DOCUMENT

DE TRAVAIL

N° 490

SPECIFICATION ANALYSIS OF INTERNATIONAL

TREASURY YIELD CURVE FACTORS

Fulvio Pegoraro, Andrew F. Siegel and Luca Tiozzo 'Pezzoli

June 2014

DIRECTION GÉNÉRALE DES ÉTUDES ET DES RELATIONS INTERNATIONALES

DIRECTION GÉNÉRALE DES ÉTUDES ET DES RELATIONS INTERNATIONALES

SPECIFICATION ANALYSIS OF INTERNATIONAL

TREASURY YIELD CURVE FACTORS

Fulvio Pegoraro, Andrew F. Siegel and Luca Tiozzo ’Pezzoli

June 2014

Specification Analysis of International Treasury Yield Curve Factors

Fulvio PEGORARO* Andrew F. SIEGEL† Luca TIOZZO "PEZZOLI"‡

*Banque de France, Financial Economics Research Service [DGEI-DEMFI-RECFIN; E-mail: Fulvio.PEGORARO@banque-france.fr] and CREST, Laboratoire de Finance-Assurance [E-mail: pegoraro@ensae.fr].

†Grant I. Butterbaugh Professor, Foster School of Business, University of Washington, Department of Finance and Business Economics, Department of Information Systems and Operations Management, Department of Statistics [asiegel@u.washington.edu], Box 353200, Seattle, WA 98195.

‡Université Paris I Pantheon - Sorbonne [Luca.Tiozzo-Pezzoli@univ-paris1.fr] and Paris Dauphine University.

Abstract: Étant donnée une approche modèles espace-état linéaires Gaussiens pour décrire la dynamique jointe des courbes de taux de plusieurs pays et, en utilisant une méthodologie d'estimation par maximum de vraisemblance, nous montrons comment extraire simultanément les facteurs communs (qui affectent tous les pays) et locaux (spécifiques à un pays seulement) qui caractérisent notre modèle. Cette extraction jointe demande le développement d'une nouvelle procédure de normalisation qui va au delà de celles classiques dans la littérature des modèles à facteurs. De plus, cela nous permet d'éviter les effets d'une estimation séquentielle des facteurs qui peut expliquer le manque de consensus en littérature pas seulement sur le nombre totale des facteurs nécessaires pour expliquer la dynamique jointe des courbes de taux, mais aussi le nombre des facteurs communs et locaux. Grace à une base de données journalière des courbes de taux de bonds du Trésor pour les États-Unis, l'Allemagne, l'Angleterre et le Japon, observés de Janvier 1986 à Décembre 2009, nous trouvons en général (à la fois pour des taux en niveau et en différence) qu'un modèle avec deux facteurs communs et trois facteurs locaux corrélés est préféré à un modèle (de complexité similaire) qui inclus un seul facteur commun ou à un modèle avec seulement des facteurs locaux corrélés. De plus, chaque facteur commun imite fortement (ou bien est similaire à) un facteur local obtenu à partir d'un modèle avec seulement des facteurs locaux. Nous constatons aussi que la dépendance entre courbes des taux internationales est due, principalement, par les corrélations instantanées entre facteurs locaux de différents pays et, à un moindre degré, par la matrice autorégressive (pleine) des facteurs latents et la matrice des loadings communs.

Mots-clés: courbes de taux internationales, facteurs communs et locaux, modèles espace-état, algorithme EM, algorithme de filtrage et lissage de Kalman.
Codes JEL: G12, E43, C52.

Abstract: We show how to compute patterns of variation over time, both among and within countries, that determine the international term structure of interest rates, using maximum likelihood within a linear Gaussian state-space framework. The simultaneous estimation of common factors (shared by all countries) and local factors (specific to one country) requires development of a normalization procedure beyond that of ordinary factor analysis. By jointly estimating common and local factors we avoid sequential estimation effects that may explain the lack of agreement in the multi-country term structure literature regarding not only the total number of latent factors required to explain the joint dynamics of yield curves, but also the number of common and of local factors. Using data on international yield curves of U.S., Germany, U.K. and Japan from January 1986 to December 2009, we generally find (analyzing yields in level and in difference) that a model with two common factors and three correlated local factors is preferred to a model (of similar complexity) that includes one common factor only or a model with only correlated local factors. In addition, each common factor closely mimics (or is similar to) a local factor extracted from a pure local factor model. We also reach the conclusion that dependence across international yield curves are driven, first, by the instantaneous correlation between local factors of different countries and, then, by the (full) autoregressive matrix of latent factors and by the matrix of common loadings.

Keywords: international treasury yield curves, common and local factors, state-space models, EM algorithm, Kalman Filter and Kalman Smoother.

JEL classification: G12, E43, C52.

Non-technical summary

The yield curve literature has focused not only on the specification and estimation of models explaining the term structure of interest rates in a single economy but, more recently, has been extended to the relevant problem of specifying and estimating the joint dynamics of international yield curves. In the single-country case, the estimation and implementation of dynamic yield curve models has found that three (the level, slope and curvature factors to five latent factors are required to match the dynamics and the shapes of the term structure. In the multi-country setting, in contrast, we observe substantial lack of agreement, not only about the number of latent factors that are required to explain the joint dynamics of two or more countries' yield curves, but also about the common/local nature of the factors where each common factor affects yields in all countries while each local factor affects yields in only one country.

The purpose of this paper is to respond to this lack of agreement by directly addressing the theoretical and empirical issues that naturally characterize the selection and estimation of interest rate factors that evolve over time in a multi-country setting. We propose using maximum likelihood (ML) criteria within a linear Gaussian state-space approach, to jointly and reliably estimate the preferred combination of common and local factors required to jointly explain multi-country yield curves.

Our choice of a state-space framework is a response, in part, to two main critiques of principal component-based (PC-based) approaches that have appeared in the literature. First, the purpose of principal component analysis is to extract factors that maximize the explained variance, and do not seek to distinguish between the role of common and that of local factor in the presence of multiple groups, resulting in estimated factors that jointly capture both local and common influences without distinguishing one from the other. Second, while the factor model literature has proposed several methods for selecting the number of factors, the reliability of these criteria requires the presence of weak-form serial and cross-sectional dependence in the idiosyncratic component of the factor model, as well as a large N (the cross-sectional dimension) and large T (the time-series dimension) database. These conditions are clearly not all satisfied by an international yield curve panel of data, given the strong persistence and cross-correlation of interest rates, as well as the typically small dimension of the maturity spectrum.

To be sure that the divergence in results cited in the literature was not merely due to the variety of data sets analyzed with different numbers of countries over differing sample periods, an extensive empirical analysis applies the same methods (as used in the literature) to a common set of data (same countries, same time period) and still finds lack of agreement among the PC-based approaches with the explained variance criterion. This controlled experiment (varying only the statistical estimation methods) involved developing an international Treasury yield curve database by applying common filtering and interpolation techniques across all countries. Not only are different combinations of common and local factors provided by the different methodologies, but even by the same methodology when it is applied to yields in level and to yields in difference [see Pegoraro, Siegel and Tiozzo 'Pezzoli' (2014) for further details]. These empirical findings reinforce our choice to develop here an alternative state-space based statistical technique that can jointly estimate common and local factors within a model that reflects their distinct natures.

Our empirical analysis suggests in general, across a variety of groups of countries and for both yield levels and yield differences, that a model with two common factors and three correlated local factors is preferred to a model (of similar complexity) that includes one common factor only or a model with only correlated local factors. Careful inspection of the optimally extracted factors reveals that each estimated common factor closely mimics (or is similar to) a local factor obtained from a pure local factor model. We reach the conclusion that international Treasury yield curves dependence are driven by a preferred set of two common factors and by the strong correlation between local factors of different countries, and that the former are spanned by (pure) local factors. This conclusion exhibits one of the advantages of our proposed method as compared to PC-based approaches (for which the initial factor extraction cannot consider the distinction between a local and a common factor).

1 Introduction

The yield curve literature, following the seminal papers of Vasicek (1977) and Cox, Ingersoll, and Ross (1985), has focused not only on the specification and estimation of models explaining the term structure of interest rates in a single economy but, more recently, has been extended to the relevant problem of specifying and estimating the joint dynamics of international yield curves[1].

In the single-country case, the estimation and implementation of dynamic yield curve models has found that three [e.g., the level, slope and curvature factors of Litterman and Scheinkman (1991)] to five latent factors are required to match the dynamics and the shapes of the term structure [see Dai and Singleton (2000), Dai and Singleton (2002), Dai and Singleton (2003), Duffee (2002), Cheridito, Filipovic, and Kimmel (2007), Duarte (2004), Duffee (2011) and Adrian, Crump, and Moench (2013)]. This wide degree of robustness has made this result a fundamental building block characterizing the modeling of single-country yield curves.

In the multi-country setting, in contrast, we observe substantial lack of agreement, not only about the number of latent factors that are required to explain the joint dynamics of two or more countries' yield curves, but also about the common/local nature of the factors where each common factor affects yields in all countries while each local factor affects yields in only one country. Some researchers make *a priori* assumptions about the combination of common and local factors [e.g., Backus, Foresi, and Telmer (2001), Anderson, Hammond, and Ramezani (2010), Ahn (2004)], while others reach different conclusions about the number of common and local factors based on the explained variance criterion within a principal components (PC) approach [see Leippold and Wu (2007), Diebold, Li, and Yue (2008) and Egorov, Li, and Ng (2011)].

The purpose of this paper is to respond to this lack of agreement by directly addressing the theoretical and empirical issues that naturally characterize the selection and estimation of interest

We would like to thank Torben Andersen, Gregory Connor, Pasquale Della Corte, Walter Distaso, René Garcia, Christian Julliard, Robert A. Korajczyk, Loriano Mancini, Nour Meddahi, Alain Monfort, Christophe Pérignon, Eric Renault, Jean-Paul Renne, Barbara Rossi, Olivier Scaillet, James Sefton, Fabio Trojani Raman Uppal, Grigory Vilkov and Michel van der Wel, as well as participants at Bank of Spain and Bank of Canada Workshop on "Advances in Fixed Income Modeling" 2011, NASM 2012 Conference in Evanston, CFE 2012 Conference in Oviedo, Banque de France 2013 Seminar, TSE 2013 Financial Econometrics Conference, Imperial College 2013 Finance Group Seminar, Joint Statistical Meetings (JSM) 2013 Conference in Montreal, the Large-scale factor models in Finance Conference in Lugano 2013, the 2014 Financial Risk International Forum in Paris and the International Association for Applied Econometrics 2014 Annual Conference for helpful comments, discussions and remarks. Andrew F. Siegel acknowledge support by the Fondation Banque de France. We thank Béatrice Saes-Escorbiac, Aurélie Touchais and Guilleume Retout for excellent research assistance. Any remaining errors are ours. The views expressed in this paper are ours and do not necessarily reflect the views of the Banque de France.

[1]See, among others, Duffie and Kan (1996), Dai and Singleton (2000), Dai and Singleton (2002) Dai and Singleton (2003), Bansal and Zhou (2002), Duffee (2002), Monfort and Pegoraro (2007), in the single-country literature, and Frachot (1995), Backus, Foresi, and Telmer (2001), Anderson, Hammond, and Ramezani (2010), Ahn (2004), Leippold and Wu (2007), Tang and Xia (2007), Diebold, Li, and Yue (2008), Egorov, Li, and Ng (2011), Gourieroux, Monfort, and Sufana (2010), Jotikasthira, Le, and Lundbland (2010), in the multi-country literature.

rate factors that evolve over time in a multi-country setting. We propose using maximum likelihood (ML) criteria within a linear Gaussian state-space approach, to jointly and reliably estimate the preferred combination of common and local factors required to jointly explain multi-country yield curves.

Our choice of a state-space framework is a response, in part, to two main critiques of PC-based approaches that have appeared in the literature. First, Perignon, Smith, and Villa (2007) have highlighted that the purpose of principal component analysis is to extract factors that maximize the explained variance, and do not seek to distinguish between the role of common and that of local factor in the presence of multiple groups, resulting in estimated factors that jointly capture both local and common influences without distinguishing one from the other. Second, while the factor model literature has proposed several methods for selecting the number of factors[2], the reliability of these criteria requires the presence of weak-form serial and cross-sectional dependence in the idiosyncratic component of the factor model, as well as a large N (the cross-sectional dimension) and large T (the time-series dimension) database. These conditions are clearly not all satisfied by an international yield curve panel of data, given the strong persistence and cross-correlation of interest rates, as well as the typically small dimension of the maturity spectrum; for instance, in the presence of serial dependence, the Bai and Ng (2002) criteria tend to overestimate the number of common factors, even when a first-difference filter is applied to stationary data in order to mitigate the persistence [see Greenaway-McGrevy, Han, and Sul (2012), Han and Sul (2011) for details].

To be sure that the divergence in results cited in the literature was not merely due to the variety of data sets analyzed with different numbers of countries over differing sample periods, an extensive empirical analysis [Pegoraro, Siegel, and Tiozzo Pezzoli (2012)] applies the same methods (as used in the literature) to a common set of data (same countries, same time period) and still finds lack of agreement among the PC-based approaches with the explained variance criterion. This controlled experiment (varying only the statistical estimation methods) involved developing an international Treasury yield curve database by applying common filtering and interpolation techniques across all countries. Not only are different combinations of common and local factors provided by the different methodologies, but even by the same methodology when it is applied to yields in level and to yields in difference. These empirical findings reinforce our choice to develop here an alternative state-space based statistical technique that can jointly estimate common and local factors within a model that reflects their distinct natures.

Our linear Gaussian state-space approach explains the joint dynamics of multi-country term structures using autoregressive stationary latent factors of two types (common and local) as spec-

[2] See, for instance, Connor and Korajczyk (1993), Forni, Hallin, Lippi, and Reichlin (2000), Bai and Ng (2002), Bai and Ng (2007), Stock and Watson (1991), Amengual and Watson (2007) and Onatski (2010).

ified by the measurement equation. Each common factor has loadings that are unrestricted for all countries, while each local factor is identified with just one country by restricting its loadings to zero for all other countries. We find an innovative solution to the identification problem that allows for causality across common and local factors as well as between common and local ones. We implement maximum likelihood estimation for the state-space model using the EM algorithm with the Kalman Filter and Kalman Smoother recursions [see Engle and Watson (1981), Quah and Sargent (1993), Monfort, Renne, Rüffer, and Vitale (2003), Doz, Giannone, and Reichlin (2011), Doz, Giannone, and Reichlin (2012), Jungbacker and Koopman (2008), Bork, Dewachter, and Houssa (2009)]. From among different scenarios, each specifying the numbers and combinations of common and local factors, we select the optimal combination on the basis of maximum likelihood-based model selection criteria such as the Akaike Information Criterion (AIC). In order to take into account the persistence and heteroskedasticity of interest rates we also calculate the (Nonparametric Monte Carlo) bootstrap variant of AIC ($AICb$, say) of Cavanaugh and Shumway (1997) and based on a block stationary bootstrap [see Politis and Romano (1994), Politis and White (2004), and Patton, Politis, and White (2009)]. We use the international Treasury yield curves database of Pegoraro, Siegel, and Tiozzo Pezzoli (2012) consisting of rates in four leading bond markets (U.S., Germany, U.K. and Japan) observed weekly from January 1, 1986 to December 31, 2009.

Our empirical analysis suggests in general, across a variety of groups of countries and for both yield levels and yield differences, that a model with two common factors and three correlated local factors is preferred to a model (of similar complexity) that includes one common factor only or a model with only correlated local factors. Careful inspection of the optimally extracted factors reveals that each estimated common factor closely mimics (or is similar to) a local factor obtained from a pure local factor model. We reach the conclusion that international Treasury yield curves dependence are driven by a preferred set of two common factors and by the strong correlation between local factors of different countries, and that the former are spanned by (pure) local factors. This conclusion exhibits one of the advantages of our proposed method as compared to PC-based approaches (for which the initial factor extraction cannot consider the distinction between a local and a common factor).

The paper is organized as follows. In Section 2 we introduce our Multi-Country Term Structure Model ($MCTSM$) that describes the joint dynamics of international yield curves (Section 2.1) and we specify the appropriate identification restrictions that respect the presence of common and local factors (Section 2.2). In Section 3 we describe our proposed EM-based recursive maximum likelihood estimation procedure that imposes the identification restrictions. Section 4 presents the empirical analysis, beginning with the database on international Treasury yield curves in Section 4.1, with results from various models (different groups of countries, yield levels, yield differences, various combinations of common and local factors) following in Section 4.2 which also includes model selection results based on the maximized likelihood. Section 4.3 presents $MCTSMs$ parameter estimates and factors interpretations, while the focus of Section 4.4 is on understanding the statistical dependence across international local factors. Section 5 concludes, while proofs, tables, and graphs are gathered in the Appendix.

2 The Multi-Country Term Structure Model

2.1 Modeling Framework

In this section we define our Multi-Country Term Structure Model ($MCTSM$) as a linear Gaussian state-space model with block structure adopted to describe the joint dynamics of international yield curves. The model is specified by the following assumptions:

Assumption 1 (Yields, Countries, Common and Local Factors). *We denote by* $Y_t^{(j)}$ *the* $\tau \times 1$ *vector of yields observed at time* t *for country* j, *with* $j \in \{1, \ldots, n\}$, n *being the total number of analyzed countries.* $Y_t = (Y_t^{(1)'}, \ldots, Y_t^{(n)'})'$ *denotes the* $N \times 1$ *vector of the observed international yields with* $N = \tau\, n$. *We denote by* F_t *the* $k \times 1$ *vector of latent factors at time* t *that explain the international term structures of interest rates. We assume that* $F_t = (F_t^{(c)'}, F_t^{(l)'})'$, *where* $F_t^{(c)} = (F_{1,t}^{(c)}, \ldots, F_{r_c,t}^{(c)})'$ *is the* $r_c \times 1$ *vector of factors common to all countries, and the local factors are* $F_t^{(l)} = (F_{1,t}^{(l)'}, \ldots, F_{n,t}^{(l)'})'$ *with* $F_{j,t}^{(l)} = (F_{1,j,t}^{(l)}, \ldots, F_{r_j,j,t}^{(l)})'$ *the* $r_j \times 1$ *vector of factors associated to country* j *only, for* $j \in \{1, \ldots, n\}$. *The total number of local factors across all countries is denoted* $r^{(l)}$, *so* $F_t^{(l)}$ *is a* $r^{(l)} \times 1$ *vector and* $k = r_c + r^{(l)}$.

Assumption 2 (The Multi-Country Term Structure Model $MCTSM$). *For a given* k*-dimensional latent factor* F_t *made up of* r_c *common and* $r^{(l)}$ *local factors, the joint dynamics of the* n *international yield curves* $Y_t = (Y_t^{(1)'}, \ldots, Y_t^{(n)'})'$ *is given by:*

$$\begin{cases} Y_t = \mu + \Lambda_{\mathcal{B}}\, F_t + \varepsilon_t \,, \;\; \varepsilon_t \sim IIN(0, \Omega_{\mathcal{B}}) \\ \\ F_t = \Phi\, F_{t-1} + \eta_t \,, \;\; \eta_t \sim IIN(0, \Psi_\eta) \,, \end{cases} \tag{1}$$

where μ *is an* $N \times 1$ *vector of constants and*

$$\Lambda_{\mathcal{B}} = \begin{bmatrix} \Lambda_c & \Lambda_l \end{bmatrix} \tag{2}$$

is the $N \times k$ *matrix of factor loadings partitioned in terms of the* $N \times r_c$ *matrix* $\Lambda_c = [\Lambda_{c,1}, \ldots, \Lambda_{c,r_c}]$

of common loadings and in the $N \times r^{(l)}$ *block-diagonal matrix of local loadings*

$$\Lambda_l = \begin{bmatrix} \Lambda_l^{(1)} & 0 & \dots & 0 \\ 0 & \Lambda_l^{(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Lambda_l^{(n)} \end{bmatrix}, \tag{3}$$

while $\Omega_{\mathcal{B}}$ *is the* $N \times N$ *variance-covariance matrix of the Gaussian-distributed white noise* ε_t. Φ *is the* $k \times k$ *autoregressive matrix,* Ψ_η *is the* $k \times k$ *variance-covariance matrix of the* k*-dimensional Gaussian distributed white noise* $\eta_t = \left(\eta_t^{(c)'}, \eta_t^{(1)'}, \dots, \eta_t^{(n)'}\right)'$ *and* $E(\varepsilon_t\, \eta_t') = 0$ *for all* t.

2.2 Identification Restrictions Imposed by Common and Local Factors

We now focus on the identification restrictions for this *MCTSM* model that stem from the fact that the model's fitted values remain unchanged if we transform the model specification using any $k \times k$ non-singular matrix A because $\Lambda_{\mathcal{B}}\, F_t = (\Lambda_{\mathcal{B}}\, A)\,(A^{-1}\, F_t)$. Ordinarily, for a model that does not distinguish common from local factors (i.e., $r^{(l)} = 0$), we would impose k^2 restrictions equal to the number of free parameters in A. However, in the presence of local factors with $r^{(l)} > 0$, the matrix A has to be such that the transformed loadings matrix $\Lambda_{\mathcal{B}}^* = \Lambda_{\mathcal{B}}\, A$ maintains same required block structure (i.e., the same pattern of zeros) as we required of $\Lambda_{\mathcal{B}}$.

Proposition 1 (Identification Restrictions). *The identification restrictions for MCTSM (1) with loadings matrix* $\Lambda_{\mathcal{B}} = [\Lambda_c\ \Lambda_l]$ *from Assumption 2, requires* $r^* := (r_c\, k) + \sum_{j=1}^n r_j^2$ *restrictions that we may solve by imposing:*

R.i) $E\left(\eta_t^{(c)}\, \eta_t^{(c)'}\right) = I_{r_c}$, $E\left(\eta_t^{(j)}\, \eta_t^{(j)'}\right) = I_{r_j}$ *and* $E\left(\eta_t^{(c)}\, \eta_t^{(j)'}\right) = 0$ *for all* $j \in \{1, \dots, n\}$, *that is we impose* $\Psi_\eta = \Psi_{\mathcal{B}}$ *where:*

$$\Psi_{\mathcal{B}} := \begin{bmatrix} I_{r_c} & 0 & 0 & \dots & 0 \\ 0 & I_{r_1} & \Psi_{12} & \dots & \Psi_{1n} \\ 0 & \Psi_{21} & I_{r_2} & \dots & \Psi_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \Psi_{1n} & \Psi_{2n} & \dots & I_{r_n} \end{bmatrix}, \tag{4}$$

and where, for any $i, j \in \{1, \dots, n\}$ *with* $i \neq j$, *the covariance matrix* $E\left(\eta_t^{(i)}\, \eta_t^{(j)'}\right) = \Psi_{ij}$ *is allowed to be different from zero;*

R.ii) $(\Lambda_c'\Lambda_c)$ *and* $\Lambda_l^{(j)\prime}\Lambda_l^{(j)}$ *for all* $j \in \{1, \ldots, n\}$, *are all diagonal, that is* $\Lambda_{\mathcal{B}}$ *has to be such that* $\Lambda_{\mathcal{B}}'\Lambda_{\mathcal{B}} = \Pi_{\mathcal{B}}$, *with:*

$$\Pi_{\mathcal{B}} := \begin{bmatrix} \Pi_{cc}^d & \Pi_{c1} & \Pi_{c2} & \ldots & \Pi_{cn} \\ \Pi_{1c} & \Pi_{11}^d & 0 & \ldots & 0 \\ \Pi_{2c} & 0 & \Pi_{22}^d & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Pi_{nc} & 0 & 0 & \ldots & \Pi_{nn}^d \end{bmatrix}, \tag{5}$$

and where Π_{cc}^d *and* Π_{jj}^d, *for* $j \in \{1, \ldots, n\}$, *are diagonal matrices whose diagonal entries are arranged in descending order.*

(Proof: see Appendix A).

Restrictions *R.i)* and *R.ii)* allow us to estimate model parameters and to extract common and local latent factors while keeping the autoregressive matrix Φ unconstrained *and* allowing local factors to be correlated ($\Psi_{ij} \neq 0$) for distinct countries i and j, thereby allowing causality estimation. By keeping the autoregressive matrix Φ unconstrained, we avoid arbitrarily imposing a lack of causality between common and local factors, as would occur in the classical case with Φ lower triangular. Thus, we are able to estimate the impact of any factor on any other factor, regardless of whether they are both common, one common and one local, or both local either from the same country or from different countries. By allowing correlation of local factors across countries, we can estimate instantaneous causality associated with these local factors, allowing $MCTSM$ to model the presence of factors common to only a subset of the analyzed countries (*regional factors*)[3]. Regarding restrictions *R.ii)*, it is important to highlight that we rank the Jordan decomposition-based eigenvalues in the main diagonal of Π_{cc}^d and Π_{jj}^d, for $j \in \{1, \ldots, n\}$, in decreasing order. Our procedure thus ranks the factors, within each group, in line with the classical level, slope and curvature order provided by the principal component approach.

In summary, the identification restrictions characterizing our $MCTSM$ model open the way

[3]An alternative set of identification restrictions concerning Ψ_η might be the one imposing $\Psi_{ij} \neq 0$ for all $i \neq j$, $i, j \in \{1, \ldots, n\}$, but leaving $E\left(\eta_t^{(c)}\,\eta_t^{(j)\prime}\right) \neq 0$ and still assuming $E\left(\eta_t^{(c)}\,\eta_t^{(c)\prime}\right) = I_{r_c}$ and $E\left(\eta_t^{(j)}\,\eta_t^{(j)\prime}\right) = I_{r_j}$. Our identification method is more general than this possible alternative. First, any initially estimated Ψ_η can be transformed to $\Psi_{\mathcal{B}}$ while respecting the common-local loading structure (2)-(3), while the alternative one would potentially require nonzero loadings for each country on all of the other country local factors, which would violate the principle that a given country can load only on the common factors and its own local factors. Second, this alternative set of restrictions can not represent regional factors. Third, while restrictions *R.i)* guarantee that Ψ_η be a proper variance-covariance (thus, positive definite) matrix regardless the assumption about r_c and r_j, if in the alternative set of restrictions we assume, for instance, $r_c = r_j$ and $E\left(\eta_t^{(c)}\,\eta_t^{(j)\prime}\right) = I_{r_c}$, then the matrix cannot be a covariance matrix because the first common factor is perfectly correlated with each of the first local ones, which must in turn be perfectly correlated with one another. However, these local factors all have zero correlation with one another thus contradicting the previous statement (for instance, in the case $n = 2$ and $r_c = r_j = 1$, it is easy to verify that $|\Psi_\eta| = -1$, which is clearly not possible for a variance-covariance matrix).

to three sources of dependence across international yield curves. First, the matrix Λ_c of common loadings allowing common factors $F_t^{(c)}$ to directly impact all term structures at the same time. Second, the unconstrained autoregressive matrix Φ allowing for causalities between all factors and, third, the instantaneous correlations between local factors of different countries ($\Psi_{ij} \neq 0$). The empirical analysis of Section 4 will thus focus on understanding not only the preferred combination of common and local factors but, also, on identifying which of the three above mentioned channels is the most important in driving international yield curves dependence.

3 The $MCTSM$ Recursive Maximum Likelihood Estimation Procedure

In this section we provide details of the statistical methodology to efficiently extract the optimal combination of common and local factors to represent the joint dynamics of international yield curves. This task entails specifying the estimation details to be used for each of several given combinations of common and local factors, from which the likelihood-based Akaike Information Criterion (AIC) and its bootstrap variant $AICb$ (say) will be used to select the optimal combination. In Section 3.1, to efficiently estimate the $MCTSM$ model given the number of common and local factors, we use the EM Algorithm (viewing the factor values as missing data) together with the recursive Kalman Filter and Kalman Smoother [see, e.g., Engle and Watson (1981), Quah and Sargent (1993) and Doz, Giannone, and Reichlin (2012)] to numerically seek the model's maximum likelihood. At each iteration of the EM algorithm the likelihood increases and (under regularity conditions) it converges to a maximum of the likelihood function. In Section 3.2 we present the four strategies we adopt to initialize the algorithm and the random perturbation technique of the associated estimations we use to avoid being trapped in a local maximum. The maximum likelihood estimator of any of $MCTSM$ model of interest will be the one coming from the strategy (among the four) providing the largest value of the log-likelihood function. In practice, we find that from each starting point convergence is typically obtained within 100 iterations (in a benchmark case of two common factors and two local factors for each of the four countries).

3.1 The Recursive MLE Procedure

Here is the EM-based recursive procedure to obtain maximum likelihood estimates for the set of parameters $\theta := (\mu, \Lambda_{\mathcal{B}}, \Omega_{\mathcal{B}}, \Phi, \Psi_\eta)$ of the $MCTSM$ (1) while imposing the identification restrictions *R.i*) and *R.ii*).

Proposition 2 (The $MCTSM$ Recursive Maximum Likelihood Estimation Procedure). *Each iteration of the procedure to calculate the maximum likelihood estimator denoted*

θ_T^{MLE}*, of the parameter set* θ *characterizing* $MCTSM$*, is based on the following three steps, where step (a) defines our notation for the results of the Kalman Filter and Kalman Smoother, step (b) maximizes the expected complete data log-likelihood function, conditionally to* Y^T *and given the imputed factor results from step (a), and step (c) shows how the identification restrictions are satisfied:*

(a) *For a given set of* $MCTSM$ *input parameters denoted* $\theta_{EM}^{(i)}$*, and for a given data set* $Y^T := (Y_1, \ldots, Y_T)$ *of international yield curves, one iteration of the Kalman Filter and of the Kalman Smoother provides the log-likelihood function value denoted* $\mathcal{L}\left(\widehat{\theta}_{EM}^{(i)}\right)$ *and the imputed factor results denoted* $\mathcal{F}_{t|T}^{(i)} := \left(F_{t|T}^{(i)}, P_{t|T}^{(i)}, P_{t-1,t|T}^{(i)}\right)$ *(Expectation step* i*), where:*

- $F_{t|T}^{(i)} := \mathbb{E}_{\theta_{EM}^{(i)}}\left[F_t \,|\, Y^T\right]$ *is the date-*t *vector of smoothed factors,*
- $P_{t|T}^{(i)} := \mathbb{V}_{\theta_{EM}^{(i)}}\left[F_t \,|\, Y^T\right] = \mathbb{E}_{\theta_{EM}^{(i)}}\left[\left(F_t - F_{t|T}^{(i)}\right)\left(F_t - F_{t|T}^{(i)}\right)' \,|\, Y^T\right]$ *is the date-*t *smoothed variance-covariance matrix of the factors, and*
- $P_{t-1,t|T}^{(i)} = \mathbb{C}_{\theta_{EM}^{(i)}}\left[F_{t-1}, F_t \,|\, Y^T\right] = \mathbb{E}_{\theta_{EM}^{(i)}}\left[\left(F_{t-1} - F_{t-1|T}^{(i)}\right)\left(F_t - F_{t|T}^{(i)}\right)' \,|\, Y^T\right]$ *is the date-*t *smoothed one lag autocovariance of the factors;*

(b) *Given* $\mathcal{F}_{t|T}^{(i)}$ *from the previous step, the Maximization step* $(i+1)$ *results in the following closed form estimators:*

$$\begin{aligned}
\Lambda_{\mathcal{B},T}^{(i+1)} &= \mathcal{D}_T^{(i)}\,\overline{\mathcal{C}}_T^{(i)-1} + \mathcal{K}_{\Lambda,T}^{(i)}\,, \quad \mu_T^{(i+1)} = \overline{Y}_T - \Lambda_{\mathcal{B},T}^{(i+1)}\,\overline{F}_T^{(i)} \\
\Omega_{\mathcal{B},T}^{(i+1)} &= \frac{1}{T}\left(\mathcal{E}_T^{(i)} - \mathcal{D}_T^{(i)}\,\overline{\mathcal{C}}_T^{(i)-1}\,\mathcal{D}_T^{(i)\prime} + \mathcal{K}_{\Lambda,T}^{(i)}\,\overline{\mathcal{C}}_T^{(i)}\,\mathcal{K}_{\Lambda,T}^{(i)\prime}\right) \\
\Phi_T^{(i+1)} &= \mathcal{B}_T^{(i)}\,\mathcal{A}_T^{(i)-1}\,, \quad \Psi_{\eta,T}^{(i+1)} = \frac{1}{T-1}\left(\mathcal{C}_T^{(i)} - \mathcal{B}_T^{(i)}\,\mathcal{A}_T^{(i)-1}\,\mathcal{B}_T^{(i)\prime}\right),
\end{aligned} \tag{6}$$

where:

$$\begin{array}{lll}
\mathcal{A}_T^{(i)} & := & \sum_{t=2}^T \left(F_{t-1|T}^{(i)} F_{t-1|T}^{(i)\prime} + P_{t-1|T}^{(i)} \right) \,, \quad \mathcal{B}_T^{(i)} := \sum_{t=2}^T \left(F_{t|T}^{(i)} F_{t-1|T}^{(i)\prime} + P_{t-1,t|T}^{(i)\prime} \right) \,, \\
\mathcal{C}_T^{(i)} & := & \sum_{t=2}^T \left(F_{t|T}^{(i)} F_{t|T}^{(i)\prime} + P_{t|T}^{(i)} \right) \,, \quad \overline{\mathcal{C}}_T^{(i)} := \sum_{t=1}^T \left[\left(F_{t|T}^{(i)} - \overline{F}_T^{(i)} \right) \left(F_{t|T}^{(i)} - \overline{F}_T^{(i)} \right)' + P_{t|T}^{(i)} \right] \,, \\
\mathcal{D}_T^{(i)} & := & \sum_{t=1}^T \left(Y_t - \overline{Y}_T \right) \left(F_{t|T}^{(i)} - \overline{F}_T^{(i)} \right)' \,, \quad \mathcal{E}_T^{(i)} := \sum_{t=1}^T \left(Y_t - \overline{Y}_T \right) \left(Y_t - \overline{Y}_T \right)' \,, \\
\overline{Y}_T & := & \frac{1}{T} \sum_{t=1}^T Y_t \,, \quad \overline{F}_T^{(i)} := \frac{1}{T} \sum_{t=1}^T F_{t|T}^{(i)} \,, \\
vec(\mathcal{K}_{\Lambda,T}^{(i)}) & := & (\overline{\mathcal{C}}_T^{(i)-1} \otimes \Omega_T^{(u,i)}) \, \mathcal{H}'_\Lambda \left[\mathcal{H}_\Lambda \, (\overline{\mathcal{C}}_T^{(i)-1}) \otimes \Omega_T^{(u,i)}) \, \mathcal{H}'_\Lambda \right]^{-1} \left[\kappa_\Lambda - \mathcal{H}_\Lambda \, vec(\mathcal{D}_T^{(i)} \, \overline{\mathcal{C}}_T^{(i)-1}) \right] \,, \\
\Omega_T^{(u,i)} & := & \dfrac{1}{T} \left(\mathcal{E}_T^{(i)} - \mathcal{D}_T^{(i)} \, \overline{\mathcal{C}}_T^{(i)-1} \, \mathcal{D}_T^{(i)\prime} \right) \,,
\end{array} \tag{7}$$

and where $\mathcal{H}_\Lambda$ is a $\vartheta \times Nk$ selection matrix such that:

$$\mathcal{H}_\Lambda \, vec\,(\Lambda) = \kappa_\Lambda \tag{8}$$

with Λ the unrestricted $N \times k$ matrix of loadings and with κ_Λ the ϑ-dimensional vector of zeros that enforces the block structure of $\Lambda_{\mathcal{B}}$ at each iteration of the algorithm.

(c) *Given the updated set of estimators $\theta_{EM}^{(i+1)} := \left(\mu_T^{(i+1)}, \Lambda_{\mathcal{B},T}^{(i+1)}, \Omega_{\mathcal{B},T}^{(i+1)}, \Phi_T^{(i+1)}, \Psi_{\eta,T}^{(i+1)} \right)$ from the previous step, the associated normalized estimator denoted $\theta_{EM}^{*(i+1)}$ satisfying the indentification restrictions R.i) and R.ii), is given by:*

$$\begin{array}{lll}
\Lambda_{\mathcal{B},T}^{*(i+1)} & := & \Lambda_{\mathcal{B},T}^{(i+1)} \, A^* \,, \quad \mu_T^{*(i+1)} = \mu_T^{(i+1)} \,, \quad \Omega_{\mathcal{B},T}^{*(i+1)} = \Omega_{\mathcal{B},T}^{(i+1)} \,, \\
\Phi_T^{*(i+1)} & := & (A^*)^{-1} \, \Phi_T^{(i+1)} \, A^* \,, \quad \Psi_{\eta,T}^{*(i+1)} := (A^*)^{-1} \, \Psi_{\eta,T}^{(i+1)} \, (A^*)^{-1\prime} \,,
\end{array} \tag{9}$$

where the (unique) normalization matrix A^* is:

$$A^* := \left(A_\perp \, A_{\eta,(i+1)} \, A_{c,l,(i+1)}^o \right) \,, \tag{10}$$

with

$$A_{\perp}^{-1} := \begin{bmatrix} I_{r_c} & 0 & \dots & 0 \\ -\left(\Psi_{10,T}^{c(i+1)}\right)\left[(\Psi_{00,T}^{c(i+1)})\right]^{-1} & I_{r_1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\left(\Psi_{n0,T}^{c(i+1)}\right)\left[(\Psi_{00,T}^{c(i+1)})\right]^{-1} & 0 & \dots & I_{r_n} \end{bmatrix},$$

$$\Psi_{\eta,T}^{(i+1)} := \begin{bmatrix} \Psi_{00,T}^{c(i+1)} & \Psi_{01,T}^{c(i+1)} & \dots & \Psi_{0n,T}^{c(i+1)} \\ \Psi_{10,T}^{c(i+1)} & \Psi_{11,T}^{(i+1)} & \dots & \Psi_{1n,T}^{(i+1)} \\ \vdots & \vdots & \ddots & \vdots \\ \Psi_{n0,T}^{c(i+1)} & \Psi_{n1,T}^{(i+1)} & \dots & \Psi_{nn,T}^{(i+1)} \end{bmatrix} \neq \Psi_{\eta,\mathcal{B}}\,,$$

$$A_{\eta,(i+1)}^{-1} \;:=\; diag\left[A_{\eta,c,(i+1)}^{-1}, A_{\eta,1,(i+1)}^{-1}, \dots, A_{\eta,n,(i+1)}^{-1}\right],$$

$$A_{\eta,c(i+1)}^{-1} \;:=\; \left(\mathcal{U}_{\eta,c(i+1)}\,\mathcal{D}_{\eta,c(i+1)}^{-1/2}\right)', \quad A_{\eta,j(i+1)}^{-1} := \left(\mathcal{U}_{\eta,j(i+1)}\,\mathcal{D}_{\eta,j(i+1)}^{-1/2}\right)' \;\forall\; j \in \{1,\dots,n\}\,,$$

where $\mathcal{U}_{\eta,c(i+1)}$ and $\mathcal{D}_{\eta,c(i+1)}$ are matrices of eigenvectors and eigenvalues of $\Psi_{00,T}^{c(i+1)}$, $\mathcal{U}_{\eta,j(i+1)}$ and $\mathcal{D}_{\eta,j(i+1)}$ are matrices of eigenvectors and eigenvalues of $\Psi_{jj,T}^{(i+1)}$, for $j \in \{1,\dots,n\}$, and where:

$$(A_{c,l,(i+1)}^{o})^{-1} \;:=\; diag\left[(\mathcal{U}_{c,(i+1)}^{o})^{-1}, (\mathcal{U}_{1,(i+1)}^{o})^{-1}, \dots, (\mathcal{U}_{n,(i+1)}^{o})^{-1}\right],$$

with $(\mathcal{U}_{c,(i+1)}^{o-1})$ and $(\mathcal{U}_{j,(i+1)}^{o-1})$ respectively denoting the eigenvector matrix of $\Lambda_{c,(i+1),T}^{o\prime}\Lambda_{c,(i+1),T}^{o}$ and $\Lambda_{l,(i+1),T}^{o(j)\prime}\Lambda_{l,(i+1),T}^{o(j)}$, for $j \in \{1,\dots,n\}$, where:

$$\Lambda_{\mathcal{B},(i+1),T}^{o} := \left[\Lambda_{c,(i+1),T}^{o}\;\;\Lambda_{l,(i+1),T}^{o}\right] = \Lambda_{\mathcal{B},T}^{(i+1)}\;\left(A_{\perp}\,A_{\eta,(i+1)}\right).$$

(Proof: see Appendix B).

3.2 Initialization Algorithm Descriptions

The recursive procedure presented in Proposition 2 has been initialized in four possible ways. A first method is to start with a Principal Components analysis and set $\theta_{EM}^{(0)} := \theta_T^{PF}$ where $\theta_T^{PF} := (\mu_T^{PF}, \Lambda_{\mathcal{B},T}^{PF}, \Omega_{\mathcal{B},T}^{PF}, \Phi_T^{PF}, \Psi_{\eta,T}^{PF})$ denotes the set of model parameters estimated by a three-step Principal Factor methodology (see Appendix C for details). A second method is to use randomized initial values for the factors, simulating (F_t) from a normalized k-dimensional Gaussian distribution and using these values along with the observed yields to estimate all parameters by

OLS regressions[4].

The last two methods, taking into account the possible presence of (one or two) common factors as suggested by the literature, aim at selecting initial parameter values by means of a strategy that favors the presence of such a common factors in the data. More precisely, the third one is based on the following steps: for any given set of common and local factors, and using the estimation methodology presented in Proposition 2, first we estimate the model with only common factors, on the residuals we estimate the model with only local factors and then we use the associated smoothed factors to select a new vector of parameter estimates, through OLS regressions, in order to provide the new starting condition to the MLE recursive procedure.

The fourth one is tailored for nested $MCTSMs$ with a fixed number of factors k, but different combination of common and locals (including $r_c = 0$), that we will analyze at the end of Section 4.2. The methodology is based on the following idea (for ease of presentation we consider here the case $n = 2$ countries): for any given number of factors k and given an estimated model with $r_c \geq 0$ commons and (r_1, r_2) locals, we move to the estimation of the nesting model with $r_c + 1$ commons and $(r_1 - 1, r_2)$ locals (say) by providing starting parameter values obtained from the linear combination of the r_1 local factors that best explains country-2 yields. More precisely, the methodology is based on the following steps. First, we regress the average yield (across maturities) of country-2, namely $\overline{Y}_t^{(2)} = \frac{1}{\tau}\sum_{i=1}^{\tau} Y_{j,t}^{(2)}$, on the r_1 local factors, the estimated linear combination of these factors is identified as the new common factor $F_{r_c+1,t}^{(c)} := \sum_{i=1}^{r_1} \beta_i \, F_{i,1,t}^{(l)}$ (say) and the associated noise is given by $\eta_{r_c+1,t}^{(c)} := \sum_{i=1}^{r_1} \beta_i \, \eta_{i,t}^{(1)}$. Second, we regress the first $(r_1 - 1)$ variables $\eta_{i,t}^{(1)}$ on $\eta_{r_c+1,t}^{(c)}$ and the noise of any of these regression is denoted $\xi_{i,t}^{(1)}$. Third, I orthonormalize the $\xi_{i,t}^{(1)}$s and then, finally, I define the associated (orthonormalized) country-1 factor as the new country-1 factors of the nesting model. With the newly specified common and country-1 factors, along with the starting r_2 local factors, we select a new vector of parameter estimates, through OLS regressions, that we adopt as starting condition to estimate the nesting model following Proposition 2.

In addition, in order to overcome the possible finding of a local (instead of the global) maximum of the log-likelihood function, we randomly perturb the estimations obtained using Proposition 2 through the following procedure. First, given the smoothed factors $\widehat{F}_{t|T}^{(i^*)} := \mathbb{E}_{\theta_{EM}^{*(i^*)}}\left[F_t \,|\, Y^T\right] = \mathbb{E}_{\widehat{\theta}_T^{MLE}}\left[F_t \,|\, Y^T\right]$ and a randomly generated number $\varepsilon^{(\sigma)}$ drawn from $N(0, \sigma^2)$, we obtain a new set of smoothed factors $\widehat{F}_{t|T}^{(\sigma)} := \widehat{F}_{t|T}^{(i^*)} + \varepsilon^{(\sigma)}$. For any given $\sigma \in \{0, 1, \dots, 5\}$, we use the associated

[4]Regarding the control for convergence of the algorithm, we adopt the following criterion:

$$\Upsilon^{(i)} = \frac{|\,\mathcal{L}(\widehat{\theta}_{EM}^{*(i+1)}) - \mathcal{L}(\widehat{\theta}_{EM}^{*(i)})\,|}{(|\,\mathcal{L}(\widehat{\theta}_{EM}^{*(i+1)})\,| + |\,\mathcal{L}(\widehat{\theta}_{EM}^{*(i)})\,|)/2}$$

and we stop the procedure after i^* iterations if $\Upsilon^{(i^*)} < 10^{-6}$.

factors $\widehat{F}_{t|T}^{(\sigma)}$ to obtain a new set parameters estimates, through OLS regressions, that are used as new starting conditions to run again the recursive MLE procedure of Proposition 2. Second, we select across the alternatives the vector of parameter estimates leading to the largest value of the log-Likelihood function, we retrieve the associated smoothed factors and we run a second set of perturbations with $\sigma \in \{0, 0.1, 0.2, \ldots, 1\}$. We then select again the parameter estimates maximizing the log-likelihood function across the alternatives and the associated factors are used for a last set of perturbation assuming now $\sigma \in \{0, 0.01, 0.02, \ldots, 0.1\}$. Finally, the vector of parameter estimates of the model of interest is thus given by the vector $\widehat{\theta}_T^{(\sigma^*)}$ (say) leading to the largest value of the log-Likelihood function across those obtained from the last perturbation stage. The associated smoothed factors are denoted by $\widehat{F}_{t|T}^{(\sigma^*)}$.

4 Empirical Analysis

This empirical analysis presents the optimal number of common and local factors, for each of several groups of countries and for both yield levels and yield differences, selected from $MCTSM$ models estimated using various given combinations of common and local factors. Section 4.1 presents the database of Treasury yield curves for the U.S., Germany, the U.K., and Japan, measured weekly from 1986 to 2009. Model estimation results, along with the optimal model for each group of countries and type of yield measurement (levels or differences) are presented in Section 4.2 using the estimation methods of Section 3 and the Akaike Information Criterion $AIC = 2\,\Xi - 2\,\mathcal{L}(\widehat{\theta}_T^{MLE})$, where $\Xi = dim(\widehat{\theta}_T^{MLE})$ denotes the number of estimated parameters of a given model. We also calculate a bootstrap variant of AIC ($AICb$) of Cavanaugh and Shumway (1997) based on the Nonparametric Monte Carlo bootstrap for state-space models of Stoffer and Wall (1991). The kind of bootstrap that is adopted is a block stationary bootstrap able to properly taking into account the persistence and the heteroskedasticity of interest rates [see Politis and Romano (1994), Politis and White (2004), and Patton, Politis, and White (2009)]. The optimal model selection is unchanged if alternative methods are used (e.g., Bayesian Information Criterion BIC or Hannan-Quinn HQ)[5]. Section 4.3 presents $MCTSMs$ parameter estimates and factors interpretations, while Section 4.4 explores the statistical dependence across international local factors.

4.1 The International Treasury Yield Curves Database

We use the international Treasury yield curves database of Pegoraro, Siegel, and Tiozzo Pezzoli (2012) consisting of four leading bond markets: the U.S., Germany, U.K. and Japan. We adopt the criteria of Gurkaynak, Sack, and Wright (2007) to filter coupon bond Treasury raw data, to guar-

[5]Results for BIC and HQ are available upon request from the authors.

antee a uniform level of liquidity, and to interpolate the discount function using the (parsimonious smoothed) Nelson and Siegel (1987) methodology. We estimate with $T = 1252$ weekly observation of these four countries, for residual maturities from 1 to 9 years (for any country), covering the period from January 1, 1986 to December 31, 2009 [see Appendix D for yields summary statistics and graphs and Pegoraro, Siegel, and Tiozzo Pezzoli (2012) for further details].

4.2 Estimating Optimal $MCTSMs$

In this section we compare model estimation results and select the optimal combination of common and local factors, using AIC and $AICb$, for groups of 2, 3, and all 4 countries, and for both yield levels and yield differences. In each case we compare combinations of common and local factors (r_c, r_ℓ) such that the yield curve of any economy is always explained by 3 to 5 factors, following the single-country term structure literature [see Adrian, Crump, and Moench (2013) and Duffee (2011)]. Accordingly, when we assume $r_c = 0$, we fix $r_\ell = 3$, $r_\ell = 4$ and $r_\ell = 5$, while, if $r_c = 1$, we consider $r_\ell = 2$, $r_\ell = 3$ and $r_\ell = 4$ and, if $r_c = 2$, we take $r_\ell = 1$, $r_\ell = 2$ and $r_\ell = 3$.

When the number of local factors is identically r_ℓ in each of the n countries, we denote the $MCTSM$ model $\mathcal{M}_n^{r_c,r_\ell}(\Phi, \Psi_\eta)$ where r_c denotes the number of common factors. The maximum value of the log-likelihood function of each model and the associated AIC and $AICb$ values are reported in Tables 2 and 3 for yield levels, and in Tables 4 and 5 for yield differences in the Appendix E. When the number of local factors is not identical in each country, we denote the $MCTSM$ model $\mathcal{M}_n^{r_c,r_j}(\Phi, \Psi_\eta)$ and specify the list for numbers of local factors by country r_j. We include the case of unequal numbers of local factors in order to compare alternative $MCTSMs$ specifications having the same factor's dimension k but different combinations of common and local factors.

Let us focus first on the case $n = 2$, that is the classical 2-country yield curve case frequently studied in the international term structure literature [see, among others, Backus, Foresi, and Telmer (2001), Ahn (2004), Bork, Dewachter, and Houssa (2009), Mosburger and Schneider (2005), Leippold and Wu (2007) and Egorov, Li, and Ng (2011)]. As can be seen from Tables 2 and 4, if we compare $MCTSMs$ providing the same number of factors to any yield curve, we always prefer the pure local factors specification $\mathcal{M}_2^{0,r_\ell}$ and this is for any pair of countries and for both for yields in level and in difference. Then, when we consider the cases $n = 3$ and $n = 4$ (Tables 3 and 5), once again we select models $\mathcal{M}_3^{0,r_\ell}$ and $\mathcal{M}_4^{0,r_\ell}$ instead of specifications where $r_c = 1$ or $r_c = 2$.

Nevertheless, as suggested by model selection literature [see Linhart and Zucchini (1986)], the above presented selection of $MCTSMs$ might be in favor of the pure local factors case $\mathcal{M}_n^{0,r_\ell}$ simply because the latter turns out to be characterized by a factor's dimension k larger than the one of the competing models $\mathcal{M}_n^{r_c,r_\ell}$. For instance, when $n = 2$ and the yield curve of any country is explained by three factors, we have that the specification $\mathcal{M}_n^{0,3}$ implies $k = 6$, while the alternative ones $\mathcal{M}_n^{1,2}$ and $\mathcal{M}_n^{2,1}$ have $k = 5$ and $k = 4$, respectively. In order to understand which specification is required by the data, we thus compare $MCTSMs$ having the same k but different combination of common and local factors including, in particular, the case $r_c = 0$. As in the previous estimations, we consider all possible combinations of countries for both yields in level and in difference. Nevertheless, for ease of presentation, the results (presented in Table 6 in the appendix Appendix E) focus on the two pairs of countries, namely $U.S.$-$U.K.$ and $U.S.$-GER, and then on the sets $U.S.$-$U.K.$-GER and $U.S.$-$U.K.$-GER-JAP, the remaining ones providing qualitatively the same information (and available upon request from the authors). From Table 6 we observe now, across alternative sets of countries and for both yield levels and differences, that the specifications with two common factors and three correlated local factors are preferred to the case $r_c = 1$ and $r_c = 0$ (the only exception being the $U.S.$-GER. case, and only for yield differences and if we consider AIC, while $AICb$ again prefers the case $r_c = 2$)[6].

4.3 Parameter Estimates and Interpretation of the Factors

Now, at that point of the analysis, we still do not know which is the nature of the common and local (smoothed) factors that we have extracted. We do not know, for instance, if common factors originate from a single economy or if they summarize some information over and above the one provided by local factors and if this feature depends on the number and the kind of analyzed countries. Indeed, we may have that some local factor of a given country loads also on the other economies. In other words, two questions naturally stand out: first, what the local factors extracted from the preferred $\mathcal{M}_n^{2,3}$ specification look like? Second, are the common factors in reality local factors loading on the other countries or are they common factors representing yield curves driving forces other than local ones?

Before focusing on this analysis, it is important to point out the ability of our estimated $MCTSMs$ to properly share interest rates information between common and local factors. Indeed, we may figure out the (extreme) case where, assuming (for ease of presentation) $n = 2$, $r_c = 1$, $r_\ell = 1$, the two local factors have dynamics identical to the common one, being $\Phi = \varphi I$ and the correlation between the two locals equal to one. In this case the two locals would look identical to the common factor and therefore distinguishing between them would be impossible. Now, if we look at the parameter estimates of $\mathcal{M}_n^{2,3}$ (yield levels) for the same set of countries analyzed in Table 6, we observe that this possible situation is completely and strongly overcome. Indeed, from Tables 10, 11 and 12 in the Appendix F we observe the following relevant features. First, we have statistically significant parameters in the AR matrix over and above those in the main diagonal; in

[6] This choice is, in addition, confirmed by likelihood ratio tests between nested models $\mathcal{M}_n^{r_c,r_j}(\Phi, \Psi_\eta)$ with a fixed k (these results are available upon request from the authors).

addition, the latter are rather different one each other. Second, we have estimated instantaneous correlations, between international local factors, that are statistically significant but never larger than 0.4 (in absolute value) and, in general, between 0.10 and 0.25 (in absolute value).

Let us move back now to factors' interpretations. An inspection of the estimated loadings in Tables 10, 11 and 12 and of the optimally extracted (smoothed) factors, provided in Appendix G, leads to the following comments. First, the local factors of any given model $\mathcal{M}_n^{2,3}$ are precisely identified with some of the local factors extracted from the associated pure local factors specification $\mathcal{M}_n^{0,4}$. For instance, in the *U.S.*-*U.K.* case, the three local *U.S.* (*U.K.*, respectively) factors in $\mathcal{M}_n^{2,3}$ are the slope, curvature and 4^{th} (level, curvature and 4^{th}, respectively) factors of the specification $\mathcal{M}_n^{0,4}$ (see Table 10 and Figure 2). In the *U.S.*-*GER.* case, they are easily identified, for both countries, as level, slope and curvature factors (Table 10 and Figure 3). In the 3-country case, the *U.S.* local factors are the level, slope and 4^{th} *U.S.* factors in $\mathcal{M}_n^{0,4}$, while, for *U.K.* and *GER.*, they are level, slope and curvature factors; we always find level, slope and curvature factors also in the 4-country case (see Tables 11 and 12, and Figures 4 and 5, respectively).

Second, the two common factors of any given model $\mathcal{M}_n^{2,3}$ tend in general to track quite closely two of the remaining (from the above mentioned identification) local factors obtained from the associated pure local factor model $\mathcal{M}_n^{0,4}$. For instance, in the case *U.S.*-*U.K.*, the two common factors closely track the *U.S.* level and the *U.K.* slope factors obtained from the specification $\mathcal{M}_2^{0,4}$. In we consider the joint dynamics of *U.S.* and Germany yield curves, the two common factors now look like the first *U.S.* and the 4^{th} German local factors provided by $\mathcal{M}_2^{0,4}$. If we now focus on the case *U.S.*-*U.K.*-*GER*, the two commons become similar to the fourth local *GER.* and the third local *U.S.* factors, respectively. In the general 4-country case, the two commons are similar to the firth local German and *U.S.* factors.

In summary, our empirical analysis highlights, first, the preference for $MCTSM$ models with $r_c = 2$ common factors that we complete with a set of $r_\ell = 3$ local factors in order to provide to each single-country yield curve five explanatory factors as suggested by the recent works of Duffee (2011) and Adrian, Crump, and Moench (2013). Second, we find that these common factors seem to be local ones (significantly) loading on the other countries.

4.4 What Drives the Dependence Between International Yield Curves?

The identification restrictions adopted for our $MCTSM$ model, and presented in Proposition 1, set up three possible sources of dependence between international term structures: the presence of common factors $F_t^{(c)}$ having a direct impact on all yield curves through the matrix Λ_c of common loadings, the unconstrained autoregressive matrix Φ allowing for causalities between all (common and local) latent factors and the instantaneous correlations between local factors of different countries ($\Psi_{ij} \neq 0$). The purpose of this section is to empirically assess the relative importance of these three channels in explaining term structures commonality.

In order to assess the importance of the first channel, namely the role played by Λ_c, we compare (through AIC and $AICb$) a $MCTSM$ model having $r_c = 2$ (as suggested by our empirical analysis) with another one with the same k but with $r_c = 0$. As far as the relevance of the second and third channel is concerned, we estimate $MCTSMs$ in which we first assume Φ block-diagonal but with the first r_c columns unconstrained (Φ_{bd}, say) and, then, we leave Φ unconstrained but we force $\widetilde{\Psi}_\eta = I$. In the former case, we turn off Granger-causalities between local factors of different countries, given that we maintain causalities of common towards local factors in order to guarantee a normalized estimator compatible with identification restrictions. In the latter specification, we switch off only the instantaneous causalities between international local factors. Let us denote this specifications $\mathcal{M}_n^{r_c,r_\ell}(\Phi_{bd},\widetilde{\Psi}_\eta)$ and $\mathcal{M}_n^{r_c,r_\ell}(\Phi, I)$, respectively. It is easily seen, following the same steps as in Appendix B, that the EM-based estimator of Φ_{bd} and $\widetilde{\Psi}_\eta$ are given by:

$$\Phi_{bd,T}^{(i+1)} = \mathcal{B}_T^{(i)}\,\mathcal{A}_T^{(i)-1} + \mathcal{K}_{\Phi,T}^{(i)}\,, \quad \widetilde{\Psi}_{\eta,T}^{(i+1)} = \tfrac{1}{T-1}\left(\mathcal{C}_T^{(i)} - \mathcal{B}_T^{(i)}\,\mathcal{A}_T^{(i)-1}\,\mathcal{B}_T^{(i)\prime} + \mathcal{K}_{\Phi,T}^{(i)}\,\mathcal{A}_T^{(i)}\,\mathcal{K}_{\Phi,T}^{(i)\prime}\right), \tag{11}$$

with:

$$vec(\mathcal{K}_{\Phi,T}^{(i)}) \;:=\; (\mathcal{A}_T^{(i)-1} \otimes \Psi_{\eta,T}^{(i+1)})\,\mathcal{H}_\Phi' \left[\mathcal{H}_\Phi\,(\mathcal{A}_T^{(i)-1} \otimes \Psi_{\eta,T}^{(i+1)})\,\mathcal{H}_\Phi'\right]^{-1} \left[\kappa_\Phi - \mathcal{H}_\Phi\, vec(\mathcal{B}_T^{(i)}\,\mathcal{A}_T^{(i)-1})\right], \tag{12}$$

where $\Psi_{\eta,T}^{(i+1)}$ is given in Proposition 2, and where $\mathcal{H}_\Phi$ is a $d \times k^2$ selection matrix such that:

$$\mathcal{H}_\Phi\, vec\,(\Phi) = \kappa_\Phi\,, \tag{13}$$

with Φ the unrestricted $k \times k$ autoregressive matrix and with κ_Φ the d-dimensional vector of zeros that guarantees to satisfy the above described structure of Φ_{bd} at each iteration of the EM algorithm[7].

Regarding the role played by matrix of common loadings, if we look at Table 6 (focusing on yield levels), and we compare the specification $r_c = 2$ with the one with $r_c = 0$ (and $k = 8$), in the *U.S.-U.K.* case AIC rises from -359198 to -356674, and in the *U.S.-GER.* one it rises from -372030 to -371226. In the 3-country case ($k = 11$), AIC moves from -529322 to -526322 while, in the 4-country case ($k = 14$), it increases from -691060 to -688300 (we reach similar conclusions if we consider $AICb$).

As far as the role played by the autoregressive matrix Φ is concerned, the results obtained for the case $\mathcal{M}_n^{r_c,r_\ell}(\Phi_{bd},\widetilde{\Psi}_\eta)$, presented in Table 7, are compared to those of $\mathcal{M}_n^{r_c,r_\ell}(\Phi,\Psi_\eta)$, in order to assess how much international local factors' dependencies are of Granger-causality kind. Let us focus again on the comparison between $r_c = 2$ and $r_c = 0$ and let us consider AIC, first. If we

[7]It is straightforward to verify that the normalization matrix A^* of equation (10) automatically preserve the structure of Φ_{bd}.

look at the 2-country $U.S.$-$U.K.$ case ($U.S.$-$GER.$ case, respectively), we observe that AIC rises of only 232 (598, respectively) while, when we close the first channel, the variation is 2524 (804, respectively). If we move to the $U.S.$-$U.K.$-$GER.$ case, AIC rises of 2164 while, when we force $\Lambda_c = 0$ (and Φ unconstrained) the variation is 3000. In the 4-country case the magnitude of this positive variation is 2482 instead of 2760. A different picture stand out if we take into account interest rate persistence using $AICb$. Indeed, while we reach the same conclusion in the $U.S.$-$U.K.$ case (the first channel is more important than the second one), we end up with an opposite result in the other cases. More precisely, in the $U.S.$-$GER.$ case, $AICb$ rises of 5864 while, when we close the first channel, the variation is 1050. In the 3-country case, $AICb$ rises of 3952 while, forcing common loadings equal to zero, induce a variation of 2232. Lastly, in the 4-country case, the size of this positive variation is 11924 instead of 4092 when we turn off common loadings. In other words, once the large interest rate dependence is taken into account through the lens of the bootstrap variant of AIC, a full AR matrix Φ seems to be (in general, but not systematically) more important than Λ_c.

Once we move to the case $\mathcal{M}_n^{r_c,r_\ell}(\Phi, I)$ (see Table 8), in order to assess the role played by the correlation terms Ψ_{ij}, we observe that, across the different number and set of countries, for several combination of common and local factors and regardless the fact to use AIC or $AICb$, the specification $\mathcal{M}_n^{r_c,r_\ell}(\Phi, \Psi_\eta)$ is strongly preferred. Indeed, in the $U.S.$-$U.K.$ case the AIC ($AICb$) difference is now 11524 (7966), and in the in the $U.S.$-$GER.$ case it is as big as 12486 (9938). If we consider the $U.S.$-$U.K.$-$GER.$ and the 4-country case, the magnitude of the AIC ($AICb$) difference is 16802 (13678) and 25308 (19096), respectively. In addition, the fact to turn off the instantaneous causalities across locals, namely to impose $\Psi_{ij} = 0$, not only strongly induces the model to significantly reduce its ability to match the data, but it provides reductions much larger (in absolute value) than the ones we have when the two other channels of dependence are closed.

These results seem therefore to suggest that dependence between international yield curves are, first of all, driven by the instantaneous correlations between international yield curves local factors, while the second most important channel seems to be provided more by the factors' (full) autoregressive matrix Φ than by the matrix Λ_c of common loadings. In other words, common factors seem to be a scarce resource that are needed to represent truly global correlations, but they may be insufficient in their ability to represent all of the regional covariance structure.

5 Conclusions and Further Developments

The purpose of this paper has been to specify, exploiting a linear Gaussian state-space approach, the preferred combination of common and local factors that are required to explain international yield curves dynamics, and to efficiently estimate by Kalman Filter the factor scores when small cross-sectional and large time-series dimensions, as well as strong serial and cross-sectional dependence, characterize the database of interest.

Our extensive empirical analysis on $MCTSMs$, allowing for Granger-causalities and instantaneous causalities across factors and exploiting a fast and powerful MLE approach based on the EM algorithm and Kalman Filter-Smoother recursions, finds that the specification with $r_c = 2$ and $r_\ell = 3$ seems to be preferred to alternative ones (of similar complexity) with $r_c = 1$ and $r_c = 0$. We also find that each common factor closely mimics (or is similar to) a local factor extracted from a pure local factor model. This result comes from an inspection of the (optimally) extracted time series of common and local factors. Indeed, the common factors turns out to be almost identical, or to closely track, local factors extracted from the associated pure local factor model. We also reach the conclusion that dependence across international yield curves are driven, first, by the instantaneous correlation between local factors of different countries and, then, by the (full) autoregressive matrix of latent factors and by the matrix of common loadings.

The purpose of future research works will be to exploit what we have learned from this work in the specification and implementation of no-arbitrage international affine term structure models with latent factors and/or with macro-financial variables.

Appendix A Proof of Proposition 1

Consider the identification problem induced by a non-singular matrix A for which the model fitted values remain unchanged when modifying factors and loadings because $\Lambda_{\mathcal{B}}\, F_t = (\Lambda_{\mathcal{B}}\, A)\,(A^{-1}\, F_t)$. Because the matrix of loadings $\Lambda_{\mathcal{B}} = [\Lambda_c\ \Lambda_l]$ must preserve the following block structure:

$$\Lambda_c = [\Lambda_{c,1}\ \ldots, \Lambda_{c,r_c}] = \begin{bmatrix} \Lambda^{(1)}_{c,1} & \Lambda^{(1)}_{c,2} & \ldots & \Lambda^{(1)}_{c,r_c} \\ \Lambda^{(2)}_{c,1} & \Lambda^{(2)}_{c,2} & \ldots & \Lambda^{(2)}_{c,r_c} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda^{(n)}_{c,1} & \Lambda^{(n)}_{c,2} & \ldots & \Lambda^{(n)}_{c,r_c} \end{bmatrix}, \qquad \Lambda_l = \begin{bmatrix} \Lambda^{(1)}_{l} & 0 & \ldots & 0 \\ 0 & \Lambda^{(2)}_{l} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \Lambda^{(n)}_{l} \end{bmatrix}. \tag{A.1}$$

the matrix A has to be such that $\Lambda^*_{\mathcal{B}} = \Lambda_{\mathcal{B}}\, A$ has the same block structure (i.e., the same pattern of zeros) as $\Lambda_{\mathcal{B}}$. More precisely, if we partition the matrix A into blocks of size $r_c, r_1, \ldots, r_n$ as follows:

$$A = \begin{bmatrix} A_{cc} & A_{c1} & \ldots & A_{cn} \\ A_{1c} & A_{11} & \ldots & A_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{nc} & A_{n1} & \ldots & A_{nn} \end{bmatrix}.$$

where the subscript c denotes entries that impact common factors, then the condition $\Lambda^*_{\mathcal{B}} = \Lambda_{\mathcal{B}}\, A$ forces A to be of the form:

$$A = \begin{bmatrix} A_{cc} & 0 & 0 & \ldots & 0 \\ A_{1c} & A_{11} & 0 & \ldots & 0 \\ A_{2c} & 0 & A_{22} & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{nc} & 0 & 0 & \ldots & A_{nn} \end{bmatrix}, \tag{A.2}$$

and therefore the number of free parameters is now given by $r^* := (r_c)^2 + r_c \left(\sum_{j=1}^n r_j\right) + \sum_{j=1}^n r_j^2 = (r_c\, k) + \sum_{j=1}^n r_j^2$. Accordingly, r^* is also the number of constraints we have to impose on the latent factors and/or the loadings in order to solve the identification problem of the international yield curve model (1)-(2)-(3) and obtain a unique model representation. The restrictions we impose are the following:

- $E\left(\eta^{(c)}_t\, \eta^{(c)\prime}_t\right) = I_{r_c}$ and $E\left(\eta^{(j)}_t\, \eta^{(j)\prime}_t\right) = I_{r_j}$ for all $j \in \{1, \ldots, n\}$, and from this set of conditions we obtain $\frac{r_c(r_c+1)}{2} + \sum_{j=1}^n \frac{r_j(r_j+1)}{2}$ restrictions;
- $E\left(\eta^{(c)}_t\, \eta^{(j)\prime}_t\right) = 0$ for all $j \in \{1, \ldots, n\}$, and here the number of restrictions is $r_c\,(\sum_{j=1}^n r_j)$;
- $(\Lambda_c^\prime \Lambda_c)$ and $\Lambda_l^{(j)\prime}\Lambda_l^{(j)}$ for all $j \in \{1, \ldots, n\}$, have to be all diagonal; these conditions imply

$\frac{r_c(r_c-1)}{2} + \sum_{j=1}^{n} \frac{r_j(r_j-1)}{2}$ restrictions.

The total number of restrictions is thus exactly r^*. The first two sets of conditions force Ψ_η to satisfy relation (4), while the last one implies $\Lambda'_{\mathcal{B}}\Lambda_{\mathcal{B}} = \Pi_{\mathcal{B}}$.

Appendix B Proof of Proposition 2

(b)

Denoting $\widetilde{Y}_t = Y_t - \mu$, the joint log-Likelihood function of Y_t and F_t (i.e., the complete data likelihood function) can be written in the following way:

$$\begin{aligned}
\ln L(Y^T, F^T) &= -\tfrac{T}{2}\ln|\Omega_{\mathcal{B}}| - \tfrac{1}{2}\textstyle\sum_{t=1}^{T}(\widetilde{Y}_t - \Lambda_{\mathcal{B}}F_t)'\Omega_{\mathcal{B}}^{-1}(\widetilde{Y}_t - \Lambda_{\mathcal{B}}F_t) \\
&\quad -\tfrac{T-1}{2}\ln|\Psi_\eta| - \tfrac{1}{2}\textstyle\sum_{t=2}^{T}(F_t - \Phi F_{t-1})'\Psi_\eta^{-1}(F_t - \Phi F_{t-1}) \\
&\quad -\tfrac{1}{2}\ln|V_1| - \tfrac{1}{2}(F_1 - \pi_1)'V_1^{-1}(F_1 - \pi_1) - \tfrac{T(N+k)}{2}\ln(2\,\pi)
\end{aligned} \tag{A.3}$$

Using the following identities (with $F_{t|T} = \mathbb{E}_\theta\left[F_t \,|\, Y^T\right]$):

$$\begin{aligned}
\widetilde{Y}_t - \Lambda_{\mathcal{B}}F_t &= \widetilde{Y}_t - \Lambda_{\mathcal{B}}F_{t|T} + \Lambda_{\mathcal{B}}(F_{t|T} - F_t) \\
F_t - \Phi F_{t-1} &= F_{t|T} - \Phi F_{t-1|T} - (F_{t|T} - F_t) + \Phi(F_{t-1|T} - F_{t-1})
\end{aligned} \tag{A.4}$$

the conditional expectation $\mathbb{E}[\ln L(Y^T, F^T) \,|\, Y^T]$, namely, the criterion maximized by the EM algorithm, is given by:

$$
\begin{aligned}
& \mathbb{E}[\ln L(Y^T, F^T)|Y^T] \\
= \; & -\tfrac{1}{2}\ln|V_1| - \tfrac{1}{2}(F_1-\pi_1)'V_1^{-1}(F_1-\pi_1) - \tfrac{T(N+k)}{2}\ln(2\,\pi) - \tfrac{T}{2}\ln|\Omega_{\mathcal{B}}| - \tfrac{T-1}{2}\ln|\Psi_\eta| \\
& -\tfrac{1}{2}\mathbf{Tr}\Bigg\{\Omega_{\mathcal{B}}^{-1}\Bigg[\Bigg(\textstyle\sum_{t=1}^T \widetilde{Y}_t\,\widetilde{Y}_t'\Bigg) - \Bigg(\textstyle\sum_{t=1}^T \widetilde{Y}_t\,F_{t|T}'\Bigg)\Lambda_{\mathcal{B}}' - \Lambda_{\mathcal{B}}\Bigg(\textstyle\sum_{t=1}^T F_{t|T}\,\widetilde{Y}_t'\Bigg) \\
& +\Lambda_{\mathcal{B}}\Bigg(\textstyle\sum_{t=1}^T F_{t|T}F_{t|T}' + P_{t|T}\Bigg)\Lambda_{\mathcal{B}}'\Bigg]\Bigg\} \qquad \text{(A.5)} \\
& -\tfrac{1}{2}\mathbf{Tr}\Bigg\{\Psi_\eta^{-1}\Bigg[\Bigg(\textstyle\sum_{t=2}^T F_{t|T}F_{t|T}' + P_{t|T}\Bigg) + \Phi\Bigg(\textstyle\sum_{t=2}^T F_{t-1|T}\,F_{t-1|T}' + P_{t-1|T}\Bigg)\Phi' \\
& \quad -\Bigg(\textstyle\sum_{t=2}^T F_{t|T}\,F_{t-1|T}' + P_{t-1,t|T}'\Bigg)\Phi' - \Phi\Bigg(\textstyle\sum_{t=2}^T F_{t-1|T}F_{t|T}' + P_{t-1,t|T}\Bigg)\Bigg]\Bigg\}.
\end{aligned}
$$

If we consider, for a given $\Lambda_{\mathcal{B}}$, the first order conditions $\frac{\partial\mathbb{E}[\ln L(Y^T,F^T)|Y^T]}{\partial\,\mu} = 0$ and $\frac{\partial\mathbb{E}[\ln L(Y^T,F^T)|Y^T]}{\partial\,\Omega_{\mathcal{B}}} = 0$, we find

$$
\begin{aligned}
\mu_T &= \overline{Y}_T - \Lambda_{\mathcal{B}}\,\overline{F}_T\,, \;\text{ where }\; \overline{Y}_T := \tfrac{1}{T}\Bigg(\textstyle\sum_{t=1}^T\, Y_t\Bigg), \overline{F}_T := \tfrac{1}{T}\Bigg(\textstyle\sum_{t=1}^T\, F_{t|T}\Bigg) \\
& \qquad \text{(A.6)} \\
\Omega_{\mathcal{B},T} &= \tfrac{1}{T}\Bigg\{\Bigg[\textstyle\sum_{t=1}^T(Y_t-\overline{Y}_T)\,(Y_t-\overline{Y}_T)'\Bigg] - \Lambda_{\mathcal{B},T}\Bigg[\textstyle\sum_{t=1}^T(F_{t|T}-\overline{F}_T)(Y_t-\overline{Y}_T)'\Bigg]\Bigg\}.
\end{aligned}
$$

The solution of the maximization problem with respect to Φ and Ψ_η provides the Φ_T and $\Psi_{\eta,T}$ presented in equation (6). Then, given μ_T, if we focus on the matrix of factor loadings, the term of (A.5) that depends on $\Lambda_{\mathcal{B}}$ only, can be written in the following way:

$$
\begin{aligned}
& -\tfrac{1}{2}\mathbf{Tr}\Bigg\{\Omega_{\mathcal{B}}^{-1}\Bigg[-\Bigg(\textstyle\sum_{t=1}^T \widetilde{Y}_t\,F_{t|T}'\Bigg)\Lambda_{\mathcal{B}}' - \Lambda_{\mathcal{B}}\Bigg(\textstyle\sum_{t=1}^T F_{t|T}\,\widetilde{Y}_t'\Bigg) \\
& \qquad +\Lambda_{\mathcal{B}}\Bigg(\textstyle\sum_{t=1}^T F_{t|T}\,F_{t|T}'\Bigg)\Lambda_{\mathcal{B}}' + \textstyle\sum_{t=1}^T \Lambda_{\mathcal{B}}P_{t|T}\Lambda_{\mathcal{B}}'\Bigg]\Bigg\} \qquad \text{(A.7)} \\
= \; & -\tfrac{1}{2}\mathbf{Tr}\Bigg\{\Omega_{\mathcal{B}}^{-1}\Bigg[-\mathcal{D}\,\Lambda_{\mathcal{B}}' - \Lambda_{\mathcal{B}}\,\mathcal{D}' + \Lambda_{\mathcal{B}}\,\overline{\mathcal{C}}\,\Lambda_{\mathcal{B}}'\Bigg]\Bigg\},
\end{aligned}
$$

where:

$$\overline{\mathcal{C}} \;=\; \textstyle\sum_{t=1}^{T}(F_{t|T}-\overline{F}_T)\,(F_{t|T}-\overline{F}_T)'+P_{t|T}\,,\quad \mathcal{D}=\textstyle\sum_{t=1}^{T}(Y_t-\overline{Y}_T)\,(F_{t|T}-\overline{F}_T)'\,, \tag{A.8}$$

and the EM-based estimator of $\Lambda_{\mathcal{B}}$ is typically determined by solving the following problem:

$$\min_{\Lambda_{\mathcal{B}}} \quad \tfrac{1}{2}\mathbf{Tr}\left\{\Omega_{\mathcal{B}}^{-1}\left[-\mathcal{D}\,\Lambda_{\mathcal{B}}'-\Lambda_{\mathcal{B}}\,\mathcal{D}'+\Lambda_{\mathcal{B}}\,\overline{\mathcal{C}}\,\Lambda_{\mathcal{B}}'\right]\right\}. \tag{A.9}$$

Nevertheless, this problem can be equivalently solved as a minimization problem with respect to the unconstrained matrix of loadings Λ, that we partition (with obvious notation) as follows:

$$\Lambda=\begin{bmatrix} \Lambda_{c1} & \Lambda_{11} & \Lambda_{12} & \dots & \Lambda_{1n} \\ \Lambda_{c2} & \Lambda_{21} & \Lambda_{22} & \dots & \Lambda_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Lambda_{cn} & \Lambda_{n1} & \dots & \dots & \Lambda_{nn} \end{bmatrix}, \tag{A.10}$$

under the equality constraint:

$$\mathcal{H}_{\Lambda}\,vec\,(\Lambda)=\kappa_{\Lambda} \tag{A.11}$$

where $\mathcal{H}_{\Lambda}$ is a $(\vartheta\,\times\,Nk)$ selection matrix that select from $vec\,(\Lambda)$ only the matrices Λ_{ij} such that $i\neq j$, $i,j\in\{1,\dots,n\}$ and κ_{Λ} is a ϑ-dimensional vector of zeros that forces Λ to be equal to $\Lambda_{\mathcal{B}}$. The Lagrangian function is:

$$L(\Lambda):=\tfrac{1}{2}\,\mathbf{Tr}\left\{\Omega^{-1}\left[-\mathcal{D}\,\Lambda'-\Lambda\,\mathcal{D}'+\Lambda\,\overline{\mathcal{C}}\,\Lambda'\right]\right\}-\lambda'\left[\mathcal{H}_{\Lambda}\,vec\,(\Lambda)-\kappa_{\Lambda}\right], \tag{A.12}$$

and the associated first order conditions are:

$$\begin{cases} \left[vec\left\{\left[\left(\overline{\mathcal{C}}\,\Lambda'-\mathcal{D}\right)\,\Omega^{-1}\right]'\right\}\right]'-\lambda'\,\mathcal{H}_{\Lambda}=0\,, \\ \\ \mathcal{H}_{\Lambda}\,vec\,(\Lambda)=\kappa_{\Lambda}\,. \end{cases} \tag{A.13}$$

If we rewrite the first equation in (A.13) as follows:

$$\left(\overline{\mathcal{C}}\otimes\Omega^{-1}\right)vec\Big(\Lambda\Big)-vec\Big(\Omega^{-1}\,\mathcal{D}\Big)=\mathcal{H}_{\Lambda}'\,\lambda\,, \tag{A.14}$$

we pre-multiply it by $\mathcal{H}_{\Lambda}\left(\overline{\mathcal{C}}^{-1}\otimes\Omega\right)$ and then we substitute from the second equation in (A.13) we find:

$$\lambda = \left[\mathcal{H}_\Lambda \left(\overline{\mathcal{C}}^{-1} \otimes \Omega\right) \mathcal{H}'_\Lambda\right]^{-1} \left\{\kappa_\Lambda - \mathcal{H}_\Lambda \left(\overline{\mathcal{C}}^{-1} \otimes \Omega\right) vec\left(\Omega^{-1}\mathcal{D}\right)\right\}. \tag{A.15}$$

We now substitute (A.15) in (A.14):

$$\begin{aligned}\left(\overline{\mathcal{C}} \otimes \Omega^{-1}\right) vec\left(\Lambda\right) - vec\left(\Omega^{-1}\mathcal{D}\right) = \mathcal{H}'_\Lambda \left[\mathcal{H}_\Lambda \left(\overline{\mathcal{C}}^{-1} \otimes \Omega\right) \mathcal{H}'_\Lambda\right]^{-1} \times \\ \left\{\kappa_\Lambda - \mathcal{H}_\Lambda \left(\overline{\mathcal{C}}^{-1} \otimes \Omega\right) vec\left(\Omega^{-1}\mathcal{D}\right)\right\},\end{aligned} \tag{A.16}$$

and let us rewrite $vec(\Omega^{-1}\mathcal{D}) = (\mathcal{D}' \otimes \Omega^{-1})\, vec(I_N)$.
Then, the term $(\overline{\mathcal{C}}^{-1} \otimes \Omega)\, vec(\Omega^{-1}\mathcal{D})$ can be written as follows:

$$\begin{aligned}\left(\overline{\mathcal{C}}^{-1} \otimes \Omega\right)(\mathcal{D}' \otimes \Omega^{-1})\, vec(I_N) &= \left(\overline{\mathcal{C}}^{-1}\mathcal{D}' \otimes \Omega\,\Omega^{-1}\right) vec(I_N) \\ &= \left(\overline{\mathcal{C}}^{-1}\mathcal{D}' \otimes I\right) vec(I_N) \\ &= vec\left(\mathcal{D}\,\overline{\mathcal{C}}^{-1}\right).\end{aligned} \tag{A.17}$$

If we substitute (A.17) in (A.16) and solve for $vec(\Lambda)$ we find that the estimator of $\Lambda_{\mathcal{B}}$ is:

$$vec\left(\Lambda_{\mathcal{B},T}\right) = vec\left(\mathcal{D}\,\overline{\mathcal{C}}^{-1}\right) + (\overline{\mathcal{C}}^{-1} \otimes \Omega)\,\mathcal{H}'_\Lambda \left[\mathcal{H}_\Lambda\,(\overline{\mathcal{C}}^{-1} \otimes \Omega)\,\mathcal{H}'_\Lambda\right]^{-1} \left[\kappa_\Lambda - \mathcal{H}_\Lambda\, vec(\mathcal{D}\,\overline{\mathcal{C}}^{-1})\right]. \tag{A.18}$$

Now, if we substitute (A.18) into (A.6) we find the estimator of the variance-covariance matrix of the measurement noise:

$$\Omega_{\mathcal{B},T} = \tfrac{1}{T}\left(\mathcal{E}_T - \mathcal{D}_T\,\overline{\mathcal{C}}_T^{-1}\,\mathcal{D}'_T + \mathcal{K}_{\Lambda,T}\,\overline{\mathcal{C}}_T\,\mathcal{K}'_{\Lambda,T}\right), \tag{A.19}$$

where $vec(\mathcal{K}_{\Lambda,T}) = (\overline{\mathcal{C}}^{-1} \otimes \Omega)\,\mathcal{H}'_\Lambda \left[\mathcal{H}_\Lambda\,(\overline{\mathcal{C}}^{-1} \otimes \Omega)\,\mathcal{H}'_\Lambda\right]^{-1} \left[\kappa_\Lambda - \mathcal{H}_\Lambda\, vec(\mathcal{D}\,\overline{\mathcal{C}}^{-1})\right]$.

(c)

For any given full rank $k \times k$ matrix A satisfying the structure (A.2), for a given factor F_t and associated smoothed value $F_{t|T} = \mathbb{E}_\theta \left[F_t \,|\, Y^T \right]$ we have the following re-parameterizations:

- $F_t^* = A^{-1} F_t$ and $F_{t|T}^* := \mathbb{E}_\theta \left[F_t^* \,|\, Y^T \right] = \mathbb{E}_\theta \left[A^{-1} F_t \,|\, Y^T \right] = A^{-1} F_{t|T}$;
- $P_{t|T}^* := \mathbb{E}[(F_t^* - F_{t|T}^*)(F_t^* - F_{t|T}^*)' \,|\, Y^T] = A^{-1} P_{t|T} (A^{-1})'$ and $P_{t-1,t|T}^* = A^{-1} P_{t-1,t|T} (A^{-1})'$;
- $\Lambda_{\mathcal{B},T}^* := \Lambda_{\mathcal{B},T} A$, $\Phi_T^* := A^{-1} \Phi_T A$, $\Psi_{\eta,T}^* := A^{-1} \Psi_{\eta,T} (A^{-1})'$, $\mu_T^* = \mu_T$ and $\Omega_{\mathcal{B},T}^* = \Omega_{\mathcal{B},T}$.

Let us assume now to have a given set of input parameters $\theta_{EM}^{(i)}$, satisfying the identification restrictions $R.i)$ and $R.ii)$, that we use to obtain $\mathcal{F}_{t|T}^{(i)} := (F_{t|T}^{(i)}, P_{t|T}^{(i)}, P_{t-1,t|T}^{(i)})$ from the Kalman Filter and Kalman Smoother (Expectation step i). Given $\mathcal{F}_{t|T}^{(i)}$, from the maximization step we obtain $\theta_{EM}^{(i+1)}$ but, at the same time, the updated parameter values do not satisfy the identification conditions anymore. More precisely, we have $\Psi_{\eta,T} \neq \Psi_{\mathcal{B}}$ and $\Lambda_{\mathcal{B},T}^{(i+1)\prime} \Lambda_{\mathcal{B},T}^{(i+1)} \neq \Pi_{\mathcal{B}}$. This means that we have to intervene in the EM recursions is such a way to guarantee, at each iteration, that the identification conditions be satisfied. This requirement is satisfied by means of the following steps:

- orthogonalizing common and local factor residuals: here we force common-factor autoregressive residuals to be uncorrelated with local-factor residuals in such a way to have the same patterns of zeros as $\Psi_{\mathcal{B}}$. Let us define the following matrix:

$$A_\perp^{-1} := \begin{bmatrix} I_{r_c} & 0 & \dots & 0 \\ -\left(\Psi_{10,T}^{c(i+1)}\right) \left[(\Psi_{00,T}^{c(i+1)})\right]^{-1} & I_{r_1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\left(\Psi_{n0,T}^{c(i+1)}\right) \left[(\Psi_{00,T}^{c(i+1)})\right]^{-1} & 0 & \dots & I_{r_n} \end{bmatrix}. \tag{A.20}$$

  such that

$$\Psi_{\eta,T}^{o(i+1)} := (A_\perp^{-1}) \, \Psi_{\eta,T}^{(i+1)} \, (A_\perp^{-1})' = \begin{bmatrix} \Psi_{00,T}^{c(i+1)} & 0 & \dots & 0 \\ 0 & \Psi_{11,T}^{o(i+1)} & \dots & \Psi_{1n,T}^{o(i+1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \Psi_{n1,T}^{o(i+1)} & \dots & \Psi_{nn,T}^{o(i+1)} \end{bmatrix}, \tag{A.21}$$

  has the desired form, that is the same blocks of zeros as $\Psi_{\mathcal{B}}$.

- orthogonalizing and normalizing within autoregressive residual blocks: now we force the matrices in the main diagonal of (A.21) to be the identity matrix by applying a Jordan decomposition to each of them. Let us denote by $\mathcal{U}_{\eta,c(i+1)}$ and $\mathcal{D}_{\eta,c(i+1)}$ the matrix of eigenvectors and eigenvalues of $\Psi_{00,T}^{c(i+1)}$, respectively. Let us define the rotation matrix $A_{\eta,c(i+1)}^{-1} :=$

$\left(\mathcal{U}_{\eta,c(i+1)}\,\mathcal{D}^{-1/2}_{\eta,c(i+1)}\right)'$ and thus we have $A^{-1}_{\eta,c(i+1)}\,\Psi^{c(i+1)}_{00,T}\,\left(A^{-1}_{\eta,c(i+1)}\right)' = I_{r_c}$. Let us now denote by $\mathcal{U}_{\eta,i(i+1)}$ and $\mathcal{D}_{\eta,i(i+1)}$ the matrix of eigenvectors and eigenvalues of $\Psi^{o(i+1)}_{jj,T}$, respectively, for any $j\in\{1,\ldots,n\}$. Let us define the rotation matrix $A^{-1}_{\eta,j(i+1)} := \left(\mathcal{U}_{\eta,j(i+1)}\,\mathcal{D}^{-1/2}_{\eta,j(i+1)}\right)'$ and thus we have $A^{-1}_{\eta,j(i+1)}\,\Psi^{o(i+1)}_{jj,T}\,\left(A^{-1}_{\eta,j(i+1)}\right)' = I_{r_j}$. We define the rotation matrix for $\Psi^{o(i+1)}_{\eta,T}$ as the block diagonal matrix
$A^{-1}_{\eta,(i+1)} := diag\left[A^{-1}_{\eta,c(i+1)}, A^{-1}_{\eta,1(i+1)},\ldots,A^{-1}_{\eta,n(i+1)}\right]$ such that:

$$\Psi^{oo(i+1)}_{T} := (A^{-1}_{\eta,(1)})\,\Psi^{o(i+1)}_{T}\,(A^{-1}_{\eta,(i+1)})' = \begin{bmatrix} I_{r_c} & 0 & 0 & \ldots & 0\\ 0 & I_{r_1} & \Psi^{oo(i+1)}_{12,T} & \ldots & \Psi^{oo(i+1)}_{1n,T}\\ 0 & \Psi^{oo(i+1)}_{21,T} & I_{r_2} & \ldots & \Psi^{oo(i+1)}_{2n,T}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & \Psi^{oo(i+1)}_{n1,T} & \Psi^{oo(i+1)}_{n2,T} & \ldots & I_{r_n}\end{bmatrix} = \Psi_{\mathcal{B}}\,. \tag{A.22}$$

In a more compact form, we define the factor's noise rotation matrix $(A^{o}_{\eta,(i+1)})^{-1} := A^{-1}_{\eta,(i+1)}\,A^{-1}_{\perp} = (A_{\perp}\,A_{\eta,(i+1)})^{-1}$ such that $(A^{o}_{\eta,(i+1)})^{-1}\,\Psi^{(i+1)}_{\eta,T}\,[(A^{o}_{\eta,(i+1)})^{-1}]' = \Psi_{\mathcal{B}}$.

- forcing orthogonality within blocks of loadings: here we intervene in the matrix of factor loadings, given that $\Lambda^{(i+1)\prime}_{\mathcal{B},T}\Lambda^{(i+1)}_{\mathcal{B},T} \neq \Pi_{\mathcal{B}}$. We know that, given the previously defined matrix $(A^{o}_{\eta,(i+1)})^{-1}$ rotating factor's noise, the associated rotation of the loadings is given by $\Lambda^{o(i+1)}_{\mathcal{B},T} := \Lambda^{(i+1)}_{\mathcal{B},T}\,A^{o}_{\eta,(i+1)}$, where $A^{o}_{\eta,(i+1)} = A_{\perp}\,A_{\eta,(i+1)}$.
  The matrix $\Lambda^{o(i+1)\prime}_{\mathcal{B},T}\Lambda^{o(i+1)}_{\mathcal{B},T}$ has the same blocks of zeros as $\Pi_{\mathcal{B}}$ but the matrices in the main diagonal are not diagonal matrices. Let us perform a Jordan decomposition of each of them. Let us diagonalize first the positive definite symmetric matrix $\Lambda^{o(i+1)\prime}_{c,\mathcal{B},T}\Lambda^{o(i+1)}_{c,\mathcal{B},T}$ associated to common factors:
$$\begin{cases} \mathcal{U}^{o\prime}_{c}\left(\Lambda^{o(i+1)\prime}_{c,\mathcal{B},T}\Lambda^{o(i+1)}_{c,\mathcal{B},T}\right)\mathcal{U}^{o}_{c} = \mathcal{D}^{o}_{c}\\[2ex] \mathcal{U}^{o\prime}_{c}\mathcal{U}^{o}_{c} = \mathcal{U}^{o}_{c}\mathcal{U}^{o\prime}_{c} = I_{r_c}\,,\ \ \mathcal{U}^{o\prime}_{c} = (\mathcal{U}^{o}_{c})^{-1}\,,\end{cases}$$
  and any positive definite symmetric matrix $\Lambda^{o(i+1)\prime}_{j,l,\mathcal{B},T}\Lambda^{o(i+1)}_{j,l,\mathcal{B},T}$ associated to the local factors of any country $j\in\{1,\ldots,n\}$:
$$\begin{cases} \mathcal{U}^{o\prime}_{j}\left(\Lambda^{o(i+1)\prime}_{j,l,\mathcal{B},T}\Lambda^{o(i+1)}_{j,l,\mathcal{B},T}\right)\mathcal{U}^{o}_{j} = \mathcal{D}^{o}_{j}\\[2ex] \mathcal{U}^{o\prime}_{j}\mathcal{U}^{o}_{j} = \mathcal{U}^{o}_{j}\mathcal{U}^{o\prime}_{j} = I_{r_j}\,,\ \ \mathcal{U}^{o\prime}_{j} = (\mathcal{U}^{o}_{j})^{-1}\,,\end{cases}$$
  and let us define the matrix $(A^{o}_{c,l,(i+1)})^{-1} := diag\left[(\mathcal{U}^{o}_{c,(i+1)})^{-1},(\mathcal{U}^{o}_{1,(i+1)})^{-1},\ldots,(\mathcal{U}^{o}_{n,(i+1)})^{-1}\right]$

such that $(A^{o}_{c,l,(i+1)})^{-1} = A^{o\prime}_{c,l,(i+1)}$. We define

$$\Lambda^{*(i+1)}_{\mathcal{B},T} := \Lambda^{o(i+1)}_{\mathcal{B},T}\, A^{o}_{c,l,(i+1)} = \Lambda^{(i+1)}_{\mathcal{B},T}\, A^{o}_{\eta,(i+1)}\, A^{o}_{c,l,(i+1)} = \Lambda^{(i+1)}_{\mathcal{B},T}\, \left(A_{\perp}\, A_{\eta,(i+1)}\, A^{o}_{c,l,(i+1)}\right) \tag{A.23}$$

and we have

$$\Lambda^{*(i+1)\prime}_{\mathcal{B},T}\Lambda^{*(i+1)}_{\mathcal{B},T} = A^{o\prime}_{c,l,(i+1)}\, \left(\Lambda^{o(i+1)\prime}_{j,l,\mathcal{B},T}\Lambda^{o(i+1)}_{j,l,\mathcal{B},T}\right)\, A^{o}_{c,l,(i+1)} = \Pi_{\mathcal{B}}\,. \tag{A.24}$$

This matrix $A^{o}_{c,l,(i+1)}$ does not perturb the structure already imposed on the factor's noise variance-covariance matrix given that, by block orthogonality, we have

$$\Psi^{*(i+1)}_{\eta,T} := (A^{o}_{c,l,(i+1)})^{-1}\, \Psi^{oo(1)}_{\eta,T}\, [(A^{o}_{c,l,(i+1)})^{-1}]' = A^{o\prime}_{c,l,(i+1)}\, \Psi^{oo(i+1)}_{\eta,T}\, A^{o}_{c,l,(i+1)} = \Psi_{\mathcal{B}}\,. \tag{A.25}$$

Thus, the normalization matrix $A^{*} := \left(A_{\perp}\, A_{\eta,(i+1)}\, A^{o}_{c,l,(i+1)}\right)$ is such that $\Lambda^{*(i+1)}_{\mathcal{B},T}$ and $\Psi^{*(i+1)}_{\eta,T}$ satisfy the indentification restrictions $R.i)$ and $R.ii)$, respectively. Moreover, it implies the factor's rotation $F^{*}_{t|T} := (A^{*})^{-1}\, F_{t|T}$ and the rotated AR matrix $\Phi^{*(i+1)}_{T} := (A^{*})^{-1}\, \Phi^{(i+1)}_{T}\, A^{*}$.
The uniqueness of the normalization matrix A^{*} is given by the fact that the diagonal entries in (5) are arranged in descending order and are distinct within each block. The latter condition is satisfied by estimates from data produced by a continuous non-degenerate model.

Appendix C A 3-Step Principal Factor Estimation Procedure

The purpose of this appendix is to briefly present a Principal Factor (PF) estimation procedure adapted to a linear Gaussian state-space model with a block structure characterizing the matrix of factor loadings (i.e., in presence of VAR distributed common and local factors). This estimation methodology is based on the following three steps:

First Step: we estimate Λ_c and $F_t^{(c)}$ by PF assuming $\Lambda_l = 0$. More precisely, denoting $\widetilde{Y}_t = Y_t - \mu$ and for any $t \in \{1,\ldots,T\}$, we have $F_t^{*(c)} := D_c^{-1/2} P_c' \widetilde{Y}_t$ and $\Lambda_{c,T}^{PF} := P_c D_c^{1/2}$ where $D_c = \text{diag}(\lambda_1^{(c)},\ldots,\lambda_{r_c}^{(c)})$ is the diagonal matrix of eigenvalues (in decreasing order of magnitude) of the variance-covariance matrix denoted $\mathcal{S}$ of the centered data $\widetilde{Y}_t$, and where $P_c = (p_1^{(c)},\ldots,p_{r_c}^{(c)})$ is the $N \times r_c$ orthogonal matrix of associated unitary eigenvectors. Given $F_t^{*(c)}$ and $\Lambda_{c,T}$, we calculate the errors $\widetilde{Y}_t^e := Y_t - \mu_T^{PF} - \Lambda_{c,T}^{PF} F_t^{*(c)}$, with $\mu_T^{PF} := \frac{1}{T}\sum_{t=1}^T Y_t$, and the associated variance-covariance matrix denoted $\mathcal{S}_j^e$ for any country $j \in \{1,\ldots,n\}$.

Second Step: we estimate $\Lambda_l^{(j)}$ and $F_{j,t}^{(l)}$ by PF on $\mathcal{S}_j^e$ and for any $j \in \{1,\ldots,n\}$. We obtain $F_{j,t}^{*(l)} := (D_l^{(j)})^{-1/2} P_l^{(j)\prime} \widetilde{Y}_t^{(j)}$ and $\Lambda_{j,l,T}^{PF} := P_l^{(j)} (D_l^{(j)})^{1/2}$ where $D_l^{(j)} := \text{diag}(\lambda_{1,l}^{(j)},\ldots,\lambda_{r_j,l}^{(j)})$ denotes the diagonal matrix of eigenvalues (in decreasing order of magnitude), and $P_l^{(j)} = (p_{1,l}^{(j)},\ldots,p_{r_j,l}^{(j)})$ the $\tau \times r_j$ orthogonal matrix of associated unitary eigenvectors of $\mathcal{S}_j^e$.

Third Step: given $F_t^* = \left(F_t^{*(c)\prime}, F_t^{*(l)\prime}\right)'$, where $F_t^{*(l)} = \left(F_{1,t}^{*(l)\prime},\ldots,F_{n,t}^{*(l)\prime}\right)'$, μ_T^{PF} and $\Lambda_{\mathcal{B},T}^{PF} := [\Lambda_{c,T}\ \Lambda_{l,T}]$, we obtain $\Omega_{\mathcal{B},T}^{PF}$ from yield errors $Y_t - \mu_T^{PF} - \Lambda_{\mathcal{B},T}^{PF} F_t^*$ while, from the regression of F_t^* on F_{t-1}^*, we estimate Φ_T^{PF} and then $\Psi_{\eta,T}^{PF}$ from associated model residuals.

Observe that, with this estimation procedure, the identification restrictions may be taken to be:

- $\mathbb{E}\left(F_t^{(c)} F_t^{(c)\prime}\right) = I_{r_c}$ and $\mathbb{E}\left(F_{j,t}^{(l)} F_{j,t}^{(l)\prime}\right) = I_{r_j}$ for all $j \in \{1,\ldots,n\}$;
- $\mathbb{E}\left(F_t^{(c)} F_{j,t}^{(l)\prime}\right) = 0$ for all $j \in \{1,\ldots,n\}$;
- $(\Lambda_c'\Lambda_c)$ and $\Lambda_l^{(j)\prime}\Lambda_l^{(j)}$, for all $j \in \{1,\ldots,n\}$, have to be diagonal.

That is, we naturally require to the marginal variance-covariance matrix of the latent factors, denoted Ψ_f, to be equal to $\Psi_{\mathcal{B}}$.

Appendix D International Treasury Yields Summary Statistics and Graphs

Maturity (Months)	Median	Mean	St. Dev.	Min	Max	$\rho(1)$	$\rho(4)$	$\rho(12)$	$\rho(36)$
U.S.									
12	5.00	4.71	2.17	0.26	9.74	0.99	0.98	0.94	0.80
24	5.18	5.00	2.10	0.48	9.70	0.99	0.98	0.93	0.80
36	5.40	5.25	2.01	0.69	9.65	0.99	0.98	0.93	0.81
48	5.56	5.45	1.94	0.94	9.60	0.99	0.98	0.93	0.82
60	5.64	5.61	1.87	1.21	9.54	0.99	0.98	0.93	0.83
72	5.74	5.75	1.81	1.51	9.56	0.99	0.98	0.93	0.83
84	5.82	5.85	1.76	1.75	9.64	0.99	0.98	0.93	0.84
96	5.88	5.94	1.72	1.95	9.70	0.99	0.98	0.93	0.84
108	5.93	6.01	1.69	2.14	9.74	0.99	0.98	0.94	0.85
Germany									
12	3.89	4.63	2.02	0.67	9.11	0.99	0.99	0.96	0.85
24	4.09	4.53	1.90	1.16	8.80	0.99	0.99	0.95	0.84
36	4.41	4.73	1.79	1.59	8.81	0.99	0.99	0.95	0.84
48	4.65	4.91	1.71	1.94	8.80	0.99	0.99	0.95	0.85
60	4.91	5.07	1.65	2.24	8.79	0.99	0.99	0.95	0.86
72	5.08	5.20	1.60	2.46	8.81	0.99	0.99	0.96	0.86
84	5.20	5.32	1.57	2.64	8.84	0.99	0.99	0.96	0.87
96	5.36	5.41	1.54	2.78	8.86	0.99	0.99	0.96	0.87
108	5.43	5.50	1.52	2.85	8.88	0.99	0.99	0.96	0.88
U.K.									
12	5.86	6.57	2.92	0.56	14.36	0.99	0.98	0.93	0.82
24	6.19	6.62	2.67	1.18	13.74	0.99	0.98	0.93	0.83
36	6.23	6.69	2.54	1.72	13.34	0.99	0.98	0.93	0.84
48	6.28	6.74	2.47	2.07	13.09	0.99	0.98	0.94	0.85
60	6.22	6.78	2.43	2.28	12.93	0.99	0.98	0.94	0.86
72	6.14	6.81	2.40	2.45	12.81	0.99	0.98	0.95	0.87
84	6.09	6.83	2.39	2.63	12.69	0.99	0.98	0.95	0.88
96	6.09	6.84	2.37	2.82	12.57	0.99	0.98	0.95	0.89
108	6.07	6.84	2.35	3.00	12.43	0.99	0.98	0.96	0.89
Japan									
12	0.59	1.91	2.24	0.01	8.35	0.99	0.99	0.97	0.91
24	0.78	2.01	2.17	0.01	8.28	0.99	0.99	0.97	0.90
36	0.99	2.18	2.12	0.07	8.21	0.99	0.99	0.96	0.90
48	1.25	2.36	2.07	0.10	8.13	0.99	0.99	0.96	0.90
60	1.47	2.52	2.03	0.15	8.06	0.99	0.99	0.96	0.90
72	1.62	2.68	1.99	0.20	7.98	0.99	0.99	0.96	0.90
84	1.74	2.83	1.96	0.26	7.90	0.99	0.99	0.96	0.90
96	1.88	2.96	1.93	0.33	7.82	0.99	0.99	0.96	0.91
108	1.99	3.08	1.90	0.40	7.74	0.99	0.99	0.96	0.91

Table 1: Summary Statistics for bond yields of U.S., Germany, U.K. and Japan daily yields. $\rho(\ell)$ denotes the sample autocorrelation for a number of lags ℓ measured in days. The sample period is from January 1, 1986 to December 31, 2009. Yields are in annual basis.

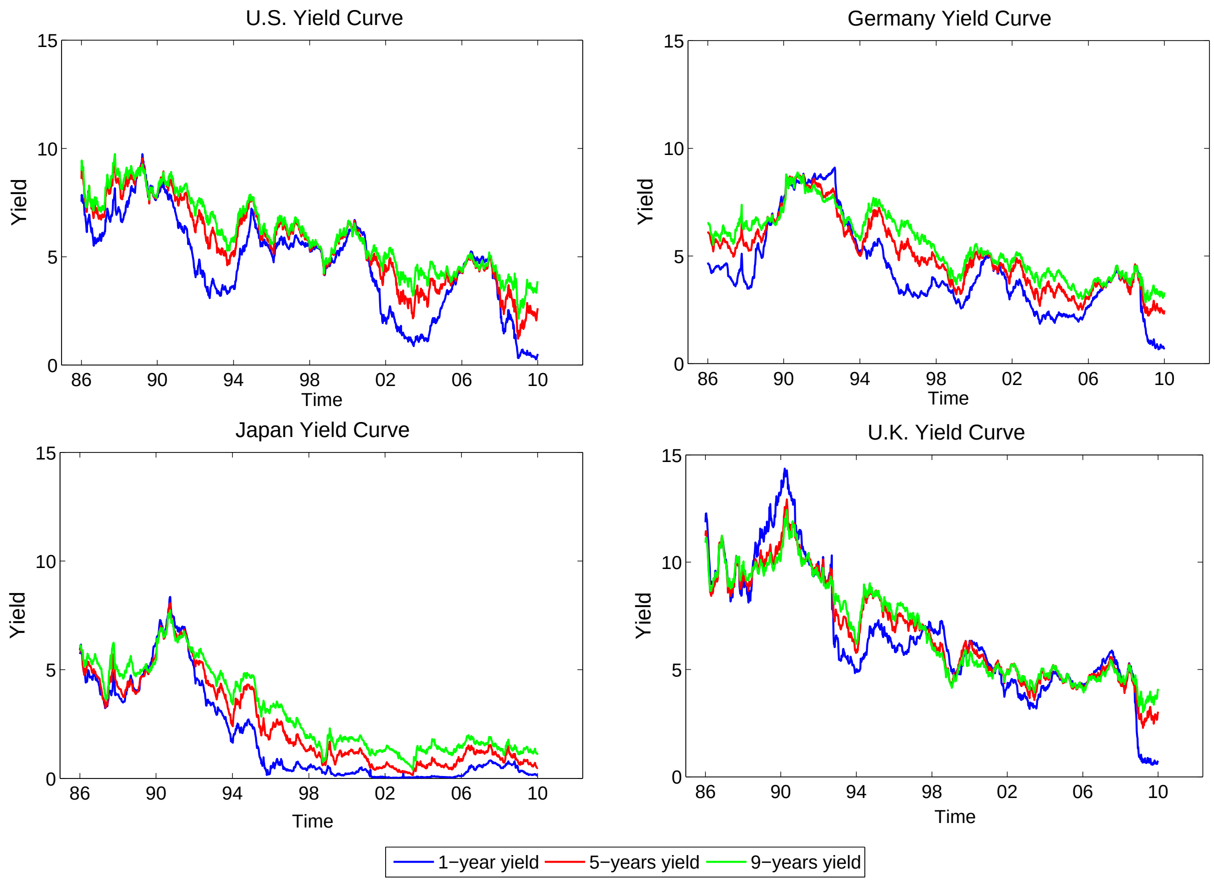


Figure 1: Treasury yield curves of U.S., Germany, Japan and U.K. and for residual maturities 1, 5 and 9 years. The term structures of interest rates of *U.S.*, Germany and Japan are taken from while *U.K.* term structures are taken from Pegoraro, Siegel, and Tiozzo Pezzoli (2012), while *U.K* term structures are taken from the Bank of England data set.

.

Appendix E Maximum Log-Likelihood of $MCTSMs$ and Model Selection

$\mathcal{M}_n^{r_c,r_\ell}(\Phi,\Psi_\eta)$ for yield levels;
The 2-Country Case

U.S. – U.K.								*U.S. – GER.*							
r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$	r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	135	159778	-319286	-322112	0	3	3	6	135	164955	-329640	-333404
0	4	4	8	188	178525	-356674	-360092	0	4	4	8	188	185801	-371226	-375164
0	5	5	10	251	191826	-383150	-386914	0	5	5	10	251	198543	-396584	-401122
1	2	2	5	119	149652	-299066	-301856	1	2	2	5	119	153828	-307418	-310800
1	3	3	7	166	170927	-341522	-344418	1	3	3	7	166	176150	-351968	-356308
1	4	4	9	223	186159	-371872	-375132	1	4	4	9	223	192456	-384466	-388910
2	1	1	4	107	138919	-277624	-280736	2	1	1	4	107	140508	-280802	-283416
2	2	2	6	148	160764	-321232	-324218	2	2	2	6	148	165476	-330656	-334980
2	3	3	8	199	179798	-359198	-361954	2	3	3	8	199	186214	-372030	-376214

U.S. – JAP.								*U.K. – GER.*							
r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$	r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	135	162186	-324102	-325156	0	3	3	6	135	160199	-320128	-322470
0	4	4	8	188	184013	-367650	-369030	0	4	4	8	188	177953	-355530	-358312
0	5	5	10	251	197486	-394470	-397430	0	5	5	10	251	190553	-380604	-384090
1	2	2	5	119	151547	-302856	-302432	1	2	2	5	119	151689	-303140	-305824
1	3	3	7	166	173371	-346410	-348280	1	3	3	7	166	170884	-341436	-344394
1	4	4	9	223	190845	-381244	-383178	1	4	4	9	223	184642	-368838	-371960
2	1	1	4	107	140173	-280132	-280060	2	1	1	4	107	138761	-277308	-280490
2	2	2	6	148	162576	-324856	-326524	2	2	2	6	148	161144	-321992	-325470
2	3	3	8	199	184357	-368316	-369678	2	3	3	8	199	179642	-358886	-361306

GER. – JAP.								*U.K. – JAP.*							
r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$	r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	135	162666	-325062	-328264	0	3	3	6	135	157484	-314698	-315300
0	4	4	8	188	183493	-366610	-369434	0	4	4	8	188	176205	-352034	-350272
0	5	5	10	251	195508	-390514	-394096	0	5	5	10	251	188779	-377056	-378448
1	2	2	5	119	151143	-302048	-304832	1	2	2	5	119	147837	-295436	-297934
1	3	3	7	166	173732	-347132	-350520	1	3	3	7	166	169090	-337848	-339070
1	4	4	9	223	189533	-378620	-382858	1	4	4	9	223	182946	-365446	-364368
2	1	1	4	107	140653	-281092	-281184	2	1	1	4	107	136423	-272632	-274402
2	2	2	6	148	162903	-325510	-328752	2	2	2	6	148	158907	-317518	-319896
2	3	3	8	199	183775	-367152	-370498	2	3	3	8	199	177459	-354520	-352702

Table 2: For any given set of $n = 2$ countries and for any given number of latent factors k, shared between r_c common factors and r_ℓ local factors, we provide the number of parameters (Ξ), the maximum value of the log-likelihood function ($\mathcal{L}(\widehat{\theta}_T^{MLE})$), the associated Akaike Information Criterion (AIC), and its bootstrap variant ($AICb$) of $MCTSMs$ $\mathcal{M}_n^{r_c,r_\ell}(\Phi,\Psi_\eta)$. We use for any country weekly yields (in level) observed from January 1, 1986 to December 31, 2009 (1252 observations) and with residual maturities from 1 to 9 years.

The 3-Country and 4-Country Case

U.S. – U.K. – GER.									U.S. – U.K. – JAP.								
r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$	r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	3	9	243	242733	-484980	-491432	0	3	3	3	9	243	239937	-479388	-483898
0	4	4	4	12	354	271533	-542358	-549436	0	4	4	4	12	354	269563	-538418	-541068
0	5	5	5	15	489	290810	-580642	-587562	0	5	5	5	15	489	288959	-576940	-583108
1	2	2	2	7	196	223047	-445702	-452496	1	2	2	2	7	196	219448	-438504	-442416
1	3	3	3	10	289	254223	-507868	-514970	1	3	3	3	10	289	251755	-502932	-507008
1	4	4	4	13	406	278610	-556408	-564154	1	4	4	4	13	406	276865	-552918	-556256
2	1	1	1	5	163	197180	-394034	-397888	2	1	1	1	5	163	195551	-390776	-393626
2	2	2	2	8	238	234589	-468702	-476422	2	2	2	2	8	238	231298	-462120	-468380
2	3	3	3	11	337	264618	-529322	-535750	2	3	3	3	11	337	262946	-525218	-528064

U.S. – GER. – JAP.									U.K. – GER. – JAP.								
r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$	r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	3	9	243	245072	-489658	-496280	0	3	3	3	9	243	240351	-480216	-486776
0	4	4	4	12	354	276840	-552972	-559126	0	4	4	4	12	354	269122	-537536	-541712
0	5	5	5	15	489	295646	-590314	-597880	0	5	5	5	15	489	287695	-574412	-580700
1	2	2	2	7	196	222296	-444200	-449490	1	2	2	2	7	196	219371	-438350	-445270
1	3	3	3	10	289	256369	-512160	-519226	1	3	3	3	10	289	252061	-503544	-509000
1	4	4	4	13	406	283476	-566140	-572512	1	4	4	4	13	406	275945	-551078	-556400
2	1	1	1	5	163	198341	-396350	-399524	2	1	1	1	5	163	196220	-392114	-397642
2	2	2	2	8	238	234520	-468564	-475800	2	2	2	2	8	238	232278	-464080	-471688
2	3	3	3	11	337	267502	-534330	-541276	2	3	3	3	11	337	262646	-524618	-530336

U.S. – U.K. – GER. – JAP.									
r_c	r_1	r_2	r_3	r_4	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	3	3	12	378	323397	-646038	-655300
0	4	4	4	4	16	568	362485	-723834	-733018
0	5	5	5	5	20	802	388039	-774474	-783988
1	2	2	2	2	9	285	290509	-580448	-590858
1	3	3	3	3	13	439	334773	-668668	-678682
1	4	4	4	4	17	637	369489	-737704	-747770
2	1	1	1	1	6	222	255090	-509736	-516604
2	2	2	2	2	10	340	303748	-606816	-618450
2	3	3	3	3	14	502	346032	-691060	-700900

Table 3: For any given set of $n = 3$ and $n = 4$ countries and for any given number of latent factors k, shared between r_c common factors and r_ℓ local factors, we provide the number of parameters (Ξ), the maximum value of the log-likelihood function ($\mathcal{L}(\widehat{\theta}_T^{MLE})$), the associated Akaike Information Criterion (AIC), and its bootstrap variant ($AICb$), of $MCTSMs$ $\mathcal{M}_n^{r_c,r_\ell}(\Phi, \Psi_\eta)$. We use for any country weekly yields (in level) observed from January 1, 1986 to December 31, 2009 (1252 observations) and with residual maturities from 1 to 9 years.

$\mathcal{M}_n^{r_c,r_\ell}(\Phi, \Psi_\eta)$ for yield differences

The 2-Country Case

U.S. – U.K.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	135	178797	-357324	-363034
0	4	4	8	188	192625	-384874	-388406
0	5	5	10	251	201177	-401852	-405164
1	2	2	5	119	172329	-344420	-350652
1	3	3	7	166	186301	-372270	-376756
1	4	4	9	223	197519	-394592	-398148
2	1	1	4	107	164085	-327956	-332634
2	2	2	6	148	179358	-358420	-361948
2	3	3	8	199	192670	-384942	-388464

U.S. – GER.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	135	180839	-361408	-367448
0	4	4	8	188	194919	-389462	-392958
0	5	5	10	251	205097	-409692	-413870
1	2	2	5	119	172591	-344944	-349458
1	3	3	7	166	188330	-376328	-381174
1	4	4	9	223	199857	-399268	-403090
2	1	1	4	107	164296	-328378	-331552
2	2	2	6	148	180856	-361416	-367574
2	3	3	8	199	194948	-389498	-393066

U.S. – JAP.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	135	180717	-361164	-366994
0	4	4	8	188	194520	-388664	-391706
0	5	5	10	251	204786	-409070	-412776
1	2	2	5	119	172463	-344688	-349080
1	3	3	7	166	188187	-376042	-380574
1	4	4	9	223	199491	-398536	-401942
2	1	1	4	107	164319	-328424	-331868
2	2	2	6	148	180722	-361148	-367048
2	3	3	8	199	194558	-388718	-391732

U.K. – GER.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	135	178714	-357158	-362968
0	4	4	8	188	191642	-382908	-386264
0	5	5	10	251	200250	-399998	-403584
1	2	2	5	119	172174	-344110	-350110
1	3	3	7	166	185104	-369876	-374526
1	4	4	9	223	196562	-392678	-396470
2	1	1	4	107	163212	-326210	-332166
2	2	2	6	148	178676	-357056	-362906
2	3	3	8	199	191685	-382972	-386440

GER. – JAP.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	135	180588	-360906	-366592
0	4	4	8	188	193529	-386682	-389722
0	5	5	10	251	203834	-407166	-411104
1	2	2	5	119	171453	-342668	-347142
1	3	3	7	166	187193	-374054	-378512
1	4	4	9	223	198503	-396560	-399956
2	1	1	4	107	164216	-328218	-331524
2	2	2	6	148	180611	-360926	-366600
2	3	3	8	199	193539	-386680	-389714

U.K. – JAP.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	135	178545	-356820	-362294
0	4	4	8	188	191247	-382118	-385060
0	5	5	10	251	199937	-399372	-402502
1	2	2	5	119	172003	-343768	-347496
1	3	3	7	166	184927	-369522	-373960
1	4	4	9	223	196218	-391990	-395446
2	1	1	4	107	163342	-326470	-332160
2	2	2	6	148	178629	-356962	-362546
2	3	3	8	199	191271	-382144	-385144

Table 4: For any given set of $n = 2$ countries and for any given number of latent factors k, shared between r_c common factors and r_ℓ local factors, we provide the number of parameters (Ξ), the maximum value of the log-likelihood function ($\mathcal{L}(\widehat{\theta}_T^{MLE})$), the associated Akaike Information Criterion (AIC), and its bootstrap variant ($AICb$), of $MCTSMs$ $\mathcal{M}_n^{r_c,r_\ell}(\Phi, \Psi_\eta)$. We use for any country weekly yields (in difference) observed from January 1, 1986 to December 31, 2009 (1252 observations) and with residual maturities from 1 to 9 years.

The 3-Country and 4-Country Case

U.S. – U.K. – GER.									U.S. – U.K. – JAP.								
r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	*AIC*	*AICb*	r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	*AIC*	*AICb*
0	3	3	3	9	243	269345	-538204	-547028	0	3	3	3	9	243	269141	-537796	-546420
0	4	4	4	12	354	289903	-579098	-584214	0	4	4	4	12	354	289437	-578166	-582962
0	5	5	5	15	489	303370	-605762	-611596	0	5	5	5	15	489	302932	-604886	-610362
1	2	2	2	7	196	254623	-508854	-516378	1	2	2	2	7	196	254330	-508268	-515792
1	3	3	3	10	289	276859	-553140	-560646	1	3	3	3	10	289	276627	-552676	-558732
1	4	4	4	13	406	295056	-589300	-594850	1	4	4	4	13	406	294216	-587620	-592986
2	1	1	1	5	163	238002	-475678	-480106	2	1	1	1	5	163	238033	-475740	-480960
2	2	2	2	8	238	262827	-525178	-534116	2	2	2	2	8	238	262595	-524714	-533586
2	3	3	3	11	337	283347	-566020	-572388	2	3	3	3	11	337	283067	-565460	-571926

U.S. – GER. – JAP.									U.K. – GER. – JAP.								
r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	*AIC*	*AICb*	r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	*AIC*	*AICb*
0	3	3	3	9	243	271224	-541962	-550812	0	3	3	3	9	243	269102	-537718	-546316
0	4	4	4	12	354	291724	-582740	-587990	0	4	4	4	12	354	288506	-576304	-581190
0	5	5	5	15	489	306868	-612758	-618998	0	5	5	5	15	489	301962	-602946	-608724
1	2	2	2	7	196	253829	-507266	-513396	1	2	2	2	7	196	253412	-506432	-513882
1	3	3	3	10	289	278698	-556844	-564380	1	3	3	3	10	289	275679	-550780	-557998
1	4	4	4	13	406	296510	-592208	-597602	1	4	4	4	13	406	293252	-585692	-591232
2	1	1	1	5	163	238028	-475730	-479152	2	1	1	1	5	163	237111	-473896	-478184
2	2	2	2	8	238	262981	-525486	-533022	2	2	2	2	8	238	262469	-524462	-533278
2	3	3	3	11	337	285356	-570038	-576552	2	3	3	3	11	337	282110	-563546	-569884

U.S. – U.K. – GER. – JAP.									
r_c	r_1	r_2	r_3	r_4	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	*AIC*	*AICb*
0	3	3	3	3	12	378	359853	-718950	-730844
0	4	4	4	4	16	568	386765	-772394	-779328
0	5	5	5	5	20	802	403614	-805624	
1	2	2	2	2	9	285	336575	-672580	-680120
1	3	3	3	3	13	439	367232	-733586	-744314
1	4	4	4	4	17	637	391400	-781526	-788674
2	1	1	1	1	6	222	311688	-622932	-628596
2	2	2	2	2	10	340	344911	-689142	-699856
2	3	3	3	3	14	502	373710	-746416	-755034

Table 5: For any given set of $n = 3$ and $n = 4$ countries and for any given number of latent factors k, shared between r_c common factors and r_ℓ local factors, we provide the number of parameters (Ξ), the maximum value of the log-likelihood function ($\mathcal{L}(\widehat{\theta}_T^{MLE})$), the associated Akaike Information Criterion (AIC), and its bootstrap variant ($AICb$), of $MCTSMs$ $\mathcal{M}_n^{r_c,r_\ell}(\Phi, \Psi_\eta)$. We use for any country weekly yields (in difference) observed from January 1, 1986 to December 31, 2009 (1252 observations) and with residual maturities from 1 to 9 years.

$\mathcal{M}_n^{r_c,r_j}(\Phi,\Psi_\eta)$ *for both yield levels and differences*

yield levels

U.S. – U.K.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	135	159778	-319286	-322112
0	4	4	8	188	178525	-356674	-360092
1	3	2	6	141	160650	-321018	-323982
1	3	4	8	193	179576	-358766	-361692
2	2	2	6	148	160764	-321232	-324218
2	3	3	8	199	179798	-359198	-361954

U.S. – GER.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	135	164955	-329640	-333404
0	4	4	8	188	185801	-371226	-375164
1	2	3	6	141	165279	-330276	-334518
1	4	3	8	193	186180	-371974	-375812
2	2	2	6	148	165476	-330656	-334980
2	3	3	8	199	186214	-372030	-376214

U.S. – U.K. – GER.

r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	2	3	8	211	232124	-463826	-469930
0	4	3	4	11	314	263475	-526322	-533518
1	3	2	2	8	224	234184	-467920	-475558
1	4	3	3	11	325	264395	-528140	-535298
2	2	2	2	8	238	234589	-468702	-476422
2	3	3	3	11	337	264618	-529322	-535750

U.S. – U.K. – GER. – JAP.

r_c	r_1	r_2	r_3	r_4	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	2	2	3	3	10	299	301081	-601564	-611526
0	4	3	3	4	14	467	344617	-688300	-696808
1	2	2	2	3	10	319	302975	-605312	-616618
1	4	3	3	3	14	484	345486	-690004	-697722
2	2	2	2	2	10	340	303748	-606816	-618450
2	3	3	3	3	14	502	346032	-691060	-700900

yield differences

U.S. – U.K.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	135	178797	-357324	-363034
0	4	4	8	188	192625	-384874	-388406
1	3	2	6	141	179330	-358378	-361856
1	3	4	8	193	192656	-384926	-388290
2	2	2	6	148	179358	-358420	-361948
2	3	3	8	199	192670	-384942	-388464

U.S. – GER.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	135	180839	-361408	-367448
0	4	4	8	188	194919	-389462	-392958
1	2	3	6	141	180849	-361416	-367618
1	4	3	8	193	194944	-389502	-393040
2	2	2	6	148	180856	-361416	-367574
2	3	3	8	199	194948	-389498	-393066

U.S. – U.K. – GER.

r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	2	3	8	211	262322	-524222	-532094
0	4	4	3	11	314	283103	-565578	-572062
1	2	2	3	8	224	262349	-524250	-532146
1	3	3	4	11	325	283327	-566004	-571762
2	2	2	2	8	238	262827	-525178	-534116
2	3	3	3	11	337	283347	-566020	-572388

U.S. – U.K. – GER. – JAP.

r_c	r_1	r_2	r_3	r_4	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	2	2	3	3	10	299	343934	-687270	-695888
0	4	4	3	3	14	467	372925	-744916	-753436
1	2	2	2	3	10	319	343969	-687300	-696162
1	4	3	3	3	14	484	373007	-745046	-753830
2	2	2	2	2	10	340	344911	-689142	-699856
2	3	3	3	3	14	502	373710	-746416	-755034

Table 6: For any given set of n countries and for any given number of latent factors k, shared between r_c common factors and r_j local factors, we provide the number of parameters (Ξ), the maximum value of the log-likelihood function ($\mathcal{L}(\widehat{\theta}_T^{MLE})$), the associated Akaike Information Criterion (AIC), and its bootstrap variant ($AICb$), of $MCTSMs$ $\mathcal{M}_n^{r_c,r_j}(\Phi,\Psi_\eta)$. We use for any country weekly yields observed from January 1, 1986 to December 31, 2009 (1252 observations) and with residual maturities from 1 to 9 years.

$\mathcal{M}_n^{r_c,r_j}(\Phi_{bd},\widetilde{\Psi}_\eta)$ for both yield levels and differences

yield levels

U.S. – U.K.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	117	159679	-319124	-319660
0	4	4	8	156	178508	-356704	-357864
1	3	2	6	124	160300	-320352	-320978
1	3	4	8	162	179378	-358432	-360822
2	2	2	6	132	160734	-321204	-323554
2	3	3	8	169	179652	-358966	-361694

U.S. – GER.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	117	164845	-329456	-328330
0	4	4	8	156	182945	-365578	-363390
1	2	3	6	124	165101	-329954	-330694
1	4	3	8	162	185875	-371426	-370624
2	2	2	6	132	165264	-330264	-332096
2	3	3	8	169	185885	-371432	-370350

U.S. – U.K. – GER.

r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	2	3	8	169	232058	-463778	-464598
0	4	3	4	11	234	263179	-525890	-526390
1	3	2	2	8	185	233904	-467438	-469188
1	4	3	3	11	249	263198	-525898	-527638
2	2	2	2	8	202	234361	-468318	-471578
2	3	3	3	11	265	263844	-527158	-531798

U.S. – U.K. – GER. – JAP.

r_c	r_1	r_2	r_3	r_4	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	2	2	3	3	10	225	300958	-601466	-602048
0	4	3	3	4	14	321	342878	-684358	-686292
1	2	2	2	3	10	250	302700	-604900	-606586
1	4	3	3	3	14	345	342904	-685118	-686426
2	2	2	2	2	10	276	303387	-606222	-609128
2	3	3	3	3	14	370	343561	-686382	-688976

yield differences

U.S. – U.K.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	117	178030	-355826	-361132
0	4	4	8	156	192515	-384718	-388068
1	2	3	6	124	178088	-355928	-361286
1	4	3	8	162	192548	-384772	-387962
2	2	2	6	132	178165	-356066	-360794
2	3	3	8	169	192557	-384776	-388092

U.S. – GER.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	117	180602	-360970	-366920
0	4	4	8	156	194788	-389264	-392662
1	2	3	6	124	180615	-360982	-366858
1	3	4	8	162	194820	-389316	-392664
2	2	2	6	132	180736	-361208	-367344
2	3	3	8	169	194833	-389328	-392806

U.S. – U.K. – GER.

r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	2	3	8	169	262213	-524088	-531752
0	4	4	3	11	234	282981	-565494	-571710
1	2	2	3	8	185	262258	-524146	-531976
1	4	3	3	11	249	283066	-565634	-571934
2	2	2	2	8	202	262256	-524108	-531940
2	3	3	3	11	265	283093	-565656	-572090

U.S. – U.K. – GER. – JAP.

r_c	r_1	r_2	r_3	r_4	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	2	2	3	3	10	225	344013	-687576	-696454
0	4	3	4	3	14	321	373285	-745928	-753908
1	2	2	2	3	10	250	344029	-687558	-696668
1	4	3	3	3	14	345	373389	-746088	-755074
2	2	2	2	2	10	276	344310	-688068	-697324
2	3	3	3	3	14	370	373448	-746156	-755336

Table 7: For any given set of n countries and for any given number of latent factors k, shared between r_c common factors and r_j local factors, we provide the number of parameters (Ξ), the maximum value of the log-likelihood function ($\mathcal{L}(\widehat{\theta}_T^{MLE})$), the associated Akaike Information Criterion (AIC), and its bootstrap variant ($AICb$), of $MCTSMs$ $\mathcal{M}_n^{r_c,r_j}\left(\Phi_{bd},\widetilde{\Psi}_\eta\right)$. We use for any country weekly yields observed from January 1, 1986 to December 31, 2009 (1252 observations) and with residual maturities from 1 to 9 years.

$\mathcal{M}_n^{r_c,r_j}(\Phi, I)$ for both yield levels and differences

yield levels

U.S. – U.K.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	126	153344	-306436	-311544
0	4	4	8	172	171278	-342212	-348900
1	2	3	6	135	154842	-309414	-314440
1	3	4	8	181	173022	-345682	-352214
2	2	2	6	144	155261	-310234	-315958
2	3	3	8	190	174027	-347674	-353988

U.S. – GER.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	126	158113	-315974	-322558
0	4	4	8	172	178253	-356702	-363412
1	2	3	6	135	159761	-319252	-325756
1	4	3	8	181	179397	-358432	-365276
2	2	2	6	144	159847	-319406	-325208
2	3	3	8	190	179962	-359544	-366276

U.S. – U.K. – GER.

r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	2	3	8	190	222523	-444666	-452714
0	4	3	4	11	274	252780	-505012	-513968
1	3	2	2	8	208	226558	-452700	-463024
1	3	3	4	11	292	255021	-509458	-520720
2	2	2	2	8	226	227664	-454876	-465198
2	3	3	3	11	310	256370	-512120	-522072

U.S. – U.K. – GER. – JAP.

r_c	r_1	r_2	r_3	r_4	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	2	2	3	3	10	262	288629	-576734	-588374
0	4	3	3	4	14	394	330392	-659996	-673336
1	2	2	3	2	10	289	292316	-584054	-596682
1	4	3	3	3	14	421	333407	-665972	-681088
2	2	2	2	2	10	316	295071	-589510	-603812
2	3	3	3	3	14	448	333732	-666568	-681804

yield differences

U.S. – U.K.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	126	176873	-353494	-357296
0	4	4	8	172	190640	-380936	-384030
1	2	3	6	135	176903	-353536	-357594
1	4	3	8	181	190743	-381124	-384290
2	2	2	6	144	176969	-353650	-357594
2	3	3	8	190	190812	-381244	-384260

U.S. – GER.

r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	126	179222	-358192	-362544
0	4	4	8	172	193020	-385696	-388548
1	2	3	6	135	179240	-358210	-362712
1	3	4	8	181	193064	-385766	-388718
2	2	2	6	144	179339	-358390	-362722
2	3	3	8	190	193149	-385918	-389134

U.S. – U.K. – GER.

r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	2	3	8	190	259813	-519246	-525386
0	4	4	3	11	274	280156	-559764	-565166
1	2	2	3	8	208	259908	-519400	-525790
1	4	3	3	11	292	280233	-559882	-565368
2	2	2	2	8	226	260168	-519884	-526274
2	3	3	3	11	310	280387	-560154	-565728

U.S. – U.K. – GER. – JAP.

r_c	r_1	r_2	r_3	r_4	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	2	2	3	3	10	262	341160	-681796	-689238
0	4	4	3	3	14	394	369721	-738654	-746008
1	2	2	2	3	10	289	341173	-681768	-689366
1	4	3	3	3	14	421	370294	-739746	-747654
2	2	2	2	2	10	316	341592	-682552	-689568
2	3	3	3	3	14	448	370300	-739704	-747754

Table 8: For any given set of n countries and for any given number of latent factors k, shared between r_c common factors and r_j local factors, we provide the number of parameters (Ξ), the maximum value of the log-likelihood function ($\mathcal{L}(\widehat{\theta}_T^{MLE})$), the associated Akaike Information Criterion (AIC), and its bootstrap variant ($AICb$), of $MCTSMs$ $\mathcal{M}_n^{r_c,r_j}(\Phi, I)$. We use for any country weekly yields observed from January 1, 1986 to December 31, 2009 (1252 observations) and with residual maturities from 1 to 9 years.

$\mathcal{M}_n^{r_c,r_j}(\Phi_{bd}, I)$ *for both yield levels and differences*

yield levels

U.S. – U.K.							
r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	108	153180	-306144	-307770
0	4	4	8	140	171073	-341866	-346398
1	2	3	6	118	154329	-308422	-310210
1	4	3	8	150	173132	-345964	-347736
2	2	2	6	128	155585	-310914	-311542
2	3	3	8	160	174076	-347832	-351980

U.S. – GER.							
r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	108	158086	-315956	-317854
0	4	4	8	140	177248	-354216	-357754
1	3	2	6	118	159600	-318964	-319624
1	3	4	8	150	179799	-359298	-362646
2	2	2	6	128	160279	-320302	-319996
2	3	3	8	160	180114	-359908	-359762

U.S. – U.K. – GER.								
r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	2	3	8	148	222160	-444024	-445862
0	4	3	4	11	194	252354	-504320	-506624
1	2	2	3	8	169	225683	-451028	-453308
1	4	3	3	11	216	255333	-510234	-512588
2	2	2	2	8	190	228490	-456600	-460682
2	3	3	3	11	238	256089	-511702	-517102

U.S. – U.K. – GER. – JAP.									
r_c	r_1	r_2	r_3	r_4	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	2	2	3	3	10	188	287904	-575432	-575928
0	4	3	3	4	14	248	329806	-659116	-661400
1	3	2	2	2	10	220	292286	-584132	-585020
1	4	3	3	3	14	282	333440	-666316	-688284
2	2	2	2	2	10	252	295940	-591376	-595346
2	3	3	3	3	14	316	333843	-667054	-671976

yield differences

U.S. – U.K.							
r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	108	176441	-352666	-356844
0	4	4	8	140	190637	-380994	-383978
1	2	3	6	118	176481	-352726	-356956
1	4	3	8	150	190715	-381130	-384290
2	2	2	6	128	177060	-353864	-358022
2	3	3	8	160	190927	-381534	-384816

U.S. – GER.							
r_c	r_1	r_2	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	3	6	108	179148	-358080	-362512
0	4	4	8	140	192944	-385608	-388492
1	3	2	6	118	179210	-358184	-362492
1	4	3	8	150	192973	-385646	-388658
2	2	2	6	128	179440	-358624	-363000
2	3	3	8	160	193298	-386276	-389320

U.S. – U.K. – GER.								
r_c	r_1	r_2	r_3	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	3	2	3	8	148	259654	-519012	-524880
0	4	4	3	11	194	280166	-559944	-564860
1	2	2	3	8	169	259675	-519012	-525056
1	4	3	3	11	216	280232	-560032	-565346
2	2	2	2	8	190	259777	-519174	-525200
2	3	3	3	11	238	280800	-561124	-566678

U.S. – U.K. – GER. – JAP.									
r_c	r_1	r_2	r_3	r_4	k	Ξ	$\mathcal{L}(\widehat{\theta}_T^{MLE})$	AIC	$AICb$
0	2	2	3	3	10	188	340965	-681554	-688634
0	4	3	3	4	14	248	369494	-738492	-744938
1	2	2	2	3	10	220	340968	-681496	-688788
1	3	4	3	3	14	282	369536	-738508	-745832
2	2	2	2	2	10	252	341371	-682238	-690458
2	3	3	3	3	14	316	370744	-740856	-748418

Table 9: For any given set of n countries and for any given number of latent factors k, shared between r_c common factors and r_j local factors, we provide the number of parameters (Ξ), the maximum value of the log-likelihood function ($\mathcal{L}(\widehat{\theta}_T^{MLE})$), the associated Akaike Information Criterion (AIC), and its bootstrap variant ($AICb$), of $MCTSMs$ $\mathcal{M}_n^{r_c,r_j}(\Phi_{bd}, I)$. We use for any country weekly yields observed from January 1, 1986 to December 31, 2009 (1252 observations) and with residual maturities from 1 to 9 years.

Appendix F Parameters Estimates

U.S.–U.K.

$\Lambda_{\mathcal{B}} \times 10^{-3}$

$\Lambda_{c,1}^{(1)}$	$\Lambda_{c,1}^{(2)}$	$\Lambda_{c,2}^{(1)}$	$\Lambda_{c,2}^{(2)}$	$\Lambda_{l}^{(1)}$			$\Lambda_{l}^{(2)}$		
0.9057**	0.4906**	-0.0543*	-1.0331**	-0.8897**	-0.3545**	-0.0471**	1.3687**	0.5412**	0.0426**
(15.24)	(4.23)	(-1.94)	(-7.15)	(-20.29)	(-20.02)	(-11.48)	(16.93)	(16.20)	(9.06)
1.1490**	0.6400**	0.0145	-0.8467**	-0.8104**	-0.0110**	0.0800**	1.3282**	0.0800**	0.0307**
(19.64)	(6.91)	(0.62)	(-7.59)	(-24.56)	(-3.17)	(12.94)	(20.47)	(12.60)	(8.80)
1.2921**	0.6708**	0.0545	-0.6077**	-0.6830**	0.1471**	0.0566**	1.3312**	-0.1219**	0.0433**
(20.17)	(8.16)	(0.12)	(-6.03)	(-22.67)	(15.54)	(12.73)	(20.62)	(-18.11)	(15.89)
1.3674**	0.6661**	0.0887	-0.3760**	-0.5494**	0.1890**	-0.0042	1.3321**	-0.1787**	0.0546**
(19.98)	(8.89)	(0.75)	(-3.75)	(-19.71)	(22.18)	(-0.97)	(20.42)	(-18.40)	(14.25)
1.4026**	0.6530**	0.1208	-0.1701	-0.4252**	0.1695**	-0.0524**	1.3198**	-0.1633**	0.0495**
(19.55)	(9.40)	(1.36)	(-1.39)	(-17.35)	(20.65)	(-14.68)	(20.48)	(-18.42)	(14.45)
1.4152**	0.6397**	0.1503*	0.0030	-0.3146**	0.1211**	-0.0717**	1.2937**	-0.1169**	0.0227**
(19.01)	(9.78)	(1.93)	(0.79)	(-15.98)	(16.59)	(-15.51)	(20.69)	(-18.97)	(13.55)
1.4155**	0.6269**	0.1770**	0.1412**	-0.2177**	0.0615**	-0.0608**	1.2579**	-0.0638**	-0.0247**
(18.34)	(10.03)	(2.46)	(2.66)	(-16.06)	(12.03)	(-15.37)	(20.89)	(-15.26)	(-16.07)
1.4094**	0.6130**	0.2006**	0.2460**	-0.1331**	-0.0006	-0.0239**	1.2177**	-0.0169**	-0.0875**
(17.52)	(10.14)	(2.93)	(4.15)	(-19.95)	(-0.63)	(-13.45)	(20.94)	(-2.55)	(-16.82)
1.4001**	0.5967**	0.2214**	0.3221**	-0.0589**	-0.0609**	0.0337**	1.1774**	0.0177*	-0.1585**
(16.52)	(10.13)	(3.34)	(5.26)	(-7.86)	(-16.48)	(13.62)	(20.75)	(1.67)	(-16.91)

Φ								Ψ_{12}		
0.9838**	-0.0167*	-0.0124**	0.0207**	-0.0027	-0.0092	-0.0143*	0.0046	0.0809**	0.0222	-0.0387
(162.69)	(-1.90)	(-2.30)	(2.55)	(0.99)	(-0.56)	(-1.70)	(1.31)	(2.16)	(0.21)	(-0.71)
0.0075*	0.9551**	-0.0095	0.0099	-0.0110	0.0043	-0.0202	0.0236	0.0297	0.0619**	-0.0137
(1.73)	(165.20)	(-0.53)	(0.28)	(-1.09)	(1.16)	(-0.18)	(1.43)	(1.09)	(2.20)	(-0.65)
-0.0223**	0.0031	0.9862**	0.0139	0.0554**	-0.0165**	-0.0032	0.0039	0.0145	-0.0884**	-0.0310
(-2.29)	(1.49)	(154.90)	(-0.0320)	(3.67)	(-2.79)	(0.63)	(0.03)	(0.88)	(-2.45)	(-1.14)
-0.0059	0.0239**	0.0249**	0.9374**	-0.0201	0.0039	0.0244**	-0.0159**			
(-1.56)	(2.69)	(2.50)	(144.80)	(-0.78)	(-0.25)	(2.10)	(-2.73)			
0.0113	0.0056	0.0066	0.0111	0.8702**	0.0068	-0.0031	-0.0073			
(-0.08)	(0.70)	(-0.32)	(0.91)	(118.12)	(0.24)	(-0.20)	(-0.69)			
-0.0205	0.0049	0.0049	-0.0125*	-0.0132	0.9789**	0.0110	0.0026			
(-0.82)	(1.04)	(1.59)	(-1.66)	(-0.98)	(181.93)	(0.82)	(-1.62)			
-0.0035	-0.0288	0.0001	0.0095	-0.0116	0.0084	0.9606**	0.0021			
(-0.39)	(0.24)	(1.54)	(-1.18)	(-0.91)	(0.54)	(147.41)	(-1.58)			
0.0065	0.0132	-0.0058	0.0090	0.0026	0.0116	0.0048	0.9742**			
(0.56)	(-0.57)	(-1.32)	(1.11)	(0.02)	(0.66)	(-0.19)	(237.28)			

U.S.–GER.

$\Lambda_{\mathcal{B}} \times 10^{-3}$

$\Lambda_{c,1}^{(1)}$	$\Lambda_{c,1}^{(2)}$	$\Lambda_{c,2}^{(1)}$	$\Lambda_{c,2}^{(2)}$	$\Lambda_{l}^{(1)}$			$\Lambda_{l}^{(2)}$		
-0.5830**	-0.2665**	0.1469**	-0.0173	1.0964**	-0.3662**	0.2188**	-0.7448**	0.5043**	-0.3220**
(-12.07)	(-7.52)	(3.52)	(-0.0721)	(23.77)	(-14.09)	(14.55)	(-14.14)	(13.25)	(-14.25)
-0.9648**	-0.3580**	0.0632	-0.2441**	0.9561**	-0.3349**	0.0110**	-0.8369**	0.4106**	-0.0073*
(-18.74)	(-9.42)	(1.63)	(-5.32)	(23.43)	(-17.23)	(3.32)	(-18.96)	(21.86)	(-1.91)
-1.1296**	-0.4110**	0.0629*	-0.2820**	0.9013**	-0.1975**	-0.1108**	-0.8913**	0.2556**	0.1486**
(-20.12)	(-10.18)	(1.67)	(-5.55)	(21.05)	(-13.37)	(-19.79)	(-21.32)	(12.88)	(13.29)
-1.1867**	-0.4338**	0.0726**	-0.2520**	0.8737**	-0.0414**	-0.1495**	-0.9139**	0.0955**	0.1885**
(-20.22)	(-10.57)	(2.08)	(-4.47)	(18.21)	(-5.45)	(-21.27)	(-22.19)	(5.94)	(19.09)
-1.1986**	-0.4422**	0.0802**	-0.2183**	0.8462**	0.0963**	-0.1314**	-0.9127**	-0.0475	0.1568**
(-19.86)	(-10.73)	(2.53)	(-3.49)	(15.85)	(6.21)	(-20.09)	(-21.96)	(-0.37)	(19.83)
-1.1959**	-0.4459**	0.0864**	-0.2058**	0.8097**	0.2045**	-0.0787**	-0.8952**	-0.1676**	0.0842**
(-19.15)	(-10.69)	(2.99)	(-2.97)	(14.09)	(18.81)	(-16.49)	(-20.89)	(-9.14)	(13.73)
-1.1923**	-0.4497**	0.0933**	-0.2207**	0.7624 **	0.2830**	-0.0062**	-0.8673**	-0.2655**	-0.0098
(-18.12)	(-10.47)	(3.46)	(-2.88)	(12.73)	(23.74)	(-2.33)	(-19.28)	(-17.32)	(-0.37)
-1.1932**	-0.4555**	0.1022**	-0.2608**	0.7057**	0.3358**	0.0768**	-0.8329**	-0.3443**	-0.1133**
(-16.85)	(-10.07)	(3.89)	(-3.04)	(11.59)	(22.32)	(13.77)	(-17.46)	(-16.87)	(-9.26)
-1.1999**	-0.4636**	0.1135**	-0.3215**	0.6419**	0.3677**	0.1644**	-0.7951**	-0.4076**	-0.2193**
(-15.50)	(-9.54)	(4.21)	(-3.29)	(10.54)	(19.53)	(18.01)	(-15.63)	(-14.34)	(-13.16)

Φ								Ψ_{12}		
0.9610**	-0.0068	-0.0283*	-0.0239	0.0104**	-0.0008	0.0088	0.0146	-0.2381**	0.0346	-0.0449
(128.49)	(-0.47)	(-1.95)	(-1.42)	(1.96)	(-0.41)	(1.45)	(1.04)	(-5.48)	(-0.96)	(-1.06)
-0.0084	0.8113**	0.0059	0.0082	0.0063	0.0045	0.0168	-0.0110	-0.0448	-0.2279**	-0.0017
(0.16)	(83.69)	(1.01)	(1.06)	(-0.14)	(1.15)	(1.38)	(-1.25)	(-0.85)	(-6.13)	(0.23)
-0.0662**	-0.0051	0.9443**	-0.0502**	0.0258	0.0129**	-0.0009	-0.0128**	0.0343	-0.1104**	-0.0656*
(-3.57)	(-0.37)	(114.86)	(-2.88)	(-0.27)	(3.18)	(-1.35)	(-1.99)	(0.52)	(-2.67)	(-1.78)
-0.0074**	0.0210	-0.0213**	0.9725**	-0.0091**	-0.0145**	0.0034	0.0194**			
(3.66)	(1.20)	(2.43)	(149.37)	(-2.87)	(-4.13)	(1.10)	(3.07)			
0.0199	0.0192	0.0449**	0.0246	0.9329**	-0.0017	-0.0030	0.0066			
(0.34)	(0.29)	(2.44)	(1.49)	(118.10)	(0.92)	(-0.74)	(-0.08)			
0.0450 **	0.0224	0.0113	0.0200*	-0.0057	0.9777**	-0.0142	0.0112			
(2.17)	(-0.84)	(0.27)	(1.66)	(0.80)	(217.36)	(-1.24)	(-0.74)			
0.0035	0.0278	0.0104	-0.0008	-0.0056	-0.0005	0.9924**	0.0018			
(-1.30)	(-0.11)	(0.47)	(-0.90)	(-0.45)	(1.43)	(223.13)	(0.26)			
-0.0004	-0.0240	-0.0067	-0.0102	0.0001	0.0120	-0.0095	0.9521**			
(-0.07)	(-1.30)	(-1.04)	(-0.83)	(0.26)	(-0.08)	(-1.08)	(182.31)			

Table 10: We report the maximum likelihood estimates of $\Lambda_{\mathcal{B}}$, Φ and Ψ_η parameters and the associated bootstrap t-values (in parenthesis). We use Nonparametric Monte Carlo block stationary bootstrap [see Stoffer and Wall (1991) and Politis and Romano (1994); the optimal block sizes are chosen following Politis and White (2004) and Patton, Politis, and White (2009)]. One and two asterisks denote statistical significance at 10% and 5% levels, respectively. The (statistically significant) parameter estimates of μ and Ω are not reported for ease of presentation.

U.S. – U.K. – GER.

$\Lambda_{\mathcal{B}} \times 10^{-3}$

$\Lambda_{c,1}^{(1)}$	$\Lambda_{c,1}^{(2)}$	$\Lambda_{c,1}^{(3)}$	$\Lambda_{c,2}^{(1)}$	$\Lambda_{c,2}^{(2)}$	$\Lambda_{c,2}^{(3)}$	$\Lambda_{l}^{(1)}$			$\Lambda_{l}^{(2)}$			$\Lambda_{l}^{(3)}$		
-0.5913**	-0.1354**	-0.2318**	-0.1468**	-0.3759**	0.0496	1.0749**	-0.4659**	-0.0860**	1.7600**	-0.7916**	-0.2316**	-0.8004**	0.4190**	0.3181**
(-10.96)	(-2.12)	(-6.42)	(-4.64)	(-3.99)	(1.67)	(22.11)	(-20.73)	(-11.96)	(14.90)	(-16.57)	(-14.48)	(-15.46)	(11.10)	(13.91)
-0.8316**	-0.3804**	-0.3505**	-0.0797**	-0.0225	0.3074**	1.1051**	-0.1503**	0.1105**	1.6198**	-0.4908**	0.0250**	-0.8315**	0.3738**	0.0180**
(-15.31)	(-6.20)	(-8.83)	(-3.15)	(0.75)	(8.77)	(25.81)	(-14.55)	(15.07)	(18.18)	(-17.02)	(3.67)	(-19.89)	(22.29)	(2.57)
-0.9812**	-0.4874**	-0.4248**	-0.0933**	0.0638**	0.3769**	1.0698**	0.0475**	0.1117**	1.5216**	-0.2059**	0.1428**	-0.8507**	0.2353**	-0.1395**
(-16.36)	(-8.02)	(-9.94)	(-3.69)	(2.64)	(9.86)	(25.18)	(3.23)	(16.02)	(20.54)	(-9.17)	(17.29)	(-22.32)	(13.74)	(-12.07)
-1.0781**	-0.5401**	-0.4651**	-0.1232**	0.0754**	0.3714**	0.9933**	0.1450**	0.0471**	1.4383**	0.0189	0.1662**	-0.8525**	0.0778**	-0.1861**
(-16.73)	(-8.96)	(-10.68)	(-4.74)	(2.90)	(9.61)	(23.83)	(19.36)	(12.46)	(21.53)	(1.20)	(18.99)	(-22.37)	(5.60)	(-19.39)
-1.1449**	-0.5720**	-0.4890**	-0.1539**	0.0681**	0.3523**	0.8976**	0.1769**	-0.0200**	1.3571**	0.1945**	0.1315**	-0.8363**	-0.0651	-0.1605**
(-16.93)	(-9.65)	(-11.18)	(-5.78)	(2.66)	(9.04)	(22.58)	(24.54)	(-7.63)	(21.52)	(9.17)	(18.89)	(-20.89)	(-1.40)	(-19.94)
-1.1945**	-0.5960**	-0.5077**	-0.1824**	0.0602**	0.3454**	0.7958**	0.1693**	-0.0654**	1.2740**	0.3342**	0.0632**	-0.8050**	-0.1826**	-0.0919**
(-17.08)	(-10.22)	(-11.38)	(-6.67)	(2.36)	(8.13)	(21.74)	(21.53)	(-16.65)	(20.79)	(14.30)	(14.22)	(-18.90)	(-10.35)	(-13.34)
-1.2339**	-0.6155**	-0.5261**	-0.2088**	0.0578**	0.3586**	0.6951**	0.1391**	-0.0831**	1.1900**	0.4483**	-0.0217**	-0.7626**	-0.2740**	0.0006
(-17.12)	(-10.73)	(-11.27)	(-7.29)	(2.23)	(7.16)	(21.44)	(18.20)	(-17.37)	(19.58)	(16.70)	(-6.54)	(-16.87)	(-18.79)	(-0.72)
-1.2669**	-0.6300**	-0.5462**	-0.2338**	0.0615**	0.3916**	0.5988**	0.0965**	-0.0751**	1.1084**	0.5444**	-0.1104**	-0.7128**	-0.3425**	0.1049**
(-16.92)	(-11.12)	(-10.89)	(-7.58)	(2.27)	(6.46)	(21.82)	(15.60)	(-17.46)	(18.01)	(17.22)	(-15.36)	(-14.94)	(-18.25)	(8.71)
-1.2959**	-0.6385**	-0.5686**	-0.2578**	0.0698**	0.4414**	0.5083**	0.0476**	-0.0457**	1.0322**	0.6275**	-0.1939**	-0.6584**	-0.3921**	0.2138**
(-16.43)	(-11.34)	(-10.36)	(-7.56)	(2.41)	(6.02)	(23.07)	(12.90)	(-17.12)	(16.26)	(16.88)	(-17.36)	(-13.14)	(-14.99)	(13.53)

Φ											$\Psi_{1,2}$			$\Psi_{1,3}$			$\Psi_{2,3}$		
0.9612**	0.0319	-0.0192**	-0.0248**	0.0206*	-0.0184*	-0.0207**	-0.0024	-0.0022	0.0061	-0.0008	0.2547**	-0.0879**	-0.0225	-0.2461**	0.1393**	0.0150	-0.3222**	0.0699**	-0.0080
(119.15)	(1.45)	(-3.06)	(-3.11)	(1.71)	(-1.80)	(-2.73)	(-0.75)	(-0.26)	(-0.23)	(0.96)	(6.40)	(-2.06)	(-0.85)	(-5.31)	(4.29)	(0.48)	(-8.50)	(2.46)	(-0.30)
0.0217	0.8170**	0.0108	-0.0010	0.0152*	0.0030	0.0131*	0.0278**	-0.0244**	-0.0425**	-0.0202*	-0.0088	-0.0325	0.0033	-0.0564	-0.0063	-0.0755*	-0.0003	-0.1558**	-0.0060
(0.66)	(64.03)	(0.07)	(-0.78)	(1.80)	(0.37)	(1.83)	(2.39)	(-2.06)	(-2.03)	(-1.77)	(-0.80)	(-0.23)	(0.32)	(-1.05)	(-0.52)	(-1.90)	(-1.39)	(-4.63)	(-0.11)
-0.0324**	0.0287	0.9724**	-0.0293	0.0598**	0.0121*	-0.0126	0.0096*	0.0279**	0.0003	0.0375**	0.0046	-0.0300	0.1112**	-0.0085	0.0643*	-0.0261	-0.0713	0.0172	-0.0765**
(-2.35)	(0.85)	(119.53)	(-1.19)	(3.41)	(1.75)	(-0.47)	(1.65)	(3.90)	(-0.66)	(3.32)	(0.73)	(-0.64)	(3.63)	(-0.71)	(1.87)	(-0.82)	(-1.62)	(1.00)	(-2.15)
-0.0092	0.0282	-0.0368*	0.9549**	0.0217	-0.0108**	-0.0176	-0.0092**	-0.0072**	0.0010	0.0059									
(0.74)	(0.71)	(-1.78)	(138.45)	(0.06)	(-2.64)	(-1.43)	(-2.66)	(-2.84)	(0.34)	(-1.29)									
0.0173	0.0093	0.0036	-0.0022	0.8845**	0.0122	0.0001	0.0147	0.0096*	-0.0032	0.0167**									
(-0.47)	(0.74)	(-1.47)	(-1.47)	(122.11)	(0.99)	(-0.77)	(0.92)	(1.68)	(-0.44)	(2.15)									
-0.0217	0.0117	0.0159	0.0081	0.0050	0.9725**	-0.0041	-0.0133*	0.0022*	0.0030	0.0074									
(-0.36)	(-0.19)	(1.51)	(1.38)	(-0.12)	(129.07)	(0.48)	(1.74)	(1.75)	(-0.01)	(0.62)									
-0.0104	0.0498	-0.0120	-0.0007	-0.0027	-0.0058	0.9703**	-0.0156	0.0041	0.0053	0.0182									
(0.60)	(0.58)	(-0.12)	(0.33)	(-0.47)	(-0.02)	(129.27)	(-1.07)	(0.64)	(1.47)	(0.15)									
-0.0045	0.0183	0.0081	0.0081	0.0116	-0.0326*	0.0055	0.9577**	-0.0036	0.0073	0.0062									
(-0.58)	(0.24)	(-0.50)	(-0.63)	(1.04)	(-1.65)	(0.08)	(140.83)	(0.09)	(0.30)	(0.77)									
0.0285	-0.0483	-0.0010	0.0059	-0.0251	-0.0181	0.0059	-0.0126**	0.9620**	-0.0251	-0.0283									
(1.41)	(0.57)	(-0.66)	(0.14)	(-1.60)	(-1.41)	(0.33)	(-2.05)	(114.94)	(-0.38)	(-0.10)									
0.0058	-0.0588	0.0040	-0.0015	-0.0054	-0.0053	-0.0292**	0.0083	-0.0213	0.9602**	-0.0009									
(-1.12)	(-0.05)	(0.54)	(0.02)	(0.12)	(0.17)	(-2.92)	(0.88)	(-0.18)	(143.61)	(0.68)									
-0.0042	-0.0390	0.0049	0.0017	-0.0073	-0.0156	0.0024	-0.0064	-0.0261	0.0031	0.9416**									
(-0.45)	(-1.49)	(0.35)	(-0.05)	(-0.44)	(-0.23)	(-1.15)	(0.33)	(-1.46)	(-0.64)	(128.77)									

Table 11: We report the maximum likelihood estimates of $\Lambda_{\mathcal{B}}$, Φ and Ψ_η parameters and the associated bootstrap t-values (in parenthesis). We use Nonparametric Monte Carlo block stationary bootstrap [see Stoffer and Wall (1991) and Politis and Romano (1994); the optimal block sizes are chosen following Politis and White (2004) and Patton, Politis, and White (2009)]. One and two asterisks denote statistical significance at 10% and 5% levels, respectively. The (statistically significant) parameter estimates of μ and Ω are not reported for ease of presentation.

U.S. – U.K. – GER. – JAP.

$\Lambda_{\mathcal{B}} \times 10^{-3}$

$\Lambda^{(1)}_{c,1}$	$\Lambda^{(2)}_{c,1}$	$\Lambda^{(3)}_{c,1}$	$\Lambda^{(4)}_{c,1}$	$\Lambda^{(1)}_{c,2}$	$\Lambda^{(2)}_{c,2}$	$\Lambda^{(3)}_{c,2}$	$\Lambda^{(4)}_{c,2}$		$\Lambda^{(1)}_{l}$			$\Lambda^{(2)}_{l}$			$\Lambda^{(3)}_{l}$			$\Lambda^{(4)}_{l}$	
0.0691**	-0.3359**	0.0633**	-0.0340	0.0030	-0.0469	0.0173	0.0337	1.1204**	0.6323**	-0.2475**	1.6742**	-0.9017**	-0.2467**	-0.8160**	-0.4650**	-0.3035**	-0.6280**	0.4677**	-0.2535**
(2.32)	(-2.94)	(1.72)	(-0.59)	(0.86)	(-0.30)	(0.34)	(1.00)	(21.23)	(24.04)	(-18.41)	(14.28)	(-20.29)	(-16.04)	(-16.27)	(-11.65)	(-12.28)	(-11.77)	(14.65)	(-15.49)
0.2445**	0.0765**	0.3295**	0.0497**	-0.0551	0.0005	0.0756	-0.0079	1.3004**	0.4143**	0.0588**	1.6333**	-0.5350**	0.0378**	-0.8706**	-0.4042**	-0.0217**	-0.6837**	0.4163**	-0.0580**
(5.45)	(2.49)	(6.98)	(2.49)	(-0.13)	(0.06)	(0.99)	(-0.37)	(25.21)	(26.12)	(8.16)	(18.24)	(-20.40)	(5.86)	(-19.95)	(-22.30)	(-2.69)	(-14.69)	(19.54)	(-5.82)
0.2878**	0.2027**	0.4135**	0.0843**	-0.0279	-0.0118	0.0855	-0.0180	1.4007**	0.2122**	0.1621**	1.5757**	-0.2281**	0.1554**	-0.9148**	-0.2602**	0.1325**	-0.7569**	0.2865**	0.0839**
(5.91)	(4.50)	(7.70)	(3.54)	(0.32)	(-0.20)	(0.94)	(-0.69)	(25.18)	(15.19)	(20.94)	(20.98)	(-10.25)	(19.62)	(-21.93)	(-13.24)	(11.77)	(-17.47)	(16.09)	(12.34)
0.2927**	0.2360**	0.4140**	0.0937**	0.0065	-0.0306	0.0743	-0.0158	1.4357**	0.0474**	0.1530**	1.5126**	0.0053	0.1729**	-0.9379**	-0.1019**	0.1823**	-0.8134**	0.1404**	0.1527**
(5.85)	(4.95)	(7.15)	(3.73)	(0.69)	(-0.60)	(0.84)	(-0.55)	(24.20)	(6.37)	(22.22)	(22.30)	(0.87)	(21.12)	(-22.36)	(-6.01)	(18.45)	(-18.78)	(6.89)	(20.70)
0.2903**	0.2432**	0.3951**	0.0919**	0.0184	-0.0472	0.0594	-0.0100	1.4340**	-0.0813**	0.0954**	1.4452**	0.1861**	0.1321**	-0.9414**	0.0423	0.1611**	-0.8472**	0.0019	0.1583**
(5.69)	(4.99)	(6.33)	(3.63)	(0.68)	(-1.01)	(0.72)	(-0.27)	(22.87)	(-5.72)	(21.22)	(22.64)	(9.07)	(20.93)	(-21.79)	(0.03)	(17.89)	(-19.08)	(-0.66)	(21.37)
0.2900**	0.2473**	0.3859**	0.0868**	0.0010	-0.0587	0.0483	-0.0064	1.4123**	-0.1800**	0.0236**	1.3744**	0.3309**	0.0588**	-0.9294**	0.1637**	0.0955**	-0.8603**	-0.1220**	0.1141**
(5.53)	(5.01)	(5.59)	(3.47)	(0.38)	(-1.34)	(0.59)	(-0.08)	(21.47)	(-22.72)	(9.76)	(22.24)	(14.74)	(15.08)	(-20.70)	(8.45)	(12.28)	(-18.80)	(-7.72)	(17.45)
0.2937**	0.2540**	0.3961**	0.0826**	-0.0433	-0.0622	0.0431	-0.0096	1.3806**	-0.2556**	-0.0462**	1.3024**	0.4505**	-0.0296**	-0.9066**	0.2619**	0.0038	-0.8564**	-0.2310**	0.0312**
(5.25)	(5.18)	(5.06)	(3.39)	(-0.15)	(-1.50)	(0.50)	(-0.14)	(20.17)	(-26.06)	(-10.54)	(21.25)	(14.75)	(-9.20)	(-19.38)	(17.46)	(1.36)	(-18.10)	(-14.98)	(3.95)
0.3010**	0.2617**	0.4264**	0.0815**	-0.1092	-0.0565	0.0437	-0.0225	1.3446**	-0.3136**	-0.1077**	1.2315**	0.5520**	-0.1207**	-0.8767**	0.3398**	-0.1019**	-0.8386**	-0.3269**	-0.0823**
(4.78)	(5.45)	(4.75)	(3.38)	(-0.79)	(-1.49)	(0.47)	(-0.60)	(19.06)	(-23.63)	(-16.65)	(19.79)	(18.65)	(-17.55)	(-17.96)	(16.19)	(-8.59)	(-16.87)	(-18.04)	(-11.21)
0.3110**	0.2668**	0.4742**	0.0845**	-0.1911	-0.0432	0.0494	-0.0475	1.3073**	-0.3583**	-0.1589**	1.1641**	0.6397**	-0.2059**	-0.8422**	0.4008**	-0.2138**	-0.8099**	-0.4125**	-0.2210**
(4.18)	(5.70)	(4.59)	(3.27)	(-1.34)	(-1.26)	(0.48)	(-1.37)	(18.13)	(-21.69)	(-18.57)	(18.07)	(18.55)	(-19.27)	(-16.52)	(13.03)	(-13.80)	(-15.01)	(-15.47)	(-17.29)

Φ

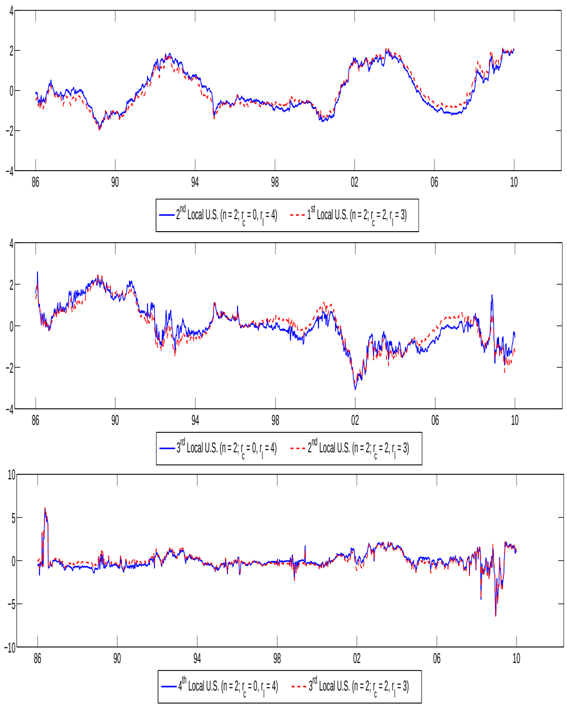
0.7992**	-0.0220	0.0304**	-0.0095	-0.0017	0.0033	-0.0022	0.0334*	0.0046	0.0369	0.0258	-0.0198	-0.0068	0.0064
(56.67)	(-0.65)	(2.21)	(-1.03)	(0.41)	(1.22)	(1.33)	(1.74)	(1.01)	(0.43)	(1.38)	(-1.61)	(-0.02)	(1.42)
-0.0069	0.8632**	0.0046	0.0006	0.0310	-0.0223	-0.0216	-0.0017	-0.0115	0.0112	0.0127	-0.0079	-0.0094	-0.0066
(-0.06)	(94.06)	(0.03)	(0.95)	(1.58)	(-1.50)	(-0.92)	(-0.40)	(-1.37)	(0.55)	(0.61)	(-0.46)	(-0.57)	(-0.76)
0.0444	-0.0108	0.9739**	0.0127	0.0174*	0.0083*	-0.0171	0.0050	0.0175**	0.0099	-0.0388*	-0.0146	-0.0059	-0.0049
(0.44)	(0.23)	(110.49)	(0.60)	(1.81)	(1.70)	(0.25)	(0.81)	(2.87)	(0.68)	(-1.78)	(-1.24)	(0.13)	(0.86)
-0.0021	-0.0378*	0.0252	0.9803**	-0.0105	0.0107	-0.0066	0.0141*	0.0220**	-0.0012	-0.0102*	-0.0030	-0.0114**	0.0061
(0.90)	(-1.67)	(1.50)	(158.41)	(0.21)	(1.48)	(-1.00)	(1.78)	(3.38)	(0.51)	(-1.84)	(-1.07)	(-1.96)	(-0.15)
0.0238	-0.0023	-0.0183*	-0.0068	0.9373**	-0.0043	-0.0238	0.0092	-0.0008	0.0106	-0.0270	-0.0039	0.0011	0.0046
(1.24)	(-0.68)	(-1.73)	(0.11)	(157.85)	(-1.29)	(-1.54)	(-0.20)	(-1.32)	(1.35)	(-1.42)	(-0.21)	(1.06)	(-0.10)
0.0174	0.0167	0.0182*	0.0048	0.0211	0.9672**	-0.0142	-0.0048	0.0005**	0.0160	-0.0262	-0.0101	0.0099*	-0.0017
(-0.13)	(0.78)	(1.66)	(0.15)	(1.51)	(93.13)	(0.78)	(1.52)	(2.24)	(0.60)	(-1.30)	(-0.25)	(1.83)	(1.31)
0.0332	0.0105	-0.0243	0.0005	0.0045	-0.0241	0.9365**	-0.0061	-0.0050	0.0189	-0.0279	-0.0303	-0.0172	-0.0074
(0.07)	(0.38)	(-0.79)	(-0.17)	(0.20)	(-0.19)	(79.12)	(-1.13)	(0.73)	(-0.27)	(-0.52)	(-1.37)	(-0.34)	(-0.93)
0.0201	-0.0080	0.0087	0.0021	0.0160	-0.0354	0.0054	0.9581**	-0.0084	-0.0037	-0.0080	0.0015	0.0025	-0.0006
(0.50)	(-0.74)	(-0.27)	(0.79)	(0.59)	(-1.63)	(-0.31)	(121.01)	(-0.65)	(0.43)	(-0.63)	(-0.33)	(0.23)	(-0.85)
-0.0482	0.0033	-0.0154**	-0.0006	-0.0171	-0.0304*	-0.0031	-0.0191**	0.9542**	0.0077	0.0511	0.0018	-0.0264*	-0.0300*
(0.37)	(0.41)	(-2.32)	(-0.02)	(-1.44)	(-1.66)	(-0.32)	(-2.14)	(100.32)	(-0.50)	(0.94)	(0.70)	(-1.86)	(-1.67)
0.0509	0.0067	-0.0101	-0.0006	0.0048	0.0331	0.0620**	-0.0235*	0.0245	0.9458**	0.0093	0.0185	0.0052	0.0231
(-0.14)	(0.26)	(-0.96)	(-0.38)	(-0.33)	(1.38)	(3.35)	(-1.83)	(1.41)	(89.70)	(1.25)	(1.24)	(-0.15)	(1.47)
0.0488*	0.0153	-0.0149	0.0107	0.0143	0.0166	-0.0103	0.0193	0.0284	0.0258	0.9038**	-0.0056	0.0325*	0.0245
(1.80)	(0.37)	(-1.52)	(1.52)	(0.95)	(0.60)	(0.87)	(0.02)	(1.12)	(0.67)	(84.86)	(-0.01)	(1.90)	(1.05)
-0.0247	-0.0028	-0.0158	-0.0004	-0.0201	-0.0161	-0.0133	0.0132	-0.0186**	0.0015	0.0396**	0.9793**	-0.0341**	-0.0234**
(-0.39)	(-0.11)	(-1.45)	(0.44)	(-1.13)	(-0.99)	(0.78)	(0.06)	(-2.48)	(-1.35)	(2.96)	(109.35)	(-2.44)	(-2.81)
-0.0296	-0.0026	0.0098	-0.0007	-0.0131	-0.0212	-0.0126	-0.0109	-0.0057	0.00004	0.0121	-0.0066	0.9714**	-0.0213
(-0.29)	(0.03)	(1.28)	(0.17)	(-0.47)	(-1.27)	(-0.01)	(-0.94)	(-0.19)	(0.36)	(-0.48)	(0.79)	(125.97)	(0.12)
0.0049	0.0146	-0.0287*	0.0226	0.0070	-0.0059	0.0043	0.0091	-0.0212	-0.0027	0.0165	0.0055	-0.0266	0.9316**
(0.18)	(-0.004)	(-1.69)	(1.40)	(0.09)	(0.35)	(0.34)	(1.03)	(-0.83)	(-0.49)	(0.62)	(0.82)	(-0.94)	(108.99)

	$\Psi_{1,2}$			$\Psi_{1,3}$			$\Psi_{1,4}$			$\Psi_{2,3}$			$\Psi_{2,4}$			$\Psi_{3,4}$	
0.4001**	0.1324**	-0.0884**	-0.4144**	0.0671	-0.0117	-0.1619**	-0.0487*	0.0054	-0.4007**	-0.0234	0.0034	-0.1504**	-0.0261	-0.0824**	0.2500**	0.0287	0.0123
(13.12)	(3.79)	(-3.94)	(-12.66)	(1.30)	(-0.15)	(-4.68)	(-1.84)	(-1.05)	(-11.45)	(-1.20)	(0.10)	(-3.79)	(-1.08)	(-3.11)	(6.28)	(1.18)	(0.50)
0.0048	-0.2008**	0.0618**	0.0434	-0.2296**	-0.0368	0.0499	0.0373	0.0203	-0.0879**	0.2301**	-0.0009	-0.0826**	-0.0403**	-0.0019	0.0033	-0.0893**	-0.0132
(1.01)	(-5.21)	(2.11)	(0.77)	(-6.35)	(-1.10)	(0.84)	(0.81)	(0.94)	(-3.54)	(6.40)	(0.02)	(-3.48)	(-2.01)	(-0.66)	(0.69)	(-2.30)	(-0.52)
-0.0138	-0.1319**	0.0861**	-0.0303	-0.1254**	0.0642*	0.0291	0.0293	0.0254	-0.0466	-0.0514**	0.0775**	-0.0128	0.0086	-0.0098	0.0829	0.0440*	0.0060
(-0.14)	(-2.88)	(2.61)	(-0.76)	(-3.14)	(1.82)	(0.50)	(0.35)	(0.81)	(-0.41)	(-2.47)	(2.64)	(0.56)	(0.36)	(0.10)	(1.51)	(1.68)	(1.07)

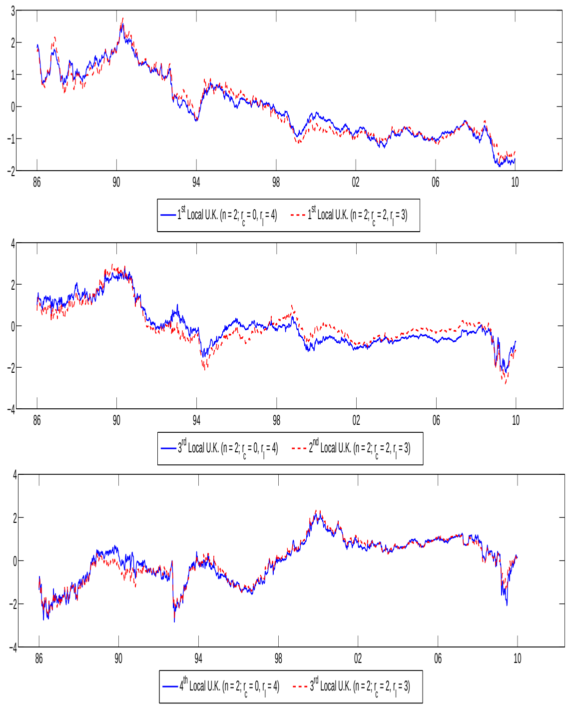
Table 12: We report the maximum likelihood estimates of $\Lambda_{\mathcal{B}}$, Φ and Ψ_η parameters and the associated bootstrap t-values (in parenthesis). We use Nonparametric Monte Carlo block stationary bootstrap [see Stoffer and Wall (1991) and Politis and Romano (1994); the optimal block sizes are chosen following Politis and White (2004) and Patton, Politis, and White (2009)]. One and two asterisks denote statistical significance at 10% and 5% levels, respectively. The (statistically significant) parameter estimates of μ and Ω are not reported for ease of presentation.

Appendix G Smoothed Common and Local Factors

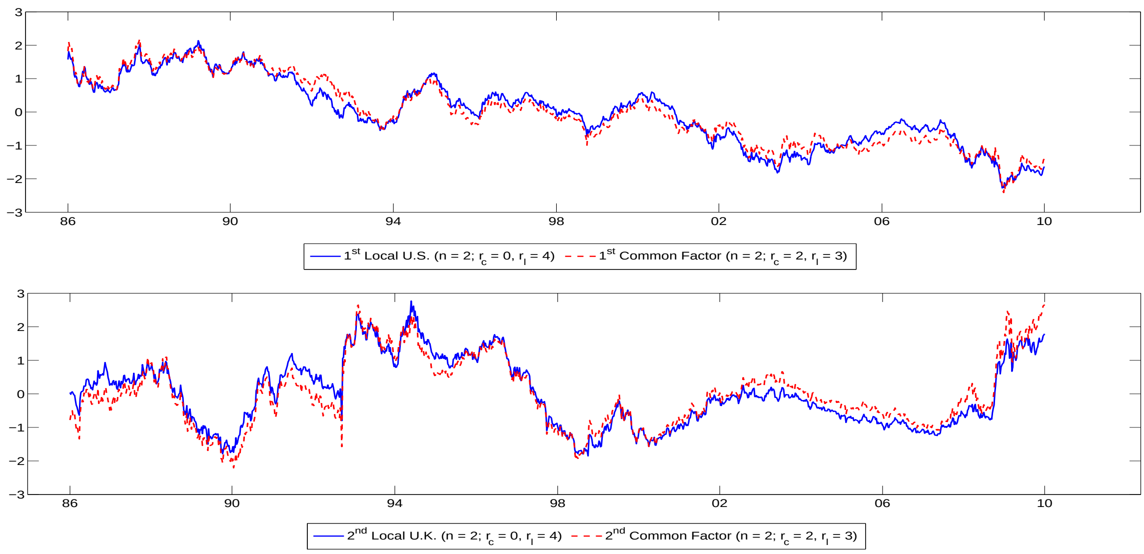
2-country case: U.S.-U.K.



(a) Local *U.S.* factors of $\mathcal{M}_2^{0,4}$ and $\mathcal{M}_2^{2,3}$



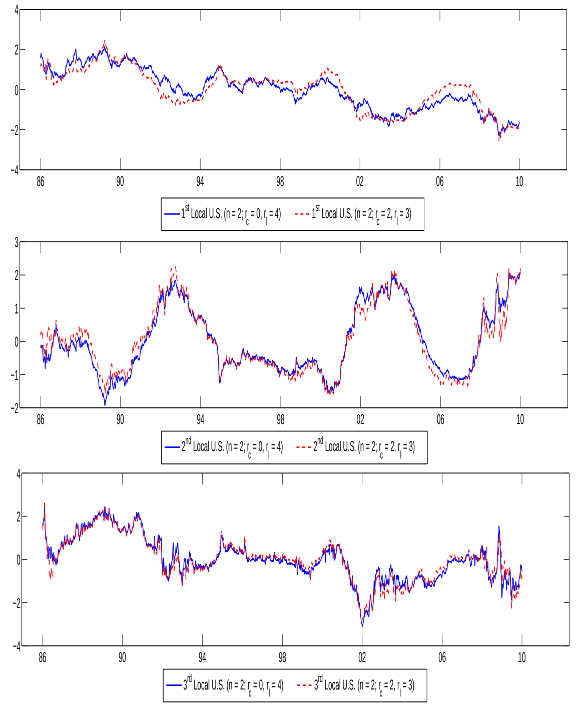
(b) Local *U.K.* factors of $\mathcal{M}_2^{0,4}$ and $\mathcal{M}_2^{2,3}$



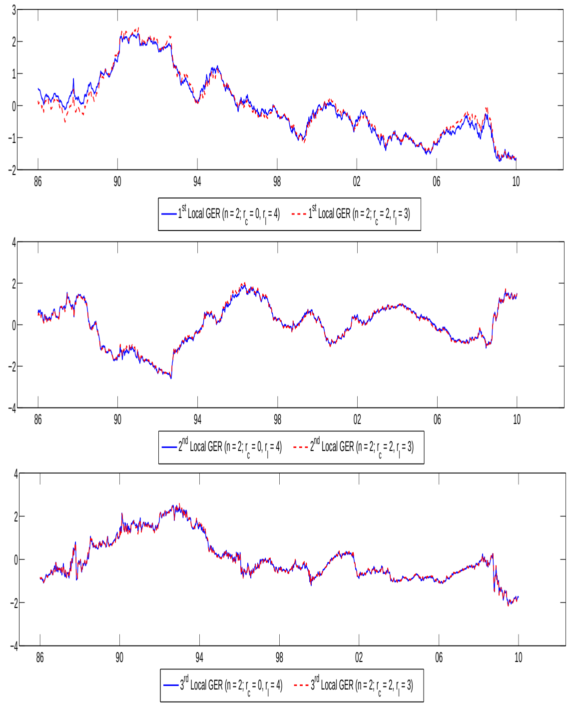
(c) 1^{st} and 2^{nd} common factor of $\mathcal{M}_2^{2,3}$

Figure 2: Smoothed factors in the 2-country *U.S.-U.K.* case when $(r_c = 0, r_\ell = 4)$ and $(r_c = 2, r_\ell = 3)$.

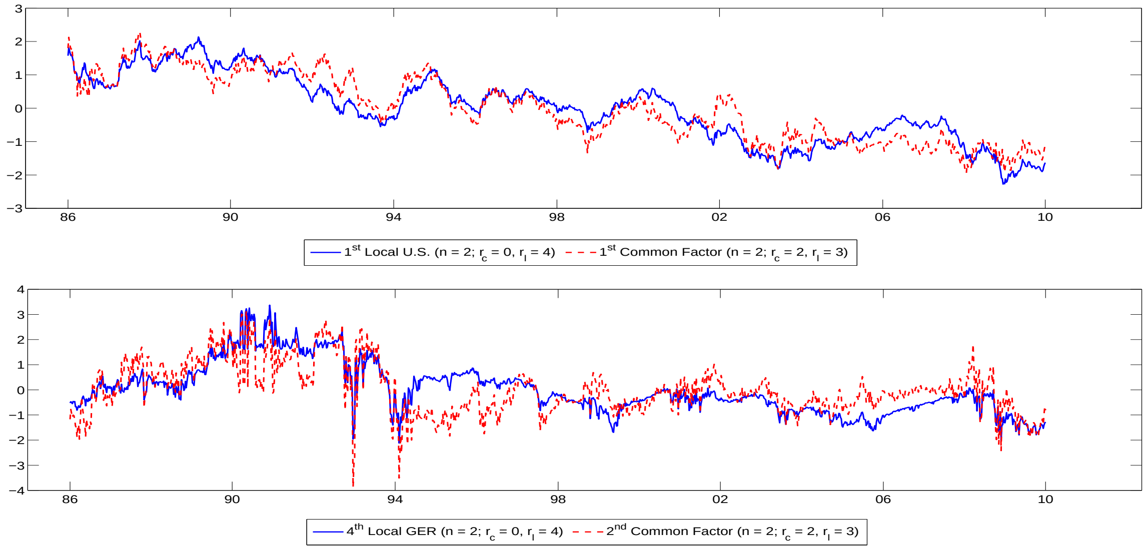
2-country case: U.S.-GER.



(a) Local *U.S.* factors of $\mathcal{M}_2^{0,4}$ and $\mathcal{M}_2^{2,3}$



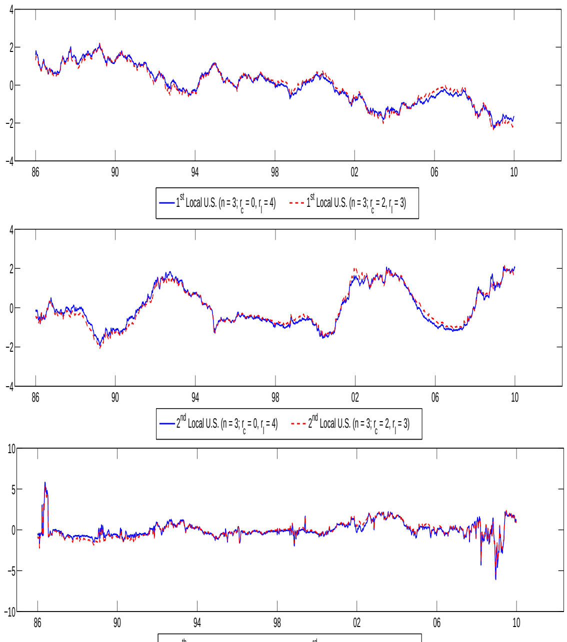
(b) Local *GER* factors of $\mathcal{M}_2^{0,4}$ and $\mathcal{M}_2^{2,3}$



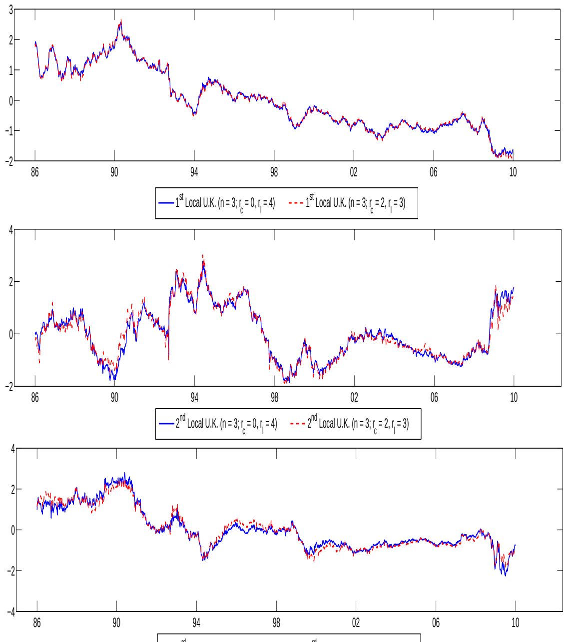
(c) 1^{st} and 2^{nd} common factor of $\mathcal{M}_2^{2,3}$

Figure 3: Smoothed factors in the 2-country *U.S.-GER* case when $(r_c = 0, r_\ell = 4)$ and $(r_c = 2, r_\ell = 3)$.

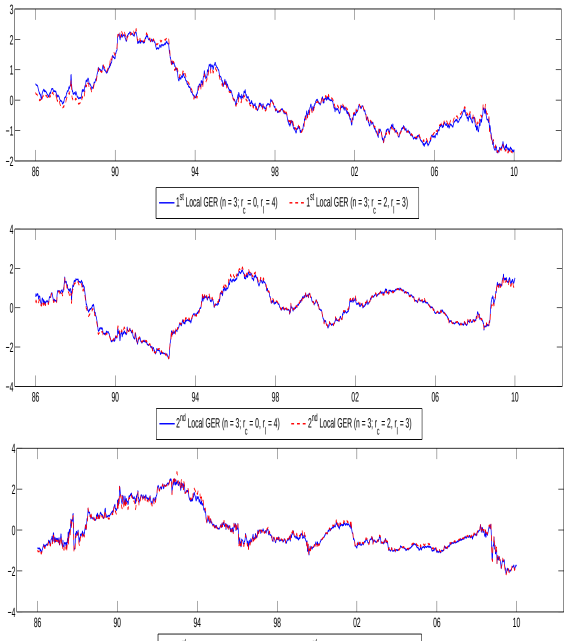
3-country case: U.S.-U.K.-GER.



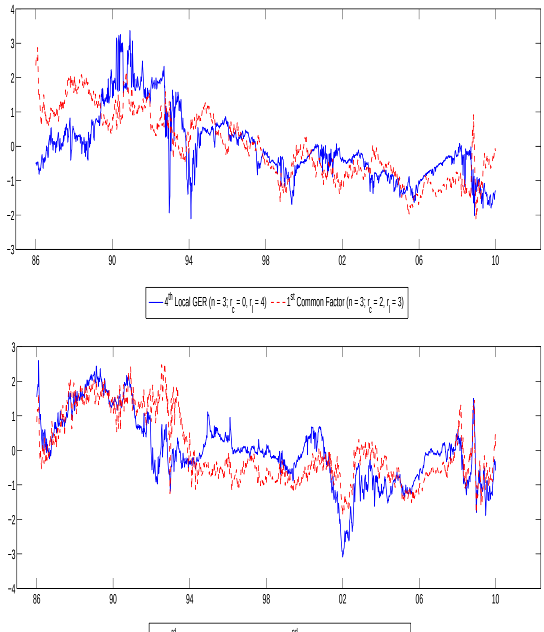
(a) Local *U.S.* factors of $\mathcal{M}_3^{0,4}$ and $\mathcal{M}_3^{2,3}$



(b) Local *U.K.* factors of $\mathcal{M}_3^{0,4}$ and $\mathcal{M}_3^{2,3}$



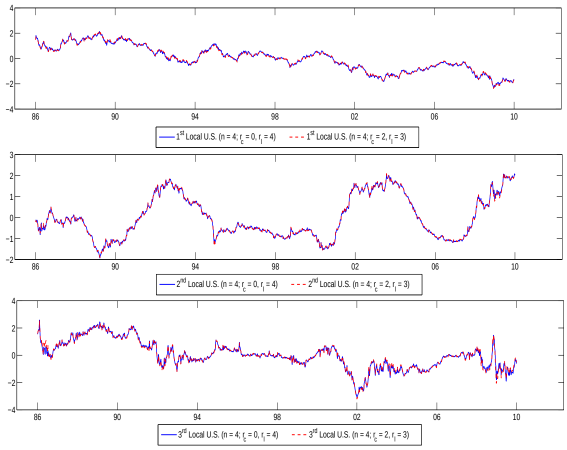
(c) Local *GER* factors of $\mathcal{M}_3^{0,4}$ and $\mathcal{M}_3^{2,3}$



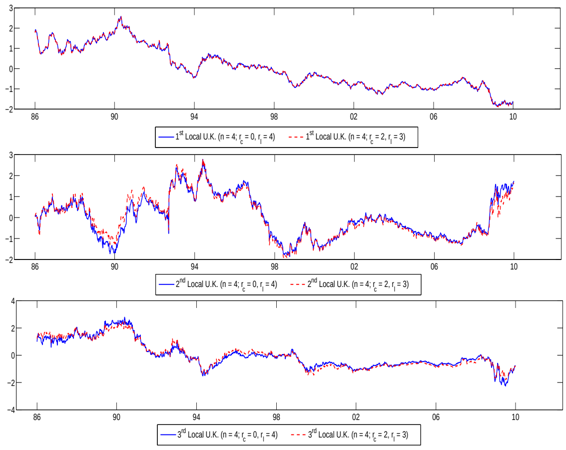
(d) 1^{st} and 2^{nd} common factor of $\mathcal{M}_3^{2,3}$

Figure 4: Smoothed factors in the 3-country case $U.S. - U.K. - GER$ when $(r_c = 0, r_\ell = 4)$ and $(r_c = 2, r_\ell = 3)$.

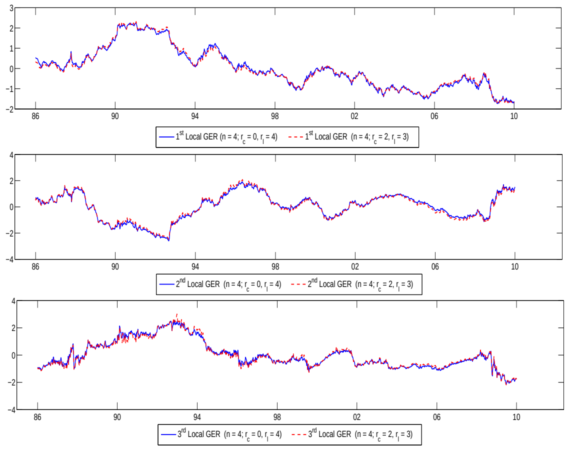
4-country case: U.S.-U.K.-GER-JAP.



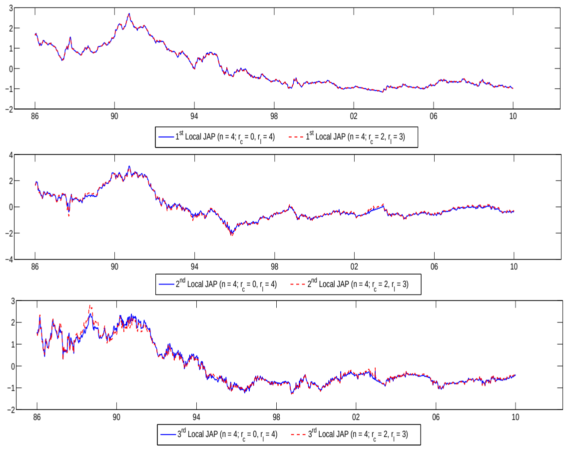
(a) Local *U.S.* factors of $\mathcal{M}_4^{0,4}$ and $\mathcal{M}_4^{2,3}$



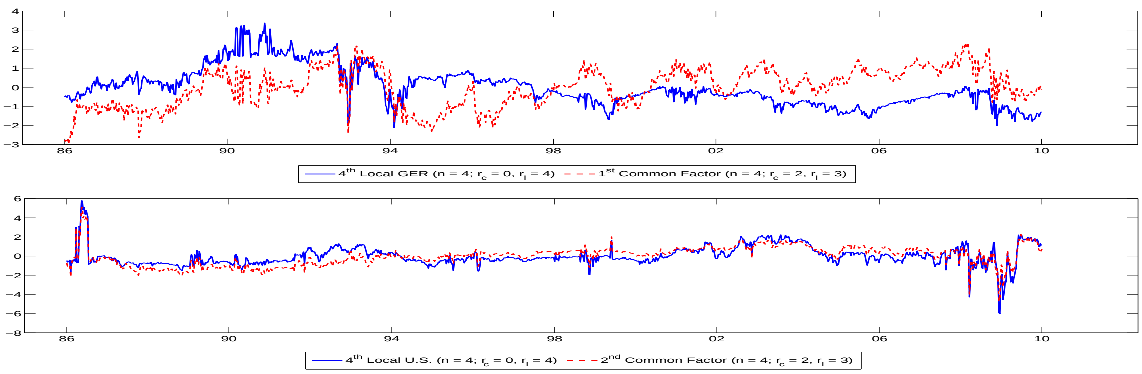
(b) Local *U.K.* factors of $\mathcal{M}_4^{0,4}$ and $\mathcal{M}_4^{2,3}$



(c) Local *GER* factors of $\mathcal{M}_4^{0,4}$ and $\mathcal{M}_4^{2,3}$



(d) Local *JAP* factors of $\mathcal{M}_4^{0,4}$ and $\mathcal{M}_4^{2,3}$



(e) 1^{st} and 2^{nd} common factor of $\mathcal{M}_4^{2,3}$

Figure 5: Smoothed factors in the 4-country case when $(r_c = 0, r_\ell = 4)$ and $(r_c = 2, r_\ell = 3)$.

References

Adrian, T., R. Crump, and E. Moench (2013). Pricing the term structure with linear regressions. *Journal of Financial Economics 110*(01), 110–138.

Ahn, D.-H. (2004). Common factors and local factors: Implications for term structures and exchange rates. *Journal of Financial and Quantitative Analysis 39*(01), 69–102.

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 & Economic Statistics 25*, 91–96.

Anderson, B., P. J. Hammond, and C. A. Ramezani (2010, October). Affine models of the joint dynamics of exchange rates and interest rates. *Journal of Financial and Quantitative Analysis 45*(05), 1341–1365.

Ang, A. and M. Piazzesi (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics 50*(4), 745–787.

Ang, A., M. Piazzesi, and M. Wei (2006). What does the yield curve tell us about gdp growth? *Journal of Econometrics 131*(1-2), 359–403.

Backus, D. K., S. Foresi, and C. I. Telmer (2001). Affine term structure models and the forward premium anomaly. *Journal of Finance 56*(1), 279–304.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*(1), 191–221.

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

Bansal, R. and H. Zhou (2002, October). Term structure of interest rates with regime shifts. *Journal of Finance 57*(5), 1997–2043.

Bork, L., H. Dewachter, and R. Houssa (2009). Identification of macroeconomic factors in large panels.

Cavanaugh, J. E. and R. Shumway (1997). A bootstrap variant of AIC for state-space model selection. *Statistica Sinica 7*.

Cheridito, P., D. Filipovic, and R. L. Kimmel (2007). Market price of risk specifications for affine models: Theory and evidence. *Journal of Financial Economics 83*(1), 123–170.

Connor, G. and R. A. Korajczyk (1993). A test for the number of factors in an approximate factor model. *Journal of Finance 48*(4), 1263–91.

Cox, J. C., J. Ingersoll, Jonathan E, and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica 53*(2), 385–407.

Dahlquist, M. and H. Hasseltoft (2013). International bond risk premia. *Journal of International Economics 90*(1), 17–32.

Dai, Q. and K. Singleton (2003). Term structure dynamics in theory and reality. *Review of Financial Studies 16*(3), 631–678.

Dai, Q. and K. J. Singleton (2000). Specification analysis of affine term structure models. *Journal of Finance 55*(5), 1943–1978.

Dai, Q. and K. J. Singleton (2002). Expectation puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics 63*(3), 415–441.

Diebold, F. X. and C. Li (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics 130*(2), 337–364.

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*(2), 351–363.

Doz, C., D. Giannone, and L. Reichlin (2011). A two-step estimator for large approximate dynamic factor models based on kalman filtering. *Journal of Econometrics 164*(1), 188–205.

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

Duarte, J. (2004). Evaluating an alternative risk preference in affine term structure models. *Review of Financial Studies 17*(2), 379–404.

Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models. *Journal of Finance 57*(1), 405–443.

Duffee, G. R. (2011). Information in (and not in) the term structure. *Review of Financial Studies 24*(9), 2895–2934.

Duffie, D. and R. Kan (1996). A yield-factor model of interest rates. *Mathematical Finance 6*(4), 379–406.

Egorov, A. V., H. Li, and D. Ng (2011). A tale of two yield curves: Modeling the joint term structure of dollar and euro interest rates. *Journal of Econometrics 162*(1), 55–70.

Engle, R. and M. Watson (1981). A one factor multivariate time series model of metropolitan wage rates. *Journal of the American Statistical Association 76*, 774–780.

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*(4), 540–554.

Frachot, A. (1995). Factor models of domestic and foreign interest rates with stochastic volatilities. *Mathematical Finance 5*(2), 167–185.

Gourieroux, C., A. Monfort, and R. Sufana (2010). International money and stock market contingent claims. *Journal of International Money and Finance 29*(8), 1727–1751.

Greenaway-McGrevy, R., C. Han, and D. Sul (2012). Estimating the number of common factors in serially dependent approximate factor models. *Economics Letters 116*(3), 531–534.

Gurkaynak, R. S., B. Sack, and J. H. Wright (2007). The U.S. treasury yield curve: 1961 to the present. *Journal of Monetary Economics 54*(8), 2291–2304.

Jotikasthira, C., A. Le, and C. Lundbland (2010). Why do term structure in different currencies comove? *University of North Carolina Mimeo*.

Jungbacker, B. and S. J. Koopman (2008). Likelihood-based analysis for dynamic factor models. *Tinbergen Institute Research Paper 08-007/4*.

Koijen, R., H. Lustig, and S. van Nieuwerburgh (2012). The cross-section and time-series of stock and bond returns. *C.E.P.R. Discussion Papers n. 9024*.

Leippold, M. and L. Wu (2007). Design and estimation of quadratic term structure models. *Review of Finance 11*, 167–207.

Linhart, H. and W. Zucchini (1986). *Model Selection*. Wiley.

Litterman, R. and J. Scheinkman (1991). Common factors affecting bond returns. *Journal of Fixed Income 1*, 54–61.

Monfort, A. and F. Pegoraro (2007). Switching varma term structure models. *Journal of Financial Econometrics 5*(1), 105–153.

Monfort, A., J.-P. Renne, R. Rüffer, and G. Vitale (2003). Is economic activity in the g7 synchronized? common shocks versus spillover effects. *C.E.P.R. Discussion Papers* (4119).

Mosburger, G. and P. Schneider (2005). Modelling international bond markets with affine term structure models. *University of Vienna, Working paper*.

Nelson, C. R. and A. F. Siegel (1987). Parsimonious modeling of yield curves. *Journal of Business 60*(4), 473–89.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–08.

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

Patton, A., D. Politis, and H. White (2009). Correction to automatic block-length selection for the dependent bootstrap by d. politis and h. white. *Econometric Reviews 28*(4), 372–375.

Pegoraro, F., A. F. Siegel, and L. Tiozzo Pezzoli (2012). International yield curves and principal components selection techniques: An empirical assessment. *Working Paper, University of Washington Business School*.

Perignon, C., D. R. Smith, and C. Villa (2007). Why common factors in international bond returns are not so common. *Journal of International Money and Finance 26*(2), 284–304.

Politis, D. and H. White (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Reviews 23*(1), 53–70.

Politis, D. N. and J. Romano (1994). The stationary bootstrap. *Journal of the American Statistical association 89*(428), 1303–1313.

Quah, D. and T. Sargent (1993). *A dynamic index model for large cross section.* In: Stock, James, Watson, Mark (Eds.), Business Cycle, Indicators and Forecasting. University of Chicago Press, pp. 285-310.

Stock, J. H. and M. W. Watson (1991). Implications of dynamic factor models for var analysis. *NBER Working paper n. 11467*.

Stoffer, D. and K. Wall (1991). Bootstrapping state-space models: gaussian maximun likelihood estimation and the kalman filter. *Journal of the american statistical association 86*(416), 1024–1033.

Tang, H. and Y. Xia (2007). An international examination of affine term structure models and the expectations hypothesis. *Journal of Financial and Quantitative Analysis 42*(01), 41–80.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics 5*(2), 177–188.

Documents de Travail

480. E. Mengus, "Honoring Sovereign Debt or Bailing Out Domestic Residents: A Theory of Internal Costs of Default," March 2014

481. C. Labonne and G. Lamé, "Credit Growth and Bank Capital Requirements: Binding or Not?," March 2014

482. S. Gilchrist and B. Mojon, "Credit Risk in the Euro area," April 2014

483. G. Vuillemey and R. Breton, "Endogenous Derivative Networks," April 2014

484. G. Cette, R. Lecat and A-O-A. Jiddou, "How do firms adjust production factors to the cycle?," April 2014

485. M. Bussière, B. Camara, F.-D. Castellani, V. Potier and J. Schmidt, " Shock Transmission Through International Banks - Evidence From France," May 2014

486. A. Monfort, J-P. Renne and G. Roussellet, "A Quadratic Kalman Filter," May 2014

487. P. Benczúr, G. Kátay, Á. Kiss and O. M. Rácz, "Income Taxation, Transfers and Labour Supply at the Extensive Margin," June 2014

488. R. Jimborean and A. Kelber, "Foreign direct investment drivers and growth in Central and Eastern Europe in the aftermath of the 2007 global financial crisis," June 2014

489. F. Pegoraro, A. F. Siegel and L. Tiozzo 'Pezzoli, "International Yield Curves and Principal Components Selection Techniques: An Empirical Assessment," June 2014

490. F. Pegoraro, A. F. Siegel and L. Tiozzo 'Pezzoli, "Specification Analysis of International Treasury Yield Curve Factors," June 2014

Pour accéder à la liste complète des Documents de Travail publiés par la Banque de France veuillez consulter le site : www.banque-france.fr

For a complete list of Working Papers published by the Banque de France, please visit the website: www.banque-france.fr

Pour tous commentaires ou demandes sur les Documents de Travail, contacter la bibliothèque de la Direction Générale des Études et des Relations Internationales à l'adresse suivante :

For any comment or enquiries on the Working Papers, contact the library of the Directorate General Economics and International Relations at the following address :

BANQUE DE FRANCE
49- 1404 Labolog
75049 Paris Cedex 01
tél : 0033 (0)1 42 97 77 24 ou 01 42 92 63 40 ou 48 90 ou 69 81
email : 1404-ut@banque-france.fr