Factoring in the Micro: a Transaction-level Dynamic Factor Approach to the Decomposition of Export Volatility Supplementary material

ANGELO CUZZOLA*[1], MATTEO BARIGOZZI[2], MARCO GRAZZI[3], and DANIELE MOSCHELLA[1]

[1]Institute of Economics, Scuola Superiore Sant'Anna

[2]Department of Economics, Università di Bologna

[3]Department of Economic Policy, Università Cattolica del Sacro Cuore

July 24, 2024

*Corresponding author. Postal address: Institute of Economics, Scuola Superiore Sant'Anna, Piazza Martiri della Libertà 33, 56127, Pisa, Italy, E-mail: angelo.cuzzola@santannapisa.it

Contents

Appendix A, Estimation strategy: technical details

In the appendices, we present the technical details of the estimation procedure for a dynamic factor model with an imposed block structure and missing data. For the technical analysis we rewrite the model adopting the vectorized notation:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{\Lambda}^G \mathbf{F}_t^G + \sum_{d=1}^{D} \mathbf{\Lambda}^d \mathbf{F}_t^d + \boldsymbol{\xi}_t \\ \mathbf{F}_t^G &= \mathbf{A}^G(L)\, \mathbf{F}_t^G + \mathbf{u}_t^G \\ \mathbf{F}_t^d &= \mathbf{A}^d(L)\, \mathbf{F}_t^d + \mathbf{u}_t^d \qquad d = 1, \dots, D \end{aligned} \tag{1}$$

The model is analyzed allowing for an arbitrary number of common factors (K) and an arbitrary number of factors per block (K_d with $d = 1, \dots, D$)[1]. We drop here the multi-index notation used in the paper and rely on two indices only, running over the number of firm-destination pairs, $i = 1, \dots, N$ and over the time steps $t = 1, \dots, T$. Each block includes n_d flows and $N = \sum_{d=1}^{D} n_d$. Defining the $(r^G \times 1)$ global factors vector $\mathbf{F}_t^G$ and the D destination specific $(r^d \times 1)$ factor vectors $\mathbf{F}^d$. $\mathbf{\Lambda^G}$ and $\mathbf{\Lambda^d}$ are loading matrices of size $(N \times r^G)$ and $(N \times r^d)$. The dynamics of the factor is encoded in the matrix polynomials $\mathbf{A}^G(L)$ and $\mathbf{A}^d(L)$ of order p^G and p^d respectively. For the sake of the synthesis, throughout we limit the exposition to the case $p^G = p^d = 1$ and no serial correlation of the idiosyncratic component.

Estimation algorithm

Given the formulation (1) and assuming gaussianity of the idiosyncratic components and the innovations $\boldsymbol{\xi}_t \sim \mathcal{N}(0, \mathbf{\Sigma}^{\boldsymbol{\xi}})$, $\boldsymbol{u}_t^c \sim \mathcal{N}(0, \mathbf{\Sigma}^{\boldsymbol{u}^{\boldsymbol{G}}})$, $\boldsymbol{u}_t^d \sim \mathcal{N}(0, \mathbf{\Sigma}^{\boldsymbol{u}^{\boldsymbol{d}}})$, together with their mutual independence, the log-likelihood given the observed series and the latent

[1] The estimation of the model (1) is obtained taking $K = K_d = 1\ \forall d$, thus restricting the analysis to one global factor and D destination-specific factors, one per each destination. Notice that while preserving the notation of the main text the index d can generally run over any partition of the cross-section into blocks. Thus in the following we will more generally refer to blocks.

factors is:

$$
\begin{aligned}
l(Y,F;\theta) \simeq & -\frac{T-1}{2}\log|\boldsymbol{\Sigma}^{\boldsymbol{u}^G}| - \frac{1}{2}\operatorname{tr}\left[\left(\boldsymbol{\Sigma}^{\boldsymbol{u}^G}\right)^{-1}\sum_{t=2}^{T}\left(\mathbf{F}_t^G - \mathbf{A}^G\mathbf{F}_{t-1}^G\right)\left(\mathbf{F}_t^G - \mathbf{A}^G\mathbf{F}_{t-1}^G\right)'\right] \\
& -\sum_{d=1}^{D}\left[\frac{T-1}{2}\log|\boldsymbol{\Sigma}^{\boldsymbol{u}^d}| + \frac{1}{2}\operatorname{tr}\left[\left(\boldsymbol{\Sigma}^{\boldsymbol{u}^d}\right)^{-1}\sum_{t=2}^{T}\left(\mathbf{F}_t^d - \mathbf{A}^d\mathbf{F}_{t-1}^d\right)\left(\mathbf{F}_t^d - \mathbf{A}^d\mathbf{F}_{t-1}^d\right)'\right]\right] \\
& -\frac{T-1}{2}\log|\boldsymbol{\Sigma}^{\boldsymbol{\xi}}| - \frac{1}{2}\operatorname{tr}\left[\left(\boldsymbol{\Sigma}^{\boldsymbol{\xi}}\right)^{-1}\sum_{t=1}^{T}\left(\mathbf{y}_t - \boldsymbol{\Lambda}\mathbf{F}_t\right)\left(\mathbf{y}_t - \boldsymbol{\Lambda}\mathbf{F}_t\right)'\right]
\end{aligned} \tag{2}
$$

Where $\boldsymbol{\Lambda}$ is the $(N \times (r^G + \sum_d r^d))$ composed loading matrix: $\boldsymbol{\Lambda} = \left(\boldsymbol{\Lambda}^G\,\boldsymbol{\Lambda}^{\mathbf{1}}\;\cdots\;\boldsymbol{\Lambda}^{\boldsymbol{D}}\right)$ related to the composed factor vector $\mathbf{F}_t = (\mathbf{F}_t^G\,\mathbf{F}_t^1\;\cdots\;\mathbf{F}_t^D)$. We recall that $[\boldsymbol{\Lambda}^{\boldsymbol{d}}]_{is} = 0$ if the index i do not belong to the block of series relative to destination d. Then, the two steps procedure is started updating sequentially i) the factors' estimates given the model's parameters (E-step), ii) the parameters estimates given the estimates of factors (M-step). This defines a sequence of increasing log-likelihood values

$$
l\left(Y, F^{(0)}, \theta^{(0)}\right) \to l\left(Y, F^{(0)}, \theta^{(1)}\right) \to l\left(Y, F^{(1)}, \theta^{(1)}\right)
$$

that needs a proper initialization and stops when an appropriate convergence condition (we adopt a standard in the literature see Bańbura and Modugno, 2014; Barigozzi and Luciani, 2019). We define the k-th increment as:

$$
\Delta l_k = \frac{|l\left(Y, F^{(k+1)}, \theta^{(k+1)}\right) - l\left(Y, F^{(k)}, \theta^{(k)}\right)|}{\left(|l\left(Y, F^{(k+1)}, \theta^{(k+1)}\right)| + |l\left(Y, F^{(k)}, \theta^{(k)}\right)|\right)/2} \tag{3}
$$

and stop the algorithm at $k = \overline{k}$ such that $\Delta l_{\overline{k}} \leq \varepsilon$, where ε is a predefined tolerance threshold. Throughout this paper, for all the estimation runs we settle $\varepsilon = 10^{-4}$. This is sufficient to get an estimation before the maximum iterations limit is reached $k_{max} = 100$.

Initialization algorithm.

The procedure is initialized computing the sequential least square estimator proposed by Breitung and Eickmeier (2014) on a 'complete' matrix of data and is composed of the following steps:

1. We fill the missing values of the original dataset with series medians, then we smooth the outcome taking the moving averages of the series so that we can work with the filled matrix $\overline{Y}$.

2. As proposed by, we apply the CCA estimator to initialize the global factors. Within a block d, this consists in estimating by PC $r_{d*} = r^G + r^d$ yielding the r_{d*}-vectors of factors $\mathbf{F^*}_{d,t}$. Then we search the global components among the r^G maximally correlated common components between the blocks: cycling all the couples of blocks v, w, we apply a Canonical Correlation Analysis to determine the linear combinations that maximizes the correlation between the quantities[2]: $\tau_v' \mathbf{F}^*_{v,t}$, $\tau_w' \mathbf{F}^*_{w,t}$. Thus, we have a first estimator of the global factors $\widehat{\mathbf{F}}_t^G = (\overline{\tau}_v^{1,\prime}\, \mathbf{F}^*_{v,t}, \ldots, \overline{\tau}_v^{r^G,\prime}\, \mathbf{F}^*_{v,t})^{\prime}$.

3. We solve a least square problem to compute the block specific factors by means of principal components. In other words, principal components is applied to the residuals of the regression $y_t \sim \mathbf{\Lambda}^G \widehat{\mathbf{F}}_t^G$.

4. A sequential least square estimator is applied starting from the estimates of the global and local factors of the previous step. Also in this case we rely on a sequential procedure iterating over two main steps. At step k, given the factor

[2] The number of possible pairs for the CCA is $D^2(D-1)/2$. The problem is solved for each pairs of blocks and then the pair that maximises the CCA is chosen.

estimates ${\mathbf{F}^{\mathbf{G}}}_t^{(k-1)}, {\mathbf{F}^{\mathbf{d}}}_t^{(k-1)}$, via the block-level regressions[3]:

$$\iota_d(\mathbf{y}_t) = \iota_d(\mathbf{\Lambda}^G)\,\widehat{\mathbf{F}}_t^{G,(k-1)} + \iota_d(\mathbf{\Lambda}^d)\,\widehat{\mathbf{F}}_t^{d,(k-1)} + \boldsymbol{\epsilon}_t^d. \tag{4}$$

one obtains the estimates of the relative factor loadings so that the estimated block-matrix can be composed $\widehat{\mathbf{\Lambda}}^{(k)}$. Then the $(k+1)$-th update of the estimators of the factors are obtained from the twin least square regression of $\mathbf{y}_t$ on $\mathbf{\Lambda}^{(k)}$.

The iteration procedure stops when a convergence condition analogous to (3) is verified. The last estimates of factors and parameters are used to start the EM algorithm.

E-step. From model parameters to factors: Kalman Filter and Smoother algorithms

Throughout, we work with the quantities $\mathbf{F}_t = (\mathbf{F}_t^G\,\mathbf{F}_t^1 \cdots \mathbf{F}_t^D)$ and $\mathbf{A} = \text{blkdiag}(\mathbf{A}^G\,\mathbf{A}^1 \cdots \mathbf{A}^D)$. At iteration k, KF and KS are two sequential procedures on the time dimension $t = 1, \ldots, T$. These are used to compute the KF estimator, $\mathbf{F}_{t|t} = \text{Proj}_\theta[\mathbf{F}_t|\mathbf{y}_t]$ and the associated mean squared error (MSE), $\mathbf{P}_{t|s} = \mathbb{E}_\theta[(\mathbf{F}_t - \mathbf{F}_{t|s})(\mathbf{F}_t - \mathbf{F}_{t|s})'|\mathbf{y}_s]$. When $s = t$ and $s = T$ we obtain respectively the KF MSE and KS MSE. First we have to state the initial conditions $\mathbf{F}_{0|0}$ and $\mathbf{P}_{0|0}$. At the very first step we set the same conditions as Barigozzi and Luciani (2019), for the following EM iterations one can settle $\mathbf{P}_{0|0}^{(k)} = \mathbf{P}_{1|T}^{(k-1)}$) (see Durbin and Koopman, 2012, for details). From now on, to deal with missing values, parameters are restricted in each time step to the portion with available information. Hence, the “NA” index or suffix denotes the matrix/vector cleaned of rows, columns or elements corresponding to NA entries at time t. Moreover to keep the notation as clean as possible we omit the step-index for the quantities associated with the factor.

---

[3] Here the operator $\iota_d(\cdot)$ is applied to an object with N rows to restricting to the n_d rows relative to block d.

We have the filtering sequential equations:

$$
\begin{aligned}
\mathbf{F}_{t|t-1} &= \mathbf{A}^{(k)}\mathbf{F}_{t-1|t-1}\ ; \qquad \mathbf{P}_{t|t-1} = \mathbf{A}^{(k)}\mathbf{P}_{t-1|t-1}\mathbf{A}^{(k)\prime} + \mathbf{\Sigma}^{\boldsymbol{u}(k)} \\
\mathbf{F}_{t|t} &= \mathbf{F}_{t|t-1} + \mathbf{P}_{t|t-1}\,\mathbf{\Lambda}^{(k)}_{\mathrm{NA},t}{}'\mathbf{G}_t^{-1}\left(\mathbf{y}_t^{\mathrm{NA}} - \mathbf{\Lambda}^{(k)}_{\mathrm{NA},t}\mathbf{F}_{t|t-1}\right) \\
\mathbf{P}_{t|t} &= \mathbf{P}_{t|t-1} + \mathbf{P}_{t|t-1}\,\mathbf{\Lambda}^{(k)}_{\mathrm{NA},t}{}'\,\mathbf{Q}_t^{-1}\,\mathbf{\Lambda}^{(k)}_{\mathrm{NA},t}\mathbf{P}'_{t|t-1}
\end{aligned}
\tag{5}
$$

where $\boldsymbol{Q}_t^{(k)} = \mathbf{\Lambda}^{(k)}_{\mathrm{NA},t}\mathbf{P}_{t|t-1}\mathbf{\Lambda}^{(k)}_{\mathrm{NA},t}{}' + \mathbf{\Sigma}^{\boldsymbol{\xi}(k)}_{\mathrm{NA},t}$. With the estimates of $\mathbf{F}_{t|t}$ and $\mathbf{P}_{t|t}$ we can initialize the KS to obtain the smoothed estimates $\mathbf{F}_{t|T}$ and ($\mathbf{P}_{t|T}$ for $t = T, \ldots, 1$ via the inverse recursion, starting from $\mathbf{F}_{T|T}$ and $\mathbf{P}_{T|T}$:

$$
\begin{aligned}
\mathbf{F}_{t|T} &= \mathbf{F}_{t|t} + \mathbf{P}_{t|t}\,\mathbf{A}^{(k)'}\,\mathbf{P}^{-1}_{t+1|t}\left(\mathbf{F}_{t+1|T} - \mathbf{F}_{t+1|t}\right) \\
\mathbf{P}_{t|T} &= \mathbf{P}_{t|t} + \mathbf{P}_{t|t}\,\mathbf{A}'\,\mathbf{P}^{-1}_{t+1|t}\,\left(\mathbf{P}_{t+1|T} - \mathbf{P}_{t+1|t}\right)\left(\mathbf{P}_{t+1|t}\,\mathbf{A}^{(k)'}\,\mathbf{P}^{-1}_{t|t}\right)'
\end{aligned}
\tag{6}
$$

To set up the M-step we will use the KS estimates only, in order to keep the notation as concise as possible we will denote $\widetilde{\mathbf{F}}_t = \mathbf{F}_{t|T}$, $\widetilde{\mathbf{P}}_t = \mathbf{P}_{t|T}$ and

$$
\widetilde{\mathbf{P}}_{-1,t} = \mathbf{P}_{t|t}\left(\mathbf{P}_{t|t}\mathbf{A}^{(k)'}\mathbf{P}^{-1}_{t|t}\right)' + \left(\mathbf{P}_{t|t}\mathbf{A}^{(k)'}\mathbf{P}^{-1}_{t+1|t}\right)\left(\mathbf{P}_{t|T} - \mathbf{A}^{(k)}\mathbf{P}_{t|t}\right)\left(\mathbf{P}_{t|t}\mathbf{A}^{(k)'}\mathbf{P}^{-1}_{t|t}\right) \tag{7}
$$

The latter three quantities are then used to compute[4]:

$$
\mathbb{E}_{\theta^{(k)}}\left[\mathbf{F}_t\mathbf{F}_{t-1}{}'|\Omega_T\right] = \sum_{t=2}^{T}\widetilde{\boldsymbol{F}}_t\widetilde{\boldsymbol{F}}'_{t-1} + \widetilde{\boldsymbol{P}}_{-1,t}\ ; \quad \mathbb{E}_{\theta^{(k)}}\left[\mathbf{F}_t\mathbf{F}_{t-1}{}'|\Omega_T\right] = \sum_{t=1}^{T}\widetilde{\boldsymbol{F}}_t\widetilde{\boldsymbol{F}}'_{t-1} + \widetilde{\boldsymbol{P}}_{-1,t} \tag{8}
$$

M-step. From latent factors to model parameters

Given the initial values, the EM algorithm is started. The proposed solution consist in the maximixation of the expectation of the loglikelihood, given an ansatz of the

[4] In the following we will equally refer to the quantities restricted to the space of global and local factors: $\widetilde{\boldsymbol{F}_t^G}, \widetilde{\boldsymbol{F}_t^d}, \widetilde{\boldsymbol{P}^G}_t, \widetilde{\boldsymbol{P}^d}_t, \widetilde{\boldsymbol{P}^G}_{-1,t}, \widetilde{\boldsymbol{P}^d}_{-1,t}$

parameters[5]. In practice, $\theta^{(k+1)}$ are the solutions to the system of first order conditions $\frac{\partial}{\partial\theta}\mathbb{E}_{\theta^{(k)}}\left[l(Y, F^{(k)}, \theta)|\Omega_T\right] = 0$ where Ω_T denotes the available information that in our application is constrained by the presence of missing values in the observed data. From the explicit form of l we get[6]:

$$\widehat{\boldsymbol{A}}^{G,(k+1)} = \left(\sum_{t=2}^{T} \widetilde{\boldsymbol{F}_t^G}\widetilde{\boldsymbol{F}^G}'_{t-1} + \widetilde{\boldsymbol{P}^G}_{-1,t}\right)\left(\sum_{t=2}^{T} \widetilde{\boldsymbol{F}^G}_{t-1}\widetilde{\boldsymbol{F}^G}'_{t-1} + \widetilde{\boldsymbol{P}^G}_t\right)^{-1} \quad (9)$$

$$\widehat{\boldsymbol{Q}}^{G,(k+1)} = \frac{1}{T}\left(\sum_{t=2}^{T} \widetilde{\boldsymbol{F}^G}_t\widetilde{\boldsymbol{F}^G}'_t + \widetilde{\boldsymbol{P}^G}_t\right) - \frac{\widehat{\boldsymbol{A}}^{G,(k+1)}}{T}\left(\sum_{t=1}^{T} \widetilde{\boldsymbol{F}_t^G}\widetilde{\boldsymbol{F}^G}'_{t-1} + \widetilde{\boldsymbol{P}^G}_{-1,t}\right)' \quad (10)$$

The derivation of the updated estimates for the loadings matrix is quite a complex task, mainly because the matrix with the incomplete observations enters the solution of the first-order condition. One way to see this step is to calculate $\boldsymbol{\Lambda}$ as the NA-corrected OLS solutions of the block-by-block regressions (see Bańbura and Modugno, 2014, pag. 138, eq. (11))[7][8]

$$\mathbf{y}_t^b = \iota_b(\mathbf{y}_t) = \iota_b(\boldsymbol{\Lambda}^G)\widetilde{\mathbf{F}^G}_t + \iota(\boldsymbol{\Lambda}^d)\widetilde{\mathbf{F}}_t^d + \mathbf{v}_t \qquad d = 1, \ldots, D \quad (11)$$

$$\text{vec}\left(\iota_d(\boldsymbol{\Lambda}^G)|\iota_d(\boldsymbol{\Lambda}^d))\right) = \left(\sum_{t=1}^{T} \widetilde{\boldsymbol{F}}_t^{Gd}\widetilde{\boldsymbol{F}}_t^{Gd\,\prime} \otimes \text{Ind}_t^{\text{NA}}\right)^{-1} \text{vec}\left(\sum_{t=1}^{T} \mathbf{y}_t^{d,\text{NA}}\widetilde{\boldsymbol{F}}_t'\right) \quad (12)$$

where the matrix $\widetilde{\boldsymbol{F}_t^{Gd}} = (\widetilde{\boldsymbol{F}_t^G}\widetilde{\boldsymbol{F}_t^d})$. This step concludes with the covariance matrix of idiosyncratic terms that is estimated only in its diagonal elements (in line with the

[5] Throughout the section we give some fundamental formulas without deriving it. The explicit derivation are generalization of Bańbura and Modugno (2014) and Barigozzi and Luciani (2019) to the case of block-DFMs.

[6] We give the formulas only for the common factors, yet those for the local factors are equivalent.

[7] In line with the notation introduced above, Ind_t^{NA} denotes a diagonal matrix with ones when the corresponding element is available in the cross section t and a zero when it is not.

[8] Here again we denote $\iota_d(\boldsymbol{\Lambda}^d)$ as $\boldsymbol{\Lambda}^d$

eq.(12) at p. 138 of Bańbura and Modugno (2014)):

$$\begin{aligned}\operatorname{diag}\widehat{\boldsymbol{\Sigma}^{\boldsymbol{\xi}}}^{(k+1)} =& \frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{y}_t^{\mathrm{NA}} - \mathrm{Ind}_t^{\mathrm{NA}}\widehat{\boldsymbol{\Lambda}}^{(k+1)}\,\widetilde{\boldsymbol{F}}_t\right)\left(\mathbf{y}_t^{\mathrm{NA}} - \mathrm{Ind}_t^{\mathrm{NA}}\widehat{\boldsymbol{\Lambda}}^{(k+1)}\,\widetilde{\boldsymbol{F}}_t\right)' + \\ & \mathrm{Ind}_t^{\mathrm{NA}}\,\widehat{\boldsymbol{\Lambda}}^{(k+1)}\,\widetilde{\boldsymbol{P}}_t\,\widehat{\boldsymbol{\Lambda}}^{(k+1)\prime}\,\mathrm{Ind}_t^{\mathrm{NA}} + \left(\boldsymbol{I}_n - \mathrm{Ind}_t^{\mathrm{NA}}\right)\operatorname{diag}\widehat{\boldsymbol{\Sigma}^{\boldsymbol{\xi}}}^{(k)}\left(\boldsymbol{I}_n - \mathrm{Ind}_t^{\mathrm{NA}}\right)\end{aligned} \tag{13}$$

Appendix B, Estimation strategy: finite sample properties from Monte Carlo simulations

The theoretical properties of the QML estimator for DFMs have been thoroughly explored in the seminal work of Doz et al. (2012) and in more recent papers (see Barigozzi and Luciani, 2019). Those studies openly state the possibility to model a block structure for DFMs, but devote less attention to the implications of modelling a block structure for datasets with arbitrary patterns of missing values, on the lines Bańbura and Modugno (2014) for ordinary DFMs. In this appendix, we answer some of the natural questions regarding the estimation methodology applied in this paper and assess its performances varying the structural conditions of the data generating process (also DGP in the following). Throughout, the data is simulated according to (1) with a set of benchmark specifications, summarized in these few points:

- We limit the evolution of the r^G global factors to a VAR process of order one ($p^G = 1$). The matrix ruling the process is generated as in Barigozzi and Luciani (2019) according to the expression $\mathbf{A} = \kappa\tilde{\mathbf{A}}\left(||\tilde{\mathbf{A}}||\right)^{-1}$, where $[\tilde{\mathbf{A}}]_{jj} \sim U[0.5, 0.8]$, while $[\tilde{\mathbf{A}}]_{jk} \sim U[0, 0.3]$, and $\kappa = 0.5$[9]. The dynamics of the r^b local factors is simulated with the same characteristics.

- We simulate the data taking either homogeneous or heterogeneous block dimensions ($\{n^d\}$ for $d = 1 \ldots, D$). With the former specification we intend that all the n^d are equal when possible, i.e. when $n = 0 \pmod{n^d}$, or differ of few units.

[9] Here $||\cdot||$ denotes the Frobenius norm.

Heterogeneous blocks are generated selecting a D-tuple at random among all the $(n^1, \dots, n^D)$ s.t. $\sum_d n^d = n$.

- Both factor loadings relative to the global and the local factors are simulated from a normal distribution. This means that $\left[\mathbf{\Lambda}^G\right]_{ij} \overset{iid}{\sim} \mathcal{N}(0,1)$, while $\left[\mathbf{\Lambda}^d\right]_{ij} \overset{iid}{\sim} \mathcal{N}(0,1)$ if $i \in I_d$ and $\left[\mathbf{\Lambda}^d\right]_{ij} = 0$ if $i \notin I_d$.

- The innovations to global and block-specific factors are simulated with normal, Student-t, Laplace or Asymmetric Laplace distributions[10]. Namely, $\mathbf{u}_t^G \overset{iid}{\sim} \mathcal{N}\left(\mathbf{0}_{r^G}, \mathbf{I}_{r^G}\right)$ and $\mathbf{u}_t^d \overset{iid}{\sim} \mathcal{N}\left(\mathbf{0}_{r^d}, \mathbf{I}_{r^d}\right), \forall d$, or $\mathbf{u}_t^G \overset{iid}{\sim} t_3\left(\mathbf{0}_{r^G}, \mathbf{I}_{r^G}\right)$ and $\mathbf{u}_t^d \overset{iid}{\sim} t_3\left(\mathbf{0}_{r^d}, \mathbf{I}_{r^d}\right) \forall d$, or $\mathbf{u}_t^G \overset{iid}{\sim} \mathcal{L}\left(\mathbf{0}_{r^G}, \mathbf{I}_{r^G}\right)$ and $\mathbf{u}_t^d \overset{iid}{\sim} \mathcal{L}\left(\mathbf{0}_{r^d}, \mathbf{I}_{r^d}\right), \forall d$, or $\mathbf{u}_t^G \overset{iid}{\sim} \mathcal{AL}\left(\mathbf{0}_{r^G}, \mathbf{I}_{r^G}, \kappa\right)$ and $\mathbf{u}_t^d \overset{iid}{\sim} \mathcal{AL}\left(\mathbf{0}_{r^d}, \mathbf{I}_{r^d}, \kappa\right), \forall d$.

- The idiosyncratic components can be cross-correlated but the autocorrelation is not modelled. To measure the cross correlation the idiosyncratic components distributes as $\boldsymbol{\xi}_t \overset{iid}{\sim} \mathcal{N}\left(\mathbf{0}_n, \mathbf{\Sigma}^{\boldsymbol{\xi}}\right)$ or $\overset{iid}{\sim} t_3\left(\mathbf{0}_n, \mathbf{\Sigma}^{\boldsymbol{\xi}}\right)$ (or $\boldsymbol{\xi}_t \overset{iid}{\sim} \mathcal{L}\left(\mathbf{0}_n, \mathbf{\Sigma}^{\boldsymbol{\xi}}\right)$, or $\boldsymbol{\xi}_t \overset{iid}{\sim} \mathcal{AL}\left(\mathbf{0}_n, \mathbf{\Sigma}^{\boldsymbol{\xi}}, \kappa\right)$), where the elements of the covariance matrix are such that $\left[\mathbf{\Sigma}^{\boldsymbol{\xi}}\right]_{ii} \sim U[0.5, 1.5]$ and $\left[\mathbf{\Sigma}^{\boldsymbol{\xi}}\right]_{ii} \sim \tau^{|i-j|}$, defining a Toeplitz matrix where the magnitude of the paired correlations is ruled by τ.

- The noise-to-signal ratio is controlled by scaling the common component $\widetilde{\boldsymbol{\chi}}_{i,t} = \frac{1}{\underline{\omega_i}} \boldsymbol{\chi}_{i,t} \sqrt{\frac{\mathrm{Var}(\boldsymbol{\xi}_{i,t})}{\mathrm{Var}(\boldsymbol{\chi}_{i,t})}}$ where $\underline{\omega_i}$ is drawn from a uniform distribution centered around the parameter ω $(\mathcal{U}([\omega - 0.2, \omega + 0.2]))$.

- After constructing the data matrix Y we remove at random a fraction δ (in $[0, 1)$) of elements.

For each experiment we consider a fixed number of replications simulating a fixed number of data matrices and then run the estimation procedure. As already noticed

[10] While the first distributions are common, it might be useful to explicitly define the Asymmetric Laplace distribution $\mathcal{AL}(\boldsymbol{\mu}, \mathbf{\Sigma}, \kappa)$ through its probability density function $f_{\boldsymbol{\mu},\mathbf{\Sigma},\kappa}(\mathbf{x})(x_1, \dots, x_k) = \frac{2e^{\mathbf{x}'\mathbf{\Sigma}^{-1}\mu}}{(2\pi)^{\kappa/2}|\mathbf{\Sigma}|^{0.5}} \left(\frac{\mathbf{x}'\mathbf{\Sigma}^{-1}\mathbf{x}}{2+\boldsymbol{\mu}'\mathbf{\Sigma}^{-1}\boldsymbol{\mu}}\right)^{(2-\kappa)/4} K_{\frac{2-\kappa}{2}}\left(\sqrt{(2+\boldsymbol{\mu}'\mathbf{\Sigma}^{-1}\boldsymbol{\mu})(\mathbf{x}'\mathbf{\Sigma}^{-1}\mathbf{x})}\right)$, with $K_{\frac{2-\kappa}{2}}$ is the modified Bessel function of the second kind.

even if the data are generated from a range of distributions and the covariance matrix has the generic Toeplitz form defined above, we estimate the model under gaussianity and independence of the idiosyncratic terms, which is a misspecified model. This will also serve to test its robustness with respect to the misspecification.
In order to evaluate the performance of the estimator we analyze the estimated factors, the estimated factor loadings or the estimated common components $\widehat{\chi}_{it} = \left[\widehat{\mathbf{\Lambda}}'\widehat{\mathbf{F}}_t\right]_i$ and compute the trace statistics against the true values and take the averages over the replications. For the vectors of the common components we take into account the mean standard error with respect to the true value: $\text{MSE}_{\boldsymbol{\chi}} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\widehat{\boldsymbol{\chi}}_{it} - \boldsymbol{\chi}_{it}\right)^2$ Notice that the trace statistic takes values in $[0, 1]$ and that the closer it is to one the better is the approximation of the true value. Clearly, the interpretation of the MSE goes in the opposite direction, as it indicates a better approximation when its value is low in the domain of the positive reals.

In the context outlined, we first compare the estimated outputs with the true values. Table 1 shows that it is possible to get satisfactory estimates of the factor space even in extreme cases (high shares of missing values) for large cross-sections and as the number series is high within each block. In a sense, the blessing of dimensionality helps contain the negative effects of missing information. The presence of non-modelled cross-correlation seems to have a mild impact on the estimations, even for high levels of τ, biting more as the number of series in some of the blocks is limited.

Block-estimators

Here we study our estimator, hereafter BQML estimator in short, in comparison with the estimator proposed by Breitung and Eickmeier (2014) (EB estimator in the following), looking at the performances of the two methods under different model's design: spanning from an excess of residual cross-correlation of the idiosyncratic component to the characteristics of the block structure.
We furthermore show the benefits of the possibility to run the estimation procedure

TABLE 1

Simulation results - BQML estimates

T	n	D	$\delta = 0\ \tau = 0.1$			$\delta = 0\ \tau = 0.4$			$\delta = 0\ \tau = 0.8$		
			$\text{TS}_{\boldsymbol{\Lambda}}$	$\text{TS}_{\boldsymbol{F}}$	$\text{MSE}_{\boldsymbol{\chi}}$	$\text{TS}_{\boldsymbol{\Lambda}}$	$\text{TS}_{\boldsymbol{F}}$	$\text{MSE}_{\boldsymbol{\chi}}$	$\text{TS}_{\boldsymbol{\Lambda}}$	$\text{TS}_{\boldsymbol{F}}$	$\text{MSE}_{\boldsymbol{\chi}}$
100	100	2	0.7152	0.9781	0.7899	0.7364	0.9813	0.7631	0.7441	0.9719	0.7403
100	1000	2	0.7142	0.9959	0.7614	0.7156	0.9962	0.7574	0.7220	0.9963	0.7338
100	1000	10	0.7402	0.9900	0.7171	0.7447	0.9896	0.6966	0.7528	0.9878	0.6761
100	1000	20	0.7435	0.9836	0.7232	0.7474	0.9831	0.7081	0.7547	0.9737	0.7019
100	2000	2	0.7213	0.9969	0.7469	0.7138	0.9971	0.7566	0.7281	0.9968	0.7250
100	2000	10	0.7384	0.9936	0.7069	0.7433	0.9935	0.6920	0.7509	0.9930	0.6807
100	2000	20	0.7404	0.9905	0.7109	0.7449	0.9905	0.6917	0.7516	0.9885	0.6775
150	100	2	0.7255	0.9787	0.7674	0.7295	0.9824	0.7722	0.7324	0.9687	0.7747
150	1000	2	0.7138	0.9972	0.7635	0.7196	0.9973	0.7476	0.7175	0.9969	0.7538
150	1000	10	0.7422	0.9910	0.7084	0.7495	0.9907	0.6898	0.7571	0.9887	0.6669
150	1000	20	0.7476	0.9838	0.7134	0.7508	0.9831	0.6961	0.7589	0.9730	0.6903
150	2000	2	0.7070	0.9980	0.7702	0.7134	0.9979	0.7637	0.7118	0.9981	0.7625
150	2000	10	0.7435	0.9949	0.6966	0.7478	0.9950	0.6828	0.7556	0.9945	0.6631
150	2000	20	0.7450	0.9912	0.6927	0.7486	0.9911	0.6805	0.7569	0.9890	0.6653
200	100	2	0.7268	0.9812	0.7792	0.7236	0.9812	0.7913	0.7470	0.9743	0.7298
200	1000	2	0.7131	0.9977	0.7669	0.7156	0.9978	0.7584	0.7151	0.9976	0.7522
200	1000	10	0.7458	0.9914	0.6971	0.7512	0.9911	0.6840	0.7600	0.9893	0.6583
200	1000	20	0.7476	0.9837	0.7061	0.7526	0.9829	0.6958	0.7595	0.9729	0.6893
200	2000	2	0.7083	0.9985	0.7744	0.7132	0.9986	0.7632	0.7144	0.9985	0.7560
200	2000	10	0.7445	0.9954	0.6906	0.7499	0.9953	0.6741	0.7583	0.9948	0.6566
200	2000	20	0.7472	0.9915	0.6898	0.7516	0.9913	0.6730	0.7603	0.9894	0.6567

T	n	D	$\delta = 0\ \tau = 0.1$			$\delta = 0.75\ \tau = 0.1$			$\delta = 0.9\ \tau = 0.1$		
			$\text{TS}_{\boldsymbol{\Lambda}}$	$\text{TS}_{\boldsymbol{F}}$	$\text{MSE}_{\boldsymbol{\chi}}$	$\text{TS}_{\boldsymbol{\Lambda}}$	$\text{TS}_{\boldsymbol{F}}$	$\text{MSE}_{\boldsymbol{\chi}}$	$\text{TS}_{\boldsymbol{\Lambda}}$	$\text{TS}_{\boldsymbol{F}}$	$\text{MSE}_{\boldsymbol{\chi}}$
100	100	2	0.7407	0.9656	0.7781	0.7106	0.9235	0.9095	0.4151	0.5646	2.5081
100	1000	2	0.7336	0.9943	0.7165	0.7159	0.9875	0.8708	0.6057	0.9573	1.7641
100	1000	10	0.7349	0.9815	0.7400	0.7161	0.9608	0.8286	0.6018	0.8523	1.5459
100	1000	20	0.7358	0.9678	0.7616	0.7182	0.9314	0.8659	0.5740	0.7557	2.1724
100	2000	2	0.7296	0.9961	0.7277	0.7091	0.9889	1.1536	0.5851	0.9608	2.8379
100	2000	10	0.7321	0.9895	0.7357	0.7156	0.9793	0.7940	0.6151	0.9194	1.6059
100	2000	20	0.7329	0.9826	0.7404	0.7158	0.9639	0.8187	0.5892	0.8625	8.6763
150	100	2	0.7401	0.9672	0.7760	0.7333	0.9262	0.8673	0.6252	0.7331	32.3138
150	1000	2	0.7376	0.9957	0.7074	0.7270	0.9913	0.7422	0.6860	0.9765	0.9080
150	1000	10	0.7402	0.9827	0.7248	0.7275	0.9636	0.7891	0.6787	0.8836	1.0466
150	1000	20	0.7420	0.9675	0.7481	0.7310	0.9315	0.8259	0.6699	0.7930	1.3204
150	2000	2	0.7366	0.9972	0.7099	0.7278	0.9956	0.7416	0.6864	0.9803	0.9436
150	2000	10	0.7378	0.9908	0.7146	0.7280	0.9812	0.7582	0.6834	0.9403	0.9841
150	2000	20	0.7402	0.9832	0.7236	0.7279	0.9648	0.7859	0.6788	0.8895	1.0593
200	100	2	0.7475	0.9666	0.7440	0.7377	0.9289	0.8345	0.6899	0.7927	1.2493
200	1000	2	0.7402	0.9962	0.6998	0.7309	0.9900	0.7368	0.7004	0.9754	0.8543
200	1000	10	0.7442	0.9829	0.7063	0.7353	0.9644	0.7646	0.7023	0.8952	0.9493
200	1000	20	0.7457	0.9672	0.7371	0.7364	0.9315	0.8064	0.6976	0.8072	1.0764
200	2000	2	0.7346	0.9979	0.7131	0.7341	0.9961	0.7177	0.7040	0.9866	0.8241
200	2000	10	0.7417	0.9911	0.7055	0.7339	0.9820	0.7393	0.7046	0.9472	0.8724
200	2000	20	0.7433	0.9832	0.7110	0.7346	0.9650	0.7611	0.7012	0.8984	1.3568

Note: The estimation evaluated with respect to the true model's parameters. The simulation parameters not explicitly stated are: $\mu = 0.5$, $\boldsymbol{\xi_t} \sim \mathcal{N}(0, \boldsymbol{\Sigma}^{\boldsymbol{\xi}})$, $\omega = 0.5$, $r^G = 1$, $r^d = 1\,\forall d$, $\eta = 0.2$. Blocks are homogeneous.

in the presence of missing values. We compare the BQML estimator with the EB estimated on a dataset where the few missing observations are imputed or the series deleted from the dataset. Notice that the EB estimator with imputed data is in fact used to initialize the BQML algorithm (see Appendix A). When dealing with missing values in the dataset, rather than a fair horse race between competing methodologies, this test aims to measure the effective gains of applying the EM algorithm as a correction to the EB estimator on imputed datasets.

The following steps compose the simulation experiment: i) the data are simulated from the benchmark model varying along a set of parameters: the number of series, the number of observations, the number of blocks and their composition, the level of cross-correlation, the number of factors and the share of missing values. ii) Then the EB estimator and the BQML estimator are applied to the generated dataset and their results are compared. Notice that we take the ratio between the trace statistics of the BQML estimators (against the true model) with respect to the trace statistics (against the true model) of the EB estimator. To ease readability the ratio for the MSE of the common components is inverted so that the direction of the change has an analogous interpretation for the three considered indicators.

Table 2 compares EB and BQML performances for different values of idiosyncratic cross-correlation. Here the parameter τ models the correlation in excess to the one generated by local factors. We see that the two models reconstruct the factors space with the same level of accuracy for different time series and blocks dimensions. In a sense, the gain we obtain running the EM algorithm on top of an initialization based on the EB estimator is limited. However, the advantages become more evident when it comes to the estimation of the factor loadings and the common component, improving the estimates of 30% or more in most of the cases (reaching 60% at the highest points obtained for very large cross-sections). These observations almost equally apply to a context of homogeneous and heterogeneous, with BQML improving the results as the cross-section increases and the structure of the blocks is heterogeneous. If we look at table 3 having in mind the absolute efficiency measures of table 1, we infer that our methodology provides good approximations of the factors and the loadings even in the most extreme cases, while any estimate based on data imputation or removal would fail. A few exceptions are the outcomes at the bottom-right of table 3, signalling that the 90% of missing values is the very limit of the methodology, at least for a setting with a number of observations and variables analogous to that considered. This fact is relevant for practical applications, for which time series frequency and numerosity are not linearly linked to the missing values share: in that context, increasing the number of variables or taking the maximal frequency might inflate the value of δ above the critical threshold. These considerations explain the choices made while preparing the dataset as presented in Section 3.

Fat-tailed symmetric and asymmetric distributions.

A major constraint to the theoretical derivation of the BQML estimator is that in the maximization step, an explicit form of the maximum likelihood is derived under gaussianity of the idiosyncratic component and of the innovations of the autoregressive processes ruling the factors' dynamics. This assumption, together with the misspecification of the covariance matrix of the idiosyncratic terms, is crucial in order to get a closed-form solution for the estimators of the parameters, to avoid proliferation of the parameters and to reduce consistently the computational complexity. On the latter, however, in many cases of interest, the gaussianity of the idiosyncratic components and the factors' innovations is not granted. Therefore, we present two simulation exercises using the benchmark model: the first one imposing the distribution of the idiosyncratic components to be normal, Student-t or Laplace distributed; the second one, drawing the idiosyncratic components from an Asymmetric Laplace distribution with varying skewness ($\kappa = 1.5,2,3$). The efficiency of the estimator in this context is analyzed both in absolute terms and relative to the EB estimator (see table 4 and 5).

As for the first, the estimate of the factors' space seems not to be affected by the excess of mass in the tails of the generating distributions, since the trace statistics outcomes are comparable to those obtained under gaussianity (as in table 1), and the gains with

TABLE 2

Simulation results - BQML vs EB estimates with no missing values

One global factor ($r^G = 1$)

T	n	D	Homogeneous blocks TSR$_{\mathbf{\Lambda}}$	TSR$_{\mathbf{F}}$	MSER$_{\boldsymbol{\chi}}$	Heterogeneous blocks TSR$_{\mathbf{\Lambda}}$	TSR$_{\mathbf{F}}$	MSER$_{\boldsymbol{\chi}}$
100	1000	5	1.3579	1.0005	1.7358	1.3276	0.9999	1.6897
100	1000	10	1.4548	1.0006	1.8851	1.3953	0.9987	1.7800
100	1000	20	1.4334	1.0002	1.8648	1.4593	0.9978	1.8840
100	5000	5	1.4354	1.0001	1.8506	1.3322	1.0001	1.6900
100	5000	10	1.4488	1.0001	1.8578	1.4447	1.0001	1.8556
100	5000	20	1.5202	1.0002	1.9382	1.5208	1.0001	1.9490
100	5000	50	1.5211	1.0002	1.9658	1.4997	1.0002	1.9399
150	1000	5	1.4188	1.0006	1.8455	1.3356	1.0000	1.7318
150	1000	10	1.3850	1.0005	1.8064	1.3791	0.9990	1.7816
150	1000	20	1.3919	1.0002	1.8052	1.4009	0.9978	1.8259
150	5000	5	1.4469	1.0001	1.8980	1.3046	1.0001	1.6697
150	5000	10	1.5096	1.0002	1.9828	1.3831	1.0001	1.8151
150	5000	20	1.4772	1.0002	1.9409	1.5116	1.0001	1.9794
150	5000	50	1.4479	1.0002	1.8909	1.4958	1.0002	1.9440

Two global factors ($r^G = 2$)

T	n	D	Homogeneous blocks TSR$_{\mathbf{\Lambda}}$	TSR$_{\mathbf{F}}$	MSER$_{\boldsymbol{\chi}}$	Heterogeneous blocks TSR$_{\mathbf{\Lambda}}$	TSR$_{\mathbf{F}}$	MSER$_{\boldsymbol{\chi}}$
100	1000	5	1.3571	1.0008	1.5358	1.2694	0.9997	1.4142
100	1000	10	1.4117	1.0012	1.5561	1.3488	0.9988	1.5166
100	1000	20	1.3885	1.0008	1.5511	1.3973	0.9978	1.5653
100	5000	5	1.3668	1.0002	1.5312	1.2583	1.0001	1.3900
100	5000	10	1.4226	1.0003	1.5981	1.3512	1.0001	1.5003
100	5000	20	1.4619	1.0004	1.6295	1.4085	1.0002	1.5314
100	5000	50	1.4475	1.0005	1.6460	1.4708	1.0004	1.6508
150	1000	5	1.3537	1.0008	1.5042	1.2751	1.0000	1.4216
150	1000	10	1.4131	1.0012	1.6059	1.3373	0.9989	1.4987
150	1000	20	1.3826	1.0008	1.5647	1.3917	0.9976	1.5354
150	5000	5	1.3648	1.0002	1.5542	1.2837	1.0001	1.4570
150	5000	10	1.4264	1.0003	1.5925	1.3379	1.0001	1.4951
150	5000	20	1.4595	1.0004	1.6194	1.4104	1.0002	1.5679
150	5000	50	1.4476	1.0004	1.6527	1.4573	1.0004	1.6387

Note: Results from Monte Carlo simulations comparing the BQML estimator with the EB estimator. No missing values. The ratio of the BQML over the EB indicators: trace statistics for the factor and factor loadings, MSE for the common component. For example: $\text{TSR}_{\mathbf{\Lambda}} = \text{TS}_{\mathbf{\Lambda}}^{\text{BQML}}/\text{TS}_{\mathbf{\Lambda}}^{\text{EB}}$ and $\text{MSER}_{\boldsymbol{\chi}} = \text{MSE}_{\boldsymbol{\chi}}^{\text{EB}}/\text{MSE}_{\boldsymbol{\chi}}^{\text{BQML}}$. The other parameters are fixed to $\mu = 0.5$, $\boldsymbol{\xi_t} \sim \mathcal{N}(0, \mathbf{\Sigma}^{\boldsymbol{\xi}})$, $\tau = 0.1$, $\delta = 0$, $\omega = 0.5$, $r^G = 1$ and $r^d = 1\, \forall d$.

TABLE 3

Simulation results - BQML vs EB estimates in the presence of missing values

		$\tau = 0.3$ $\delta = 0$			$\tau = 0.3$ $\delta = 0.4$			$\tau = 0$ $\delta = 0.6$		
T	n	$\text{TSR}_{\boldsymbol{\Lambda}}$	$\text{TSR}_{\boldsymbol{F}}$	$\text{MSER}_{\boldsymbol{\chi}}$	$\text{TSR}_{\boldsymbol{\Lambda}}$	$\text{TSR}_{\boldsymbol{F}}$	$\text{MSER}_{\boldsymbol{\chi}}$	$\text{TSR}_{\boldsymbol{\Lambda}}$	$\text{TSR}_{\boldsymbol{F}}$	$\text{MSER}_{\boldsymbol{\chi}}$
100	100	1.2237	0.9978	1.4729	1.2290	1.0480	1.9618	1.3438	1.1282	1.9618
100	500	1.2604	1.0009	1.5247	1.2397	1.0308	2.0388	1.3112	1.0767	2.0388
100	500	1.3240	1.0008	1.6668	1.2683	1.0431	2.0952	1.3918	1.1014	2.0952
100	500	1.4017	1.0008	1.7920	1.5295	1.0521	2.1467	1.5236	1.1308	2.1467
100	1000	1.1805	1.0003	1.3684	1.2652	1.0285	2.0123	1.3077	1.0693	2.0123
100	1000	1.3142	1.0006	1.6423	1.3811	1.0348	2.1176	1.4061	1.0857	2.1176
100	1000	1.4323	1.0008	1.8266	1.4039	1.0386	2.1482	1.4712	1.0978	2.1482
150	100	1.2318	1.0030	1.4899	1.2172	1.0494	1.8864	1.2340	1.1022	1.8864
150	500	1.2157	1.0007	1.4368	1.2300	1.0284	2.0466	1.2676	1.0697	2.0466
150	500	1.3677	1.0005	1.7538	1.3134	1.0420	2.2103	1.4200	1.0974	2.2103
150	500	1.4368	1.0011	1.8464	1.4318	1.0531	2.1417	1.4226	1.1193	2.1417
150	1000	1.1577	1.0003	1.3241	1.2444	1.0288	2.0978	1.2787	1.0646	2.0978
150	1000	1.3473	1.0008	1.7405	1.3203	1.0337	2.1493	1.3477	1.0799	2.1493
150	1000	1.4244	1.0009	1.8410	1.4725	1.0404	2.2453	1.4614	1.0930	2.2453

		$\tau = 0.6$ $\delta = 0$			$\tau = 0.6$ $\delta = 0.4$			$\tau = 0.6$ $\delta = 0.6$		
T	n	$\text{TSR}_{\boldsymbol{\Lambda}}$	$\text{TSR}_{\boldsymbol{F}}$	$\text{MSER}_{\boldsymbol{\chi}}$	$\text{TSR}_{\boldsymbol{\Lambda}}$	$\text{TSR}_{\boldsymbol{F}}$	$\text{MSER}_{\boldsymbol{\chi}}$	$\text{TSR}_{\boldsymbol{\Lambda}}$	$\text{TSR}_{\boldsymbol{F}}$	$\text{MSER}_{\boldsymbol{\chi}}$
100	100	1.1929	0.9972	1.4305	1.2397	1.0420	1.6502	1.2901	1.1215	1.9083
100	500	1.2069	0.9998	1.4396	1.2333	1.0311	1.6848	1.2971	1.0781	2.0672
100	500	1.2769	0.9954	1.5939	1.3518	1.0361	1.8373	1.4727	1.0989	2.1972
100	500	1.3818	0.9905	1.7483	1.4149	1.0412	1.9117	1.4436	1.1113	2.1231
100	1000	1.2082	1.0002	1.4365	1.2853	1.0286	1.7733	1.2788	1.0676	2.0592
100	1000	1.3243	0.9981	1.6885	1.3303	1.0320	1.8611	1.3779	1.0814	2.1592
100	1000	1.4354	0.9973	1.8451	1.3723	1.0351	1.8752	1.4945	1.0931	2.2238
150	100	1.2184	0.9992	1.4694	1.2153	1.0456	1.6521	1.2516	1.1060	1.9080
150	500	1.2125	0.9984	1.4486	1.2315	1.0293	1.6932	1.2718	1.0690	2.0797
150	500	1.3094	0.9971	1.6684	1.3398	1.0388	1.8943	1.3965	1.0886	2.2488
150	500	1.3857	0.9888	1.7839	1.4462	1.0435	2.0126	1.4160	1.1047	2.1749
150	1000	1.2366	1.0003	1.4993	1.2349	1.0283	1.7111	1.2649	1.0637	2.1154
150	1000	1.3130	0.9992	1.6967	1.3726	1.0328	1.9508	1.4062	1.0770	2.2765
150	1000	1.4359	0.9971	1.8765	1.4245	1.0353	2.0250	1.5187	1.0872	2.3329

Note: Results from Monte Carlo simulations comparing the BQML estimator with the EB estimator. The ratios of the BQML over the EB indicators are defined as for Table 2. The other parameters are fixed to be $\mu = 0.5$, $\boldsymbol{\xi_t} \sim \mathcal{N}(0, \boldsymbol{\Sigma^{\xi}})$, $B = 5$, $\omega = 0.5$, $r^G = 1$ and $r^d = 1 \, \forall d$.

respect to the EB estimator are not significant. The major difference is observed for the factor loadings that are anyway estimated efficiently while being between 4 and 8 percentage points below the reference values of table 1.

Regarding the simulations with the asymmetric tails, the estimates seem not to be affected by the degree of skewness of the distribution: i) the trace statistics outcomes are comparable to those obtained under symmetric distributions with a minimal reduction around 0.01 for the estimates (as in table 4); ii) there is no observable pattern of dependence on the skewness level parametrized by κ, namely that results are comparable if we take $\kappa = 2$ or $\kappa = 3$. The gains with respect to the EB estimator are analogous to those observed for symmetric fat tails, setting above the 30% for the loadings and the 5% for factor estimates in the presence of missing values.

In summary, for application to the dataset with Laplace distributed observations, symmetric of equivalently asymmetric, the estimator of factor models via BQML seems not to be problematic if the limiting conditions for the estimator hold true both considering the whole cross-section ($n \to \infty$) and only the series relative to each block ($n^d \to \infty$).

TABLE 4

Simulation results - BQML estimates under fat-tailed distributions

Cross-correlations and missing values: $\tau = 0.1 \quad \delta = 0$

		Laplace		Student-t (3)		Laplace		Student-t (3)	
T	n	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$
100	100	0.7560	0.9237	0.7145	0.9240	1.2841	1.0093	1.3000	1.0258
100	1000	0.7426	0.9889	0.6932	0.9877	1.3248	1.0008	1.3382	1.0089
100	2000	0.7415	0.9934	0.6877	0.9932	1.3288	1.0003	1.3625	1.0056
150	100	0.7589	0.9281	0.7253	0.9255	1.3345	1.0104	1.3010	1.0234
150	1000	0.7470	0.9905	0.7022	0.9884	1.3017	1.0009	1.3449	1.0080
150	2000	0.7453	0.9945	0.6985	0.9930	1.3530	1.0004	1.3732	1.0050

Cross-correlations and missing values: $\tau = 0.1 \quad \delta = 0.5$

		Laplace		Student-t (3)		Laplace		Student-t (3)	
T	n	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$
100	100	0.7456	0.8635	0.7150	0.8526	1.3642	1.1451	1.3960	1.2143
100	1000	0.7328	0.9782	0.6847	0.9785	1.3991	1.0540	1.4048	1.0924
100	2000	0.7352	0.9884	0.6836	0.9890	1.3649	1.0452	1.4193	1.0651
150	100	0.7552	0.8608	0.7164	0.8546	1.3301	1.1412	1.3677	1.2045
150	1000	0.7419	0.9816	0.6948	0.9811	1.4028	1.0525	1.4091	1.0784
150	2000	0.7406	0.9897	0.6995	0.9894	1.3682	1.0443	1.3962	1.0629

Cross-correlations and missing values: $\tau = 0.5 \quad \delta = 0$

		Laplace		Student-t (3)		Laplace		Student-t (3)	
T	n	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$
100	100	0.7611	0.9114	0.7323	0.9012	1.3176	0.9886	1.3460	0.9975
100	1000	0.7487	0.9871	0.7126	0.9865	1.3264	0.9997	1.3468	1.0030
100	2000	0.7464	0.9924	0.7053	0.9931	1.3415	1.0000	1.3392	1.0040
150	100	0.7648	0.9009	0.7357	0.9006	1.3302	0.9901	1.2801	0.9936
150	1000	0.7516	0.9893	0.7182	0.9887	1.3143	0.9999	1.3579	1.0041
150	2000	0.7501	0.9940	0.7151	0.9940	1.3690	1.0000	1.3074	1.0032

Cross-correlations and missing values: $\tau = 0.5 \quad \delta = 0.5$

		Laplace		Student-t (3)		Laplace		Student-t (3)	
T	n	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$
100	100	0.7528	0.8466	0.7234	0.8477	1.4040	1.1300	1.3672	1.1743
100	1000	0.7385	0.9789	0.7040	0.9778	1.3408	1.0525	1.4390	1.0735
100	2000	0.7393	0.9877	0.7021	0.9875	1.3599	1.0443	1.4216	1.0571
150	100	0.7591	0.8501	0.7332	0.8480	1.4125	1.1236	1.3575	1.1742
150	1000	0.7465	0.9803	0.7157	0.9811	1.3349	1.0510	1.3569	1.0726
150	2000	0.7462	0.9891	0.7116	0.9894	1.4181	1.0439	1.3635	1.0564

Note: Monte Carlo simulations with data generated under fat-tailed distributions. TS denotes the trace statistics of the BQML estimates against the true model. TSR is the ratio of the BQML indicators over the EB indicators as for Table 2. Under Laplace: $\boldsymbol{\xi_t} \sim \mathcal{L}(0, \mathbf{\Sigma}^{\boldsymbol{\xi}})$ and $\boldsymbol{u_t} \sim \mathcal{L}(0, \mathbf{\Sigma}^{\boldsymbol{u}})$ (both for global and local factors). Under Student-t (3): $\boldsymbol{\xi_t} \sim t_3\left(\mathbf{0}_n, \mathbf{\Sigma}^{\boldsymbol{\xi}}\right)$ and $\boldsymbol{u_t} \sim t_3(0, \mathbf{\Sigma}^{\boldsymbol{u}})$ (both for global and local factors). The other parameters are fixed to be $\mu = 0.5$, $B = 5$, $\omega = 0.5$, $r^G = 1$ and $r^d = 1 \, \forall d$.

TABLE 5

Simulation results - BQML estimates under skewed distributions

Cross-correlations and missing values: $\tau = 0.1 \quad \delta = 0$

		Asym.Lapl. (k=1.5)		Asym.Lapl. (k=2)		Asym.Lapl. (k=3)		Asym.Lapl. (k=1.5)		Asym.Lapl. (k=2)		Asym.Lapl. (k=3)	
T	n	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$
100	500	0.7316	0.9802	0.7327	0.9810	0.7290	0.9795	1.2886	1.0025	1.3632	1.0030	1.3331	1.0024
100	1000	0.7327	0.9892	0.7274	0.9874	0.7288	0.9877	1.3613	1.0013	1.3314	0.9999	1.3291	1.0016
100	2000	0.7306	0.9927	0.7267	0.9936	0.7277	0.9932	1.3145	1.0003	1.3395	1.0006	1.2943	1.0007
150	500	0.7399	0.9805	0.7399	0.9818	0.7355	0.9796	1.3106	1.0025	1.3268	1.0027	1.3073	1.0020
150	1000	0.7390	0.9906	0.7392	0.9903	0.7355	0.9892	1.2912	1.0012	1.3577	1.0013	1.3252	1.0014
150	2000	0.7347	0.9941	0.7363	0.9941	0.7320	0.9938	1.2766	1.0006	1.3088	1.0006	1.3244	1.0006

Cross-correlations and missing values: $\tau = 0.1 \quad \delta = 0.5$

		Asym.Lapl. (k=1.5)		Asym.Lapl. (k=2)		Asym.Lapl. (k=3)		Asym.Lapl. (k=1.5)		Asym.Lapl. (k=2)		Asym.Lapl. (k=3)	
T	n	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$
100	500	0.7268	0.9628	0.7248	0.9643	0.7220	0.9612	1.4216	1.0759	1.4170	1.0729	1.4144	1.0789
100	1000	0.7224	0.9788	0.7228	0.9784	0.7217	0.9783	1.3957	1.0567	1.3874	1.0581	1.4136	1.0586
100	2000	0.7231	0.9876	0.7211	0.9885	0.7206	0.9888	1.3872	1.0471	1.3879	1.0466	1.3489	1.0461
150	500	0.7356	0.9638	0.7317	0.9603	0.7351	0.9626	1.3654	1.0711	1.3795	1.0749	1.3328	1.0734
150	1000	0.7335	0.9815	0.7316	0.9814	0.7320	0.9772	1.3655	1.0527	1.3131	1.0547	1.3350	1.0541
150	2000	0.7336	0.9895	0.7323	0.9898	0.7297	0.9899	1.3591	1.0455	1.2722	1.0448	1.3881	1.0447

Cross-correlations and missing values: $\tau = 0.5 \quad \delta = 0$

		Asym.Lapl. (k=1.5)		Asym.Lapl. (k=2)		Asym.Lapl. (k=3)		Asym.Lapl. (k=1.5)		Asym.Lapl. (k=2)		Asym.Lapl. (k=3)	
T	n	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$
100	500	0.7444	0.9759	0.7389	0.9732	0.7438	0.9733	1.3552	0.9985	1.3249	0.9973	1.3385	0.9983
100	1000	0.7415	0.9868	0.7366	0.9871	0.7396	0.9865	1.3170	0.9999	1.3228	0.9997	1.2840	0.9998
100	2000	0.7379	0.9930	0.7400	0.9930	0.7370	0.9929	1.3035	1.0001	1.3520	1.0002	1.3187	1.0003
150	500	0.7466	0.9735	0.7476	0.9762	0.7481	0.9755	1.3327	0.9970	1.3594	0.9984	1.3600	0.9987
150	1000	0.7477	0.9887	0.7462	0.9889	0.7425	0.9892	1.3061	0.9999	1.3765	0.9999	1.3382	0.9999
150	2000	0.7464	0.9942	0.7417	0.9947	0.7434	0.9938	1.3331	1.0002	1.3391	1.0002	1.3141	1.0000

Cross-correlations and missing values: $\tau = 0.5 \quad \delta = 0.5$

		Asym.Lapl. (k=1.5)		Asym.Lapl. (k=2)		Asym.Lapl. (k=3)		Asym.Lapl. (k=1.5)		Asym.Lapl. (k=2)		Asym.Lapl. (k=3)	
T	n	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TS_{\mathbf{\Lambda}}$	$TS_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$	$TSR_{\mathbf{\Lambda}}$	$TSR_{\boldsymbol{F}}$
100	500	0.7300	0.9573	0.7336	0.9597	0.7331	0.9629	1.3051	1.0697	1.3463	1.0661	1.4236	1.0667
100	1000	0.7333	0.9784	0.7335	0.9788	0.7318	0.9773	1.4062	1.0521	1.3629	1.0541	1.3530	1.0529
100	2000	0.7341	0.9884	0.7318	0.9882	0.7306	0.9871	1.3899	1.0444	1.4060	1.0454	1.3692	1.0450
150	500	0.7420	0.9600	0.7444	0.9631	0.7382	0.9609	1.3766	1.0668	1.3386	1.0645	1.3773	1.0640
150	1000	0.7410	0.9792	0.7421	0.9795	0.7383	0.9793	1.3090	1.0518	1.3361	1.0520	1.3965	1.0527
150	2000	0.7411	0.9904	0.7400	0.9896	0.7403	0.9896	1.3314	1.0437	1.3719	1.0438	1.3641	1.0446

Note: Monte Carlo simulations with data generated under left skewed fat-tailed asymmetric distributions. TS denotes the trace statistics of the BQML estimates against the true model. TSR is the ratio of the BQML indicators over the EB indicators as for Table 2. The parameters have the following characteristics (with $\mathcal{AL}$ for the Asymmetric Laplace distribution): $\boldsymbol{\xi_t} \sim \mathcal{AL}(0, \mathbf{\Sigma}^{\boldsymbol{\xi}}, \kappa)$ and $\boldsymbol{u_t} \sim \mathcal{AL}(0, \mathbf{\Sigma}^{\boldsymbol{u}}, \kappa)$ (both for global and local factors). The other parameters are fixed to be $\mu = 0.5$, $B = 5$, $\omega = 0.5$, $r^G = 1$ and $r^d = 1\,\forall d$.

Appendix C, Autocorrelation analysis

This section is devoted to assessment the impact of the autocorrelation in the export growth rate time series. Before entering the analysis and its motivations, we should mention that in the context of our paper, where the share of missing values approaches the 70% of the dataset, series autocorrelation might be ill-defined from a theoretical point of view and even more challenging to assess empirically through the estimation of the autocorrelation functions or through standard autocorrelation tests. We thus restrict the empirical exercises to those series having at least the 50% of available observations (50SDT in the following) and offer a focus on the small sample of complete time series (CSDT in the following) [11].

As the frequency of time series increases, the increasing available information facilitates the identification of dynamic factors and helps to correct potential biases that may arise from yearly analysis alone (e.g. the partial-year effect). However, while working with quarterly data enhances our ability to track fluctuations in the business cycle, it brings the issue of intra-year seasonality to the forefront.

To mitigate the impact of seasonality on the data, we have chosen to calculate quarter-to-quarter yearly growth rates, which compare the growth of a given quarter to the same quarter of the previous year. This approach aligns with the strategy employed by Bricongne et al. (2022), who recently applied a static orthogonal decomposition model as those of Gabaix (2011) and di Giovanni et al. (2014) to monthly export data, aiming to investigate the effects of macrofluctuations at a finer frequency. The advantage of this transformation lies in the possibility to capture the primary source of autocorrelation, without introducing unnecessary complexity to the model. Indeed, while other sources of autocorrelation exist, such as those specific to sectors or firms, they require individual series analysis and specialized filtering, which would extend

[11] Notice that, in this respect, the selection of the series is performed on the year-to-year quarterly dataset. Then the other two datasets are restricted to the series selected at that stage so that the comparison includes the same time-series.

beyond the current paper's scope and not necessarily clarify our understanding of volatility.

The analysis includes the following steps:

Assessment of the autocorrelations. Using the 50SDT, we analyse the distributions of the estimated autocorrelation functions for time series of the growth rate calculated at different frequencies and using different transformations (yearly, quarterly, quarter-to-quarter yearly, see Figure1). We look also at the residual serial correlation in the idiosyncratic components to see whether the autocorrelation in the data is absorbed by the common components in the estimation models (2).

DFM with different dynamic specifications. With the aim to check whether part of the autocorrelation in the data can be absorbed in the common component as we change the way the common movements are modelled, in addition to the model defined in the main text and the benchmark by di Giovanni et al. (2014), we test different specification of our DFM including more lags in the factor autoregressive process: the first specification allows the global factor to follow an AR of order four and the local factors an AR of order one(GAR(4)-LAR(1)), while the second specification allows AR processes of order four for both global and local factors (GAR(4)-LAR(4)).

Volatility analysis with non-autocorrelated vs autocorrelated samples. We show the robustness of our methodology to the presence of autocorrelation with a simple split sample exercise: we compare the aggregate volatility estimates resulting from a dataset composed by highly autocorrelated series and another one coming from non-autocorrelated series. Leveraging the empirical analysis of the first two points we split the CSDT in two subsamples, the first one composed by the series for which one would not reject the hypothesis of autocorrelation up to order four at a 95% significance level in for the Breusch–Godfrey test, the second using the remainder. Then we take 20 random subsamples of 14k firms in both samples (resulting in approximately 80k firm-destination pairs in each sample) and compare the volatility analsysis for the common component and the idiosyncratic component using bootstrapped standard errors. Notice, that this exercise is optimal to test the effect of the presence of autocorrelation on the aggregate volatility estimates, yet should not be compared with the main results of the paper, since with that subsample size we incur in a significant change of the components estimates at the firm-destination level and also of the weights used for the aggregation.

Looking at the results of the three analyses described above, we summarize the findings in the following points:

- Quarter-to-quarter yearly growth rates are effective in reducing the seasonality without inducing autocorrelation in excess in the data and offer a clear and intuitive interpretation that might not be as straightforward when applying series-specific or other frequency-specific filters.
- The remaining autocorrelation is not captured by the factor structure, even when extended to higher autoregressive orders, and by static fixed effects as those of the SODMs. All the autocorrelation in the data is then absorbed by the idiosyncratic term of the decomposition.
- As for the paper's primary findings on the aggregate volatility components, any residual autocorrelation does not pose a serious issue for the estimates of the volatility, which are proven not to depend on the share ofautocorrelated series in the sample (see Figure 5).
- As a side result, we test the robustness of our findings using different dynamic specifications of the dynamic factor models. From Figures 3 and 4 we see that the results are consitent with the main exercise, with the richer structure in the factor leading to more pronounced differences with respect to SODM in the synchronization with the international trade cycle.

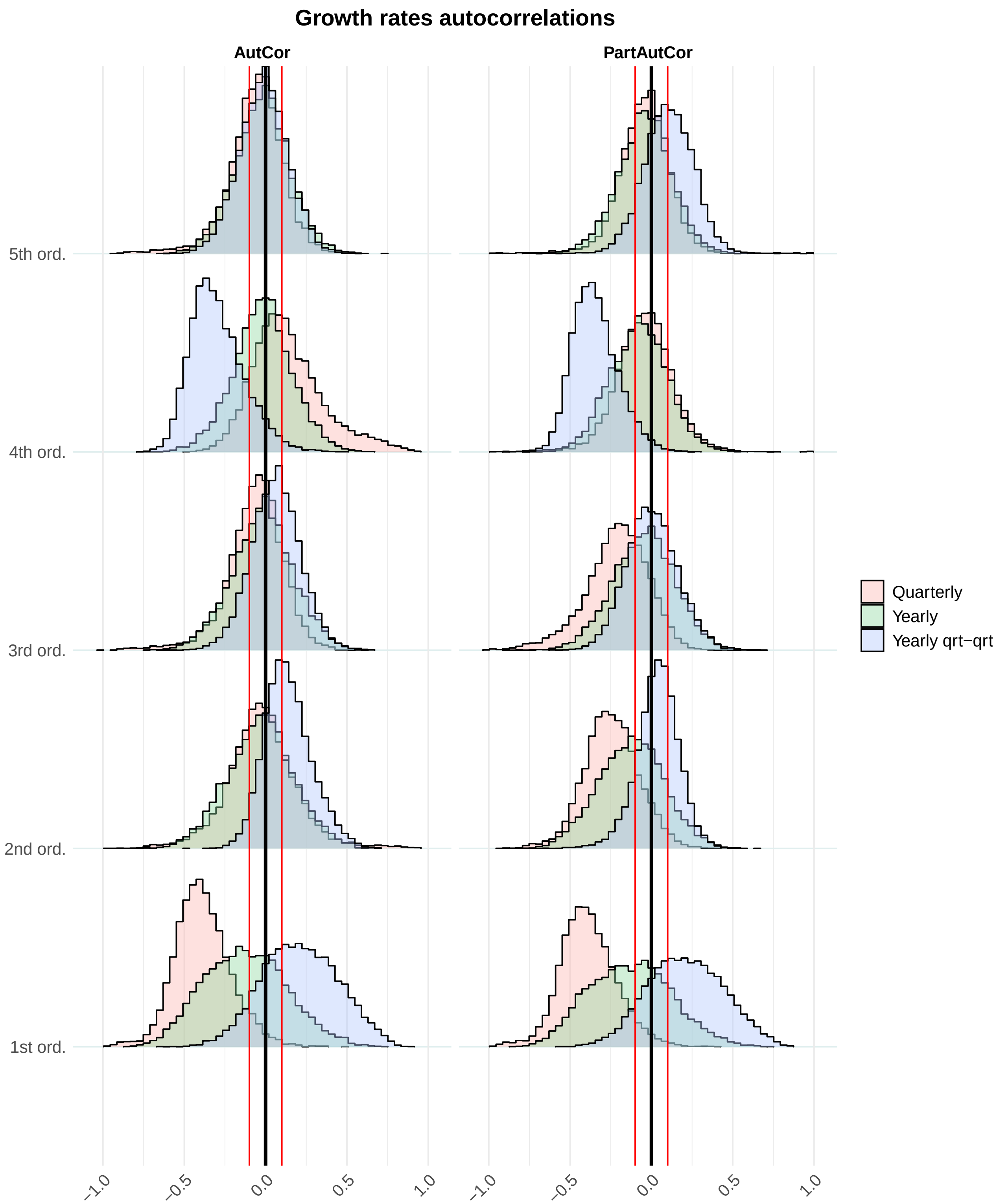

Figure 1. The autocorrelation of the growth rates
Note: The distributions of the autocorrelation functions (up to order 5) for the time series of the growth rates. Different histogram plots refer to growth rates computed from yearly and quarterly data in levels. Quarterly data are transformed into quarterly growth rates (i.e. the growth rates in a given quarter with respect to the previous quarter) or into yearly quarter-to-quarter growth rates (i.e. the growth rates in a given quarter with respect to the same quarter of the previous year). The distributions are computed restricting the dataset to the series with at least the 50% of non-missing points..

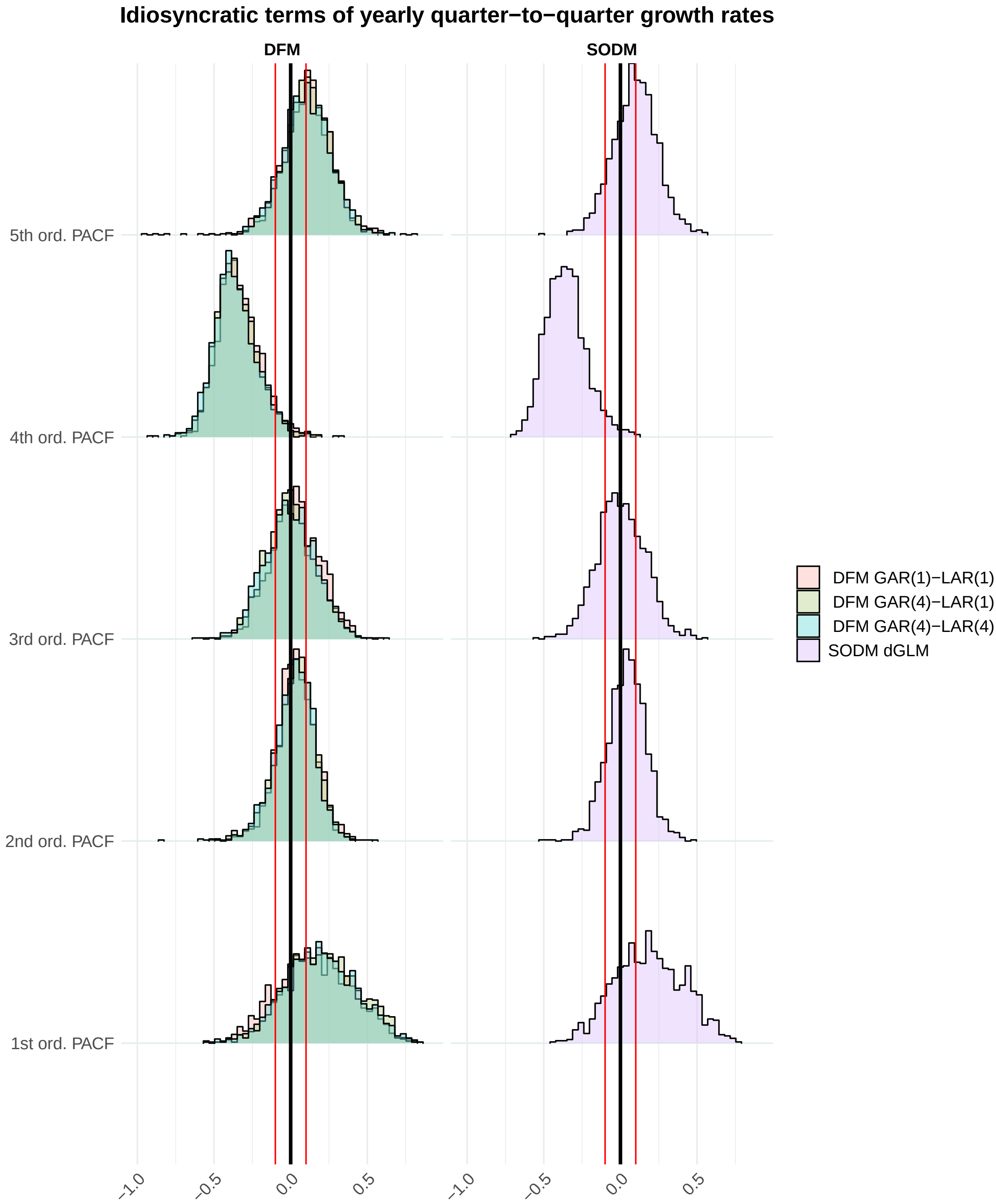

Figure 2. The autocorrelation of the idiosyncratic component

Note: The distributions of the partial autocorrelations (up to order 5) for the time series of the idiosyncratic. The histogram plots refer to the idiosyncratic components estimated on yearly quarter-to-quarter growth rates from four different models and grouped into the DFMs and SODM families: the decomposition proposed in di Giovanni et al. (2014) with destination and sector effects for the SODM vis-a-vis DFMs with global and local factors alternatively represented as AR processes of order one or four (respectively $GAR(k)$ and $LAR(k)$.

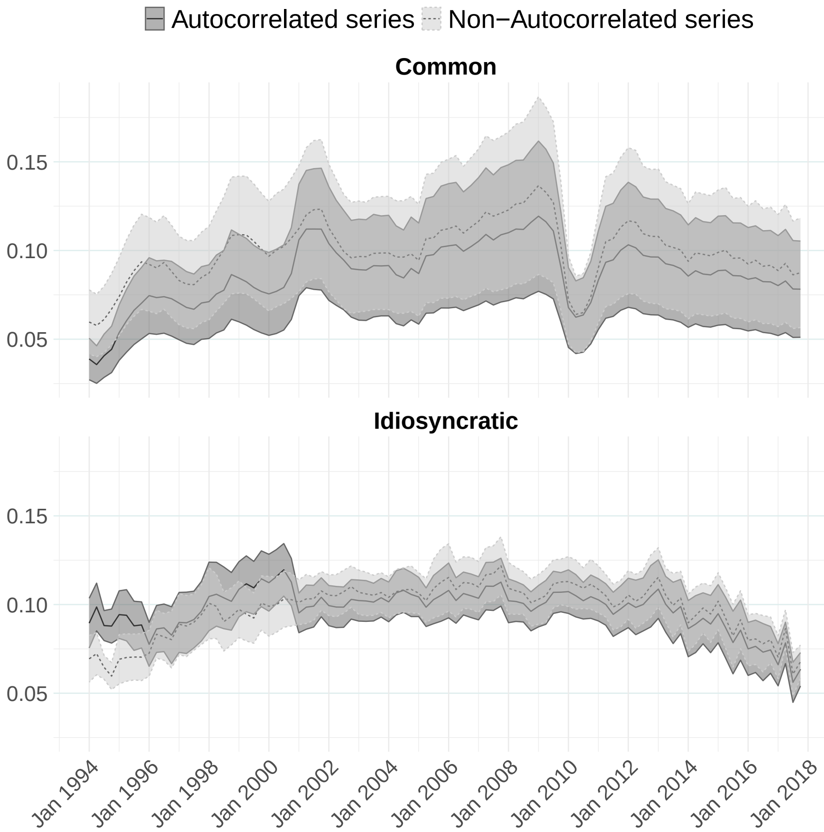

Figure 3. The volatility decomposition under growth rates autocorrelation
Note: Volatility decomposition for a split sample of the dataset of complete time series. Analyzed subsamples include only autocorrelated series and non-autocorrelated series.

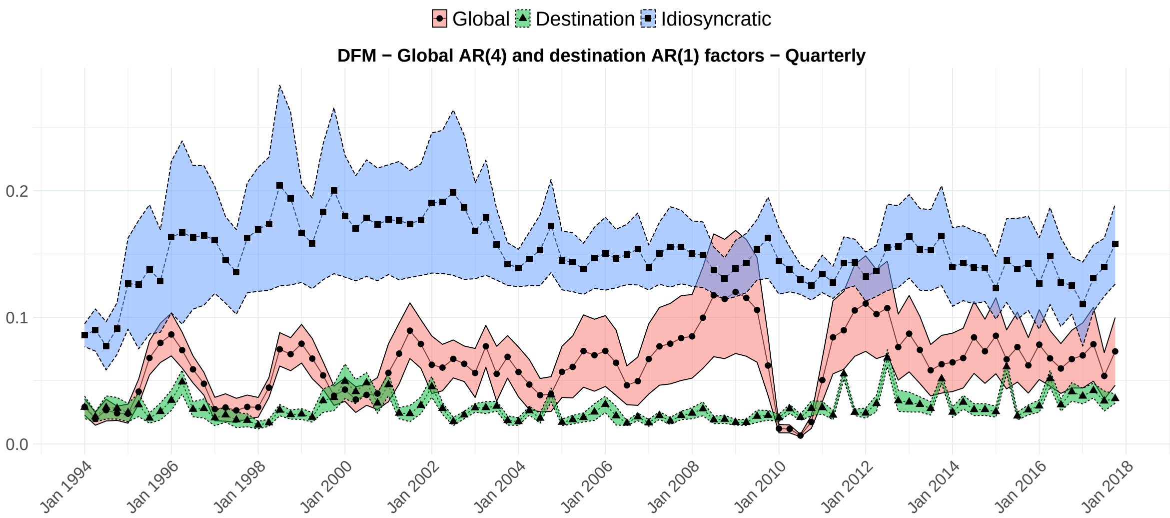

Figure 4. Volatility components under global AR(4) and local AR(1)
Note: The decomposition obtained estimating a Dynamic Factor Model with the global factor modelled as an AR(4) process and the local factors modelled as AR(1) processes.

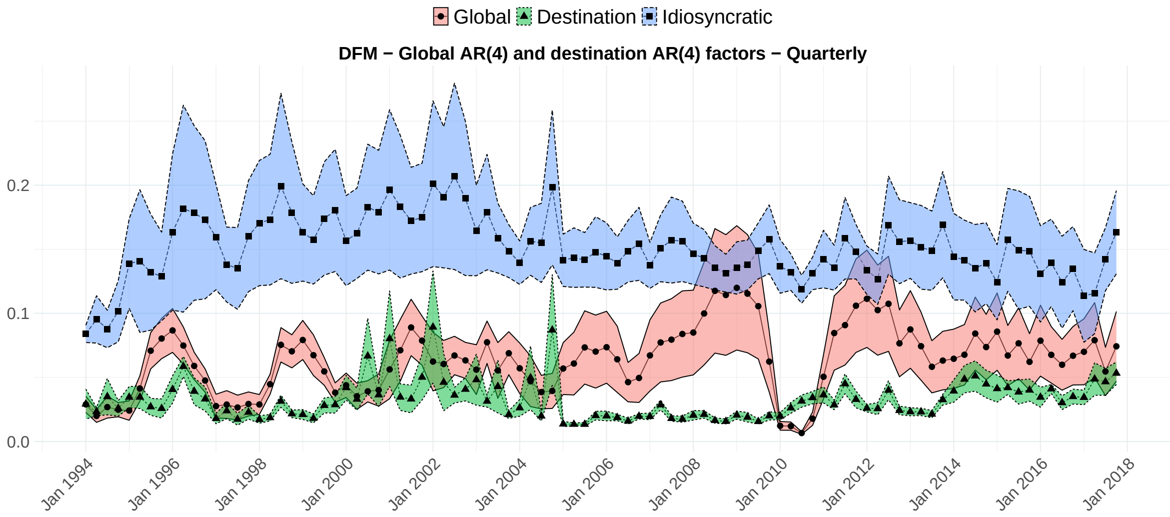

Figure 5. Volatility components under global AR(4) and local AR(4)
Note: The aggregate volatility decomposition with a Dynamic Factor Model with the global and local factors modelled as AR(4) process.

Appendix D, Destination factors analysis

This appendix complements the findings from Section 4 by diving deeper into local factors and their interpretation. While the global factor could be easily interpreted by visual inspection, the interpretation of the 67 time series representing local factors necessitates a more systematic approach. To do this, we examined the correlations of the factors extracted with the BQML estimator against a suite of economic and financial indicators.

Initial findings are presented in Table 1, where we have compared all the destination-specific factors to economic indicators. The sources of these indicators, their meanings, and any transformation to the original data we made are detailed in Table 6.

Our work suggests that our approach is effective in grasping broader economic trends by enconding destination-specific fluctuations in local factors. Empirically the

[12] For a proper comparison, the transformation imposed on the bilateral flows aims at isolating common movements on the series to mimic the correlation structure of the BDFM estimated in the main exercise.

TABLE 6

Country-specific variables sources and transformations

Variable	Source	Data	Transformation
Bilateral Export Flows	Direction of Trade Statistics (DOTS) of the IMF	Total export value Free on Board (FOB) by areas and countries of export	The values of this variable are the idiosyncratic components of a single factor DFM estimated on the quarter-to-quarter yearly growth rates series[12]
Nominal Exchange Rates	International Financial Statistics (IFS) by IMF, "Data by indicator" tab, Exchange rates selected indicators	National currency per US Dollar, period average	Ratio with France's values
Real Exchange Rates	International Financial Statistics (IFS) by IMF, "Data by indicator" tab, Exchange rates selected indicators	Real effective exchange rate, Consumer Price Index	Ratio with France's values
Interest Rates	International Financial Statistics (IFS) by IMF, "Data by indicator" tab, Interest rates selected indicators	Interest rates of Government bonds	Used as differentials w.r.t. France's values
Output Growth Rates	Quarterly National Accounts of the OECD	B1_GE: Gross domestic product expenditure approach. GYSA: Growth rate based on seasonally adjusted volume data, percentage change on the same quarter of the previous year	Used as differentials w.r.t. France's values

Note: A summary of the macroeconomic variables used in the correlation analysis for local factors. All variables are available at a quarterly frequency. When the original variable is provided in level, if growth rates are used, they are computed as the log quarter-to-quarter yearly growth rates.

correlation with bilateral export flows from independent source is strong for all the considered destinations. Building on this, we can extend the pooled linear model with a series of destination-specific regressions. In light of the insights gained from the primary analysis, we opted to exclude bilateral flow from the set of regressors and instead present the coefficients of the linear models where the dependent variables are the destination-specific residuals from a regression between the factors and the growth rates of the OECD's bilateral flows.

The results of this refined analysis are collected in Table 7 and plotted per each variable in Figures 6-9. This approach offers two main advantages. First, it maximizes the utilization of the available data to examine macroeconomic variable correlations, even for countries with incomplete datasets due to data scarcity[13]. Second, the correlation coefficients, although estimated from a limited dataset of 96 observations, provide insights into the variable-specific effects on the factors. While a comprehensive discussion of these effects is not within the scope of this paper, the identified correlation patterns highlight the economic significance of the estimated factors derived from microeconomic export flow data.

[13] Countries lacking bilateral flows data or those with insufficient data points were excluded from the analysis.

TABLE 7

Results for the destination-specific regressions

Dest.		Intercept	Int. Rate Diff.	Nom. Ex. Rate	Real Ex. Rate	Out. Gr. Diff.
AE	Estimate	0.0016	-	-0.0066	-	-
	Pr(>|t|)	0.9809	-	0.529	-	-
AR	Estimate	0.0312	-	0.0164	-	-0.0154
	Pr(>|t|)	0.3048	-	0.1145	-	0.0282
AT	Estimate	0.0119	-0.7908	0.0024	0.0014	0.009
	Pr(>|t|)	0.8975	0.1117	0.4673	0.0723	0.9316
AU	Estimate	0.126	-0.0151	0.0018	0.0023	-0.0676
	Pr(>|t|)	0.3961	0.8788	0.2499	0.1559	0.0181
BG	Estimate	4.8524	-0.1806	2.9055	-0.0035	-0.0174
	Pr(>|t|)	0.247	0.234	0.2741	0.764	0.6512
BR	Estimate	-0.0036	-	0.0264	0.0125	-0.0149
	Pr(>|t|)	0.9537	-	0.2159	0.1434	0.537
CA	Estimate	-0.0894	-0.0019	-0.0237	0.0038	0.1139
	Pr(>|t|)	0.3614	0.9852	0.0001	0.0121	0.0644
CG	Estimate	-0.027	-	0.1603	-	-
	Pr(>|t|)	0.8433	-	0.0031	-	-
CH	Estimate	-0.0605	-0.0794	-0.0255	0.0018	0.0059
	Pr(>|t|)	0.7406	0.529	0.0022	0.2766	0.9141
CM	Estimate	-0.003	-	4e-04	-0.002	-
	Pr(>|t|)	0.9219	-	0.9647	0.4096	-
CN	Estimate	0.5593	-	-0.0549	-0.0022	-0.0705
	Pr(>|t|)	0.057	-	4e-04	0.4545	0.0741
CY	Estimate	-0.0091	-0.0015	0.0028	-3e-04	-
	Pr(>|t|)	0.9689	0.986	0.7242	0.7733	-
CZ	Estimate	-0.1328	-0.2112	0.0174	0.0064	0.1025
	Pr(>|t|)	0.286	0.3225	0.4171	0.0345	0.0422
DE	Estimate	0.0049	0.0593	0.0035	-0.0012	0.1005
	Pr(>|t|)	0.8998	0.5649	0.099	0.0000	0.0001
DK	Estimate	-0.001	-0.0491	0.009	-6e-04	0.0743
	Pr(>|t|)	0.9845	0.6122	0.5918	0.178	0.0191
DZ	Estimate	0.0066	-	0.0098	0.0088	-
	Pr(>|t|)	0.9439	-	0.6126	0.2262	-
EE	Estimate	-0.3699	0.0319	-0.0498	-	0.0843
	Pr(>|t|)	0.0354	0.318	0.1189	-	0.0001
EG	Estimate	9e-04	-	-0.0106	-	-
	Pr(>|t|)	0.9905	-	0.3916	-	-
ES	Estimate	-0.2462	0.1889	0.0334	0.0038	0.1038
	Pr(>|t|)	0.1645	0.1259	0.0001	0.0000	0.1743
FI	Estimate	-0.0135	0.0145	0.0013	-0.0011	0.0156
	Pr(>|t|)	0.842	0.8989	0.1458	0.1645	0.7093
GA	Estimate	-0.0115	-	-0.0095	-0.0088	-
	Pr(>|t|)	0.824	-	0.3168	0.0000	-
GB	Estimate	0.0433	-0.1077	-0.0096	0.0018	0.0421
	Pr(>|t|)	0.5389	0.1783	0.0000	0.208	0.4095
GR	Estimate	-0.0584	0.0036	0.0123	4e-04	-0.0474
	Pr(>|t|)	0.7448	0.8822	0.8582	0.5549	0.6043
HK	Estimate	-0.0019	-	0.0101	-	-
	Pr(>|t|)	0.9757	-	0.5197	-	-
HR	Estimate	-0.0164	-	-0.0076	0.0000	0.0067
	Pr(>|t|)	0.7499	-	0.5714	0.9423	0.6945
HU	Estimate	0.0739	-0.0722	-0.2424	0.0024	-0.0039
	Pr(>|t|)	0.9189	0.5483	0.6504	0.6243	0.9486
IE	Estimate	-0.2512	0.0425	0.0169	0.0031	0.0501
	Pr(>|t|)	0.1249	0.3563	0.0000	0.0564	0.0128
IL	Estimate	0.194	-	-0.0065	0.001	-0.0717
	Pr(>|t|)	0.3008	-	0.7654	0.7816	0.2261
IN	Estimate	-1.0149	0.0981	-0.0325	-	0.0759
	Pr(>|t|)	0.0693	0.1085	0.8045	-	0.2862

Note: Regressors may vary depending on data availability and country with no available data have been excluded. Shadowed cells legend: $p < 0.01$ - Dark Grey; $0.01 \leq p < 0.05$ - Grey; $0.05 \leq p < 0.1$ - Light Grey. The table cotinues to next page.

Dest.		Intercept	Int. Rate Diff.	Nom. Ex. Rate	Real Ex. Rate	Out. Gr. Diff.
IT	Estimate	0.0429	-1e-04	0.0184	-0.0012	0.0448
	Pr(>|t|)	0.3871	0.9989	0.4768	0.3416	0.4439
JP	Estimate	0.3413	0.099	0.1364	0.002	0.123
	Pr(>|t|)	0.0033	0.1139	0.001	0.5624	0.0517
KR	Estimate	0.0085	0.0533	-0.0291	-	-0.0407
	Pr(>|t|)	0.8931	0.0013	0.0015	-	0.0000
KW	Estimate	-0.0042	-	0.0038	-	-
	Pr(>|t|)	0.9294	-	0.0675	-	-
LB	Estimate	0.0079	-	0.0373	-	-
	Pr(>|t|)	0.9275	-	0.5033	-	-
LT	Estimate	0.3556	-0.2004	-0.0126	-	-0.0026
	Pr(>|t|)	0.203	0.0091	0.627	-	0.9429
MA	Estimate	-0.0386	0.0169	-0.0412	1e-04	-
	Pr(>|t|)	0.8027	0.7481	0.0000	0.9481	-
MG	Estimate	8e-04	-	-0.0049	-	-
	Pr(>|t|)	0.9956	-	0.8679	-	-
MU	Estimate	-0.2081	0.0308	-0.0977	-	-
	Pr(>|t|)	0.6145	0.6659	0.612	-	-
MX	Estimate	-0.5625	0.1292	0.1549	0.0428	0.2004
	Pr(>|t|)	0.1261	0.0701	0.1625	0.0001	0.0000
MY	Estimate	0.0134	-0.0374	0.0197	0.0073	-
	Pr(>|t|)	0.7978	0.3529	0.2434	0.0015	-
NC	Estimate	2e-04	-	-0.0018	-	-
	Pr(>|t|)	0.9939	-	0.8052	-	-
NL	Estimate	0.0033	0.0753	0.0029	3e-04	0.0236
	Pr(>|t|)	0.9606	0.7037	0.0964	0.5894	0.5821
NO	Estimate	0.021	-0.0489	0.0174	-3e-04	-0.027
	Pr(>|t|)	0.6987	0.7025	0.0473	0.8235	0.3956
PF	Estimate	-0.0012	-	0.0086	-	-
	Pr(>|t|)	0.9786	-	0.1249	-	-
PH	Estimate	0.0551	-0.011	0.0118	0.0016	-
	Pr(>|t|)	0.5967	0.4631	0.157	0.3003	-
PL	Estimate	0.158	-0.1198	-0.1251	0.0043	0.0092
	Pr(>|t|)	0.8449	0.308	0.7836	0.8158	0.9101
PT	Estimate	-0.0611	0.0348	-0.0638	2e-04	0.0404
	Pr(>|t|)	0.5281	0.2838	0.0000	0.6865	0.3177
RO	Estimate	-0.3145	-0.149	-0.3387	-0.0099	-
	Pr(>|t|)	0.7336	0.372	0.2372	0.7668	-
RU	Estimate	-0.7817	-0.1802	-3.1608	-0.1961	-
	Pr(>|t|)	0.0808	1e-04	0.0017	0.0079	-
SA	Estimate	0.0047	-	-0.022	-0.0011	-
	Pr(>|t|)	0.9305	-	0.0563	0.631	-
SE	Estimate	-0.1403	-0.042	0.0227	0.0027	0.1484
	Pr(>|t|)	0.0154	0.2243	2e-04	0.0349	0.0000
SG	Estimate	-0.0749	-0.089	0.076	0.0103	-
	Pr(>|t|)	0.5762	0.369	0.0001	0.0097	-
SI	Estimate	-0.1013	0.0188	0.0521	-	0.0896
	Pr(>|t|)	0.38	0.6	0.154	-	9e-04
SK	Estimate	0.127	-0.0643	0.1532	-6e-04	-0.001
	Pr(>|t|)	0.1092	0.1148	0.0000	0.5472	0.9596
TH	Estimate	-0.0498	0.1392	0.0749	-	-
	Pr(>|t|)	0.6529	0.0972	0.0013	-	-
TN	Estimate	0.0013	-	-0.008	-4e-04	-
	Pr(>|t|)	0.983	-	0.0211	0.7743	-
TR	Estimate	0.0289	-	-0.0073	-	-0.0111
	Pr(>|t|)	0.7422	-	0.5303	-	0.4297
TW	Estimate	0.0055	-	-0.0332	-	-
	Pr(>|t|)	0.9014	-	0.0141	-	-
US	Estimate	-0.1207	0.1402	0.0072	-0.0097	0.0748
	Pr(>|t|)	0.1497	0.1184	0.0479	0.0046	0.1414
XU	Estimate	0.0019	-0.2353	-0.0193	-2e-04	-0.014
	Pr(>|t|)	0.9729	0.1257	0.0000	0.5155	0.6148
ZA	Estimate	-0.128	0.0191	-0.0063	-0.0025	-
	Pr(>|t|)	0.4806	0.4367	0.4807	0.3086	-

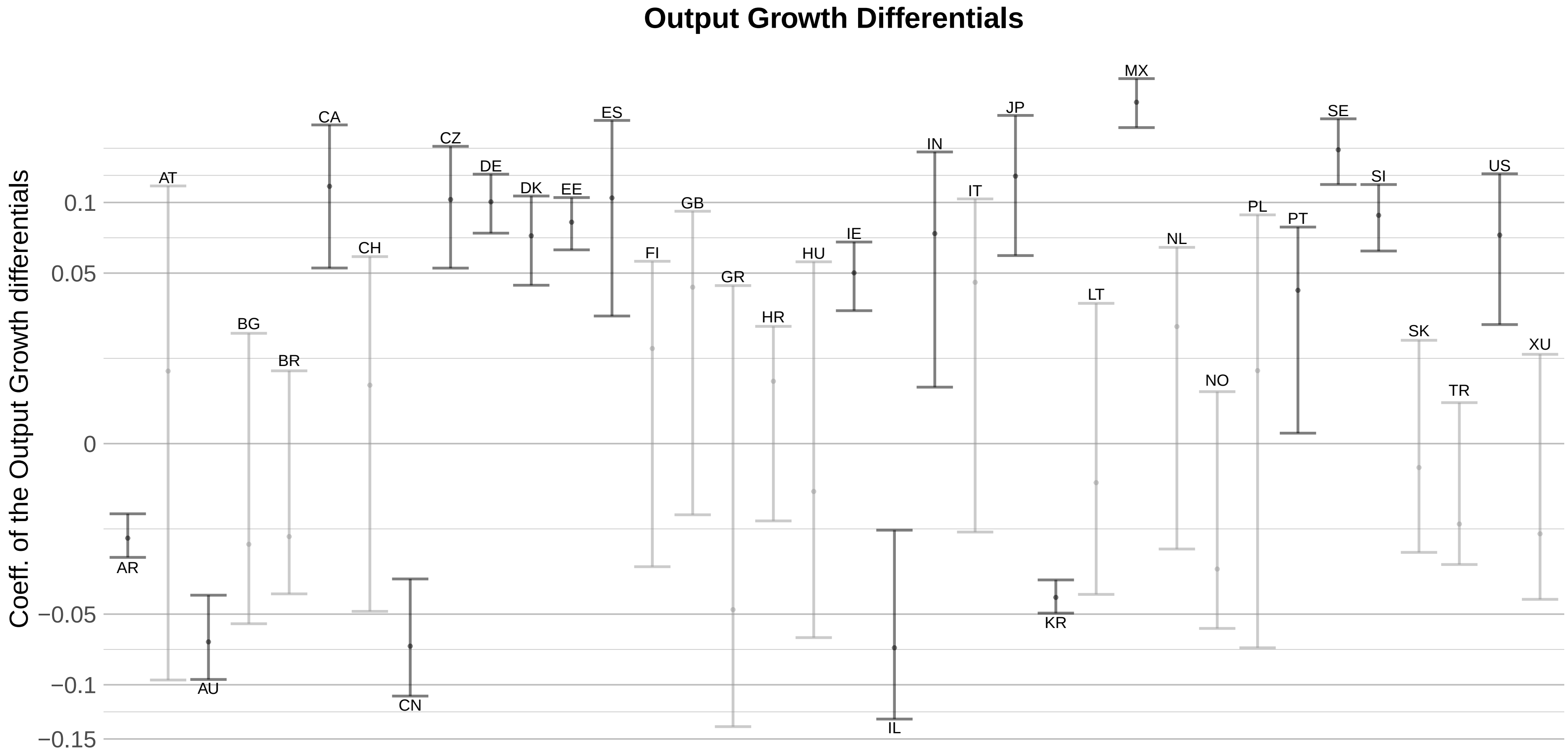

Figure 6. Output growth differentials vs destination factors

Note: Output growth differentials estimated correlations with the relative error band at the 95% confidence level for the destination-specific regressions. The dependent variable is the residual of the regression of the destination factor against growth rates of the relative bilateral flows. Darker bands denote significant values at 95%. Regressors vary depending on data availability.

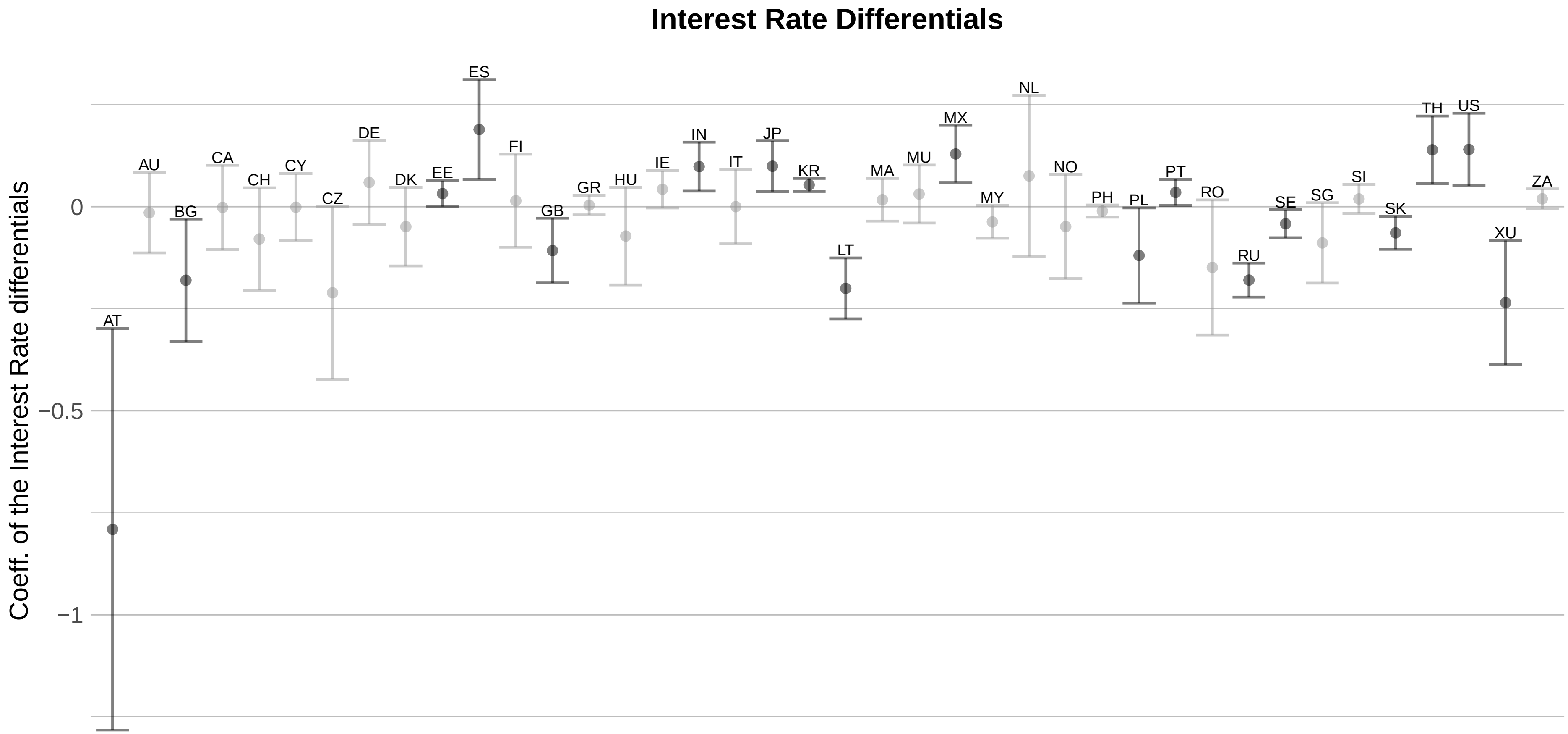

Figure 7. Interest rates differentials vs destination factors

Note: Interest rates differentials estimated correlations with the relative error band at the 95% confidence level for the destination-specific regressions. The dependent variable is the residual of the regression of the destination factor against growth rates of the relative bilateral flows. Darker bands denote significant values at 95%. Regressors vary depending on data availability.

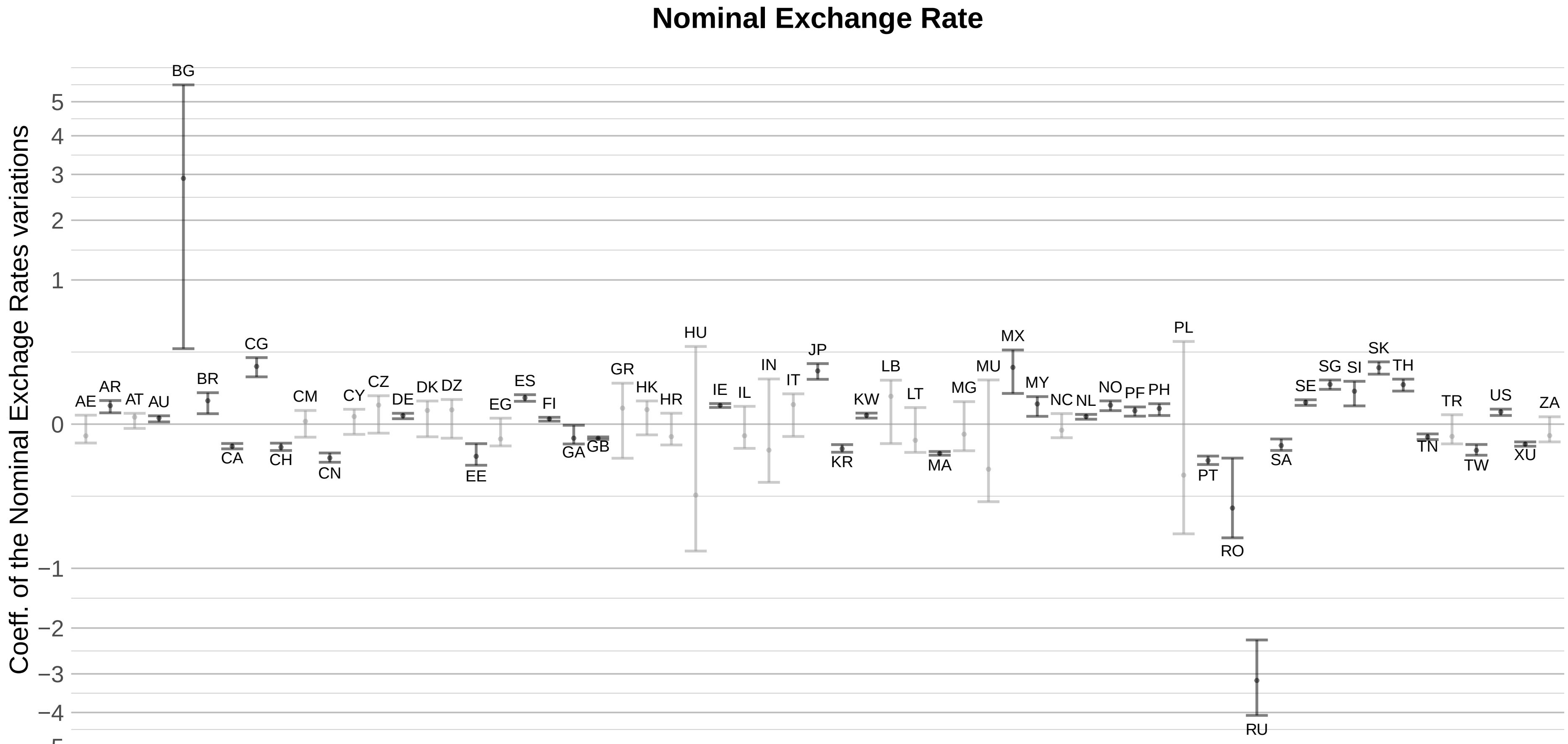

Figure 8. Nominal exchange rates vs destination factors

Note: Nominal exchange rates variations estimated correlations with the relative error band at the 95% confidence level for the destination-specific regressions. The dependent variable is the residual of the regression of the destination factor against growth rates of the relative bilateral flows. Darker bands denote significant values at 95%. Regressors vary depending on data availability.

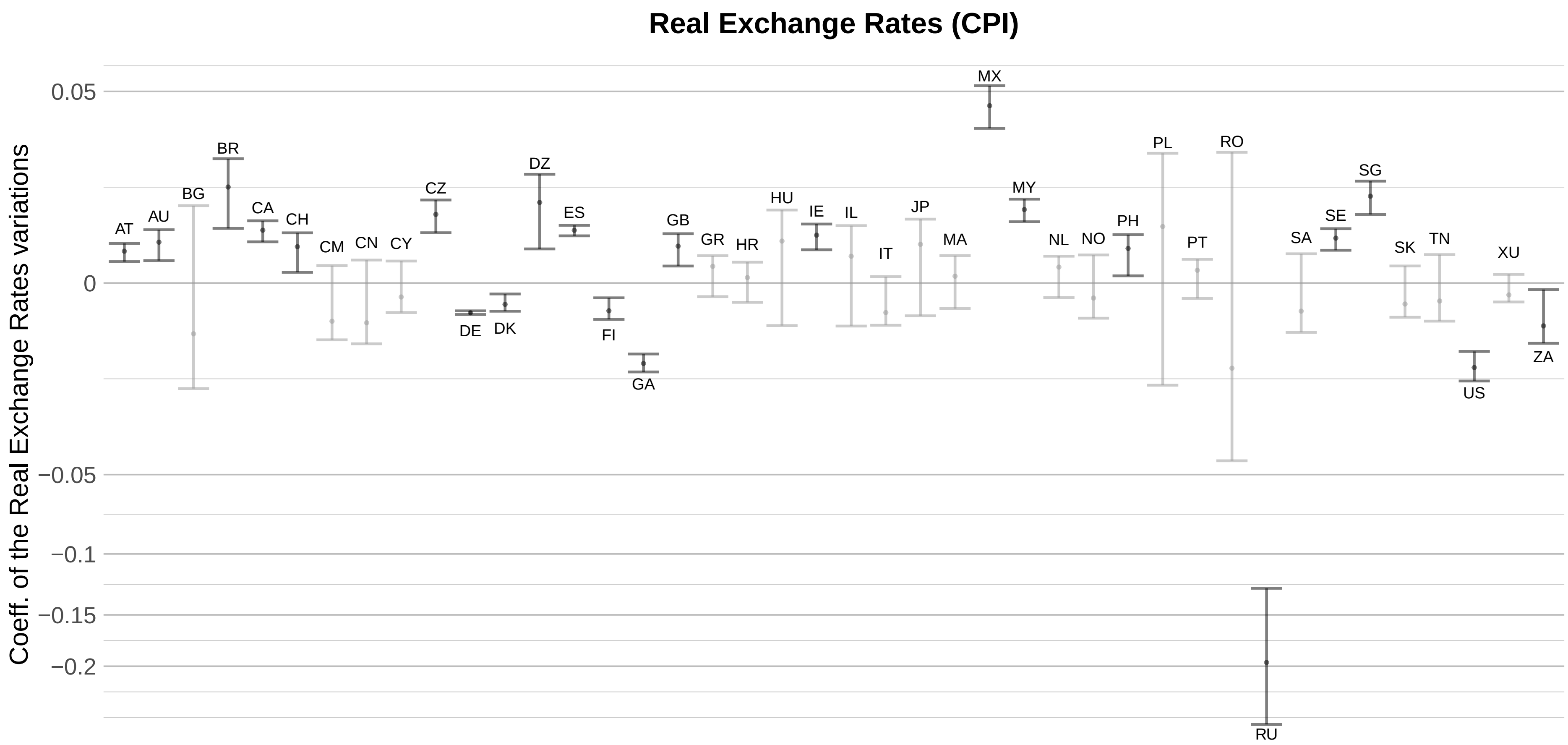

Figure 9. Real exchange rates vs destination factors

Note: Real exchange rates variations estimated correlations with the relative error band at the 95% confidence level for the destination-specific regressions. The dependent variable is the residual of the regression of the destination factor against growth rates of the relative bilateral flows. Darker bands denote significant values at 95%. Regressors vary depending on data availability.

Appendix E, Volatility with mid-point growth rates

In this Appendix we discuss the main results on the volatility estimates focusing on the interaction between the estimation strategy and the aggregation method. The aim of the paper is to choose the weights and the estimator for the main equation as to reconstruct consistently the aggregate growth rate (or a valid proxy).
Using equation 6, we exploit the estimated factors and loadings to assess the impact on the volatility of aggregate export sales of the three sources of shocks: i.e. the global, the destination-specific and the idiosyncratic part. This baseline decomposition equation is coupled with standard size-based aggregation weights to get the estimates of the volatility components. However, that aggregation strategy, defined by the equation 6, might not be optimal given the distributional properties of the growth rates and their components.

In the following, we propose a simple robustness check for our analysis, with the aim of checking whether the insights from the volatility decomposition presented in Section 4 hold when we use different growth rate calculations and the related aggregation strategy. To this end, we repeat the analysis using a standard in the trade literauture along the lines af the recent paper by (Bricongne et al., 2022): we decompose the mid-point growth rates between quarters t and $t-4$ defined by

$$\mathrm{y}_{de,t}^{MP} = \frac{x_{de,t} - x_{de,t-4}}{\frac{1}{2}(x_{de,t} + x_{de,t-4})} \tag{14}$$

where $\mathrm{x}_{de,t}$ is the level of sales of export e to destintation d at time t,and the resulting growth rate $\mathrm{y}_{de,t}^{MP}$ is constrained within the range of -2 to $+2$. It assumes a value of -2 when there is an exit ($x_{de,t} = 0$ and $x_{de,t-4} > 0$), and a value of $+2$ when there is an entry ($x_{de,t} > 0$ and $x_{de,t-4} = 0$).

Mid-point growth rates offer a significant advantage when analyzing detailed trade data, particularly in scenarios of high turnover in the dataset. Such turnover is prevalent in highly disaggregated trade data, occurring not only during large crisis Bernard et al. (2009). In such context, this approach offers a convenient way of capturing both intensive and extensive margin changes in a single measure. This property makes them particularly useful for analyzing small changes, as they provide a reliable approximation without the extremes often associated with year-on-year variations in exports.

Another advantageous characteristic of mid-point growth rates is their exact aggregation property. Unlike logarithmic changes, which may require approximations, aggregate growth rates can be expressed as a weighted average of transaction-level growth rates without any approximation:

$$\mathrm{Y}_t = \frac{X_t - X_{t-4}}{\frac{1}{2}(X_t + X_{t-4})} = \sum_{de} \omega_{de,t}^{MP} \mathrm{y}_{de,t}^{MP} \tag{15}$$

where capital letters denote aggregate quantities and the weights $\omega_{de,t}^{MP}$ are determined by the relative magnitudes of $x_{de,t}$ and $x_{de,t-4}$ compared to the total exports at time t and $t-4$, namely $\omega_{de,t}^{MP} = \frac{x_{de,t}+x_{de,t-4}}{X_t+X_{t-4}}$

This exact aggregation property is particularly valuable in settings characterized by high fluctuations in year-on-year firm-level quarterly exports. Given the substantial volatility in exports within firms over time, convering the cases when the weighted average of firm-level logarithmic changes provides a weak approximation for aggregate changes.

Once defined the dataset of the midpoint growth rates, we run two decompositions — using the DFM in (1) and the SODM of di Giovanni et al. (2014). — on the complete dataset and on a subsample capturing the intensive margin of trade only. Figure 10 presents the results of the exercise. There is no major difference in the SODM when applied to logarithmic or midpoint growth rates. Some differences emerge for the DFM decomposition[14], yet the main insight from Figure 8 is confirmed: the common component estimated through the DFM captures the considerable increase in volatility

[14] In this respect, notice that the volatility over the phases of the cycle are less pronounced, especially when the extensive margin is included.

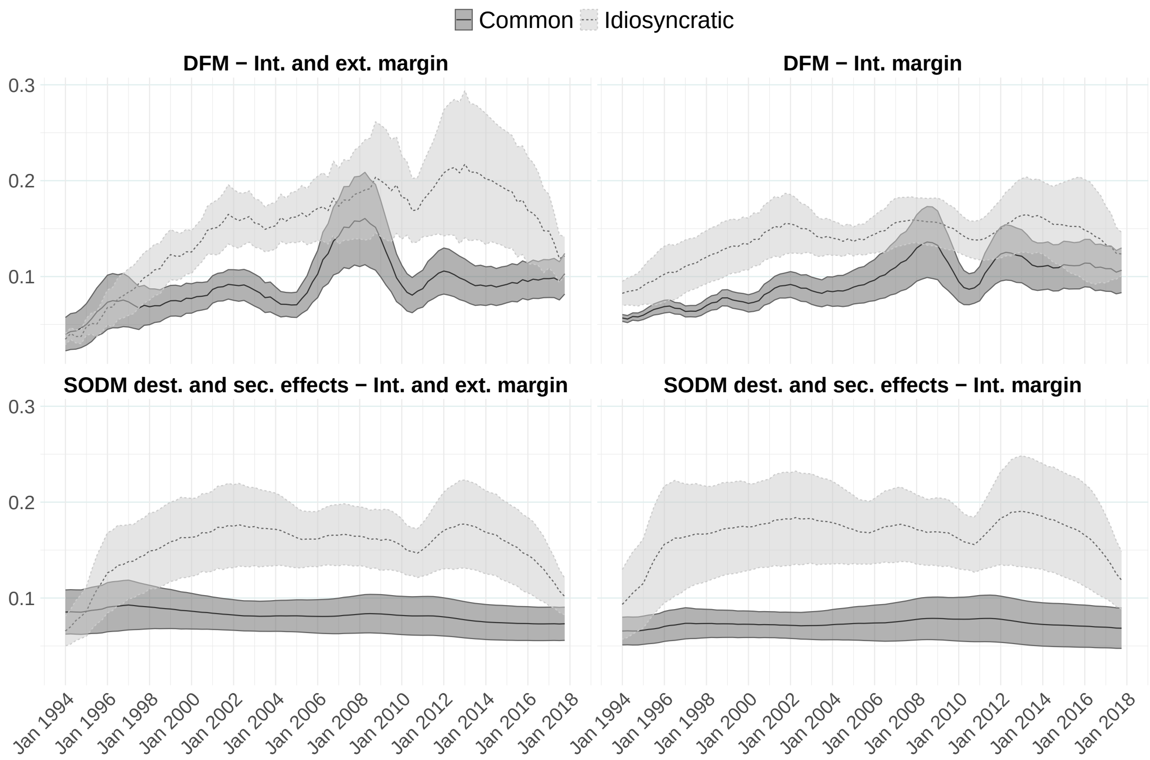

Figure 10. Volatility decomposition for the mid-point growth rates
Note: A comparison of the common and idiosyncratic component of the aggregate volatility derived from a decomposition through the DFM of (1) and the SODM of di Giovanni et al. (2014), on the dataset of the midpoint growth rates and a subsample including the intensive margin only. Error bands are constructed using bootstrapped standard errors at the 95% confidence level.

of the trade collapse. At that point in time, as in the standard decomposition, the idiosyncratic and the common components almost match. This is true both if we include the extensive and the intensive margin.

References

Bańbura, M. and M. Modugno (2014). 'Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data', *Journal of Applied Econometrics*, 29, 133–160.

Barigozzi, M. and M. Luciani (2019). 'Quasi maximum likelihood estimation and inference of large approximate dynamic factor models via the EM algorithm', *arXiv preprint arXiv:1910.03821*.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2009). 'The Margins of US Trade', *American Economic Review*, 99, 487–93.

Breitung, J. and S. Eickmeier (2014). 'Analyzing Business and Financial Cycles Using Multi-level Factor Models', *Bundesbank Discussion Paper*.

Bricongne, J.-C., J. Carluccio, L. G. Fontagné, G. Gaulier, and S. Stumpner (2022). 'From Macro to Micro: Large Exporters Coping with Common Shocks', *CESifo Working Paper*.

di Giovanni, J., A. A. Levchenko, and I. Mejean (2014). 'Firms, Destinations, and Aggregate Fluctuations', *Econometrica*, 82, 1303–1340.

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

Durbin, J. and S. J. Koopman (2012). *Time series analysis by state space methods*, vol. 38, Oxford University Press, Oxford.

Gabaix, X. (2011). 'The Granular Origins of Aggregate Fluctuations', *Econometrica*, 79, 733–772.