

INSTITUTE
OF ECONOMICS



Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM

WORKING PAPER SERIES

Factoring in the micro: a transaction-level dynamic factor approach to the decomposition of export volatility

Matteo Barigozzi ^a
Angelo Cuzzola ^b
Marco Grazzi ^c
Daniele Moschella ^b

^a Department of Economics, Università di Bologna, Italy.

^b Institute of Economics and Department EMbeDS, Scuola Superiore Sant'Anna, Pisa, Italy.

^c Department of Economic Policy, Università Cattolica del Sacro Cuore, Milan, Italy.

2021/22

June 2021

ISSN(ONLINE) 2284-0400

Factoring in the Micro: a Transaction-level Dynamic Factor Approach to the Decomposition of Export Volatility^{*,†}

Angelo Cuzzola^{‡1}, Matteo Barigozzi², Marco Grazzi³, and Daniele
Moschella¹

¹Inst. of Economics, Scuola Superiore Sant'Anna

²Dept. of Economics, Università di Bologna

³Dept. of Economic Policy, Università Cattolica del Sacro Cuore

*Declarations of interest: none.

[†]This paper is the result of a long process that over time, at different stages, benefited from helpful comments at several conferences, including IAAE 2021 annual conference and CAED 2021. We are also indebted to Julien Martin and Marco Cococcioni for their insightful comments. This work has been partly supported by the European Commission under the H2020, GROWINPRO, Grant Agreement 822781 and by the Italian Ministry of Education and Research under the PRIN-2017 Programme, project code 201799ZJSN. This work is also supported by a public grant overseen by the French National Research Agency (ANR) as part of the 'Investissements d'avenir' program (reference: ANR-10-EQPX-17, Centre d'accès sécurisé aux données, CASD). M. Barigozzi acknowledges financial support from MIUR (PRIN 2017, Grant 2017TA7TYC). The usual disclaimers apply.

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

Abstract

This paper analyzes the sources of export volatility estimating a dynamic factor model on transaction-level data. Using an exhaustive dataset covering all French export transactions over the period 1993-2017, we reconstruct the latent factor space associated with global and destination-specific macroeconomic shocks through a Quasi-Maximum likelihood approach that allows to accommodate both the high share of missing values and the high dimensionality of the microeconomic time series. We then use the estimated model to provide a decomposition of the volatility of both aggregate and firm-level export growth rates, highlighting structural spatial patterns and drawing attention to geographical diversification's role in mitigating risks related to firms' export activities.

Keywords: Dynamic Factor Models, International Trade Volatility, Export Diversification

JEL classification: C38, L25, F14

1 Introduction

Finely disaggregated data have fostered a fast-growing body of research on microeconomic trade flows and their influence on firm-level and aggregate growth rates, with many empirical studies exploiting the granular information to investigate possible avenues of the micro-to-macro channel. However, as pointed by Armenter and Koren (2014), a proper model of the trade activities of firms and countries should always take into account some key structural properties of transaction-level datasets: the pervasive sparsity, intended as the number of zeros in trade flows at the microeconomic level, and the skewness of cross-sectional distributions. Those aspects have fundamental implications for the design and the interpretation of trade models and their stylized facts, gaining even more relevance when the focus lies on the dynamics of economic interactions. This is certainly the case of the volatility models based on the micro-level decomposition of the growth rates (di Giovanni et al., 2014; Kelly et al., 2013; Kramarz et al., 2020).

In this context, our paper presents a novel approach to the decomposition of the growth rates, based on the identification and estimation of global and local comovements for high dimensional time series with arbitrary patterns of missing data. The method requires the estimation of an approximate dynamic factor model (DFM) with a block structure (or block-DFMs, see e.g. Hallin and Liška, 2011; Moench et al., 2013) for the growth rates of export value at the firm-destination level for the universe of French exporting firms. DFMs endowed with a block structure feature both global and local factors to capture respectively the comovements common to all the time series and comovements within block of series. Thus, associating a block to each served destination, we provide an additive decomposition of the growth rates into three orthogonal terms: the global component, embodying the contribution of a global latent factor and the related loading, the destination-specific component, endowed again with a related loading and capturing the influence of the local factor associated to the target destination, and an idiosyncratic irreducible component.

We estimate the model via a Quasi-maximum Likelihood (QML) approach implemented through the Expectation Maximization (EM) algorithm (Watson and Engle, 1983; Doz et al., 2012), in line with the applications proposed by Bańbura and Modugno (2014). We extend their work giving the explicit derivation of the M-step for models with a block structure and proposing a suitable initialization procedure based on the sequential least square estimator of Breitung and Eickmeier (2015), recovering consistent estimates of the global and local factors, the related loadings and the idiosyncratic terms for highly dimensional time series in presence of missing values. Macroeconomic applications of a similar estimation technique based on a block structure of the data are Coroneo et al. (2016) and Delle Chiaie et al. (2022) (see below for details).

This allows to improve on the extant export volatility decompositions along three main dimensions. First, in contrast to di Giovanni et al. (2014); Kramarz et al. (2020), the macroeconomic determinants of the volatility — driven by global or destination-specific factors — are explicitly modelled as autoregressive stochastic processes using efficiently the information available at the microeconomic level. Second, DFMs capture the relevance of macroeconomic shocks, not only *per se* but also as drivers of heterogeneous responses at the firm level. Indeed, global and destination-specific components are defined as the product between the latent factors and the so-called factor loadings, i.e. numeric time-invariant coefficients specific to each firm-destination cell. These parameters represent the elasticity of firm-destination growth rates to common movements encoded in the global and local factors. Third, this exercise can be performed by increasing the time-frequency of the data without restricting the dataset to persistent exporters only, because the estimation technique efficiently tames the increasing number of missing values. We use here quarterly data on firm-destination transactions for 144k firms exporting to 67 destinations from 1993 to 2017, a period including relevant macroeconomic events (e.g. the trade collapse) and different phases of the French export cycles. These yearly quarter-to-quarter growth rates allow working on long time series, at the same time mitigating the bias of partial-year effects for

first-year export growth rates (Bernard et al., 2017).

We contribute to several streams of literature. *First*, we present a novel application of dynamic factor models to firm-level and transaction-level data. To the best of our knowledge, our work is the first to estimate a DFM on such disaggregated data, which allows us to identify the interactions between macroeconomic factors and heterogeneous agents. Our approach is aligned with the literature on block-DFMs, postulating global and local factors to capture hierarchical correlation structures within economic and financial datasets. Existing papers in this area have applied similar models to assess the relative importance of world, regional, and country-specific factors on countries' business cycles (Kose et al., 2012; Mumtaz and Surico, 2012; Miranda-Agrippino and Rey, 2020), to decompose the variance of commodity price indexes taming the local cross-correlation within groups (Delle Chiaie et al., 2022), and to impose restrictions on the loadings of nominal and real variables (Coroneo et al., 2016). However, our application differs in that it requires the estimation of a cross-section that is four orders of magnitude larger and accommodates a high share of missing values, reaching approximately 70% of the dataset. Thanks to the methodology we propose, the incomplete firm-destination time series need not be excluded or imputed and the growth rates' comovements can be estimated by exploiting all available information overcoming the concerns on the varying distribution of missing values at different time steps.

Second, acknowledging the prominent role of trade flows in contributing to GDP volatility (see di Giovanni and Levchenko, 2009 and Figure 1a and 1b), we provide new quantitative estimates of the *granular* component of the volatility of French exports, thanks to the decomposition of transaction-level growth rates. Starting from the seminal work by Gabaix (2011), a rich stream of literature has shown that in a *granular economy*, with a fat-tailed firm size distribution, idiosyncratic shocks to individual firms explain a significant part of the aggregate movements (Acemoglu et al., 2012; Carvalho and Gabaix, 2013; Carvalho and Grassi, 2019)¹. Those effects be-

¹ First highlighted for firms' size and aggregate GDP, that intuition applies to several economic phenomena (see e.g. Mésonnier and Stevanovic, 2017; Amiti and Weinstein, 2018, for the macroe-

come increasingly relevant in international trade (di Giovanni and Levchenko, 2009; di Giovanni et al., 2014; Di Giovanni et al., 2018; Bricongne et al., 2022), whereby the exporters’ size distribution is even more skewed (among the others, Bernard et al., 2009, 2016). In such a context, we use the estimated BDFM to provide a decomposition of the aggregate volatility into global, destination-specific and idiosyncratic components. The adopted methodology based on the Kalman smoother aggregates data in an optimal way (in mean squared sense) by taking cross-sectional and dynamic weighted averages of all observed time series. This is fundamentally different from taking sectoral (or destination-specific) averages to identify macroeconomic effects and isolate the granular component (Gabaix (2011); Carvalho and Gabaix (2013); di Giovanni et al. (2014); Bricongne et al. (2022)) since by dynamically aggregating the data we allow for leading-lagging relations among time series which account for feedback effects from the level of global and local factors to the data and viceversa (see Figure 2). This is directly reflected in a better approximation of the time-varying behaviour of the volatility measure itself. In fact, when considering the average volatility over the time span, our results confirm the general wisdom that attributes a dominant role to idiosyncratic shocks in explaining aggregate volatility; at the same time, our analysis highlights how the common component of the volatility can be as relevant as the idiosyncratic one during the economic downturns (e.g. the trade collapse).

Third, using the same decomposition, we analyze the volatility at the firm level providing new insights on the volatility-diversification nexus. Several contributions find a dampening effect of diversification on volatility (see e.g. Herskovic et al., 2020). In particular, firms selling multiple products to multiple destinations are those responsible for the largest share of a country’s export flows (Eaton et al., 2004; Bernard et al., 2012) and they could reduce their volatility by diversifying their portfolio (di Giovanni et al., 2014; Kramarz et al., 2020). We measure the effects of the three distinct components on the firms’ volatility distribution, showing how the global and the destination-specific terms generate a significant part of the risks inherent to ex-

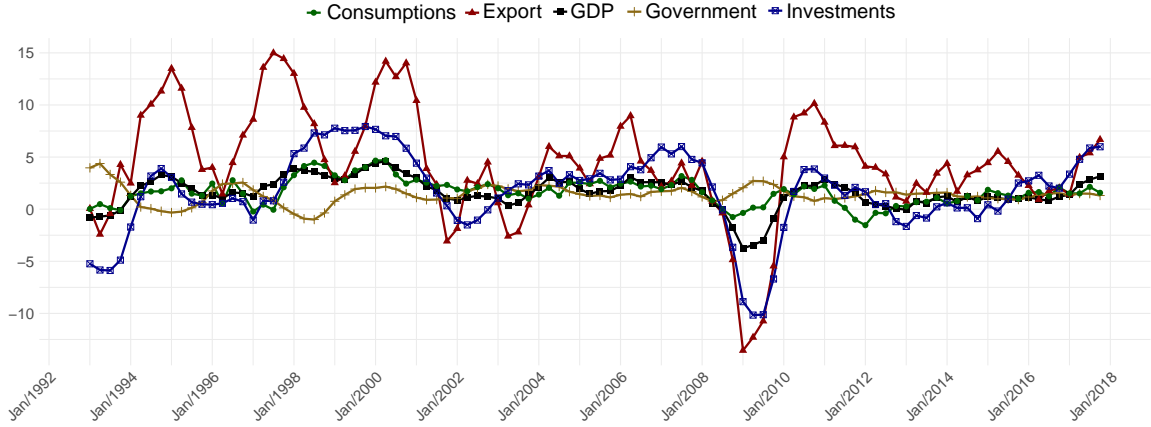
conomic effects of microeconomic shocks in the banking sector).

port growth, even though the impact of the idiosyncratic non-reducible components is relatively higher. The decomposition is then used to understand how and to what extent geographical diversification strategies help dampen export volatility overall and the single components on their own. We find that while firms diversifying across different destinations succeed in mitigating the risks associated with the macroeconomic cycle, a reverse U-shaped relation between diversification and idiosyncratic volatility suggests that the same strategies do not dampen idiosyncratic risks until a certain level of diversification is reached.

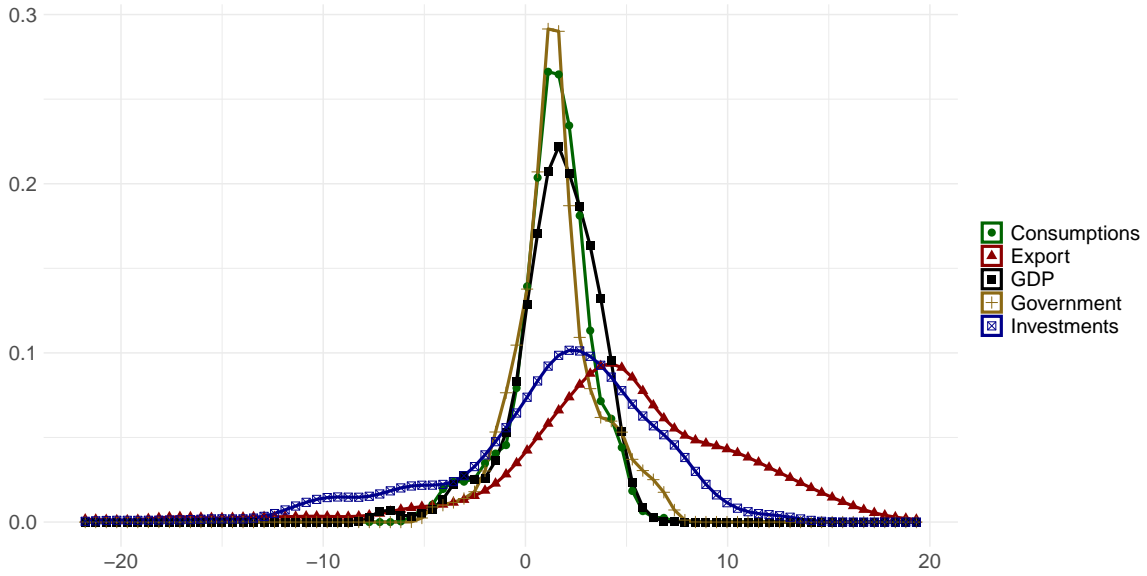
The remainder of the paper is structured as follows. Section 2 introduces the model and gives a brief and concise description of the methodology. Section 3 offers a bird's eye view of the dataset characteristics, focusing on sparsity, the distribution across destinations, and some firm-level statistics. Section 4 presents the reconstruction of the latent factor space and the volatility decomposition at the aggregate and firm levels. Section 5 concludes.

2 Model and methodology

This work provides an econometric framework to identify different sources of shocks affecting international trade flows. The methodology guarantees a high degree of flexibility as the structure that we impose *a priori* is kept to a minimum and is ultimately derived from the fundamental composition of the disaggregated transaction data (see Figure 3 below). The approach allows for identifying global and destination-specific components of the growth rates of the exports and their influence on aggregate and firm-level volatility. We build upon a widespread class of models that typically represent the growth rate of exporters as the sum of orthogonal terms, including macro shocks, estimated statically from each cross-section, and micro perturbations (see e.g. di Giovanni et al., 2014; Kramarz et al., 2020; Bricongne et al., 2022). In line with these static orthogonal decomposition models (SODMs), we recover the growth rates of the export sales directed to a given destination as the sum of three parts: two



(a) Time-series.



(b) Kernel densities.

Figure 1: Our elaboration on the OECD Main Economic Indicators (MEI) dataset, including quarterly deseasonalized growth rates (measure GYSA) for the GDP and its main components: GDP [NAEXKP01], Export [NAEXKP06], Investments [NAEXKP04], Governments [NAEXKP03], Consumptions [NAEXKP02]. Figure (a): time series of France’s quarterly deseasonalized growth rates. Figure (b): Gaussian kernel estimates of the pooled distribution of quarterly GRs for France, Italy, Germany and Spain from 1993 to 2018.

terms accounting for the macroeconomic effects of a global and a destination-specific component and a third microeconomic component that is specific to firm-destination pairs.

The proposed methodology improves upon existing SODMs in the characterization of the macroeconomic effects: we assume that firms-to-destination trade varies in re-

sponse to macroeconomic movements common to all the firms (global) or affecting only the firms exporting to a given destination (destination-specific). Those movements are represented by latent global and destination-specific factors and come endowed with their own dynamic specification. Moreover, the model is considerably richer since the dynamics of the microeconomic growth rates are driven by heterogeneous responses to the factors through elasticities (factor loadings) that depend on the specific match between firms and destinations. In practical terms, this decomposition is achieved by estimating a dynamic factor model with a block structure or block-DFM, induced by geographical export patterns. Once estimated, the growth rate decomposition is scaled up at different levels of aggregation, recovering the volatility decomposition for the total export, the export to each specific destination, and the firms' export.

Our proposed application of DFM also improves upon existing techniques in yet another way: the estimation of factors is based on the Kalman smoother, which aggregates data optimally (in mean squared sense) by taking cross-sectional and dynamic weighted averages of all observed time series. This is fundamentally different from taking simple averages as in Gabaix (2011), Carvalho and Gabaix (2013), di Giovanni et al. (2014), Bricongne et al. (2022) since by dynamically aggregating the data we allow for leading-lagging relations among time series which account for feedback effects from the level of global and local factors to the data and vice versa. Consider Figure 2, representing the information flow and the implied dynamics within the two frameworks. For DFMs, along the vertical dimension, information flows are bidirectional. In a two-step sequential estimation procedure, the estimated factors are used to determine the loadings that are then exploited to update the factor estimates until convergence. This interplay between the horizontal and vertical dimensions allows for a full-fledged dynamic decomposition and a much more efficient handling of the available information.

Model's equations Working with flows of exports at the firm-country level, we denote with $y_{de,t}$ the growth rate of the flow towards destination d by exporter e . We



Figure 2: Two stylized examples for the SODMs and the DFMs methodologies. The diagrams compare the information flowing directly (red arrows) or indirectly (blue arrows) in orthogonal decomposition models (ODMs) and dynamic factor models (DFMs) for a generic time series of growth rates $\{y_{i,t}\}$ (with i running over the number of flows). The symbols f_t and λ_i denote respectively a common latent factor, whose dynamics are described by an $AR(1)$ process, and the associated loadings. Common components in ODMs are estimated using the information from single cross-sections only, while in DFMs sequential time steps jointly concur to the formation of factors and loadings estimates.

specify a simple model, featuring both macroeconomic shocks and destination-exporter specific shocks, similar to di Giovanni et al. (2014). The model is described by the set of equations:

$$y_{de,t} = \lambda_{de} f_t + \rho_{de} g_{d,t} + \xi_{de,t} \quad (1)$$

$$f_t = a_f f_{t-1} + u_{f,t} \quad (2)$$

$$g_{d,t} = a_d g_{d,t-1} + u_{d,t} \quad (3)$$

With equation (1) we postulate the influence on $y_{de,t}$ of one latent factor common to all the flows, f_t , and one $g_{d,t}$ specific to the destination d . Factors are assumed to be orthogonal, and we model the dynamics of each factor as autoregressive processes of order one. The loadings λ_{de} and ρ_{de} reflect the response to the factor-related shocks, $u_{f,t}$ and $u_{d,t}$. Such loadings are specific to each firm-destination pair, thus embodying the possible heterogeneous response of exporters to global and destination-specific shocks.² Notice that, while the complete model features one global factor and D destination factors, if we naively restrict the model to the space spanned by the series directed to a given destination the model becomes a simple DFM with two factors, where the variance explained by the second factor (the destination-specific one in the complete

² The estimation of a model with AR processes of higher orders and with multiple common factors or multiple destination-specific factors is possible and technically feasible. Nevertheless, we have chosen a simpler structure to avoid parameters' proliferation.

representation) is residual with respect to that explained by the first one.

Estimation Dynamic factor models have an established tradition in macroeconomics. Their early applications were based on the exact factor structure assumption that all cross-correlations in the data could depend on a few common factors while the idiosyncratic noise remained cross-sectionally uncorrelated (Quah et al., 1993; Sargent et al., 1977). Unfortunately, this becomes an unrealistic hypothesis in the case of interest to our research question, where the dataset’s cross-sections are very large, favouring the emergence of idiosyncratic cross-sectional correlations. Estimation of such approximate factor structures has been originally proposed in Stock and Watson (2002) and Bai (2003) using classical principal component analysis and only global factors. Estimation of local factors via principal components has been studied by Onatski (2012), Breitung and Eickmeier (2015), and Freyaldenhoven (2021), among others. These works show how in presence of a factor structure, a high-dimensional dataset with high-dimensional blocks is precisely what allows to consistently estimate the global and local factors and their loadings even in presence of (limited) cross-correlation among idiosyncratic components, thus transforming a potential curse of dimensionality into a blessing.

More recently, Doz et al. (2012) proposed to estimate DFMs via Quasi-Maximum Likelihood (QML) implemented via the Expectation Maximization (EM) algorithm, a method originally proposed by Watson and Engle (1983), Shumway and Stoffer (1982), and Quah et al. (1993) for small size state space models. This approach is particularly suited to deal with missing data (Mariano and Murasawa, 2003; Marcellino and Schumacher, 2010; Bańbura and Modugno, 2014), yet it is widely used generally on financial and economic applications (see e.g. Coroneo et al., 2016; Delle Chiaie et al., 2022). Building upon the literature on EM estimation of DFMs, we extend the framework dealing with missing values and block structures simultaneously. We also propose an *ad hoc* initialization procedure, based on a least square sequential estimator proposed by Breitung and Eickmeier (2014, 2015) (see below and the Appendix A for the

details). In addition, notice that, since we work with huge cross-sections (approaching an order of magnitude of $\sim 10^6$ units) and tens of factors, the estimation entails a consistent memory overload, which cannot be tamed without an efficient handling of the sparse matrices of input data and loadings.³ In this respect, the missing values share, reaching the 70% (see Section 3 below for details), in the dataset allows to work with a portion of the dataset actually reduced from those implied by the formal definition of the model.

The EM algorithm consists of an iterated estimation procedure that converges towards the QML estimates of the parameters in a sequence of steps. We briefly recall here its fundamentals, referring to the Appendix A for a general description. Adopting the synthetic notation, where F denotes the factors, the other parameters⁴ as θ and the matrix of data as Y , we write the joint log-likelihood of data and factors as $l(Y, F, \theta)$, and iterate the following two steps:

1. Given an estimate of the parameters $\theta^{(k-1)}$, estimate the factors $F^{(k-1)}$ (jointly with their covariance) via the Kalman smoother, then compute the expected joint log-likelihood of Y and F :

$$L(\theta; \theta^{(k-1)}) = \mathbb{E}_{\theta^{(k-1)}} [l(Y, F^{(k-1)}, \theta) | \Omega_T],$$

where Ω_T denotes the σ -field generated by Y and the expectation is taken using the conditional distribution of the factors given Ω_T and the estimated parameters $\theta^{(k-1)}$.

2. Obtain an update of the parameters by solving:

$$\theta^{(k)} = \arg \max_{\theta} L(\theta; \theta^{(k-1)}).$$

³ See <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/ZKNTUA> for a publicly sample code.

⁴ Namely, the loadings matrix, the VAR coefficients and the relevant variance-covariance matrices.

This cycle defines a sequence of increasing log-likelihood values

$$l(Y, F^{(0)}, \theta^{(0)}) \rightarrow l(Y, F^{(0)}, \theta^{(1)}) \rightarrow l(Y, F^{(1)}, \theta^{(1)})$$

and stops when an appropriate convergence condition is fulfilled. While typically the algorithm is initialized by principal component estimates of loadings and factors, in order to account for block-specific factors, we propose to initialize it using the following iterative procedure: first, missing values are imputed using time-series medians and moving average smoother, then a sequential least square estimator proposed by Breitung and Eickmeier (2014, 2015) is applied on the ‘completed’ matrix to obtain the block-by-block parameters initialization. In this respect, whereas applied on imputed data, we can exploit the well established asymptotic properties of the initializing estimator (see propositions 2 and 3 of Choi et al., 2018).

Throughout, we employ a misspecified likelihood where the idiosyncratic components are treated as if they were both cross-sectionally and serially uncorrelated and normally distributed. This makes estimation fast and easy, and allows to have analytical expressions for the solutions at the maximization step. Nevertheless, it can be shown that, as both the total number of series, $\sum_{d=1}^D n_d$, and the sample size, T , grow to infinity, the consistency and efficiency of the estimated loadings and factors are not affected by such misspecifications. Moreover, the estimated factors are likely to be more efficient than those recovered by principal component analysis. We refer to Bai and Li (2016) and Barigozzi and Luciani (2019) for more details on the asymptotic properties of the estimators. Furthermore, we stress the robustness of the methodology to deviations from Gaussianity of the idiosyncratic terms and the factors’ innovations, collecting Monte Carlo evidence on the finite sample properties under this and other misspecifications (see Appendix B) and joining several empirical and numerical applications to leptokurtic or asymmetric distributed data (see e.g. Reis and Watson, 2010). These results justify the application to the growth rates of the export transactions whose distributional properties will be explored in the

following section. Finally, notice that the estimation is performed on standardized and demeaned series, then the estimated values are remapped to the original scales before proceeding with aggregation.

Dataset subsampling. For this specific application of DFMs, given the significant share of missing values, it is important to check that available information is sensibly conveyed for the estimation and identification of the factor space and the other aggregate volatility measures. Hence, we estimate the model on random subsamples of the original dataset. We first build a predefined number of reduced datasets (H), selecting at random a fixed number of firms (N_h) and keeping all the time-series associated with those firms. Then we apply the estimation procedure to each reduced dataset, obtaining H different estimates of the factors and the aggregate volatility coming from the decomposition (1). The final estimates for the factors and the aggregate volatility are thus averages of the H factors estimates and H volatility estimates and come along with the relative confidence intervals.

Volatility estimates and decompositions Our approach aims to the identification of idiosyncratic and macroeconomic shocks to export sales growth rates and to the estimation of their impact on the volatility of the aggregate. While the model outlined in (1)-(3) differs from the existing identifying methodologies adopted by Gabaix (2011), di Giovanni et al. (2014) and Kramarz et al. (2020), we apply an analogous strategy for the mapping between the microeconomic decomposition and the chosen aggregate.⁵ We report here a quick overview of the main lines.

Once estimated, the model main equation provides a decomposition of the logarithmic growth rates of each exporter-destination cell

$$y_{de,t} = \widehat{\lambda}_{de} \widehat{f}_t + \widehat{\rho}_{de} \widehat{g}_{d,t} + \widehat{\xi}_{de,t}. \quad (4)$$

⁵ For a thorough review of the possible strategies to recover the aggregates from microeconomic flows decompositions, see Amiti and Weinstein (2018).

These estimates will be used to assess the impact of any of the terms (or a combination of two) on the aggregate fluctuations. In order to recover the aggregate we will make use of size-based weights, encoding the share of the single exporter-destination cell within a given aggregation level:

$$\omega_{de,t}^{\text{agg}} = \frac{y_{de,t-1}}{\sum_{d,e} y_{de,t-1}}. \quad (5)$$

While summing up the contribution of the single flows to the aggregate, these weights can be chosen to be fixed (di Giovanni et al., 2014) or time-varying (Kramarz et al., 2020). Then one can recover the aggregate time series as:

$$\gamma_{t|\tau}^{\text{agg}} = \sum_{d,e} \omega_{de,\tau}^{\text{agg}} \left(\hat{\lambda}_{de} \cdot \hat{f}_t + \sum_d \hat{\rho}_{de,i} \cdot \hat{g}_{d,t} + \hat{\xi}_{de,t} \right) \quad (6)$$

$$\gamma_t^{\text{agg}} = \sum_{d,e} \omega_{de,t}^{\text{agg}} \left(\hat{\lambda}_{de} \cdot \hat{f}_t + \sum_d \hat{\rho}_{de,i} \cdot \hat{g}_{d,t} + \hat{\xi}_{de,t} \right) \quad (7)$$

where in the first equation we set up the weights to be fixed at a given time step τ , while in the second they are allowed to vary along the time series together with the growth rates components. To construct a proxy of the aggregate volatility, we will work with the variances and standard deviations of the quantities in (6) and (7). In particular we define the actual or aggregate variance as $\sigma_{\text{agg},\tau}^2 = \text{Var} \left(\gamma_{t|\tau}^{\text{agg}} \right)$ and those of the components as:

$$\begin{aligned} \sigma_{\text{glob},\tau}^2 &= \text{Var} \left(\sum_{d,e} \omega_{de,\tau}^{\text{agg}} \cdot \hat{\lambda}_{de} \cdot \hat{f}_t \right) & \sigma_{\text{glob}}^2 &= \text{Var} \left(\sum_{d,e} \omega_{de,t}^{\text{agg}} \cdot \hat{\lambda}_{de} \cdot \hat{f}_t \right) \\ \sigma_{\text{dest},\tau}^2 &= \text{Var} \left(\sum_{d,e} \omega_{de,\tau}^{\text{agg}} \cdot \sum_d \hat{\rho}_{de,i} \cdot \hat{g}_{d,t} \right) & \sigma_{\text{dest}}^2 &= \text{Var} \left(\sum_{d,e} \omega_{de,t}^{\text{agg}} \cdot \sum_d \hat{\rho}_{de,i} \cdot \hat{g}_{d,t} \right) \\ \sigma_{\text{idio},\tau}^2 &= \text{Var} \left(\sum_{d,e} \omega_{de,\tau}^{\text{agg}} \cdot \hat{\xi}_{de,t} \right) & \sigma_{\text{idio}}^2 &= \text{Var} \left(\sum_{d,e} \omega_{de,t}^{\text{agg}} \cdot \hat{\xi}_{de,t} \right) \end{aligned}$$

Let us emphasize a few points on the characteristics of these aggregate variances. First, note that the aggregation with fixed weights provides T different estimates of

the volatility, depending on the weight selected, so we will consider the time average as the point estimate for the whole time span. Second, the variance of the aggregate cannot be recovered as the simple sum of variances of the components, because of the covariances of the paired terms. In this respect, non-null covariances between the estimated components might emerge even though orthogonal decomposition models (dynamic or static) assume their independence in population. Following the convention in the literature, throughout the paper we measure volatility as the standard deviation and define the *relative standard deviations* or the *relative volatility* of the global, destination-specific and idiosyncratic components as, respectively, the ratios of the form $\sigma_{\text{glob},\tau}/\sigma_{\text{agg},\tau}$, $\sigma_{\text{dest},\tau}/\sigma_{\text{agg},\tau}$ and $\sigma_{\text{idio},\tau}/\sigma_{\text{agg},\tau}$.

In analogy with the formulas for the aggregation over all the export transactions, it is possible to generalize to different levels of aggregation. In this paper, we work both with destination-specific and firm aggregates. The former is obtained by aggregating the series targeting a specific destination d , thus the weighting can be restricted to the set, I_d , of flows targeting d ,

$$\omega_{e,t}^{(d)} = \frac{y_{de,t-1}}{\sum_{e \in I_d} y_{de,t-1}} . \quad (8)$$

On the same line, firm-specific volatility can be obtained as the standard deviation associated to the sum of the transactions of each exporter directed to any destination. Formally, taking the portfolio of the destinations for the exporter e (I_e), the weights become

$$\omega_{d,t}^{(e)} = \frac{y_{de,t-1}}{\sum_{d \in I_e} y_{de,t-1}} . \quad (9)$$

Both for destination and firm-level aggregates the same observations on the dynamic and static weighting apply.

3 Data and stylized facts

To estimate the outlined model, we rely on transaction-level exports recorded by the French customs office (*Direction Générale des Douanes et des Droits Indirects*,

DGDDI).⁶ The dataset contains detailed information on export flows on a monthly basis for each year from 1993 to 2017 for all French exporters. A unique official identification number identifies each exporter (SIREN code) and transactions report export value, quantity, country of destination, and an 8-digit product code following the European Union’s Combined Nomenclature (CN8). Our analysis relies on export values at the firm-country level. We start by applying standard cleaning methodologies described in Bergounhon et al. (2018). They include the harmonization of product codes, constructing a coherent chain of HS system’s labels, and homogenization of registered transactions. As to the latter, since the registration of the transactions below the threshold of 1000 euros (or 1000 Kgs) was not compulsory before 2010, we opted for the deletion of all the transactions below the threshold before and after 2010. In total, we dropped around 1.5 millions of firm-product-destination-month tetrads per year, accounting for a tiny fraction of the total export value (around 0.5%). This basic cleaning leaves an average value of export per year of euros 340.99 billions and, after aggregating along the product dimension, 3.2 millions of firm-destination pairs, which constitute the units of our analysis. We then aggregate monthly data into quarterly data and transform the panel of transactions into a matrix of time series, one for each firm-destination pair. Notice that the dataset includes the universe of (legal) transactions and it is originally provided in the so-called ‘long format’, where the id of each firm gets repeated as many times as the number of transactions in a given year. When creating a panel, we artificially generate missing values which we proceed to fill with zeros. Figure 3 offers a visual representation of this operation. Let us notice that, since our analysis focuses on the intensive margin of export flows, when estimating the model on logarithmic growth rates, the imputed zeros generate NAs. Their incidence and distribution need to be analyzed in order to proceed with estimation.

⁶ The data are directly provided to researchers by the DGDDI upon the approval of a research proposal by the *Comité du Secret Statistique*.

FirmID	Dest.	Value	Year	Quarter
00215	DE	6126	1993	04
00215	DE	8114	1994	02
00215	DE	1134	1994	03
00215	IT	474469	1993	03
00215	IT	127050	1993	04
00215	IT	55780	1994	03
00215	IT	357415	1994	04

Time flow

FirmID	Dest.	...	1993.03	1993.04	1994.01	1994.02	1994.03	1994.04	...
00215	DE	...	-	6126	-	8114	1134	-	...
00215	IT	...	474469	127050	-	-	55780	357415	...

Figure 3: The transformed series. From the table at the bottom, yearly growth rates are calculated on four points per years on a yearly basis, taking quarter-to-quarter logarithmic ratios.

Missing values and skewness along the country dimension The first relevant issue arising in the estimation of the model (1) regards the sparsity of the dataset: 89.99% of all observations are missing. Moreover, available points are unevenly distributed across firm-destination pairs: Figure 4 shows that on a log-log scale the distribution of available information follows a Pareto-like distribution. Within our interval span the share of missing values in the cross-section has minor variations over time, with a slightly improving situation in the more recent years. These changes along the time dimension do not pose serious issues for the estimation because the maximum observed spread is around 3%.

The second problem concerns the skewness along the country dimension. Among the 259 countries included in the original dataset, only a few of them are relevant for our analysis. For example, during the considered period only 67 countries report at least 1000 active flows at any cross-section (out of potential 3.2 million of flows).

Time frequency and country restrictions Given the properties of the dataset, the choice of the optimal time frequency and the selection of destination countries are of strategic relevance. Concerning the frequency, we construct yearly growth rates by taking quarter-to-quarter logarithmic growth rates on quarterly data in an attempt to enhance the standard volatility analysis in two different aspects. First, quarterly

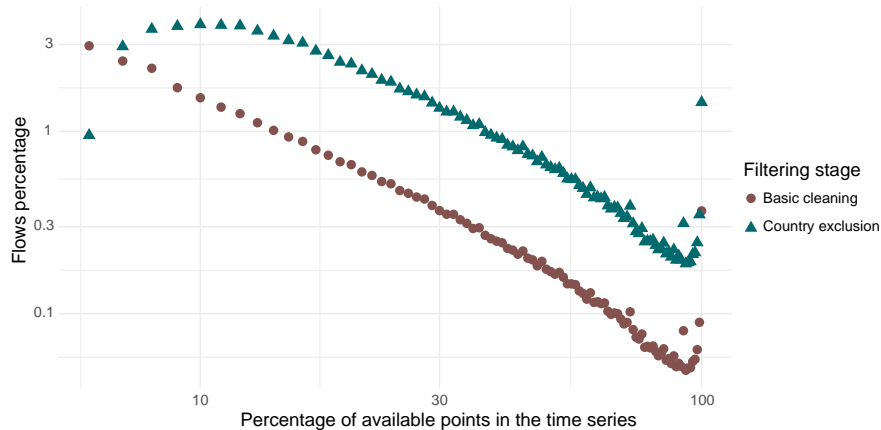


Figure 4: Percentage of active flows (y axis in \log_{10} scale) vs the percentage of available points in the time span, i.e. number of active quarters (x axis in \log_{10} scale). Two curves are proposed to compare the missing value distribution before and after the filtering procedure implemented along the country dimension.

data allow dealing with sufficiently long time series (96 points in yearly quarter-to-quarter growth rates) providing the proper dimensionality for the identification of common and destination factors. Notice that yearly quarter-to-quarter growth rates are also functional in removing at best the seasonality in each time series without adopting additional filtering procedures. Second, taking quarters in place of years reduces possible biases due to the so-called partial-year effect (see e.g. Bernard et al., 2017), which might lead to the overestimation of the growth rates between the first and the second year (and therefore of the associated volatility) because firms start exporting at different months during the first year of activity.

We next consider how to restrict the number of countries to exclude those least interested by French flows of exports. In this respect, a robust and consistent estimation of the destination-specific effects requires that factors have measurable impacts both at the macroeconomic and microeconomic levels. We keep in the dataset those countries that: i) are sufficiently represented in the firms' portfolios; ii) are relevant in terms of export value as a share of the total export. This leaves us with 67 destinations, accounting for 88.25% of the total export value. As we are ultimately interested in growth rates, we further drop firm-country flows that over the whole span report data on three points or less (over 96). We are finally left with

86.44% of the total export value and close to 900 thousand firm-destination pairs. Not surprisingly, the operated spatial restriction induces a reduction in micro-level sparsity to 70% (see Figure 4).⁷

4 Results

In this section, we outline an overview of the main results of our analysis, grouped into two main categories. First, at the aggregate level, we show how dynamic factor models provide a robust identification of macroeconomic comovements that, together with firm-destination specific loadings, serve as the primary tool for the volatility decomposition. Using the decomposition, we show how the volatility associated with specific destinations distributes along geographical patterns typical of gravity models for export flows. Second, at the firm level, we look at the distributions of the components of firms' volatility and then characterize the linkages between geographical diversification and volatility trends.

4.1 Volatility at the aggregate level

Factor space reconstruction. The global factor is shown in Figure 5 where it is compared with yearly quarter-to-quarter growth rates for French export from the FRED database. Simple visual inspection suggests a good level of agreement between the two independent measures of export growth, confirming that the comovements of the microeconomic export flows encode enough information to reconstruct the behaviour of aggregate statistics. Before moving ahead, we complete the consistency check implied by the sampling procedure applied. To check whether factors' sample estimates describe the same factor space we compute pairwise trace statistics with the

⁷ After imposing the restrictions, the number of firms in the dataset shrinks from around 167 thousands to 140 thousands.

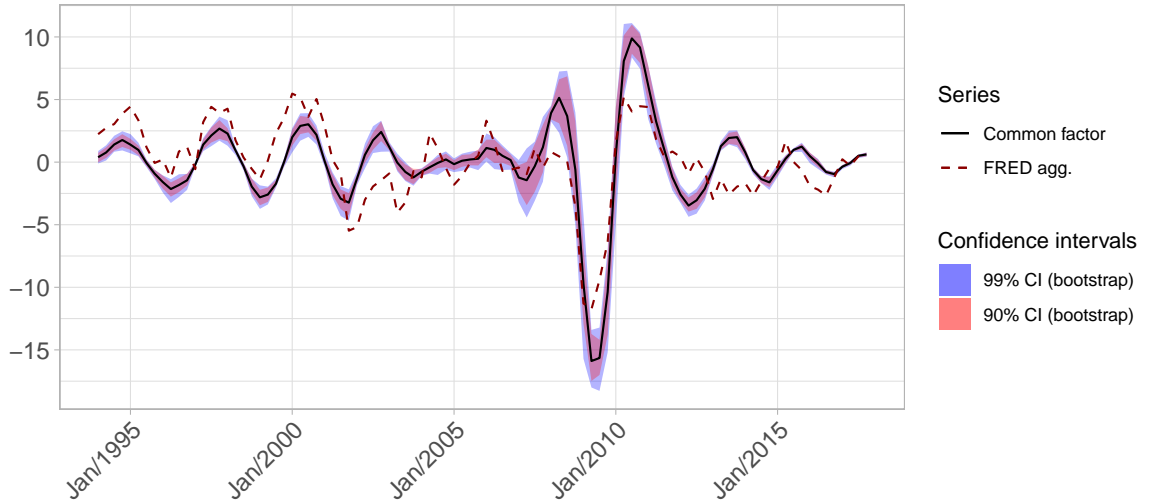


Figure 5: The identified global factor (black solid line) with 90% and 99% confidence intervals, compared with the aggregate growth rate of the French export from an independent source: our elaboration of the series from Organization for Economic Co-operation and Development, Exports: Value Goods for France [XTEXVA01FRQ664N], retrieved from FRED, Federal Reserve Bank of St. Louis; <https://fred.stlouisfed.org/series/XTEXVA01FRQ664N>, December 3, 2020.

formula⁸:

$$\text{Tr}_{(k,h)} = \frac{\text{Tr} \left(\widehat{F}^{(k)'} \widehat{F}^{(h)} \left(\widehat{F}^{(h)} \widehat{F}^{(h)'} \right)^{-1} \widehat{F}^{(h)'} \widehat{F}^{(k)} \right)}{\text{Tr} \left(\widehat{F}^{(k)'} \widehat{F}^{(k)} \right)}$$

The range of the trace statistics is $[0, 1]$ and different factor spaces tend to be closer when $\text{Tr}_{(k,h)}$ approaches the right limit. Trace statistics test for the equivalence of the factor spaces estimated by running the EM algorithm on each sample. The estimated matrices of factors from two different samples are compared by taking the related trace statistics. The pairwise computed values for 20 samples have a minimum of 0.96 and a maximum of 0.98, confirming a very good coherence of the different estimates.

Aggregate volatility and granularity. Before providing the point estimates of the aggregate volatility components, we present a comparison between the growth rates decomposition obtained via SODMs and DFMs. As a benchmark for the former we

⁸ F defines a $(D+1) \times T$ -matrix containing both the global and the destination-specific factors. With a slight abuse of notation, we use the indices k and h to denote different samples of the dataset, not to be confused with the indices denoting the steps of the EM steps.

replicate on our dataset the decomposition methodology of di Giovanni et al. (2014).⁹ For a proper comparison, we analyze the common term of our decomposition, joining the global and local factor and loadings (i.e the first two terms of equation 1), vis a vis the sector-destination shocks of di Giovanni et al. (2014) (the first term in equation 5, pag. 1309).¹⁰ Looking at Figure 6 we see in which direction our decomposition enriches the analysis of the variation in the export sales growth rates. Both estimates are in line with previous findings highlighting that the variation induced by the idiosyncratic shocks is dominant in magnitude over the common shocks for most of the time span. Nevertheless, it is worth emphasizing two relevant differences: i) the common components derived as the interplay between factors and the related firm responses are i) significantly more volatile than common components extracted as sectoral and destination averages; ii) this evidence is even more relevant across specific subsets of the time span, with some cases in which the common component isolated through DFMs approaches the idiosyncratic shocks in magnitude. These results highlight the flexibility of DFMs in capturing non-trivial effects of the macroeconomic components on the export growth rates at microeconomic level. This evidence leads to a partial review of the standard results that most shocks hitting firms are firm-specific, suggesting instead that also shocks common to all the firms can generate significant variations in the growth rates because firms' reactions are highly heterogeneous. Not surprisingly, such a mechanism becomes particularly relevant during deep downturns and rapid hypes.

After a first exploration of the decompositions of the growth rates, we move to the analysis of the aggregate volatility to assess the differential impact of the decompositions. Using equation (6) the aggregate volatility estimates are determined not only by the variation in growth rates components but also by the possible synchronization

⁹ The algorithms are adapted from those available at the link https://julian.digiovanni.ca/Papers/FirmGranular_replication.zip.

¹⁰ In the following we will always specify if the calculations include both sectors and destinations effect or only one of the two.

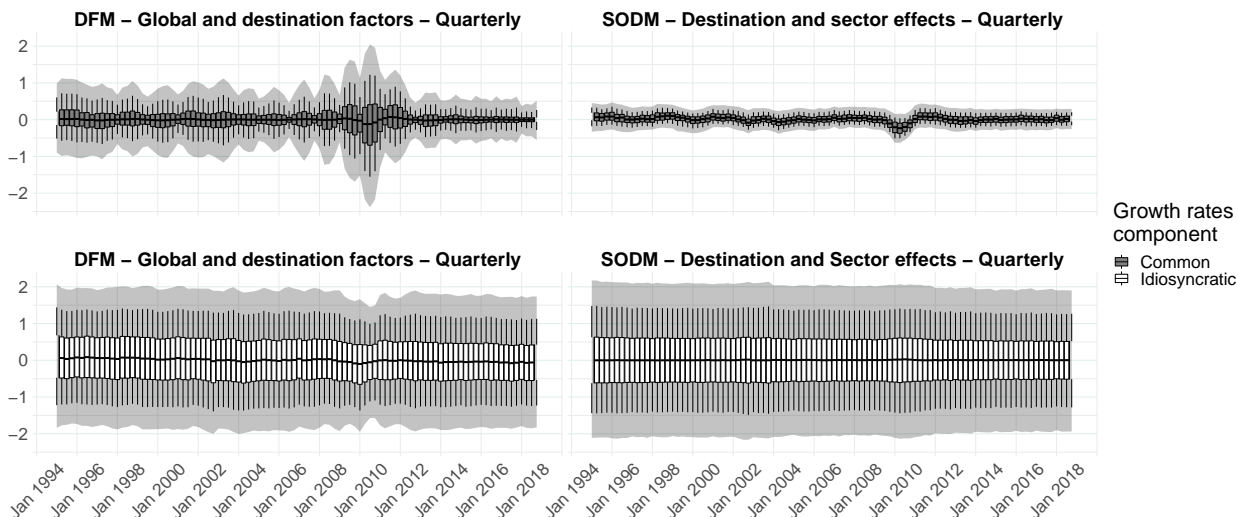


Figure 6: The direct comparison of the aggregate volatility common (top) and idiosyncratic (bottom) components obtained via dynamic factor models (top left) or the static orthogonal decomposition models of di Giovanni et al. (2014) under different specifications, varying frequency and macroeconomic effects. Solid lines denote the point estimate at each time step, while confidence intervals are constructed using the components’ median (box centers), the first and third quartiles (box extremes) and the first and last deciles (background shadowed area).

with the changes in the distribution of the size-based weights. In this respect, Figure 7 and Table 1 provide the aggregate volatility estimates respectively at each time step and considering all the time span. The firm-destination specific component accounts for the 0.84 of the actual volatility; the global and destination-specific volatility components for the 0.30 and the 0.18 respectively. When both the business cycle terms are combined to form the common component they reach a relative share of 0.37 over the time span, in line with the outcomes of the SODM, setting around the 0.36 and 0.31 depending if both sector and destination effects are included or destination effects only. As before, the comparison shows a significant difference when we zoom into the details of the time variations. In fact, in our model the dynamic of the volatility is not only driven by the variation of the relative size of the exporters or the exporter-destination cells (as measured by the weights), as in di Giovanni et al. (2014), but also by the significant variation in the growth rates and in particular by their synchronization. This might induce relevant changes in the estimates of the aggregate volatility as the time span changes: note, for example, how the common component of the volatility surged during the trade collapse matching in magnitude the idiosyncratic one. These

differences do not depend on the time-frequency nor on the set of effects that are included in the benchmark decomposition (compare the four plots in Figure 7).

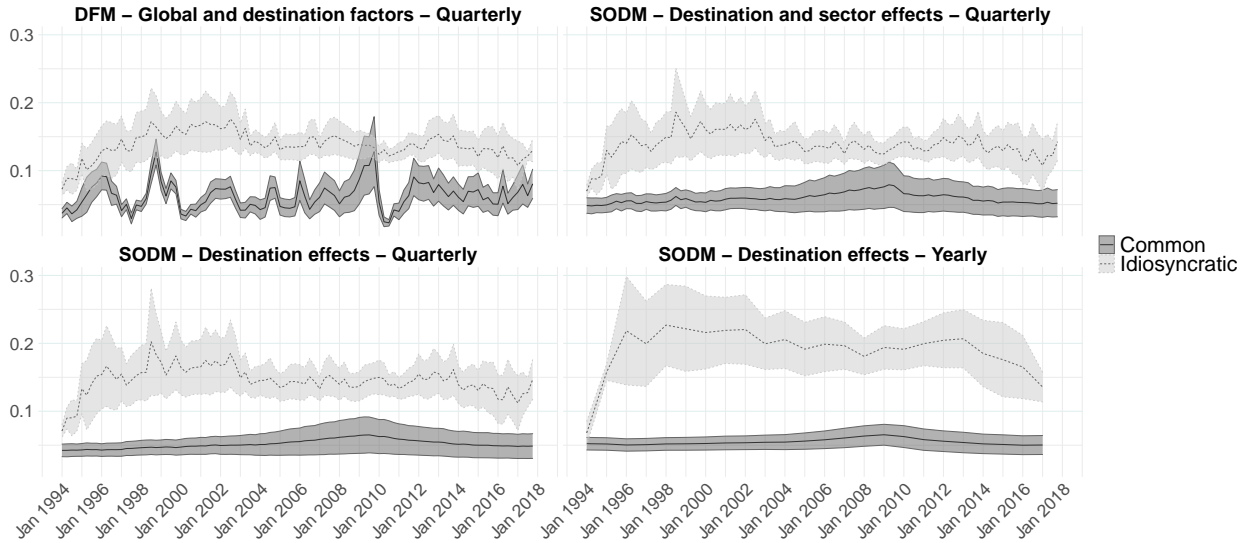


Figure 7: The direct comparison of the aggregate volatility estimates obtained decomposing the growth rates via DFMs or the SODMs. The latter is presented under two different specifications, namely including destination and sector terms in the common component, or destination effects only. The SODM is estimated both on quarterly and yearly data to highlight the impact on the estimates of a change in the measured frequency. Shaded areas outline the 95% analytical confidence intervals as provided in di Giovanni et al. (2014).

Figure 8 reports the main result of the paper: the time series of the estimates of the components of the volatility associated to growth rates decomposition outlined in (1)-(3). The common component is replaced by its subcomponents, the global and destination specific, and presented together with the idiosyncratic one. Destination and global components are comparable in magnitudes except for a few spikes of the global component standing out during downturns of the cycles. The idiosyncratic component is dominant in magnitude, as mentioned before. This pattern is reflected in the point estimates of the volatility associated to the time-span in table 1, top panel.

Destination-specific patterns. Looking at the destination-specific volatility associated to the aggregate French export, we can identify gravity-like patterns showing that country specific effects mitigate the risks attached to high volatile trade connections. Figure 9 shows a simple bivariate relation between the measure of volatility associated to a given destination and its GDP level. In the right panel, where the

Constant weighted aggregation					
	DFM (sampled)	SODM			
	Dest. (Quarterly)	Dest. (Quarterly)	Dest. (Yearly)	Dest. + Sec. (Quarterly)	Dest. + Sec. (Yearly)
Common	0.3753 (0.3579,0.3913)	0.3173	0.2608	0.3630	0.2913
Global	0.3049 (0.2594,0.3405)				
Destination	0.1800 (0.1539,0.2063)				
Idiosyncratic	0.8444 (0.8369,0.8506)	0.8915	0.9073	0.8498	0.8722

Dynamic weighted aggregation					
	DFM (sampled)	SODM			
	Dest. (Quarterly)	Dest. (Quarterly)	Dest. (Yearly)	Dest. + Sec. (Quarterly)	Dest. + Sec. (Yearly)
Common	0.7853 (0.7431,0.8188)	0.6484	0.5414	0.7171	0.5786
Global	0.684 (0.5992,0.7651)				
Destination	0.3683 (0.3150,0.4399)				
Idiosyncratic	0.7873 (0.7653,0.8106)	0.7904	0.5367	0.7235	0.5108

Table 1: The volatility explained by each component in relation to the actual aggregate volatility as measured by the DFMs and the benchmark SODMs. The statistics and confidence intervals for the DFM are computed out of 20 estimations of the same model over random subsamples of the original dataset (each subsample is constructed selecting 80k firms at random).

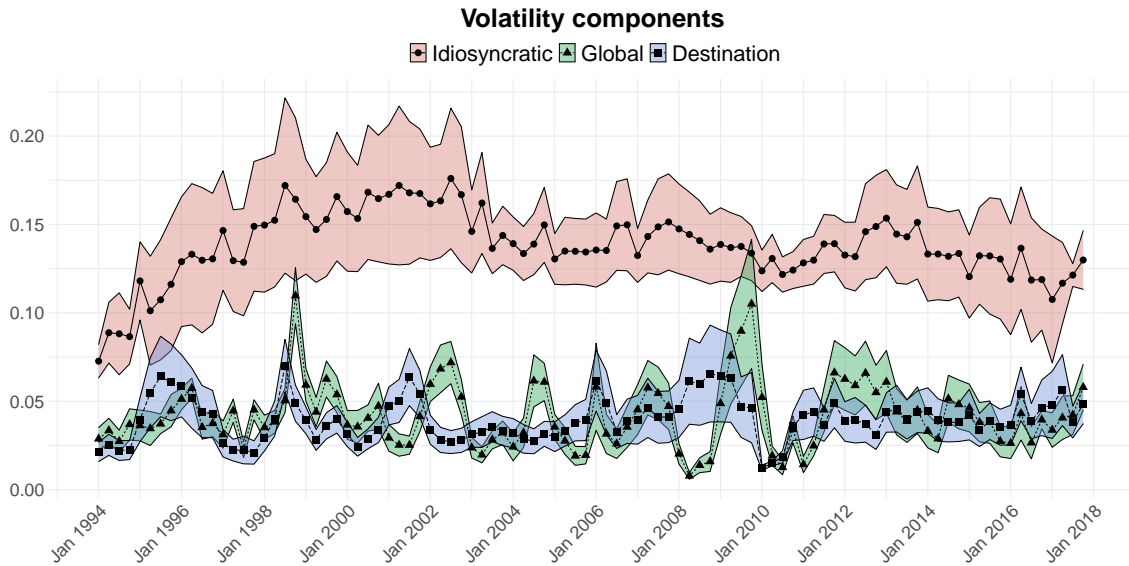


Figure 8: The estimates of the aggregate volatility components, global, destination-specific and idiosyncratic components obtained via dynamic factor models, aggregating with constant weights defined at different time steps (95% theoretical confidence interval derived as in di Giovanni et al. (2014)).

analysis is restricted to EU countries, a clear inverse correlation emerges between volatility and GDP. Indeed, the outliers, if any, are EU commercial partners joining

the European Union in the second half of our time window.¹¹

Similar considerations apply to the patterns on the left panel: an inverse relationship is apparent among the Extra-EU countries, which is however obscured by the presence of specific group of countries which are likely to experience less volatile trade flows from French firms, for a given level of GDP. This is the case, for example, of former French colonies status, which reduce the risks associated to trade relationships with low income countries.

To provide a more quantitative assessment of these trends, we ran a simple OLS regression of destination specific volatility on GDP, further adding controls accounting for geographical distance, free trade agreements, former French colonies, EU countries. Results, presented in Table 2, confirm the main finding from Figure 9: when accounting for additional covariates, the relationship between (log) GDP and (log) destination-specific volatility becomes negative and significant (see columns 2 and 3).

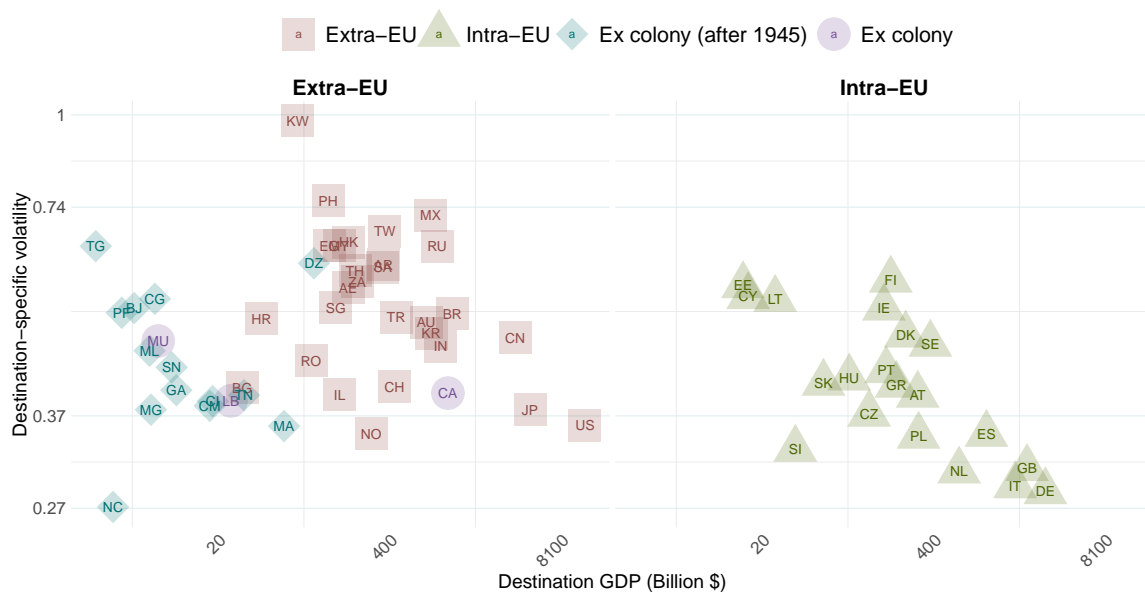


Figure 9: Destination-specific volatility vs destination GDP for extra-EU (left panel) and intra-EU (right panel) trade relations.

¹¹ Slovakia and Lithuania joined the EU single market in 2004, while the analyzed time window spans from 1993 and 2017.

	Log. dest. vol.		
	(1)	(2)	(3)
Log. dist.	0.123*** (0.029)	0.096*** (0.034)	0.097*** (0.031)
Log. GDP	0.004 (0.014)	-0.053*** (0.017)	-0.040** (0.017)
EU+Colonies controls	No	Yes	No
FTA+Colonies controls	No	No	Yes
Constant	-1.994*** (0.444)	-0.172 (0.553)	-0.513 (0.486)
Observations	65	65	65
R ²	0.231	0.460	0.501
Adjusted R ²	0.206	0.415	0.430
Residual Std. Error	0.218 (df = 62)	0.187 (df = 59)	0.185 (df = 56)
F Statistic	9.309*** (df = 2; 62)	10.064*** (df = 5; 59)	7.029*** (df = 8; 56)

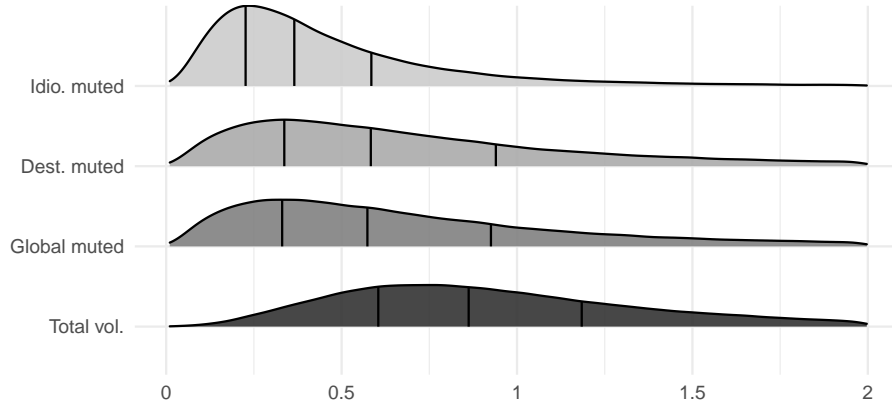
Note: *p<0.1; **p<0.05; ***p<0.01

Table 2: OLS estimates of gravity-like regressions for the volatility associated to each destination. GDP values and distances are the variables *distw* and *gdp_d* with France taken as origin country from the CEPII GeoDist database (available at http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id6 and described in Head et al., 2010; Mayer and Zignago, 2011). EU+Col controls include the variables *eu_d*, *col45* and, *colony*, whereas for FTA+Col *eu_d* is replaced by *fta_bb*; the latter includes information on the participation in a common market or in the same currency union.

4.2 Volatility at the firm level

In order to establish the role of the different components of volatility at the firm level, Figure 10 shows the distributions of firm-level volatility where one or two terms of Equation 4 have been set to zero before aggregating using the dynamic weights of the form (9), in the spirit of Kramarz et al. (2020). The global and destination-specific components have similar impacts on firms' volatility distribution, both in terms of magnitude and direction: once muted singularly, there is a visible shift to the left of the second and third quartile threshold and, to a lesser extent, of the first one; the effect is almost identical for both components. On the other hand, if we exclude the idiosyncratic component, we observe a relevant left-shift of all quartile thresholds and a substantial narrowing of the right tail. More precisely, the median reduction obtained removing the macroeconomic components is around 50% (from 0.78 to 0.37 or 0.39). In contrast, the impact of the microeconomic (idiosyncratic) contribution amounts to a dampening of 83% (from 0.78 to 0.13), confirming the prominent role

of the idiosyncratic component, yet showing that global and destination-specific terms have a non-negligible effect as drivers of volatility. This means that exporters, even though mostly exposed to idiosyncratic risks, face also the risk of relevant global and destination-specific shocks. In the remainder, the focus will move on the role of diversification in mitigating the effects of these shocks.



	10th pct.	25th pct.	50th pct	75th pct.	90th pct.
Total vol.	0.1754	0.3723	0.7835	1.5758	2.9223
Glo. muted	0.0351 (-80%)	0.1149 (-69%)	0.3714 (-53%)	1.1051 (-29%)	3.0645 (+5%)
Dest. muted	0.0371 (-79%)	0.1198 (-68%)	0.3881 (-50%)	1.1588 (-26%)	3.2063 (+ 9%)
Idio. muted	0.0182 (-89%)	0.0478 (-87%)	0.1301 (-83%)	0.3497 (-78%)	0.8725 (-70%)

Figure 10: Gaussian kernel densities for ‘counterfactual’ volatility distributions compared with the volatility distribution. On each panel a fit of the empirical distributions is obtained silencing the idiosyncratic component (top panel) and the macroeconomic effects: destination-specific component and global component (middle panels). The table below provides a quantitative assessment of the shift of the quantile thresholds for the same distributions.

Measuring diversification. To investigate the relationship between export growth volatility and firm diversification, we start by defining a set of diversification measures that have been used in the literature: the number of destination markets (Dest. Mkts.), the share of firm exports accounted for by the most important market (Top Share), and the Herfindahl-Hirschman Index (HHI) of export shares (see, among many others, Braakmann and Wagner, 2011; di Giovanni et al., 2014; Vannoorenberghe et al., 2016; Kramarz et al., 2020). We add to this indices also a diversification measure constructed using the inverse of the HHI. For each of these variables, we compute the firm average over time on a quarterly basis and then report basic descriptive statistics (Table 3).

We observe a consistent level of skewness in all the relevant distributions, in line with the previous evidence of Eaton et al. (2004) on French exporters.

	N Obs.	Mean	SD	Skew.	Kurt.	Median	1 st pc.	10 th pc.	90 th pc.	99 th pc.
Dest. Mkts.	143194	2.12	4.98	6.89	76.52	0.58	0.07	0.13	5.09	24.21
Top Share	143194	0.28	0.18	1.01	0.54	0.24	0.05	0.09	0.55	0.81
HH Ind.	143194	0.74	0.17	-1.16	1.21	0.79	0.23	0.50	0.92	0.96
Div. Ind.	143194	1.69	1.30	3.43	16.75	1.12	1.00	1.00	3.14	7.25

Table 3: Summary statistics for the distributions of four key diversification indicators: the number of destination markets, the share (of value) of the principal destination market in the firm’s portfolio, the Herfindahl–Hirschman Index (HHI) and a diversification index constructed using the inverse of the HHI. The synthetic information is computed as an average over the active time span for each single exporter.

	σ	Dest. Mkts.	Top Share	HH Ind.
Dest. Mkts.	-0.1627			
Top Share	-0.0893	0.1592		
HH Ind.	0.0489	-0.0292	-0.9824	
Div. Ind.	-0.2230	0.8022	0.0072	0.1268

Table 4: Correlation matrix of the diversification indicators. Different diversification measures and log weighted volatility.

Volatility components and diversification The preceding analysis has one main implication: the idiosyncratic component of the firm-level volatility has a prominent role compared to the macroeconomic ones, i.e the global component and the destination-specific component. How do diversification strategies help firms reduce trade risks related to the different components? Start by noting that standard portfolio theory would imply a negative relationship between firm-level volatility and the degree of diversification in the destination markets (Hirsch and Lev, 1971; Vannoorenberghe et al., 2016). Table 4, showing pairwise correlations among the different diversification indicators and firm-level volatility, provides a descriptive confirmation that more diversified firms tend indeed to experience less volatile export growth patterns.

To dig deeper into these correlations, we look at the distributional properties of the volatility components for classes of firms grouped by diversification quantiles, in Figures 11 (for all firms) and 12 (focusing only high and low volatile firms). First, the macroeconomic volatility components move together with a downward trend in

logs. Second, the idiosyncratic component moves along two opposite trends if we look at exporters that diversify below or above the median. Indeed, looking at the first half of the diversification spectrum, the volatility lies on a steady path, whereas on the second half moves along the expected inverse linear path, meaning that risk mitigation becomes relevant only after a certain threshold. This diversification limit is indeed relatively high for the population of French exporters, corresponding to a diversification index of around 3. On our data, the median value of the diversification index is 1.12, and only the 10% most diversified firms reach that limit (See Table 3, bottom). Figure 12, which focuses only on the most and the least volatile firms, additionally confirms that the idiosyncratic component dwarfs the other two in magnitude.

One concern raised by Figures 11 and 12 is that they do not control for size effects: as more diversified firms are also larger, the negative relationship between diversification and volatility could be due to an underlying size effect. To account for this, Figure 13 shows the relationship between volatility and its components and the residuals from a OLS regression of each of these components on size percentiles. The main message does not change: there is a sharp negative relationship between diversification and the common and destination-specific components of volatility. On the other hand, the idiosyncratic component is much less responsive to diversification.

Summing up, the risk exposure is reduced for firms that diversify their activities on the destination markets. Log-linear risk dampening effects seem to work only for shocks originated by macroeconomic induced fluctuations, with no remarkable difference between global and destination-specific shocks. On the contrary, the idiosyncratic component of the growth rate generates a volatility distribution at the firm level that does not change substantially while firms diversify more until a certain level. Beyond the threshold, diversification strategies give a consistent reduction helping firms to approach less volatile growth paths.

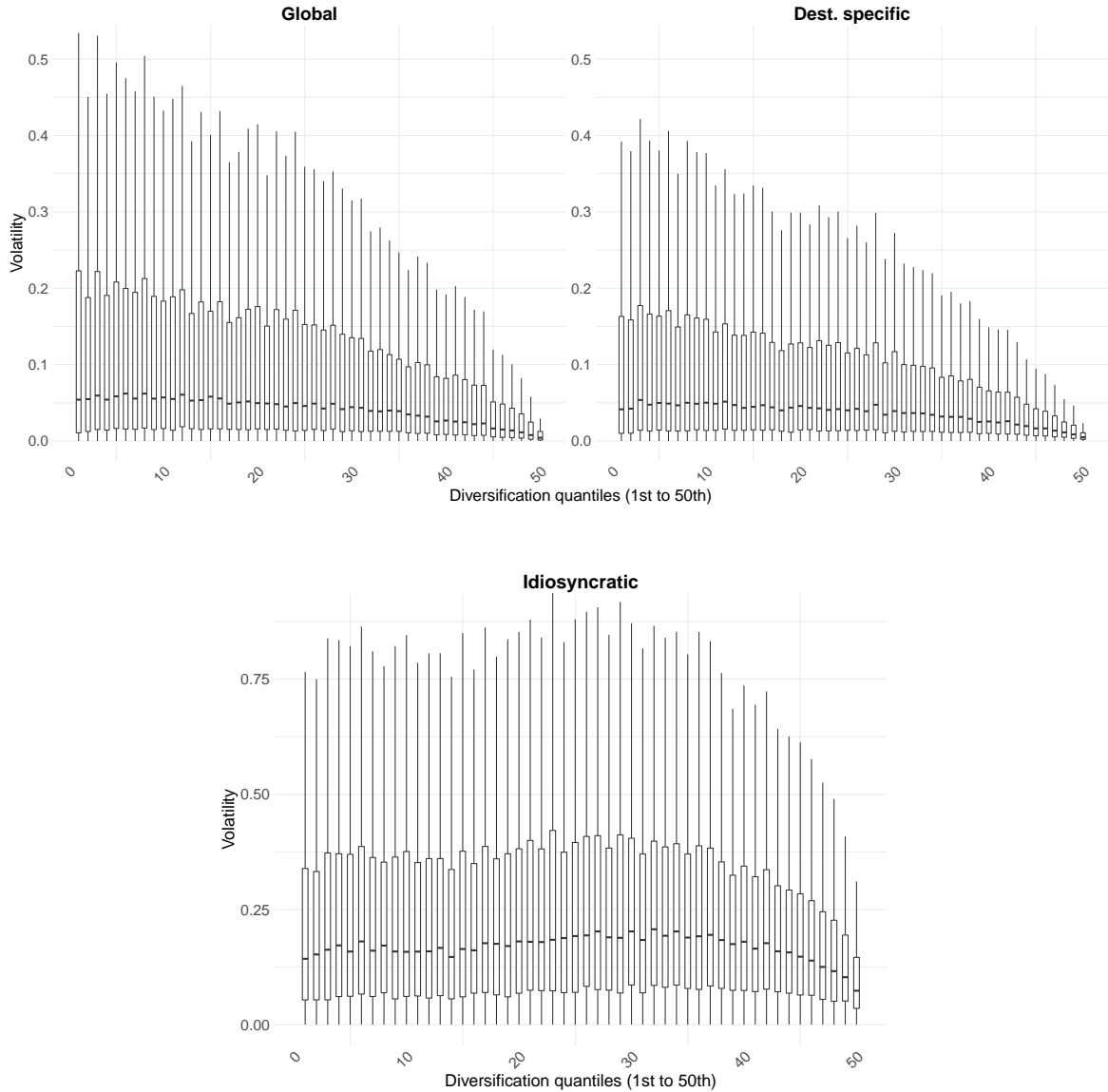


Figure 11: Diversification versus volatility components. The population of firms is divided into 50 groups, one for each 50tile of diversification indicator (Div. Ind. as per Table 3). The graphs display the boxplot associated with the distribution of the volatility components of each group, ordered from the least to the most diversified.

5 Conclusions

This paper proposes a dynamic factor model approach to the decomposition of aggregate and firm-level volatility. This allows to reconstruct the latent space of macroeconomic factors and decompose the growth rate of firm-destination cells into three orthogonal components: a global component, a destination-specific component and an idiosyncratic component. This provides the first application of dynamic factor models

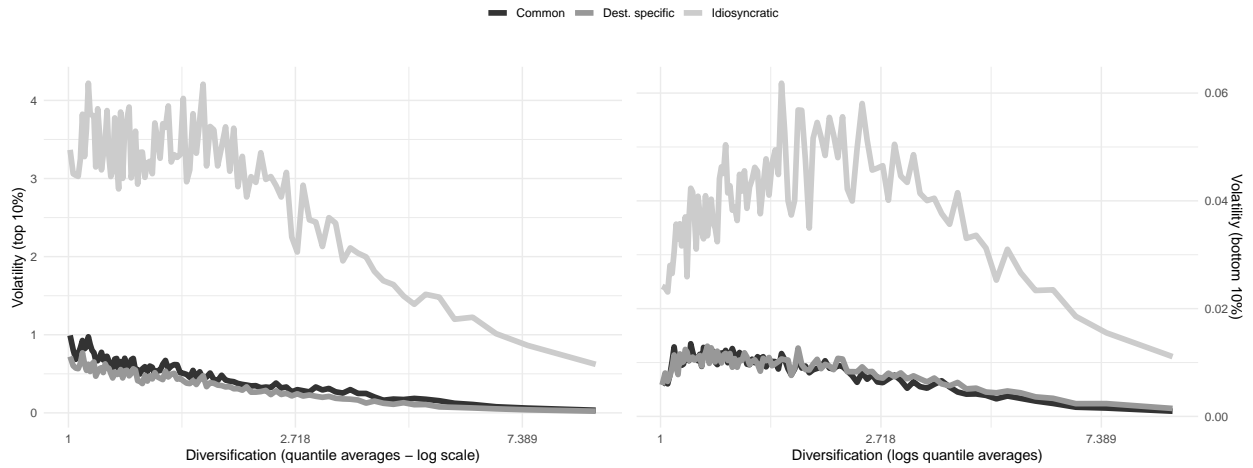


Figure 12: Diversification versus volatility, quantile-quantile plot. On the left panel, restricted to the most volatile firms (top 10%), the plot shows the three different volatility components vs percentiles of inverse Herfindahl index (average on the time window). On the right panel, the analogous graph for the least volatile firms (bottom 10%).

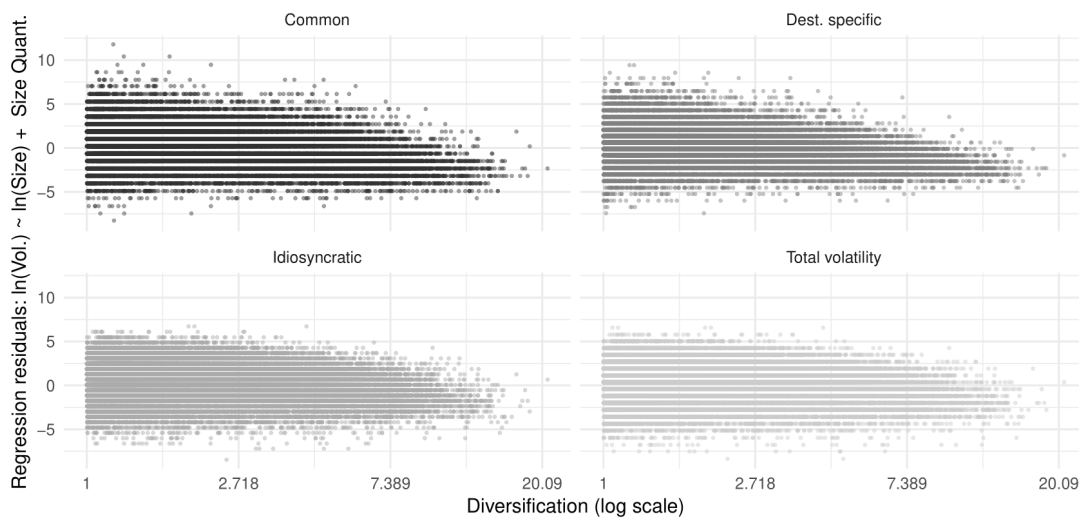


Figure 13: The residuals of the regression $\ln(\text{Vol.}) \sim \ln(\text{Size}) + \text{Size percentiles}$ plotted against the diversification index. As explained variable we take each component of the volatility and the volatility itself at the firm-level.

to transaction level data and requires an estimation based on the Expectation Maximization algorithm to accommodate both the prevalence of missing values and the high number of time series.

The decomposition is then mapped at the aggregate and at the firm level measuring the contribution of the three components to the total export's and firms' volatility. Our method gives new insights on the impact of the granular component of the aggregate export volatility and its measurement. In particular, we find that macroeconomic

shocks play a bigger role in generating aggregate volatility than it is usually recognized.

When analyzing the volatility associated with firms' growth, we show that the global and destination-specific components have a comparable effect on the first and second moments of the volatility distribution and show how diversification across destination markets can protect firms from shocks coming from macroeconomic events but seems to have contrasting effects on the idiosyncratic shocks.

References

- ACEMOGLU, D., V. M. CARVALHO, A. OZDAGLAR, AND A. TAHBAZ-SALEHI (2012): “The network origins of aggregate fluctuations,” *Econometrica*, 80, 1977–2016.
- AMITI, M. AND D. E. WEINSTEIN (2018): “How much do idiosyncratic bank shocks affect investment? Evidence from matched bank-firm loan data,” *Journal of Political Economy*, 126, 525–587.
- ARMENTER, R. AND M. KOREN (2014): “A balls-and-bins model of trade,” *American Economic Review*, 104, 2127–51.
- BAI, J. (2003): “Inferential theory for factor models of large dimensions,” *Econometrica*, 71, 135–171.
- BAI, J. AND K. LI (2016): “Maximum likelihood estimation and inference for approximate factor models of high dimension,” *Review of Economics and Statistics*, 98, 298–309.
- BAÑBURA, 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*.
- BERGOUNHON, F., C. LENOIR, AND I. MEJEAN (2018): “A guideline to French firm-level trade data,” Tech. rep., mimeo Polytechnique.
- BERNARD, A. B., E. A. BOLER, R. MASSARI, J.-D. REYES, AND D. TAGLIONI (2017): “Exporter dynamics and partial-year effects,” *American Economic Review*, 107, 3211–28.

- 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.
- (2012): “The Empirics of Firm Heterogeneity and International Trade,” *Annual Review of Economics*, 4, 283–313.
- (2016): “Global Firms,” CEP Discussion Papers dp1420, Centre for Economic Performance, LSE.
- BRAAKMANN, N. AND J. WAGNER (2011): “Product diversification and stability of employment and sales: first evidence from German manufacturing firms,” *Applied Economics*, 43, 3977–3985.
- BREITUNG, J. AND S. EICKMEIER (2014): “Analyzing Business and Financial Cycles Using Multi-Level Factor Models,” *Bundesbank Discussion Paper*.
- (2015): “Analyzing business cycle asymmetries in a multi-level factor model,” *Economics Letters*, 127, 31–34.
- BRICONGNE, J.-C., J. CARLUCCIO, L. G. FONTAGNÉ, G. GAULIER, AND S. STUMPFNER (2022): “From Macro to Micro: Large Exporters Coping with Common Shocks,” *CESifo Working Paper*.
- CARVALHO, V. AND X. GABAIX (2013): “The Great Diversification and Its Undoing,” *American Economic Review*, 103, 1697–1727.
- CARVALHO, V. M. AND B. GRASSI (2019): “Large firm dynamics and the business cycle,” *American Economic Review*, 109, 1375–1425.
- CHOI, I., D. KIM, Y. J. KIM, AND N.-S. KWARK (2018): “A multilevel factor model: Identification, asymptotic theory and applications,” *Journal of Applied Econometrics*, 33, 355–377.
- CORONEO, L., D. GIANNONE, AND M. MODUGNO (2016): “Unspanned macroeconomic factors in the yield curve,” *Journal of Business & Economic Statistics*, 34, 472–485.

- DELLE CHIAIE, S., L. FERRARA, AND D. GIANNONE (2022): “Common factors of commodity prices,” *Journal of Applied Econometrics*, 37, 461–476.
- DI GIOVANNI, J., A. LEVCHENKO, AND I. MEJEAN (2014): “Firms, Destinations, and Aggregate Fluctuations,” NBER Working Papers 20061, National Bureau of Economic Research, Inc.
- DI GIOVANNI, J. AND A. A. LEVCHENKO (2009): “Trade Openness and Volatility,” *The Review of Economics and Statistics*, 91, 558–585.
- DI GIOVANNI, J., A. A. LEVCHENKO, AND I. MEJEAN (2014): “Firms, Destinations, and Aggregate Fluctuations,” *Econometrica*, 82, 1303–1340.
- DI GIOVANNI, J., A. A. LEVCHENKO, AND I. MEJEAN (2018): “The micro origins of international business-cycle comovement,” *American Economic Review*, 108, 82–108.
- 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*, Oxford university press.
- EATON, J., S. KORTUM, AND F. KRAMARZ (2004): “Dissecting Trade: Firms, Industries, and Export Destinations,” *American Economic Review*, 94, 150–154.
- FREYALDENHOVEN, S. (2021): “Factor models with local factors—Determining the number of relevant factors,” *Journal of Econometrics*.
- GABAIX, X. (2011): “The Granular Origins of Aggregate Fluctuations,” *Econometrica*, 79, 733–772.
- HALLIN, M. AND R. LIŠKA (2011): “Dynamic factors in the presence of blocks,” *Journal of Econometrics*, 163, 29–41.

- HEAD, K., T. MAYER, AND J. RIES (2010): “The erosion of colonial trade linkages after independence,” *Journal of international Economics*, 81, 1–14.
- HERSKOVIC, B., B. KELLY, H. LUSTIG, AND S. VAN NIEUWERBURGH (2020): “Firm volatility in granular networks,” *Journal of Political Economy*, 128, 4097–4162.
- HIRSCH, S. AND B. LEV (1971): “Sales Stabilization Through Export Diversification,” *The Review of Economics and Statistics*, 53, 270–77.
- KELLY, B., H. LUSTIG, AND S. V. NIEUWERBURGH (2013): “Firm Volatility in Granular Networks,” NBER Working Papers 19466, National Bureau of Economic Research, Inc.
- KOSE, M. A., C. OTROK, AND E. PRASAD (2012): “Global business cycles: convergence or decoupling?” *International economic review*, 53, 511–538.
- KRAMARZ, F., J. MARTIN, AND I. MEJEAN (2020): “Volatility in the small and in the large: The lack of diversification in international trade,” *Journal of International Economics*, 122, 103276.
- MARCELLINO, M. AND C. SCHUMACHER (2010): “Factor MIDAS for nowcasting and forecasting with ragged-edge data: A model comparison for German GDP,” *Oxford Bulletin of Economics and Statistics*, 72, 518–550.
- MARIANO, R. S. AND Y. MURASAWA (2003): “A new coincident index of business cycles based on monthly and quarterly series,” *Journal of applied Econometrics*, 18, 427–443.
- MAYER, T. AND S. ZIGNAGO (2011): “Notes on CEPII’s distances measures: The GeoDist database,” *CEPII working paper*.
- MÉSONNIER, J.-S. AND D. STEVANOVIC (2017): “The macroeconomic effects of shocks to large banks’ capital,” *Oxford Bulletin of Economics and Statistics*, 79, 546–569.

- MIRANDA-AGRIPPINO, S. AND H. REY (2020): “US monetary policy and the global financial cycle,” *The Review of Economic Studies*, 87, 2754–2776.
- MOENCH, E., S. NG, AND S. POTTER (2013): “Dynamic hierarchical factor models,” *Review of Economics and Statistics*, 95, 1811–1817.
- MUMTAZ, H. AND P. SURICO (2012): “Evolving international inflation dynamics: world and country-specific factors,” *Journal of the European Economic Association*, 10, 716–734.
- ONATSKI, A. (2012): “Asymptotics of the principal components estimator of large factor models with weakly influential factors,” *Journal of Econometrics*, 168, 244–258.
- QUAH, D., T. J. SARGENT, ET AL. (1993): “A Dynamic Index Model for Large Cross Sections,” *NBER Chapters*, 285–310.
- REIS, R. AND M. W. WATSON (2010): “Relative goods’ prices, pure inflation, and the Phillips correlation,” *American Economic Journal: Macroeconomics*, 2, 128–57.
- SARGENT, T. J., C. A. SIMS, ET AL. (1977): “Business cycle modeling without pretending to have too much a priori economic theory,” *New methods in business cycle research*, 1, 145–168.
- SHUMWAY, R. H. AND D. S. STOFFER (1982): “An approach to time series smoothing and forecasting using the EM algorithm,” *Journal of time series analysis*, 3, 253–264.
- STOCK, J. H. AND M. W. WATSON (2002): “Forecasting using principal components from a large number of predictors,” *Journal of the American statistical association*, 97, 1167–1179.
- VANNOORENBERGHE, G., Z. WANG, AND Z. YU (2016): “Volatility and diversification of exports: Firm-level theory and evidence,” *European Economic Review*, 89, 216 – 247.

WATSON, M. W. AND R. F. ENGLE (1983): "Alternative algorithms for the estimation of dynamic factor, mimic and varying coefficient regression models," *Journal of Econometrics*, 23, 385–400.

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:

$$\begin{aligned}
 \mathbf{y}_t &= \mathbf{\Lambda}^G \mathbf{F}_t^G + \sum_{d=1}^D \mathbf{\Lambda}^d \mathbf{F}_t^d + \boldsymbol{\xi}_t \\
 \mathbf{F}_t^G &= \mathbf{A}^G(L) \mathbf{F}_t^G + \mathbf{u}_t^G \\
 \mathbf{F}_t^d &= \mathbf{A}^d(L) \mathbf{F}_t^d + \mathbf{u}_t^d \quad d = 1, \dots, D
 \end{aligned} \tag{10}$$

The model is analyzed allowing for an arbitrary number of common factors (K) and an arbitrary number of factors per block (K_d with $d = 1, \dots, D$)¹². 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}_t^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.

A.1 Estimation algorithm

Given the formulation (10) and assuming gaussianity of the idiosyncratic components and the innovations $\boldsymbol{\xi}_t \sim \mathcal{N}(0, \boldsymbol{\Sigma}^\xi)$, $\mathbf{u}_t^c \sim \mathcal{N}(0, \boldsymbol{\Sigma}^{\mathbf{u}^G})$, $\mathbf{u}_t^d \sim \mathcal{N}(0, \boldsymbol{\Sigma}^{\mathbf{u}^d})$, together with their mutual independence, the log-likelihood given the observed series and the latent

¹²The 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:

$$\begin{aligned}
l(Y, F; \theta) \simeq & -\frac{T-1}{2} \log |\boldsymbol{\Sigma}^{u^G}| - \frac{1}{2} \text{tr} \left[(\boldsymbol{\Sigma}^{u^G})^{-1} \sum_{t=2}^T (\mathbf{F}_t^G - \mathbf{A}^G \mathbf{F}_{t-1}^G) (\mathbf{F}_t^G - \mathbf{A}^G \mathbf{F}_{t-1}^G)' \right] \\
& - \sum_{d=1}^D \left[\frac{T-1}{2} \log |\boldsymbol{\Sigma}^{u^d}| + \frac{1}{2} \text{tr} \left[(\boldsymbol{\Sigma}^{u^d})^{-1} \sum_{t=2}^T (\mathbf{F}_t^d - \mathbf{A}^d \mathbf{F}_{t-1}^d) (\mathbf{F}_t^d - \mathbf{A}^d \mathbf{F}_{t-1}^d)' \right] \right] \\
& - \frac{T-1}{2} \log |\boldsymbol{\Sigma}^\xi| - \frac{1}{2} \text{tr} \left[(\boldsymbol{\Sigma}^\xi)^{-1} \sum_{t=1}^T (\mathbf{y}_t - \boldsymbol{\Lambda} \mathbf{F}_t) (\mathbf{y}_t - \boldsymbol{\Lambda} \mathbf{F}_t)' \right]
\end{aligned} \tag{11}$$

Where $\boldsymbol{\Lambda}$ is the $(N \times (r^G + \sum_d r^d))$ composed loading matrix: $\boldsymbol{\Lambda} = (\boldsymbol{\Lambda}^G \boldsymbol{\Lambda}^1 \dots \boldsymbol{\Lambda}^D)$ related to the composed factor vector $\mathbf{F}_t = (\mathbf{F}_t^G \mathbf{F}_t^1 \dots \mathbf{F}_t^D)$. We recall that $[\boldsymbol{\Lambda}^d]_{is} = 0$ 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

$$l(Y, F^{(0)}, \theta^{(0)}) \rightarrow l(Y, F^{(0)}, \theta^{(1)}) \rightarrow l(Y, F^{(1)}, \theta^{(1)})$$

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:

$$\Delta l_k = \frac{|l(Y, F^{(k+1)}, \theta^{(k+1)}) - l(Y, F^{(k)}, \theta^{(k)})|}{(|l(Y, F^{(k+1)}, \theta^{(k+1)})| + |l(Y, F^{(k)}, \theta^{(k)})|) / 2} \tag{12}$$

and stop the algorithm at $k = \bar{k}$ such that $\Delta l_{\bar{k}} \leq \varepsilon$, where ε is a predefined tolerance threshold¹³

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:

¹³ 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 $k_{max} = 100$.

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 \bar{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 quantities¹⁴: $\tau_v' \mathbf{F}_{v,t}^*$, $\tau_w' \mathbf{F}_{w,t}^*$. Thus, we have a first estimator of the global factors $\hat{\mathbf{F}}_t^G = (\bar{\tau}_v^{1'}, \mathbf{F}_{v,t}^*, \dots, \bar{\tau}_v^{r^G'}, \mathbf{F}_{v,t}^*)'$.
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 $y_t \sim \mathbf{\Lambda}^G \hat{\mathbf{F}}_t^G$.
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 estimates $\mathbf{F}_t^{\mathbf{G}^{(k-1)}}$, $\mathbf{F}_t^{\mathbf{d}^{(k-1)}}$, via the block-level regressions¹⁵:

$$\iota_d(\mathbf{y}_t) = \iota_d(\mathbf{\Lambda}^G) \hat{\mathbf{F}}_t^{G,(k-1)} + \iota_d(\mathbf{\Lambda}^d) \hat{\mathbf{F}}_t^{d,(k-1)} + \boldsymbol{\epsilon}_t^d. \quad (13)$$

one obtains the estimates of the relative factor loadings so that the estimated block-matrix can be composed $\hat{\mathbf{\Lambda}}^{(k)}$. 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 $\mathbf{\Lambda}^{(k)}$.

¹⁴The number of possible pairs for the CCA is $D^2(D-1)/2$. The problem is solved for each pair of blocks and then the pair that maximises the CCA is chosen.

¹⁵Here the operator $\iota_d(\cdot)$ is applied to an object with N rows to restricting to the n_d rows relative to block d .

The iteration procedure stops when a convergence condition analogous to (12) 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 $t = 1, \dots, T$. 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.

We have the filtering sequential equations:

$$\begin{aligned} \mathbf{F}_{t|t-1} &= \mathbf{A}^{(k)} \mathbf{F}_{t-1|t-1}; & \mathbf{P}_{t|t-1} &= \mathbf{A}^{(k)} \mathbf{P}_{t-1|t-1} \mathbf{A}^{(k)'} + \Sigma^{\mathbf{u}(k)} \\ \mathbf{F}_{t|t} &= \mathbf{F}_{t|t-1} + \mathbf{P}_{t|t-1} \Lambda_{\text{NA},t}^{(k)'} \mathbf{G}_t^{-1} \left(\mathbf{y}_t^{\text{NA}} - \Lambda_{\text{NA},t}^{(k)} \mathbf{F}_{t|t-1} \right) \\ \mathbf{P}_{t|t} &= \mathbf{P}_{t|t-1} + \mathbf{P}_{t|t-1} \Lambda_{\text{NA},t}^{(k)'} \mathbf{Q}_t^{-1} \Lambda_{\text{NA},t}^{(k)} \mathbf{P}_{t|t-1}' \end{aligned} \quad (14)$$

where $\mathbf{Q}_t^{(k)} = \Lambda_{\text{NA},t}^{(k)} \mathbf{P}_{t|t-1} \Lambda_{\text{NA},t}^{(k)'} + \Sigma^{\xi(k)}$. 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 $t = T, \dots, 1$ via the inverse recursion, starting from $\mathbf{F}_{T|T}$ and $\mathbf{P}_{T|T}$:

$$\begin{aligned} \mathbf{F}_{t|T} &= \mathbf{F}_{t|t} + \mathbf{P}_{t|t} \mathbf{A}^{(k)'} \mathbf{P}_{t+1|t}^{-1} (\mathbf{F}_{t+1|T} - \mathbf{F}_{t+1|t}) \\ \mathbf{P}_{t|T} &= \mathbf{P}_{t|t} + \mathbf{P}_{t|t} \mathbf{A}^{(k)'} \mathbf{P}_{t+1|t}^{-1} (\mathbf{P}_{t+1|T} - \mathbf{P}_{t+1|t}) \left(\mathbf{P}_{t+1|t} \mathbf{A}^{(k)'} \mathbf{P}_{t|t}^{-1} \right)' \end{aligned} \quad (15)$$

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 $\tilde{\mathbf{F}}_t = \mathbf{F}_{t|T}$, $\tilde{\mathbf{P}}_t = \mathbf{P}_{t|T}$ and

$$\tilde{\mathbf{P}}_{-1,t} = \mathbf{P}_{t|t} (\mathbf{P}_{t|t} \mathbf{A}^{(k)'} \mathbf{P}_{t|t}^{-1})' + (\mathbf{P}_{t|t} \mathbf{A}^{(k)'} \mathbf{P}_{t+1|t}^{-1}) (\mathbf{P}_{t|T} - \mathbf{A}^{(k)} \mathbf{P}_{t|t}) (\mathbf{P}_{t|t} \mathbf{A}^{(k)'} \mathbf{P}_{t|t}^{-1}) \quad (16)$$

The latter three quantities are then used to compute¹⁶:

$$\mathbb{E}_{\theta^{(k)}} \left[\mathbf{F}_t \mathbf{F}_{t-1}' | \Omega_T \right] = \sum_{t=2}^T \tilde{\mathbf{F}}_t \tilde{\mathbf{F}}_{t-1}' + \tilde{\mathbf{P}}_{-1,t}; \quad \mathbb{E}_{\theta^{(k)}} \left[\mathbf{F}_t \mathbf{F}_{t-1}' | \Omega_T \right] = \sum_{t=1}^T \tilde{\mathbf{F}}_t \tilde{\mathbf{F}}_{t-1}' + \tilde{\mathbf{P}}_{-1,t} \quad (17)$$

M-step. From latent factors to model parameters. Given the initial values, the EM algorithm is started. The proposed solution consist in the maximization of the expectation of the loglikelihood, given an ansatz of the parameters¹⁷. In practice, $\theta^{(k+1)}$ are the solutions to the system of first order conditions $\frac{\partial}{\partial \theta} \mathbb{E}_{\theta^{(k)}} [l(Y, F^{(k)}, \theta) | \Omega_T] = 0$ 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 get¹⁸:

$$\hat{\mathbf{A}}^{G,(k+1)} = \left(\sum_{t=2}^T \tilde{\mathbf{F}}_t^G \tilde{\mathbf{F}}_{t-1}^{G'} + \tilde{\mathbf{P}}_{-1,t}^G \right) \left(\sum_{t=2}^T \tilde{\mathbf{F}}_{t-1}^G \tilde{\mathbf{F}}_{t-1}^{G'} + \tilde{\mathbf{P}}_{-1,t}^G \right)^{-1} \quad (18)$$

$$\hat{\mathbf{Q}}^{G,(k+1)} = \frac{1}{T} \left(\sum_{t=2}^T \tilde{\mathbf{F}}_t^G \tilde{\mathbf{F}}_t^{G'} + \tilde{\mathbf{P}}_{-1,t}^G \right) - \frac{\hat{\mathbf{A}}^{G,(k+1)}}{T} \left(\sum_{t=1}^T \tilde{\mathbf{F}}_t^G \tilde{\mathbf{F}}_{t-1}^{G'} + \tilde{\mathbf{P}}_{-1,t}^G \right)' \quad (19)$$

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

¹⁶ In the following we will equally refer to the quantities restricted to the space of global and local factors: $\tilde{\mathbf{F}}_t^G, \tilde{\mathbf{F}}_t^d, \tilde{\mathbf{P}}_{-1,t}^G, \tilde{\mathbf{P}}_{-1,t}^d, \tilde{\mathbf{P}}_{-1,t}^{Gd}, \tilde{\mathbf{P}}_{-1,t}^{dG}$

¹⁷ 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.

¹⁸ We give the formulas only for the common factors, yet those for the local factors are equivalent.

of the first-order condition. One way to see this step is to calculate $\mathbf{\Lambda}$ as the NA-corrected OLS solutions of the block-by-block regressions (see Bańbura and Modugno, 2014, pag. 138, eq. (11))¹⁹²⁰

$$\mathbf{y}_t^b = \iota_b(\mathbf{y}_t) = \iota_b(\mathbf{\Lambda}^G)\widetilde{\mathbf{F}}_t^G + \iota(\mathbf{\Lambda}^d)\widetilde{\mathbf{F}}_t^d + \mathbf{v}_t \quad d = 1, \dots, D \quad (20)$$

$$\text{vec}(\iota_d(\mathbf{\Lambda}^G)|\iota_d(\mathbf{\Lambda}^d)) = \left(\sum_{t=1}^T \widetilde{\mathbf{F}}_t^{Gd} \widetilde{\mathbf{F}}_t^{Gd'} \otimes \text{Ind}_t^{\text{NA}} \right)^{-1} \text{vec} \left(\sum_{t=1}^T \mathbf{y}_t^{d,\text{NA}} \widetilde{\mathbf{F}}_t' \right) \quad (21)$$

where the matrix $\widetilde{\mathbf{F}}_t^{Gd} = (\widetilde{\mathbf{F}}_t^G \widetilde{\mathbf{F}}_t^d)$. This step concludes with the covariance matrix of idiosyncratic terms that is estimated only in its diagonal elements (in line with the eq.(12) at p. 138 of Bańbura and Modugno (2014)):

$$\begin{aligned} \text{diag} \widehat{\mathbf{\Sigma}}^{\xi^{(k+1)}} &= \frac{1}{T} \sum_{t=1}^T \left(\mathbf{y}_t^{\text{NA}} - \text{Ind}_t^{\text{NA}} \widehat{\mathbf{\Lambda}}^{(k+1)} \widetilde{\mathbf{F}}_t \right) \left(\mathbf{y}_t^{\text{NA}} - \text{Ind}_t^{\text{NA}} \widehat{\mathbf{\Lambda}}^{(k+1)} \widetilde{\mathbf{F}}_t \right)' + \\ &\quad \text{Ind}_t^{\text{NA}} \widehat{\mathbf{\Lambda}}^{(k+1)} \widetilde{\mathbf{P}}_t \widehat{\mathbf{\Lambda}}^{(k+1)'} \text{Ind}_t^{\text{NA}} + (\mathbf{I}_n - \text{Ind}_t^{\text{NA}}) \text{diag} \widehat{\mathbf{\Sigma}}^{\xi^{(k)}} (\mathbf{I}_n - \text{Ind}_t^{\text{NA}}) \end{aligned} \quad (22)$$

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

¹⁹ 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.

²⁰ Here again we denote $\iota_d(\mathbf{\Lambda}^d)$ as $\mathbf{\Lambda}^d$

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 (10) 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 $\kappa = 0.5^{21}$. 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 $d = 1 \dots, D$). 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. 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 $[\mathbf{\Lambda}^G]_{ij} \stackrel{iid}{\sim} \mathcal{N}(0, 1)$, while $[\mathbf{\Lambda}^d]_{ij} \stackrel{iid}{\sim} \mathcal{N}(0, 1)$ if $i \in I_b$ and $[\mathbf{\Lambda}^d]_{ij} = 0$ if $i \notin I_d$.
- The innovations to global and block-specific factors are simulated with normal, Student-t or Laplace distributions. Namely, $\mathbf{u}_t^G \stackrel{iid}{\sim} \mathcal{N}(\mathbf{0}_{r^G}, \mathbf{I}_{r^G})$ and $\mathbf{u}_t^d \stackrel{iid}{\sim} \mathcal{N}(\mathbf{0}_{r^d}, \mathbf{I}_{r^d})$, $\forall b$, or $\mathbf{u}_t^G \stackrel{iid}{\sim} t_3(\mathbf{0}_{r^G}, \mathbf{I}_{r^G})$ and $\mathbf{u}_t^d \stackrel{iid}{\sim} t_3(\mathbf{0}_{r^d}, \mathbf{I}_{r^d}) \forall b$, or $\mathbf{u}_t^G \stackrel{iid}{\sim} \mathcal{L}(\mathbf{0}_{r^G}, \mathbf{I}_{r^G})$ and $\mathbf{u}_t^d \stackrel{iid}{\sim} \mathcal{L}(\mathbf{0}_{r^d}, \mathbf{I}_{r^d})$, $\forall d$.
- The idiosyncratic components can be cross-correlated but the autocorrelation is not modelled. To measure the cross correlation the idiosyncratic components distributes as $\boldsymbol{\xi}_t \stackrel{iid}{\sim} \mathcal{N}(\mathbf{0}_n, \boldsymbol{\Sigma}^\xi)$ or $\stackrel{iid}{\sim} t_3(\mathbf{0}_n, \boldsymbol{\Sigma}^\xi)$ (or $\boldsymbol{\xi}_t \stackrel{iid}{\sim} \mathcal{L}(\mathbf{0}_n, \boldsymbol{\Sigma}^\xi)$), where the elements of the covariance matrix are such that $[\boldsymbol{\Sigma}^\xi]_{ii} \sim U[0.5, 1.5]$ and $[\boldsymbol{\Sigma}^\xi]_{ii} \sim \tau^{|i-j|}$, defining a Toeplitz matrix where the magnitude of the paired correlations is ruled by τ .

²¹ Here $\|\cdot\|$ denotes the Frobenius norm.

- The noise-to-signal ratio is controlled by scaling the common component $\tilde{\chi}_{i,t} = \frac{1}{\omega_i} \chi_{i,t} \sqrt{\frac{\text{Var}(\xi_{i,t})}{\text{Var}(\chi_{i,t})}}$ where ω_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 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 $\hat{\chi}_{it} = [\hat{\Lambda}' \hat{\mathbf{F}}_t]_i$ 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: $\text{MSE}_{\chi} = \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T (\hat{\chi}_{it} - \chi_{it})^2$. 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 5 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.

T	n	D	$\delta = 0 \tau = 0.1$			$\delta = 0 \tau = 0.4$			$\delta = 0 \tau = 0.8$		
			TS_{Λ}	$TS_{\mathcal{F}}$	$MSE_{\mathcal{X}}$	TS_{Λ}	$TS_{\mathcal{F}}$	$MSE_{\mathcal{X}}$	TS_{Λ}	$TS_{\mathcal{F}}$	$MSE_{\mathcal{X}}$
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	$\delta = 0 \tau = 0.1$			$\delta = 0.75 \tau = 0.1$			$\delta = 0.9 \tau = 0.1$		
			TS_{Λ}	$TS_{\mathcal{F}}$	$MSE_{\mathcal{X}}$	TS_{Λ}	$TS_{\mathcal{F}}$	$MSE_{\mathcal{X}}$	TS_{Λ}	$TS_{\mathcal{F}}$	$MSE_{\mathcal{X}}$
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

Table 5: The estimation evaluated with respect to the true model's parameters. The simulation parameters not explicitly stated are: $\mu = 0.5$, $\xi_t \sim \mathcal{N}(0, \Sigma^{\xi})$, $\omega = 0.5$, $r^G = 1$, $r^d = 1 \forall d$, $\eta = 0.2$. Blocks are homogeneous.

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 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 6 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 7 having in mind the absolute efficiency measures of table 5, 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 7, 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 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 closed-form solution of the estimators of the parameters, avoid proliferation of the parameters and 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 simulate the benchmark model imposing the distribution of the idiosyncratic components to be normal, Student-t or Laplace distributed. The efficiency of the estimator in this context

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

T	n	D	Homogeneous blocks			Heterogeneous blocks		
			TSR $_{\Lambda}$	TSR $_{\mathbf{F}}$	MSER $_{\chi}$	TSR $_{\Lambda}$	TSR $_{\mathbf{F}}$	MSER $_{\chi}$
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			Heterogeneous blocks		
			TSR $_{\Lambda}$	TSR $_{\mathbf{F}}$	MSER $_{\chi}$	TSR $_{\Lambda}$	TSR $_{\mathbf{F}}$	MSER $_{\chi}$
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

Table 6: 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: $\text{TSR}_{\Lambda} = \text{TS}_{\Lambda}^{\text{BQML}} / \text{TS}_{\Lambda}^{\text{EB}}$ and $\text{MSER}_{\chi} = \text{MSE}_{\chi}^{\text{EB}} / \text{MSE}_{\chi}^{\text{BQML}}$. The other parameters are fixed to $\mu = 0.5$, $\xi_t \sim \mathcal{N}(0, \Sigma^{\xi})$, $\tau = 0.1$, $\delta = 0$, $\omega = 0.5$, $r^G = 1$ and $r^d = 1 \forall d$.

T	n	$\tau = 0.3 \delta = 0$			$\tau = 0.3 \delta = 0.4$			$\tau = 0 \delta = 0.6$		
		TSR $_{\Lambda}$	TSR $_{\mathcal{F}}$	MSER $_{\mathcal{X}}$	TSR $_{\Lambda}$	TSR $_{\mathcal{F}}$	MSER $_{\mathcal{X}}$	TSR $_{\Lambda}$	TSR $_{\mathcal{F}}$	MSER $_{\mathcal{X}}$
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

T	n	$\tau = 0.6 \delta = 0$			$\tau = 0.6 \delta = 0.4$			$\tau = 0.6 \delta = 0.6$		
		TSR $_{\Lambda}$	TSR $_{\mathcal{F}}$	MSER $_{\mathcal{X}}$	TSR $_{\Lambda}$	TSR $_{\mathcal{F}}$	MSER $_{\mathcal{X}}$	TSR $_{\Lambda}$	TSR $_{\mathcal{F}}$	MSER $_{\mathcal{X}}$
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

Table 7: 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 6. The other parameters are fixed to be $\mu = 0.5$, $\xi_t \sim \mathcal{N}(0, \Sigma^\xi)$, $B = 5$, $\omega = 0.5$, $r^G = 1$ and $r^d = 1 \forall d$.

is analyzed both in absolute terms and relative to the EB estimator (see table 8). 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 5), and the gains with 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 5. In summary, for application to the dataset with Laplace distributed observations the estimator of factor models via BQML seems not to be problematic if the limiting conditions for the estimator holds true both considering the whole cross-section ($n \rightarrow \infty$) and only the series relative to each block ($n^d \rightarrow \infty$).

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

T	n	Laplace		Student-t (3)		Laplace		Student-t (3)	
		TS $_{\Lambda}$	TS $_{\mathcal{F}}$	TS $_{\Lambda}$	TS $_{\mathcal{F}}$	TSR $_{\Lambda}$	TSR $_{\mathcal{F}}$	TSR $_{\Lambda}$	TSR $_{\mathcal{F}}$
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: $\tau = 0.1 \quad \delta = 0.5$

T	n	Laplace		Student-t (3)		Laplace		Student-t (3)	
		TS $_{\Lambda}$	TS $_{\mathcal{F}}$	TS $_{\Lambda}$	TS $_{\mathcal{F}}$	TSR $_{\Lambda}$	TSR $_{\mathcal{F}}$	TSR $_{\Lambda}$	TSR $_{\mathcal{F}}$
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: $\tau = 0.5 \quad \delta = 0$

T	n	Laplace		Student-t (3)		Laplace		Student-t (3)	
		TS $_{\Lambda}$	TS $_{\mathcal{F}}$	TS $_{\Lambda}$	TS $_{\mathcal{F}}$	TSR $_{\Lambda}$	TSR $_{\mathcal{F}}$	TSR $_{\Lambda}$	TSR $_{\mathcal{F}}$
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: $\tau = 0.5 \quad \delta = 0.5$

T	n	Laplace		Student-t (3)		Laplace		Student-t (3)	
		TS $_{\Lambda}$	TS $_{\mathcal{F}}$	TS $_{\Lambda}$	TS $_{\mathcal{F}}$	TSR $_{\Lambda}$	TSR $_{\mathcal{F}}$	TSR $_{\Lambda}$	TSR $_{\mathcal{F}}$
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

Table 8: 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 6. Under Laplace: $\xi_t \sim \mathcal{L}(0, \Sigma^\xi)$ and $u_t \sim \mathcal{L}(0, \Sigma^u)$ (both for global and local factors). Under Student-t (3): $\xi_t \sim t_3(\mathbf{0}_n, \Sigma^\xi)$ and $u_t \sim t_3(0, \Sigma^u)$ (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 $r^d = 1 \forall d$.