\left[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+\left(\bar{x}_{0,1}-\kappa\right)\bar{f}'_{1,0}+\left(\bar{x}_{1,0}-\kappa\right)\bar{f}'_{0,1}\right]_{(j-1)I+i,(n-1)N+m}\Big\}+$$

$$-\sum_{\tilde{n}=1}^{N}\delta_{1,\tilde{n}}\beta_{j,\tilde{n}}\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{(n-1)M+m,(\tilde{n}-1)N+1}+\rho_{i,j}^2\left[\bar{\mathbb{F}}_{1,1}\right]_{(n-1)M+m,(\tilde{n}-1)N+1}+$$

$$-\rho_{i,j}\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}'_{0,1}\right]_{(n-1)M+m,(\tilde{n}-1)N+1}\Big\}\Big\}\Bigg\}$$

(ii) $\bar{B}$

$$d_{\bar{B}}\ell()=0 \quad \forall d\bar{B} \quad \Leftrightarrow \quad \mathrm{tr}\Big\{Z(\bar{A},\bar{B},\delta,\kappa,\rho,\sigma)\left(\left[\mathbf{0}_J\ d\bar{B}\right]\otimes\mathbf{I}_M\right)\Big\}=0 \quad \forall d\bar{B}$$

where[23]

$$Z(\bar{A},\bar{B},\delta,\kappa,\rho,\sigma)=\mathrm{diag}(\delta)\Big[\bar{\mathbb{F}}_{0,0}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+\bar{\mathbb{F}}_{1,1}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'\mathrm{diag}(\rho)^2+$$

$$-\left(\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}'_{0,1}\right)\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'\mathrm{diag}(\rho)-\Big[\bar{\mathbb{W}}_{0,0}+\left(\bar{x}_{0,0}-\kappa\right)\bar{f}'_{0,0}\Big]'+$$

23. Again, in order to simplify notation, we will keep implicit the dependence of $Z()$ on observables and time-factors.

$$-\Big[\bar{\mathbb{W}}_{1,1}+(\bar{x}_{1,1}-\kappa)\,\bar{f}'_{1,1}\Big]'\mathrm{diag}(\rho)^2++\Big[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+$$

$$+(\bar{x}_{0,1}-\kappa)\,\bar{f}'_{1,0}+(\bar{x}_{1,0}-\kappa)\,\bar{f}'_{0,1}\Big]'\mathrm{diag}(\rho)\Big]\mathrm{diag}(\sigma)^{-1}\left(\mathbf{I}_J\otimes\left[\begin{array}{cc}\mathbf{1}_I & \bar{A}\end{array}\right]\right)$$

Thus, making use of the equality between traces of matrices presented in Annex A.2.2, the first order conditions of the maximization of $\ell()$ with respect to $\bar{B}$ are:

$$\Upsilon(\bar{A},\bar{B},\delta,\kappa,\rho,\sigma)=0$$

where

$$\underset{((N-1)\times J)}{\Upsilon}=\left[\begin{array}{ccccc}\mathrm{tr}\left(Z_{[2,1]}\right) & \cdots & \mathrm{tr}\left(Z_{[2,j]}\right) & \cdots & \mathrm{tr}\left(Z_{[2,J]}\right)\\ \vdots & & \vdots & & \vdots\\ \mathrm{tr}\left(Z_{[n,1]}\right) & \cdots & \mathrm{tr}\left(Z_{[n,j]}\right) & \cdots & \mathrm{tr}\left(Z_{[n,J]}\right)\\ \vdots & & \vdots & & \vdots\\ \mathrm{tr}\left(Z_{[N,1]}\right) & \cdots & \mathrm{tr}\left(Z_{[N,j]}\right) & \cdots & \mathrm{tr}\left(Z_{[N,J]}\right)\end{array}\right]$$

and $Z_{[n,j]}$ $(n=2,\cdots,N;j=1,\cdots,J)$ are blocks of $Z()$ with size $(M\times M)$ obtained after breaking down the latter matrix as follows:[24]

$$\underset{(MN\times MJ)}{Z(\bar{A},\bar{B},\delta,\kappa,\rho,\sigma)}=\left[\begin{array}{cccccc}Z_{[1,1]} & Z_{[1,2]} & \cdots & Z_{[1,j]} & \cdots & Z_{[1,J]}\\ Z_{[2,1]} & Z_{[2,2]} & \cdots & Z_{[2,j]} & \cdots & Z_{[2,J]}\\ \vdots & \vdots & & \vdots & & \vdots\\ Z_{[n,1]} & Z_{[n,2]} & \cdots & Z_{[n,j]} & \cdots & Z_{[n,J]}\\ \vdots & & \vdots & & \vdots & \\ Z_{[N,1]} & Z_{[N,2]} & \cdots & Z_{[N,j]} & \cdots & Z_{[N,J]}\end{array}\right]$$

From the expression for $Z()$, the generic bloc $Z_{[n,j]}$ may be written as:

$$Z_{[nj]}=\mathrm{diag}(\delta_{[n]})\Big\{\sum_{\tilde{n}=1}^{N}\beta_{j,\tilde{n}}\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\begin{array}{cc}\mathbf{1}_I & \bar{A}\end{array}\right]'+$$

$$+\left[\bar{\mathbb{F}}_{1,1}\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\begin{array}{cc}\mathbf{1}_I & \bar{A}\end{array}\right]'\mathrm{diag}(\rho_{[j]})^2+$$

$$-\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}'_{0,1}\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\begin{array}{cc}\mathbf{1}_I & \bar{A}\end{array}\right]'\mathrm{diag}(\rho_{[j]})\Big\}+$$

$$-\Big\{\left[\bar{\mathbb{W}}_{0,0}+(\bar{x}_{0,0}-\kappa)\,\bar{f}'_{0,0}\right]'_{[jn]}+\left[\bar{\mathbb{W}}_{1,1}+(\bar{x}_{1,1}-\kappa)\,\bar{f}'_{1,1}\right]'_{[j,n]}\mathrm{diag}(\rho_{[j]})^2+$$

$$-\left[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+(\bar{x}_{0,1}-\kappa)\,\bar{f}'_{1,0}+(\bar{x}_{1,0}-\kappa)\,\bar{f}'_{0,1}\right]'_{[j,n]}\cdot$$

$$.\mathrm{diag}(\rho_{[j]})\Big\}\Big\}\mathrm{diag}(\sigma_{[j]})^{-1}\left[\begin{array}{cc}\mathbf{1}_I & \bar{A}\end{array}\right]$$

24. Note that the first row of blocks of $Z()$ is disregarded when calculating Υ.

Taking into account that $\beta_{j1}=1$ for all j, we get:

$$Z_{[n,j]}=\mathrm{diag}(\delta_{[n]})\Bigg\{\sum_{\tilde{n}=2}^{N}\beta_{j,\tilde{n}}\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\mathbf{1}_I\ \bar{A}\right]'+$$

$$+\left[\bar{\mathbb{F}}_{1,1}\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\mathbf{1}_I\ \bar{A}\right]'\mathrm{diag}(\rho_{[j]})^2+$$

$$-\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}_{0,1}'\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\mathbf{1}_I\ \bar{A}\right]'\mathrm{diag}(\rho_{[j]})\Big\}+$$

$$-\Big\{\left[\bar{\mathbb{W}}_{0,0}+(\bar{x}_{0,0}-\kappa)\,\bar{f}_{0,0}'\right]_{[j,n]}'+\left[\bar{\mathbb{W}}_{1,1}+(\bar{x}_{1,1}-\kappa)\,\bar{f}_{1,1}'\right]_{[j,n]}'\mathrm{diag}(\rho_{[j]})^2+$$

$$-\left[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+(\bar{x}_{0,1}-\kappa)\,\bar{f}_{1,0}'+(\bar{x}_{1,0}-\kappa)\,\bar{f}_{0,1}'\right]_{[j,n]}'\mathrm{diag}(\rho_{[j]})\Big\}+$$

$$+\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{[n,1]}\mathrm{diag}(\delta_{[1]})\left[\mathbf{1}_I\ \bar{A}\right]'+\left[\bar{\mathbb{F}}_{1,1}\right]_{[n,1]}\mathrm{diag}(\delta_{[1]})\left[\mathbf{1}_I\ \bar{A}\right]'\mathrm{diag}(\rho_{[j]})^2+$$

$$-\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}_{0,1}'\right]_{[n,1]}\mathrm{diag}(\delta_{[1]})\left[\mathbf{1}_I\ \bar{A}\right]'\mathrm{diag}(\rho_{[j]})\Big\}\Bigg\}\mathrm{diag}(\sigma_{[j]})^{-1}\left[\mathbf{1}_I\ \bar{A}\right]$$

Thus, by equating to zero the trace of each block $Z_{[n,j]}$, we have (for $n=2,\cdots,N;j=1,\cdots,J$):

$$\sum_{\tilde{n}=2}^{N}\beta_{j,\tilde{n}}\mathrm{tr}\Bigg\{\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\mathbf{1}_I\ \bar{A}\right]'+$$

$$+\left[\bar{\mathbb{F}}_{1,1}\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\mathbf{1}_I\ \bar{A}\right]'\mathrm{diag}(\rho_{[j]})^2+$$

$$-\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}_{0,1}'\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\mathbf{1}_I\ \bar{A}\right]'\mathrm{diag}(\rho_{[j]})\Big\}\mathrm{diag}(\sigma_{[j]})^{-1}\left[\mathbf{1}_I\ \bar{A}\right]\mathrm{diag}(\delta_{[n]})\Bigg\}=$$

$$=\mathrm{tr}\Bigg\{\Big\{\Big\{\left[\bar{\mathbb{W}}_{0,0}+(\bar{x}_{0,0}-\kappa)\,\bar{f}_{0,0}'\right]_{[j,n]}'+\left[\bar{\mathbb{W}}_{1,1}+(\bar{x}_{1,1}-\kappa)\,\bar{f}_{1,1}'\right]_{[j,n]}'\mathrm{diag}(\rho_{[j]})^2+$$

$$-\left[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+(\bar{x}_{0,1}-\kappa)\,\bar{f}_{1,0}'+(\bar{x}_{1,0}-\kappa)\,\bar{f}_{0,1}'\right]_{[j,n]}'\mathrm{diag}(\rho_{[j]})\Big\}+$$

$$-\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{[n,1]}\mathrm{diag}(\delta_{[1]})\left[\mathbf{1}_I\ \bar{A}\right]'+\left[\bar{\mathbb{F}}_{1,1}\right]_{[n,1]}\mathrm{diag}(\delta_{[1]})\left[\mathbf{1}_I\ \bar{A}\right]'\mathrm{diag}(\rho_{[j]})^2+$$

$$-\left[\bar{\mathbb{F}}_{0,1}+\bar{\mathbb{F}}_{0,1}'\right]_{[n,1]}\mathrm{diag}(\delta_{[1]})\left[\mathbf{1}_I\ \bar{A}\right]'\mathrm{diag}(\rho_{[j]})\Big\}\Big\}\mathrm{diag}(\sigma_{[j]})^{-1}\left[\mathbf{1}_I\ \bar{A}\right]\mathrm{diag}(\delta_{[n]})\Bigg\}$$

$$\Leftrightarrow\quad \mathbf{S}_{\beta_{[j]}}\beta_{[j]}=\mathbf{c}_{\beta_{[j]}}\quad(j=1,\cdots,J)$$

with (for $n,\tilde{n}=2,\cdots,N$):

$$\left[\mathbf{S}_{\beta_{[j]}}\right]_{n-1,\tilde{n}-1}=$$

$$=\mathrm{tr}\Bigg\{\Big\{\left[\bar{\mathbb{F}}_{0,0}\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\mathbf{1}_I\ \bar{A}\right]'+\left[\bar{\mathbb{F}}_{1,1}\right]_{[n,\tilde{n}]}\mathrm{diag}(\delta_{[\tilde{n}]})\left[\mathbf{1}_I\ \bar{A}\right]'\mathrm{diag}(\rho_{[j]})^2+$$

$$- \left[\bar{\mathbb{F}}_{0,1} + \bar{\mathbb{F}}_{0,1}'\right]_{[n,\tilde{n}]} \mathrm{diag}(\delta_{[\tilde{n}]}) \left[\mathbf{1}_I\ \bar{A}\right]' \mathrm{diag}(\rho_{[j]}) \Big\} \mathrm{diag}(\sigma_{[j]})^{-1} \left[\mathbf{1}_I\ \bar{A}\right] \mathrm{diag}(\delta_{[n]}) \Big\} =$$

and

$$\left[\mathbf{c}_{\beta_{[j]}}\right]_{n-1} =$$

$$= \mathrm{tr}\Bigg\{ \Bigg\{ \Big\{ \left[\bar{\mathbb{W}}_{0,0} + (\bar{x}_{0,0} - \kappa)\,\bar{f}_{0,0}'\right]_{[j,n]}' + \left[\bar{\mathbb{W}}_{1,1} + (\bar{x}_{1,1} - \kappa)\,\bar{f}_{1,1}'\right]_{[j,n]}' \mathrm{diag}(\rho_{[j]})^2 +$$

$$- \left[\bar{\mathbb{W}}_{0,1} + \bar{\mathbb{W}}_{1,0} + (\bar{x}_{0,1} - \kappa)\,\bar{f}_{1,0}' + (\bar{x}_{1,0} - \kappa)\,\bar{f}_{0,1}'\right]_{[jn]}' \mathrm{diag}(\rho_{[j]}) \Big\} +$$

$$- \Big\{ \left[\bar{\mathbb{F}}_{0,0}\right]_{[n,1]} \mathrm{diag}(\delta_{[1]}) \left[\mathbf{1}_I\ \bar{A}\right]' + \left[\bar{\mathbb{F}}_{1,1}\right]_{[n,1]} \mathrm{diag}(\delta_{[1]}) \left[\mathbf{1}_I\ \bar{A}\right]' \mathrm{diag}(\rho_{[j]})^2 +$$

$$- \left[\bar{\mathbb{F}}_{0,1} + \bar{\mathbb{F}}_{0,1}'\right]_{[n,1]} \mathrm{diag}(\delta_{[1]}) \left[\mathbf{1}_I\ \bar{A}\right]' \mathrm{diag}(\rho_{[j]}) \Big\} \Bigg\} \mathrm{diag}(\sigma_{[j]})^{-1} \left[\mathbf{1}_I\ \bar{A}\right] \mathrm{diag}(\delta_{[n]}) \Bigg\}$$

A.1.8. σ. For all i and j:

$$d_{\sigma_{i,j}^{-1}} \ell() = 0 \quad \forall d(\sigma_{i,j}^{-1}) \quad \Leftrightarrow$$ [25]

$$\Leftrightarrow \quad \sigma_{i,j} = \frac{T-1}{T} e_{(j-1)I+i}^{(IJ)\prime} \Big\{ \Big[\bar{\mathbb{X}}_{0,0} + \frac{T}{T-1} (\bar{x}_{0,0} - \kappa)(\bar{x}_{0,0} - \kappa)' \Big] +$$

$$+ \mathrm{diag}(\rho)^2 \Big[\bar{\mathbb{X}}_{1,1} + \frac{T-2}{T-1} (\bar{x}_{1,1} - \kappa)(\bar{x}_{1,1} - \kappa)' \Big] +$$

$$-2\mathrm{diag}(\rho) \Big[\bar{\mathbb{X}}_{0,1} + (\bar{x}_{0,1} - \kappa)(\bar{x}_{1,0} - \kappa)' \Big] + \Theta(\bar{A},\bar{B}) \mathrm{diag}(\delta) \bar{\mathbb{F}}_{0,0} \mathrm{diag}(\delta) \Theta(\bar{A},\bar{B})' +$$

$$+ \mathrm{diag}(\rho)^2 \Theta(\bar{A},\bar{B}) \mathrm{diag}(\delta) \bar{\mathbb{F}}_{1,1} \mathrm{diag}(\delta) \Theta(\bar{A},\bar{B})' +$$

$$-2\mathrm{diag}(\rho) \Theta(\bar{A},\bar{B}) \mathrm{diag}(\delta) \bar{\mathbb{F}}_{0,1} \mathrm{diag}(\delta) \Theta(\bar{A},\bar{B})' +$$

$$-2 \Big[\bar{\mathbb{W}}_{0,0} + (\bar{x}_{0,0} - \kappa)\,\bar{f}_{0,0}' \Big] \mathrm{diag}(\delta) \Theta(\bar{A},\bar{B})' +$$

$$-2\mathrm{diag}(\rho)^2 \Big[\bar{\mathbb{W}}_{1,1} + (\bar{x}_{1,1} - \kappa)\,\bar{f}_{1,1}' \Big] \mathrm{diag}(\delta) \Theta(\bar{A},\bar{B})' +$$

$$+2\mathrm{diag}(\rho) \Big[\bar{\mathbb{W}}_{0,1} + \bar{\mathbb{W}}_{1,0} + (\bar{x}_{0,1} - \kappa)\,\bar{f}_{1,0}' + (\bar{x}_{1,0} - \kappa)\,\bar{f}_{0,1}' \Big].$$

$$.\mathrm{diag}(\delta) \Theta(\bar{A},\bar{B})' \Big\} e_{(j-1)I+i}^{(IJ)}$$

Stacking the equations for all i and j:

$$\sigma = \frac{T-1}{T} \mathrm{diag}\Big\{ \Big[\bar{\mathbb{X}}_{0,0} + \frac{T}{T-1} (\bar{x}_{0,0} - \kappa)(\bar{x}_{0,0} - \kappa)' \Big] + \mathrm{diag}(\rho)^2 \Big[\bar{\mathbb{X}}_{1,1} +$$

25. Note that $d_{\left(\sigma_{i,j}^{-1}\right)} \mathrm{diag}(\sigma)^{-1} = e_{(j-1)I+i}^{(IJ)} e_{(j-1)I+i}^{(IJ)\prime} d\left(\sigma_{i,j}^{-1}\right)$.

$$+\frac{T-2}{T-1}\left(\bar{x}_{1,1}-\kappa\right)\left(\bar{x}_{1,1}-\kappa\right)'\Big]-2\mathrm{diag}(\rho)\Big[\bar{\mathbb{X}}_{0,1}+\left(\bar{x}_{0,1}-\kappa\right)\left(\bar{x}_{1,0}-\kappa\right)'\Big]+$$
$$+\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{\mathbb{F}}_{0,0}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+$$
$$+\mathrm{diag}(\rho)^2\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{\mathbb{F}}_{1,1}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+$$
$$-2\mathrm{diag}(\rho)\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{\mathbb{F}}_{0,1}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+$$
$$-2\Big[\bar{\mathbb{W}}_{0,0}+\left(\bar{x}_{0,0}-\kappa\right)\bar{f}'_{0,0}\Big]\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+$$
$$-2\mathrm{diag}(\rho)^2\Big[\bar{\mathbb{W}}_{1,1}+\left(\bar{x}_{1,1}-\kappa\right)\bar{f}'_{1,1}\Big]\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+$$
$$+2\mathrm{diag}(\rho)\Big[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+\left(\bar{x}_{0,1}-\kappa\right)\bar{f}'_{1,0}+\left(\bar{x}_{1,0}-\kappa\right)\bar{f}'_{0,1}\Big]\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'\Big\}$$

A.1.9. ρ. For all i and j:

$$d_{\rho_{i,j}}\ell()=0\quad\forall d\rho_{i,j}\quad\Leftrightarrow^{26}$$
$$\Leftrightarrow\ \frac{\rho_{i,j}}{1-\rho_{i,j}^2}+\frac{(T-1)}{\sigma_{i,j}}e^{(IJ)\prime}_{(j-1)I+i}\Big\{\rho_{i,j}\Big[\bar{\mathbb{X}}_{1,1}+\frac{T-2}{T-1}\left(\bar{x}_{1,1}-\kappa\right)\left(\bar{x}_{1,1}-\kappa\right)'\Big]+$$
$$-\Big[\bar{\mathbb{X}}_{0,1}+\left(\bar{x}_{0,1}-\kappa\right)\left(\bar{x}_{1,0}-\kappa\right)'\Big]+\rho_{i,j}\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{\mathbb{F}}_{1,1}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+$$
$$-\Theta(\bar{A},\bar{B})\mathrm{diag}(\delta)\bar{\mathbb{F}}_{0,1}\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+$$
$$-2\rho_{i,j}\Big[\bar{\mathbb{W}}_{1,1}+\left(\bar{x}_{1,1}-\kappa\right)\bar{f}'_{1,1}\Big]\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'+$$
$$+\Big[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+\left(\bar{x}_{0,1}-\kappa\right)\bar{f}'_{1,0}+\left(\bar{x}_{1,0}-\kappa\right)\bar{f}'_{0,1}\Big].$$
$$.\mathrm{diag}(\delta)\Theta(\bar{A},\bar{B})'\Big\}e^{(IJ)}_{(j-1)I+i}=0\quad\Leftrightarrow$$
$$\Leftrightarrow\ \rho_{ij}=\Bigg\{\Big[\bar{\mathbb{X}}_{1,1}+\frac{T-2}{T-1}\left(\bar{x}_{1,1}-\kappa\right)\left(\bar{x}_{1,1}-\kappa\right)'\Big]_{(j-1)I+i,(j-1)I+i}+$$
$$+\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{\tilde{m}=1}^{M}\sum_{\tilde{n}=1}^{N}\delta_{m,n}\delta_{\tilde{m},\tilde{n}}\alpha_{i,m}\alpha_{i,\tilde{m}}\beta_{j,n}\beta_{j,\tilde{n}}\left[\bar{\mathbb{F}}_{1,1}\right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}}+$$
$$-2\sum_{m=1}^{M}\sum_{n=1}^{N}\delta_{m,n}\alpha_{i,m}\beta_{j,n}\Big[\left(\bar{x}_{1,1}-\kappa\right)\bar{f}'_{1,1}\Big]_{(j-1)I+i,(n-1)M+m}+$$
$$+\frac{\sigma_{i,j}}{(T-1)\left(1-\rho_{i,j}^2\right)}\Bigg\}^{-1}\Bigg\{\Big[\bar{\mathbb{X}}_{0,1}+\left(\bar{x}_{0,1}-\kappa\right)\left(\bar{x}_{1,0}-\kappa\right)'\Big]_{(j-1)I+i,(j-1)I+i}+$$

26. Note that $d_{\rho_{i,j}}\left[\mathrm{diag}(\sigma)^{-1}\mathrm{diag}(\rho)\right]=\sigma_{i,j}^{-1}e^{(IJ)}_{(j-1)I+i}e^{(IJ)\prime}_{(j-1)I+i}d\rho_{i,j}$ and $d_{\rho_{i,j}}\left[\mathrm{diag}(\sigma)^{-1}\mathrm{diag}(\rho)^2\right]=2\rho_{i,j}\sigma_{i,j}^{-1}e^{(IJ)}_{(j-1)I+i}e^{(IJ)\prime}_{(j-1)I+i}d\rho_{i,j}$.

$$+\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{\tilde{m}=1}^{M}\sum_{\tilde{n}=1}^{N}\delta_{m,n}\delta_{\tilde{m},\tilde{n}}\alpha_{i,m}\alpha_{i,\tilde{m}}\beta_{j,n}\beta_{j,\tilde{n}}\left[\bar{\mathbb{F}}_{0,1}\right]_{(n-1)M+m,(\tilde{n}-1)M+\tilde{m}}+$$

$$-\sum_{m=1}^{M}\sum_{n=1}^{N}\delta_{m,n}\alpha_{i,m}\beta_{j,n}\Big[\bar{\mathbb{W}}_{0,1}+\bar{\mathbb{W}}_{1,0}+\left(\bar{x}_{0,1}-\kappa\right)\bar{f}'_{1,0}+$$

$$\left.+\left(\bar{x}_{1,0}-\kappa\right)\bar{f}'_{0,1}\Big]_{(j-1)I+i,(n-1)M+m}\right\}$$

A.2. Useful identities on traces of matrices

Let H and C be matrices $(K_1\times K_2)$ and $(K_2Q\times K_1Q)$, respectively, with $H=\left[\begin{array}{cc}\mathbf{0}_{K_1} & \bar{H}\end{array}\right]$ and

$$\bar{H}=\begin{bmatrix} h_{1,2} & \cdots & h_{1,k_2} & \cdots & h_{1,K_2}\\ \vdots & & \vdots & & \vdots\\ h_{k_1,2} & \cdots & h_{k_1,k_2} & \cdots & h_{k_1,K_2}\\ \vdots & & \vdots & & \vdots\\ h_{K_1,2} & \cdots & h_{K_1,k_2} & \cdots & h_{K_1,K_2}\end{bmatrix}$$

A.2.1. Expressing $\mathrm{tr}\big\{C\left(\boldsymbol{I}_Q\otimes\left[\boldsymbol{0}_{K_1}\ \bar{H}\right]\right)\big\}$ *as* $\mathrm{tr}\left(\Pi\bar{H}\right)$. In order to simplify $\mathrm{tr}\left[C\left(\mathbf{I}_Q\otimes\left[\begin{array}{cc}\mathbf{0}_{K_1} & \bar{H}\end{array}\right]\right)\right]$, it is convenient to breakdown C into blocks $C_{[q_2,q_1]}$ $(q_1,q_2=1,\cdots,Q)$, each of dimension $(K_2\times K_1)$, allowing us to write:

$$C\left(\mathbf{I}_Q\otimes\left[\mathbf{0}_{K_1}\ \bar{H}\right]\right)=$$

$$=\begin{bmatrix} C_{[1,1]} & \cdots & C_{[1,q_1]} & \cdots & C_{[1,Q]}\\ \vdots & & \vdots & & \vdots\\ C_{[q_2,1]} & \cdots & C_{[q_2,q_1]} & \cdots & C_{[q_2,Q]}\\ \vdots & & \vdots & & \vdots\\ C_{[Q,1]} & \cdots & C_{[Q,q_1]} & \cdots & C_{[Q,Q]}\end{bmatrix}\cdot$$

$$\cdot\begin{bmatrix} \left[\mathbf{0}_{K_1}\ \bar{H}\right] & \cdots & \mathbf{0}_{K_1\times K_2} & \cdots & \mathbf{0}_{K_1\times K_2}\\ \vdots & & \vdots & & \vdots\\ \mathbf{0}_{K_1\times K_2} & \cdots & \left[\mathbf{0}_{K_1}\ \bar{H}\right] & \cdots & \mathbf{0}_{K_1\times K_2}\\ \vdots & & \vdots & & \vdots\\ \mathbf{0}_{K_1\times K_2} & \cdots & \mathbf{0}_{K_1\times K_2} & \cdots & \left[\mathbf{0}_{K_1}\ \bar{H}\right]\end{bmatrix}=$$

$$=\begin{bmatrix} C_{[1,1]}\left[\mathbf{0}_{K_1}\ \bar{H}\right] & \cdots & C_{[1,q_1]}\left[\mathbf{0}_{K_1}\ \bar{H}\right] & \cdots & C_{[1,Q]}\left[\mathbf{0}_{K_1}\ \bar{H}\right]\\ \vdots & & \vdots & & \vdots\\ C_{[q_2,1]}\left[\mathbf{0}_{K_1}\ \bar{H}\right] & \cdots & C_{[q_2,q_1]}\left[\mathbf{0}_{K_1}\ \bar{H}\right] & \cdots & C_{[q_2,Q]}\left[\mathbf{0}_{K_1}\ \bar{H}\right]\\ \vdots & & \vdots & & \vdots\\ C_{[Q,1]}\left[\mathbf{0}_{K_1}\ \bar{H}\right] & \cdots & C_{[Q,q_1]}\left[\mathbf{0}_{K_1}\ \bar{H}\right] & \cdots & C_{[Q,Q]}\left[\mathbf{0}_{K_1}\ \bar{H}\right]\end{bmatrix}$$

Thus,

$$\mathrm{tr}\Big\{C\left(\mathbf{I}_Q \otimes \left[\mathbf{0}_{K_1}\ \bar{H}\right]\right)\Big\} = \mathrm{tr}\Big\{\left(\sum_{q=1}^{Q} C_{[q,q]}\right)\left[\mathbf{0}_{K_1}\ \bar{H}\right]\Big\} = \mathrm{tr}\left(\Pi\bar{H}\right)$$

where Π is a matrix $((K_2-1)\times K_1)$ obtained from $\sum_{q=1}^{Q} C_{[q,q]}$ by dropping the first row of the latter matrix, i.e. $\Pi = \mathrm{trim}\left(\sum_{q=1}^{Q} C_{[q,q]}\right)$.

A.2.2. Expressing $\mathrm{tr}\big\{C\left(\left[\mathbf{0}_{K_1}\ \bar{H}\right]\otimes\mathbf{I}_Q\right)\big\}$ *as* $\mathrm{tr}\left(\Upsilon\bar{H}\right)$. In order to simplify $\mathrm{tr}\left[C\left(\left[\ \mathbf{0}_{K_1}\quad \bar{H}\ \right]\otimes\mathbf{I}_Q\right)\right]$, it is convenient to breakdown C into blocks $C_{[k_2k_1]}$ $(k_1 = 1,\cdots,K_1; k_2 = 1,\cdots,K_2)$, each of dimension $(Q\times Q)$, allowing us to write:

$$C\left(\left[\mathbf{0}_{K_1}\ \bar{H}\right]\otimes\mathbf{I}_Q\right) =$$

$$= \begin{bmatrix} C_{[1,1]} & C_{[1,2]} & \cdots & C_{[1,k_1]} & \cdots & C_{[1,K_1]} \\ C_{[2,1]} & C_{[2,2]} & \cdots & C_{[2,k_1]} & \cdots & C_{[2,K_1]} \\ \vdots & \vdots & & \vdots & & \vdots \\ C_{[k_2,1]} & C_{[k_2,2]} & \cdots & C_{[k_2,k_1]} & \cdots & C_{[k_2,K_1]} \\ \vdots & \vdots & & \vdots & & \vdots \\ C_{[K_2,1]} & C_{[K_2,2]} & \cdots & C_{[K_2,k_1]} & \cdots & C_{[K_2,K_1]} \end{bmatrix}.$$

$$\cdot\begin{bmatrix} \mathbf{0}_{Q\times Q} & h_{1,2}\mathbf{I}_Q & \cdots & h_{1,k_2}\mathbf{I}_Q & \cdots & h_{1K_2}\mathbf{I}_Q \\ \mathbf{0}_{Q\times Q} & h_{2,2}\mathbf{I}_Q & \cdots & h_{2,k_2}\mathbf{I}_Q & \cdots & h_{2,K_2}\mathbf{I}_Q \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathbf{0}_{Q\times Q} & h_{k_1,2}\mathbf{I}_Q & \cdots & h_{k_1,k_2}\mathbf{I}_Q & \cdots & h_{k_1,K_2}\mathbf{I}_Q \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathbf{0}_{Q\times Q} & h_{K_1,2}\mathbf{I}_Q & \cdots & h_{K_1,k_2}\mathbf{I}_Q & \cdots & h_{K_1,K_2}\mathbf{I}_Q \end{bmatrix} =$$

$$= \begin{bmatrix} \mathbf{0}_{Q\times Q} & \sum_{k_1=1}^{K_1} h_{k_1,2}C_{[1,k_1]} & \cdots & \sum_{k_1=1}^{K_1} h_{k_1,K_2}C_{[1,k_1]} \\ \mathbf{0}_{Q\times Q} & \sum_{k_1=1}^{K_1} h_{k_1,2}C_{[2,k_1]} & \cdots & \sum_{k_1=1}^{K_1} h_{k_1,K_2}C_{[2,k_1]} \\ \vdots & \vdots & & \vdots \\ \mathbf{0}_{Q\times Q} & \sum_{k_1=1}^{K_1} h_{k_1,2}C_{[k_2,k_1]} & \cdots & \sum_{k_1=1}^{K_1} h_{k_1,K_2}C_{[k_2,k_1]} \\ \vdots & \vdots & & \vdots \\ \mathbf{0}_{Q\times Q} & \sum_{k_1=1}^{K_1} h_{k_1,2}C_{[K_2,k_1]} & \cdots & \sum_{k_1=1}^{K_1} h_{k_1,K_2}C_{[K_2,k_1]} \end{bmatrix}$$

Thus,

$$\mathrm{tr}\Big\{C\left(\left[\mathbf{0}_{K_1}\ \bar{H}\right]\otimes\mathbf{I}_Q\right)\Big\} = \sum_{k_2=2}^{K_2}\sum_{k_1=1}^{K_1}\mathrm{tr}\left(C_{[k_2,k_1]}\right)h_{k_1,k_2} = \mathrm{tr}\left(\Upsilon\bar{H}\right)$$

where

$$\underset{((K_2-1)\times K_1)}{\Upsilon} = \begin{bmatrix} \mathrm{tr}\left(C_{[2,1]}\right) & \cdots & \mathrm{tr}\left(C_{[2,k_1]}\right) & \cdots & \mathrm{tr}\left(C_{[2,K_1]}\right) \\ \vdots & & \vdots & & \vdots \\ \mathrm{tr}\left(C_{[k_2,1]}\right) & \cdots & \mathrm{tr}\left(C_{[k_2,k_1]}\right) & \cdots & \mathrm{tr}\left(C_{[k_2,K_1]}\right) \\ \vdots & & \vdots & & \vdots \\ \mathrm{tr}\left(C_{[K_2,1]}\right) & \cdots & \mathrm{tr}\left(C_{[K_2,k_1]}\right) & \cdots & \mathrm{tr}\left(C_{[K_2,K_1]}\right) \end{bmatrix}$$

A.2.3. List of countries and products.

Country	Weight	Country	Weight
Austria	0.032	Spain	0.117
Belgium	0.035	Greece	0.028
Finland	0.017	Slovenia	0.004
France	0.204	Cyprus	0.002
Germany	0.276	Malta	0.001
Ireland	0.014	Slovakia	0.007
Italy	0.182	Estonia	0.002
Luxembourg	0.003	Latvia	0.002
Netherlands	0.051	Lithuania	0.004
Portugal	0.022		

Note: The figures correspond to the average weight over the whole sample period 2002-2019.

Table A1. List of countries

Product	Weight
Bread and cereals	0.026
Meat	0.037
Fish and seafood	0.011
Milk, cheese and eggs	0.023
Oils and fats	0.005
Fruit	0.012
Vegetables	0.016
Sugar, jam, honey, chocolate and confectionery	0.010
Other food products	0.005
Coffee, tea and cocoa	0.004
Mineral waters, soft drinks, fruit and vegetable juices	0.009
Spirits	0.003
Wine	0.007
Beer	0.006
Tobacco	0.024
Clothing	0.051
Other articles of clothing and clothing accessories	0.002
Cleaning, repair and hire of clothing	0.002
Footwear	0.014
Actual rentals for housing	0.063
Maintenance and repair of the dwelling	0.014
Water supply and miscellaneous services relating to the dwelling	0.027
Electricity, gas and other fuels	0.056
Furniture and furnishings, carpets and other floor coverings	0.026
Household textiles	0.005
Household appliances	0.011
Glassware, tableware and household utensils	0.006
Tools and equipment for house and garden	0.005
Goods and services for routine household maintenance	0.019
Medical products, appliances and equipment	0.019
Out-patient services	0.019
Purchase of vehicles	0.044
Spare parts and accessories for personal transport equipment	0.006
Fuels and lubricants for personal transport equipment	0.044
Maintenance and repair of personal transport equipment	0.029
Other services in respect of personal transport equipment	0.011
Transport services	0.024
Communications	0.032
Audio-visual, photographic and information processing equipment	0.015
Games, toys and hobbies	0.005
Equipment for sport, camping and open-air recreation	0.003
Gardens, plants and flowers	0.006
Pets and related products; veterinary and other services for pets	0.006
Recreational and sporting services	0.010
Cultural services	0.014
Newspapers, books and stationery	0.018
Package holidays	0.016
Education	0.011
Catering services	0.079
Accommodation services	0.018
Hairdressing salons and personal grooming establishments	0.012
Electrical appliances, articles and products for personal care	0.017
Other personal effects	0.010
Insurance	0.020
Other financial services	0.010

Note: The figures correspond to the average weight over the whole sample period, 2002-2019, and are normalized to sum to one (as these indices cover approximately the whole HICP).

Table A2. List of products

Working Papers

2017

1|17 The diffusion of knowledge via managers' mobility
Giordano Mion | Luca David Opromolla | Alessandro Sforza

2|17 Upward nominal wage rigidity
Paulo Guimarães | Fernando Martins | Pedro Portugal

3|17 Zooming the ins and outs of the U.S. unemployment
Pedro Portugal | António Rua

4|17 Labor market imperfections and the firm's wage setting policy
Sónia Félix | Pedro Portugal

5|17 International banking and cross-border effects of regulation: lessons from Portugal
Diana Bonfim | Sónia Costa

6|17 Disentangling the channels from birthdate to educational attainment
Luís Martins | Manuel Coutinho Pereira

7|17 Who's who in global value chains? A weighted network approach
João Amador | Sónia Cabral | Rossana Mastrandrea | Franco Ruzzenenti

8|17 Lending relationships and the real economy: evidence in the context of the euro area sovereign debt crisis
Luciana Barbosa

9|17 Impact of uncertainty measures on the Portuguese economy
Cristina Manteu | Sara Serra

10|17 Modelling currency demand in a small open economy within a monetary union
António Rua

11|17 Boom, slump, sudden stops, recovery, and policy options. Portugal and the Euro
Olivier Blanchard | Pedro Portugal

12|17 Inefficiency distribution of the European Banking System
João Oliveira

13|17 Banks' liquidity management and systemic risk
Luca G. Deidda | Ettore Panetti

14|17 Entrepreneurial risk and diversification through trade
Federico Esposito

15|17 The portuguese post-2008 period: a narrative from an estimated DSGE model
Paulo Júlio | José R. Maria

16|17 A theory of government bailouts in a heterogeneous banking system
Filomena Garcia | Ettore Panetti

17|17 Goods and factor market integration: a quantitative assessment of the EU enlargement
FLorenzo Caliendo | Luca David Opromolla | Fernando Parro | Alessandro Sforza

2018

1|18 Calibration and the estimation of macroeconomic models
Nikolay Iskrev

2|18 Are asset price data informative about news shocks? A DSGE perspective
Nikolay Iskrev

3|18 Sub-optimality of the friedman rule with distorting taxes
Bernardino Adão | André C. Silva

4|18 The effect of firm cash holdings on monetary policy
Bernardino Adão | André C. Silva

5|18 The returns to schooling unveiled
Ana Rute Cardoso | Paulo Guimarães | Pedro Portugal | Hugo Reis

6|18 Real effects of financial distress: the role of heterogeneity
Francisco Buera | Sudipto Karmakar

7|18 Did recent reforms facilitate EU labour market adjustment? Firm level evidence
Mario Izquierdo | Theodora Kosma | Ana Lamo | Fernando Martins | Simon Savsek

8|18 Flexible wage components as a source of wage adaptability to shocks: evidence from European firms, 2010–2013
Jan Babecký | Clémence Berson | Ludmila Fadejeva | Ana Lamo | Petra Marotzke | Fernando Martins | Pawel Strzelecki

9|18 The effects of official and unofficial information on tax compliance
Filomena Garcia | Luca David Opromolla Andrea Vezulli | Rafael Marques

10|18 International trade in services: evidence for portuguese firms
João Amador | Sónia Cabral | Birgitte Ringstad

11|18 Fear the walking dead: zombie firms, spillovers and exit barriers
Ana Fontoura Gouveia | Christian Osterhold

12|18 Collateral Damage? Labour Market Effects of Competing with China – at Home and Abroad
Sónia Cabral | Pedro S. Martins | João Pereira dos Santos | Mariana Tavares

13|18 An integrated financial amplifier: The role of defaulted loans and occasionally binding constraints in output fluctuations
Paulo Júlio | José R. Maria

14|18 Structural Changes in the Duration of Bull Markets and Business Cycle Dynamics
João Cruz | João Nicolau | Paulo M.M. Rodrigues

15|18 Cross-border spillovers of monetary policy: what changes during a financial crisis?
Luciana Barbosa | Diana Bonfim | Sónia Costa | Mary Everett

16|18 When losses turn into loans: the cost of undercapitalized banks
Laura Blattner | Luísa Farinha | Francisca Rebelo

17|18 Testing the fractionally integrated hypothesis using M estimation: With an application to stock market volatility
Matei Demetrescu | Paulo M. M. Rodrigues | Antonio Rubia

18|18 Every cloud has a silver lining: Micro-level evidence on the cleansing effects of the Portuguese financial crisis
Daniel A. Dias | Carlos Robalo Marques

19|18 To ask or not to ask? Collateral versus screening in lending relationships
Hans Degryse | Artashes Karapetyan | Sudipto Karmakar

20|18 Thirty years of economic growth in Africa
João Amador | António R. dos Santos

21|18 CEO performance in severe crises: the role of newcomers
Sharmin Sazedj | João Amador | José Tavares

22|18 A general equilibrium theory of occupational choice under optimistic beliefs about entrepreneurial ability
Michele Dell'Era | Luca David Opromolla | Luís Santos-Pinto

23|18 Exploring the implications of different loan-to-value macroprudential policy designs
Rita Basto | Sandra Gomes | Diana Lima

24|18 Bank shocks and firm performance: new evidence from the sovereign debt crisis
Luísa Farinha | Marina-Eliza Spaliara | Serafem Tsoukas

25|18 Bank credit allocation and productivity: stylised facts for Portugal
Nuno Azevedo | Márcio Mateus | Álvaro Pina

26|18 Does domestic demand matter for firms' exports?
Paulo Soares Esteves | Miguel Portela | António Rua

27|18 Credit Subsidies
Isabel Correia | Fiorella De Fiore | Pedro Teles | Oreste Tristani

2019

18|19 Trends and cycles under changing economic conditions
Cláudia Duarte | José R. Maria | Sharmin Sazedj

19|19 Bank funding and the survival of start-ups
Luísa Farinha | Sónia Félix | João A. C. Santos

20|19 From micro to macro: a note on the analysis of aggregate productivity dynamics using firm-level data
Daniel A. Dias | Carlos Robalo Marques

21|19 Tighter credit and consumer bankruptcy insurance
António Antunes | Tiago Cavalcanti | Caterina Mendicino | Marcel Peruffo | Anne Villamil

2020

1|20 On-site inspecting zombie lending
Diana Bonfim | Geraldo Cerqueiro | Hans Degryse | Steven Ongena

2|20 Labor earnings dynamics in a developing economy with a large informal sector
Diego B. P. Gomes | Felipe S. Iachan | Cezar Santos

3|20 Endogenous growth and monetary policy: how do interest-rate feedback rules shape nominal and real transitional dynamics?
Pedro Mazeda Gil | Gustavo Iglésias

4|20 Types of International Traders and the Network of Capital Participations
João Amador | Sónia Cabral | Birgitte Ringstad

5|20 Forecasting tourism with targeted predictors in a data-rich environment
Nuno Lourenço | Carlos Melo Gouveia António Rua

6|20 The expected time to cross a threshold and its determinants: A simple and flexible framework
Gabriel Zsurkis | João Nicolau | Paulo M. M. Rodrigues

7|20 A non-hierarchical dynamic factor model for three-way data
Francisco Dias | Maximiano Pinheiro; António Rua