DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

The dynamics of factor loadings in the cross-section of returns

Riccardo Borghi, Eric Hillebrand, Jakob Mikkelsen and Giovanni Urga

CREATES Research Paper 2018-38

Department of Economics and Business Economics
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

The dynamics of factor loadings in the cross-section of returns*

Riccardo Borghi[†] Eric Hillebrand[‡] Jakob Mikkelsen[§] Giovanni Urga[¶]

December 8, 2018

Abstract

In this paper, we propose a two-level factor model with time-varying loadings to investigate the dynamics of factor betas in the cross-section of returns of a large portfolio of 1815 firms from 54 countries over the period 2006-2016. The model contains a global observed financial factor and unobserved global and regional factors consistently estimated via principal component. When unexpected events happen globally, loadings on global factors increase. The dynamics of the global factor loadings is related to the profile of the firm. Loadings persistence is decreasing in firm size and expected returns are increasing in the variance of the loading.

JEL classification: *C32; C55; C38; G01; G15.*

Keywords: *High-dimensional factor models; financial, global and regional risk factors; time-varying loadings; systematic risk.*

*We wish to thank Giuliano De Rossi, Richard Payne, and participants in the "2018 Camp Econometrics XIII" (The Sagamore, 6-8 April 2018), in particular Ba Chu and Nicholas M. Kiefer, the "19th OxMetrics User Conference" (Paris, 11th September 2017), in particular Catherine Doz, Neil Ericsson, Siem Jan Koopman, and Sébastien Laurent, the "40th International Panel Data Conference" (Thessaloniki, 7th July 2017), the "18th OxMetrics User Conference" (London, 12th September 2016), in particular David Hendry, and the "2016 PhD Research Days" (London, 9th June 2016), in particular Ian Marsh and Enrique Schroth, for useful comments and stimulating discussions. The usual disclaimer applies. Riccardo Borghi acknowledges financial support from the "2015 - 2017 PhD Studentship in Memory of Ana Timberlake" and the Centre for Econometric Analysis.

[†]Centre for Econometric Analysis, Cass Business School, City, University of London, 106 Bunhill Row, London EC1Y 8TZ, UK. Email: Riccardo.Borghi.1@cass.city.ac.uk.

[‡]CREATES, Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark. E-mail: ehillebrand@econ.au.dk.

[§]Danmarks Nationalbank, Havnegade 5, DK-1093 Copenhagen K, E-mail: jgm@nationalbanken.dk.

[¶]Centre for Econometric Analysis, Cass Business School, City, University of London, 106 Bunhill Row, London EC1Y 8TZ, UK and Bergamo University, Italy. Email: G.Urga@city.ac.uk.

1 Introduction

The analysis of comovements between stock returns is at the heart of the empirical asset pricing literature. Portfolio managers invest internationally [Heston and Rouwenhorst (1995), Bekaert et al. (2009)], and this requires knowledge of relevant common factors and their importance for stock returns, both for risk management and for asset allocation purposes.

Searching for the most influential factors, various researchers have concluded that global, country- and region-specific factors are more important than industry factors in explaining the cross-section of expected returns [e.g., Heston and Rouwenhorst (1995), Griffin (2002), Bekaert et al. (2009), Fama and French (2017), Ando and Bai (2017), Barigozzi et al. (2018)]. Bekaert et al. (2009) show that the relevance of industry factors was a short-lived phenomenon. Nevertheless, factor loadings have been shown to vary over time, making it difficult for the investor to estimate the factor exposure of a firm. For instance, Fama and French (1997) remark that "...there is strong variation through time in the CAPM and the three-factor risk loadings of industries..." and this variation is even stronger for individual firms. Consequently, the relative importance of global versus group-specific factors is time-varying and difficult to estimate. So far the literature has assessed the time-varying relevance of global and regional factors using rolling window estimation [see, among others, Bekaert et al. (2009)], with results depending on a nuisance parameter such as the length of the window. Gagliardini et al. (2016) prove the consistency of a two-step estimation of a factor model with time-varying parameters that are function of stock specific and macroeconomic variables, as in Shanken (1990).

In the field of asset pricing, several researchers have estimated factor models with time-varying loadings but the results are based on models with observable factors (CAPM and Fama-French extensions). Some authors have assumed that the factor loading is a function of economic state variables [Robichek and Cohn (1974), Shanken (1990), Rosenberg and Guy (1995), Ferson et al. (2002), Santos and Veronesi (2004)], leading to the derivation of the conditional CAPM. Intuitively, a firm in distress is more likely to report low earnings when the economy is in a bad state. However, Ghysels (1998) and Lewellen and Nagel (2006), among others, point out that the loading estimates are highly dependent on the assumed information set and, in case of misspecification, the unconditional CAPM works better. To avoid this drawback, a recent strand of literature makes use of non-parametric estimation to retrieve factor loadings from high-frequency data [e.g. Bollerslev and Zhang (2003), Patton and Verardo (2012)]. Ang and Kristensen (2012) develop a non-parametric technique to estimate conditional betas in multifactor models, finding that also momentum and book-to-market decile portfolios have strong variation in factor loadings. However, non-parametric estimation by construction cannot capture the parameters driving the dynamics of the loadings, which have been shown to contain important information for asset prices. For instance, Armstrong et al. (2013) develop a model where uncertainty about a firm's loading is negatively related to expected returns. The main limitation of this literature is that the analysis of factor loadings highly depends on the prior identification of the factors. In particular, because the factors in the asset pricing theory are unknown, many authors focus solely on the loading of the market factor.

We introduce a two-level factor model with time-varying loadings to investigate the dynamics of factor loadings in a large global cross-section of returns. The model features global and regional as well as observable and unobservable factors, which are consistently estimated via principal component analysis. For each stock, the persistence and the variance of the factor loadings are estimated via maximum likelihood. The main aim of this paper is twofold. First, we estimate the unknown global and regional risk factors and their contribution to the overall variance in the portfolio. Thus, we contribute to the literature on time-series risk models [e.g. Ludvigson and Ng (2007)]. Second, we relate the dynamics (uncertainty and persistence) of factor loadings to expected returns and firm characteristics. Thus, we contribute to the literature on empirical asset pricing [e.g. Fama and French (2017)].

The empirical applications are based on a panel of 1815 stock returns from six world regions from January 2006 to January 2016. We formulate a factor model where stock returns are assumed to be a function of two types of factors: global (one observed financial factor and one latent non-financial factor) and region-specific (one latent factor per region). This modelling choice is motivated by Boivin and Ng (2006), who show that increasing the number of stocks, N, does not always help the estimation of the common factors when there is large cross-sectional correlation in groups of variables. Goyal et al. (2008) use the clustering of Nasdaq- and NYSE-listed stocks to identify common factors. Ando and Bai (2017) let the group membership of a stock be an unknown parameter to be estimated. The global financial factor in our model is the S&P500 Financials Index while the latent global and regional factors are estimated by principal components analysis (PCA). Furthermore, in our model the loadings of each factor vary over time, thus the model is an extension of Breitung and Eickmeier (2015). Principal components are a consistent estimator of the unknown factors in the presence of time-varying loadings for large panels with $N, T \to \infty$: Bates et al. (2013) prove average consistency in t, while Mikkelsen et al. (2018) prove uniform consistency in t if $\frac{T}{N^2} \to 0$ is satisfied.

There is a large body of empirical evidence in favour of modelling factor loadings as time-varying parameters. For instance, Stock and Watson (2009) find significant improvement in forecasting macro variables when coefficients are allowed to change after a structural break. Del Negro and Otrok (2008) and Eickmeier et al. (2015) estimate factor models with loadings modelled as random walks using large panels of macro data. Bates et al. (2013) show analytically that the principal component estimator remains consistent if the loadings are stationary experiencing a structural break, or if they are random walks of a restricted class. In this paper, we model the factor loadings as stationary processes, and we estimate them in a state-space framework following Mikkelsen et al. (2018), who prove that the maximum likelihood estimator (MLE) of the parameters of autoregressive loadings processes is consistent.

Several papers find that the factor exposure of a firm should only temporarily diverge from its long-run average. From a statistical point of view, Andersen et al. (2006) find that a stock's beta is best approximated by a stationary I(0) process, due to a cancellation in the ratio of covariance and market variance. From a corporate finance perspective, stationary fluctuations of factor loadings reconcile with theories of the firm. First, systematic risk is a decreasing function of investment because the level of investment increases with the availability of low risk projects [Berk et al. (1999), Cooper and Priestley (2011)], in alignment with pro-cyclical macroeconomic factors. Predictions from real options models [e.g. Carlson et al. (2006)] agree with this mechanism because undertaking a real investment can be considered as exercising a risky option, which decreases systematic risk. Second, the beta of a firm increases around earning announcements [Savor and Wilson (2016)] and before Seasoned Equity Offerings [Carlson et al. (2010)], while gradually decreasing afterwards. Finally, entertaining takeover bids temporarily modifies the beta of a firm according to the difference with the beta of the target, because the new entity will have an average beta of the two firms [Hackbarth and Morellec (2008)].

A final point is that factor models can be either estimated on portfolios, as proposed by Fama and French (1997), or on individual stocks. In this paper, we estimate the model at stock level, and this has two main advantages. First, it avoids the loss of information caused by grouping the stocks into portfolios when testing for the pricing of market anomalies, such as size or value [Ang et al. (2017), Gagliardini et al. (2016)]. Second, it allows to test if the dynamics of factor loadings (uncertainty and persistence) are related to expected returns. Lately, Armstrong et al. (2013) provide evidence of a negative relationship between factor loading uncertainty and future stock returns in a CAPM setting. In this paper, we extend this framework to the case of global and region latent factors.

Our empirical analysis yields a number of results that we can summarise as follows. First, using canonical correlation analysis, we find that our estimated factors are linear combinations of Fama and French's market, value and size factors. Thus, our estimated factors correctly capture the risk at which firms are exposed.[1] Second, we find that the relative importance of unobserved regional and global factors is time-varying: when

[1]We find that bigger firms have larger exposure to financial and regional factors. Thus, the global financial factor (the S&P Financials Index) is not enough to capture the relationship between firm size and its market beta, and adding the regional factor factor improves the fit of the model. This finding supports the result of Fama and French (2017) that a global version of the factor model would not be able to price the cross-section of stock returns.

unexpected events happen globally, loadings on global factors increase. For instance, Energy stocks started to be more exposed to global shocks, both during the Great Financial Crisis in 2007-2008 (GFC) and from the beginning of 2015, due to the shocks to oil prices. Third, the dynamics (persistence and variance) of the factor loading are related to the profile of a company. Expected returns are higher when the variance of financial and global factor loadings is large, while they are not so pronounced when the variance of the regional factor loading is large. This decreasing relationship is in line with the finding of Armstrong et al. (2013) for US stocks. However, our model suggests that there is a premium for holding stocks whose exposure to global systematic risk is very volatile. Finally, expected returns are decreasing in the persistence of financial and global factor loadings, implying that there is no premium for holding firms with highly persistent factor exposures.

The remainder of the paper is organised as follows. Section 2 presents the model and the estimation procedure. Section 3 describes our data base and the identification of the factors. Section 4 presents the estimation results. In Section 5, we compare the performance of our time-varying loading factor model with one with constant loadings. Section 6 uses our model to analyses the comovements between stock returns. Section 7 connects the loadings persistence and variance to the profile of the firm. Section 8 concludes.

2 Model and estimation

This paper contributes to factor models used to analyse comovements in equity markets [Bekaert et al. (2009), Bekaert et al. (2014)], to test for market integration [Flood and Rose (2005)], and for contagion across countries and asset classes [Dungey and Martin (2007), Belvisi et al. (2016)]. In this section, we introduce the two-level factor model with time-varying loadings, and the estimation procedures for factor extraction and loading estimation via maximum likelihood estimation (MLE).

2.1 A two-level factor model with time-varying loadings

We have N stocks in total. We divide them into regions $R_1, R_2, \dots, R_K$. Each region has n_k stocks, thus $\sum_{k=1}^{K} n_k = N$. The log-return $r_{i,t}$ on stock i in week t is modelled as:

$$r_{i,t} = a_{i,t}O_t + b_{i,t}G_t + \sum_{k=1}^{K} c_{i,t}F_{k,t}\mathbb{1}_{\{i\in R_k\}} + e_{i,t}, \quad e_{i,t} \sim N(0, \psi_i), \tag{1}$$

for $i = 1, \dots, N$ and $t = 1, \dots, T$. O_t is an observable global factor, G_t is an unobservable (latent) global factor and $F_{k,t}$ is an unobservable factor specific to stocks in region R_k, for

$k = 1, \dots, K$. The factor loadings vary over time according to the following specifications:

$$
\begin{aligned}
a_{i,t} &= (1-\phi_i^O)\bar{a}_i + \phi_i^O a_{i,t-1} + \eta_{i,t}^O, \quad \eta_{i,t}^O \sim iid\, N(0, q_i^O) \\
b_{i,t} &= (1-\phi_i^G)\bar{b}_i + \phi_i^G b_{i,t-1} + \eta_{i,t}^G, \quad \eta_{i,t}^G \sim iid\, N(0, q_i^G) \\
c_{i,t} &= (1-\phi_i^k)\bar{c}_i + \phi_i^k c_{i,t-1} + \eta_{i,t}^k, \quad \eta_{i,t}^k \sim iid\, N(0, q_i^k)
\end{aligned}
$$

The model could easily accommodate multiple regional factors, but to keep the structure interpretable we assume that there is one factor for each region. After a sign identification, this allows us to have a multi-factor model with interpretable statistical factors (each region factor is the main regional driver of returns). The total number of factors in the model is denoted by $m = K + 2$.

The main innovation of the model is to introduce time-varying loadings. We specify the dynamics of the factor loadings by stacking $a_{i,t}, b_{i,t}, c_{i,t}$ in the loading vector $\lambda_{i,t}$. The formulation in Eq. (1) implies a sparsity condition in the loading matrix, so that $\lambda_{i,t}$ is an m-dimensional vector that contains the same number of non-zero elements for all i. For instance $\lambda_{1t} = (a_{1t}, b_{1t}, c_{1t}, 0, \dots, 0)'$ and $\lambda_{2t} = (a_{2t}, b_{2t}, 0, c_{2t}, 0, \dots)'$. The non-zero elements of $\lambda_{i,t}$ evolve according to the following vector autoregression:

$$
\lambda_{it} = (\mathrm{I} - \Phi_i)\bar{\lambda}_i + \Phi_i \lambda_{i,t-1} + \eta_{it}, \tag{2}
$$

where $\bar{\lambda}_i = E(\lambda_{it}) = (\bar{a}_i, \bar{b}_i, \bar{c}_i)'$ is the unconditional mean vector, $\Phi_i = diag(\phi_i^O, \phi_i^G, \phi_i^k)$ is the persistence parameter matrix and the characteristic roots of Eq. (2) lie outside the unit circle. $Q_i \equiv E(\eta_{it}\eta_{it}') = diag(q_i^O, q_i^G, q_i^k)$ is the covariance matrix of the innovations η_{it}, which is a Gaussian white noise process. Thus, the loadings of stock i on the three factors evolve as independent autoregressive (AR) processes of order one, around their respective unconditional means, $\bar{a}_i, \bar{b}_i$ and $\bar{c}_i$, with AR coefficient ϕ_i^f, $f \in \{O, G, k\}$, and condition $|\phi_i^f| < 1$ satisfied for all f. The higher ϕ_i^f the higher the weight of the factor loading at $t-1$ in determining the loading today and the lower the weight on its unconditional mean. Stationarity of the loadings on market factors has been demonstrated by, among others, Andersen et al. (2006) and Patton and Verardo (2012) and we extend this to the case of global and regional factors.

Furthermore, to simplify the exposition we group the model by region:

$$
\begin{bmatrix} r_{1,\cdot t} \\ r_{2,\cdot t} \\ \vdots \\ r_{K,\cdot t} \end{bmatrix} = \begin{bmatrix} \mathcal{A}_{1t} \\ \mathcal{A}_{2t} \\ \vdots \\ \mathcal{A}_{Kt} \end{bmatrix} O_t + \begin{bmatrix} \mathcal{B}_{1t} & \mathcal{C}_{1t} & 0 & \cdots & 0 \\ \mathcal{B}_{2t} & 0 & \mathcal{C}_{2t} & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ \mathcal{B}_{Kt} & 0 & \cdots & \cdots & \mathcal{C}_{Kt} \end{bmatrix} \begin{bmatrix} G_t \\ F_{1t} \\ F_{2t} \\ \vdots \\ F_{Kt} \end{bmatrix} + \begin{bmatrix} e_{1,\cdot t} \\ e_{2,\cdot t} \\ \vdots \\ e_{K,\cdot t} \end{bmatrix}, \tag{3}
$$

where $r_{k,\cdot t}$ is the vector of returns on the n_k stocks in region R_k and $k = 1, \ldots, K$. $\mathcal{A}_{k,t}$, $\mathcal{B}_{k,t}$ and $\mathcal{C}_{k,t}$ are the $n_k \times 1$ vectors of loadings on the observable global, latent global and latent regional factor, respectively. The peculiarity of the two-level structure is that stocks in region R_k are not influenced by shocks that are specific to other regions, which render the estimation of the factors more challenging than in the case where all stocks load on all factors. Breitung and Eickmeier (2015) and Wang (2010) show that regional and global factors can be disentangled by adding sparsity conditions as in Eq. (3). Finally, the model for the N stocks can be written in a more compact form as:

$$r_t = \mathcal{A}_t O_t + \mathcal{B}_t^* F_t + e_t, \tag{4}$$

where $r_t = (r'_{1t}, \ldots, r'_{Kt})'$, $\mathcal{A}_t = (\mathcal{A}'_{1t}, \ldots, \mathcal{A}'_{Kt})'$, $\mathcal{B}_t^* = [(\mathcal{B}'_{1t}, \ldots, \mathcal{B}'_{Kt})', (\mathcal{C}'_{1t}, 0, \ldots)']$. F_t contains both the global factor G_t and K regional factors $F_{1t}, \ldots, F_{Kt}$. Finally,

$$r_t = \Lambda_t F_t^* + e_t, \tag{5}$$

where $\Lambda_t = (\mathcal{A}_t, \mathcal{B}_t^*)$, and $F_t^* = (O_t, F_t)'$. The covariance matrix of idiosyncratic errors e_t is $\Psi_0 \equiv E(e_t e'_t)$. To summarise, the N-dimensional vector of returns r_t is generated by $m \ll N$ global and region factors, time-varying factor loadings $\Lambda_t = (\lambda_{1t}, \ldots, \lambda_{it}, \ldots, \lambda_{Nt})'$ and normally distributed idiosyncratic errors $e_t = (e_{1t}, \ldots, e_{it}, \ldots, e_{Nt})'$.

With known factors, specification (1) - (2) can be written and estimated in a linear state-space form as in Harvey (1990). With unknown factors, and principal components as estimators, Mikkelsen et al. (2018) prove that the MLE of $\Phi_i, \bar{\lambda}_i, Q_i$ and ψ_i is consistent as $N, T \to \infty$. An alternative model specification is one where the loadings are static and the dynamics of the factors are estimated with the Kalman filter. However, since the Asset Pricing Theory assumes that stock prices are generated by a set of unpredictable factors, we follow the same rationale in the specification of our model. Furthermore, estimating stock-specific loadings dynamics allows one to sort stocks by loading variance or persistence, similarly to Armstrong et al. (2013).[2]

In the next section, we present the procedure for the estimation of the model.

2.2 Estimation

The estimation procedure consists of two steps. First, unobserved global and regional factors are estimated using principal components assuming constant loadings. Second, we replace the unobserved factors with principal components in the likelihood function and estimate the unknown loadings and variance parameters in Eqs. (1) - (2) via MLE.

[2]Estimating both the dynamics of the factors and of the loadings is an interesting extension of our model. This goes beyond the scope of this study and we leave this extension to future research.

2.2.1 Principal component estimation of the latent factors

The principal component estimator treats the loadings as constant over time and we use it to estimate the common factors in the model:

$$y_t = \mathcal{B}^* F_t + u_t, \tag{6}$$

where y_t is the residual from the regression of r_t against O_t. Thus, y_t is orthogonal to the observable factor and, consequently, the global and regional factors will not be influenced by the variation in O_t. This ensures that we can disentangle two risk factors that affect all assets in the portfolio: the first is represented by an observable factor and the second is an unobservable latent global factor.

The estimated principal components (PC) are consistent estimates of the latent factors also in the presence of stationary fluctuations of the loadings around a constant mean. In particular, we make use of the results in both Bates et al. (2013), who prove the average convergence in t of the PCs to the true factor space and Mikkelsen et al. (2018), who prove that the PCs uniformly converge in t when $\frac{T}{N^2} \to 0$.

The principal component estimator minimises the following sum of squared residuals:

$$\begin{aligned} S(F, \mathcal{B}^*) &= \sum_{t=1}^{T} (y_t - \mathcal{B}^* F_t)'(y_t - \mathcal{B}^* F_t) && (7) \\ &= \sum_{k=1}^{K} \sum_{i=1}^{n_k} \sum_{t=1}^{T} (y_{i,t} - b_i G_t - c_i F_{k,t} \mathbb{1}_{\{i \in R_k\}})^2. && (8) \end{aligned}$$

Since both $\mathcal{B}^*$ and F_t are unobserved, we need to impose the following identifying restrictions to obtain a unique solution.

IR1 $T^{-1} \sum_{t=1}^{T} G_t^2 = 1$ and $T^{-1} \sum_{t=1}^{T} F_{k,t}^2 = 1$.

IR2 $T^{-1} \sum_{t=1}^{T} F_{k,t} G_t' = 0$ for all k. This ensures regional factors are orthogonal to the global one.

IR3 $\sum_{t=1}^{T} F_{k,t} S_{k,t}' > 0$, where $S_{k,t}$ is the biggest country's stock market index return at time t in region k. This identifies the sign of the factors by imposing correlation with the most important stock index of the region. This procedure eliminates the rotation indeterminacy and allows to interpret the sign of the factor loadings.

IR1 and IR2 ensure that all parameters are identified, while IR3 is also used by Breitung and Eickmeier (2015). Note that we do not need to assume $T^{-1} \sum_{t=1}^{T} F_t F_t = \mathrm{I}_m$ as in standard factor analysis, thus the regional factors can be correlated with one another. The sparsity assumptions in Eq. (3) ensure that this does not create any multicollinearity issue. The following estimation algorithm follows Breitung and Eickmeier (2015):

1. Initialise the global and regional factors in Eq. (6) with suitable values[3], $\widehat{F}_{t,(0)}$, and estimate the zero-stage loadings $\widehat{\mathcal{B}}^*_{(0)}$ from N time series OLS regressions of y_t on $\widehat{F}_{t,(0)}$.
2. The estimated loadings are then used as regressors in $y_t = F_{t,(1)}\widehat{\mathcal{B}}^*_{(0)} + \widetilde{u}_t$ to get the update of the factors at stage one, $\widehat{F}_{t,(1)}$, by OLS.
3. At stage s, update $F_{t,(s)}$ by regressing y_t on $\widehat{\mathcal{B}}^*_{(s-1)}$ and estimate $\mathcal{B}^*_{(s)}$ by regressing y_t on $\widehat{F}_{t,(s)}$.
4. Step 3 is repeated until convergence of $S(\widehat{F}_{(s)}, \widehat{\mathcal{B}}^*_{(s)})$ to a minimum.

2.2.2 Maximum likelihood estimation of the time-varying loadings

We now turn to the estimation of Φ_i, λ_i, Q_i and ψ_i, for $i = 1, \ldots, N$. Using a two-step maximum likelihood estimator, Mikkelsen et al. (2018) prove that the feasible likelihood function, which replaces the unobserved factors with PCs, convergences uniformly to the infeasible one containing the unobserved factors, despite the presence of estimation error in the principal components and time-variation in the loadings.

In our model, the global and regional factors allow for cross-sectional dependence in the returns, taking into account the fact that the market is partitioned in groups [Goyal et al. (2008)]. This allows to estimate the loadings and their parameters as unobserved stationary states in a standard state-space model [Durbin and Koopman (2012)]. The measurement equation connects the unobserved loadings linearly to the principal components, which replace the factors. The MLE remains consistent in the presence of cross-sectional and temporal dependence in the errors. We refer to Mikkelsen et al. (2018) for details on the two-step estimation procedure.

Thus, conditional on the factors, we can treat r_i as uncorrelated across stocks, and the likelihood can be analysed separately for each i. Thus, if r_i is the $T \times 1$ vector of time-series observations for stock i, we can write:

$$r_i = \widehat{\mathbf{F}}^* \Lambda_i + e_i, \tag{9}$$

where $\widehat{\mathbf{F}}^* = diag(\widehat{F}_1^{*\prime}, \ldots, \widehat{F}_T^{*\prime})$ is a $T \times mT$ block-diagonal matrix that stacks the time series observations on the estimated factors, with diagonal elements $\widehat{F}_t^{*\prime} = (\widehat{O}_t, \widehat{G}_t, \widehat{F}_{1t}, \ldots, \widehat{F}_{Kt})$ representing the observations of each factor at time t. $\Lambda_i = (\lambda'_{i1}, \ldots, \lambda'_{iT})$ denotes the $Tm \times 1$ vector of time-varying loadings.

Assuming that the idiosyncratic errors are normally distributed, the likelihood function for r_i is Gaussian and, conditional on $\widehat{F}^* = (\widehat{F}_1^*, \ldots, \widehat{F}_T^*)'$, can be specified as follows:

$$\widehat{L}_T(r_i|\widehat{F}^*; \theta_i) = -\frac{1}{2}\log(2\pi) - \frac{1}{2T}\log|\Sigma_i| - \frac{1}{2T}(r_i - E(r_i))'\Sigma_i^{-1}(r_i - E(r_i)), \tag{10}$$

[3]We initialise the algorithm with the first PC of all stocks (for the global factor) and the first PC of each group of stocks clustered by region (for the regional factor).

with parameter vector $\theta_i = \{\Phi_i, \lambda_i, Q_i, \psi_i\}$ for each i. $E(r_i) = (F_1^{*\prime}\lambda_i^0, \ldots, F_T^{*\prime}\lambda_i^0)$ is the $T \times 1$ mean vector of r_i and its covariance matrix is $\Sigma_i \equiv Var(r_i) = \widehat{\mathbf{F}}^* Var(\Lambda_i)\widehat{\mathbf{F}}^* + \psi_i I_T$. Eq. (10) represents the feasible likelihood because the factors $\widehat{F}^*$ are estimated by principal components. Finally, the maximum likelihood estimator of θ_i is:

$$\widehat{\theta}_i = \underset{\theta}{\operatorname{argmax}}\, \widehat{\mathcal{L}}_T(r_i|\widehat{F}^*; \theta_i), \tag{11}$$

for each i. In practice, Eqs. (2) and (9) can be expressed as a linear state-space model and the likelihood is maximised via the Kalman filter. Theorem 1 in Mikkelsen et al. (2018) shows that $\widehat{\theta}_i \xrightarrow{p} \theta_i^0$, i.e. the estimates are consistent and converge in probability to their true values θ_i^0.

3 Data and factors identification

In this section, we present the database, the data preparation and descriptive statistics, together with some figures that outline the region and sector trends of equity markets between 2006 and 2016.

3.1 Data description

The sample period runs from Friday 13 January 2006 to Friday 8 January 2016, for a total of 521 weekly observations. The sample period contains several shocks that are interesting to analyse: the great financial crisis of 2007-2008; the European sovereign debt crisis of 2011-2012; the Arab spring of 2011; and the oil shocks of 2015. Our dataset include 1815 stocks that have been part of the main stock market indices of 54 countries during the sample period and have complete time series of prices. Table 1 reports details of the number of stocks for each country. The data are downloaded from Bloomberg and the dataset resembles the one used by Bekaert et al. (2014). For each firm, we download the following variables: share price, number of shares outstanding, total assets and total debt. Prices refer to the last transaction of the week reported by the exchange, adjusted for subsequent splits but not for subsequent dividends. Working with weekly prices is especially convenient to avoid problems caused by asynchronous trading of stocks listed in countries with different time zones. The balance sheet data is available at quarterly frequency.

We consider all stocks that entered the indices during our sample period. This procedure limits the survivorship bias, which could be particularly severe in our sample period, considering the changes in the composition of indexes that occurred in 2008. Table 1 reports detailed information on each stock market index.

[Table 1 about here.]

Prices are expressed in US dollars, and returns in the model are calculated as the first difference of the natural logarithm of the share prices. Each stock return is demeaned by subtracting the sample mean and scaled by dividing each series by the standard deviation. Before performing PCA, the data have been winsorised at the 99% level, while the dependent variable is not modified, so that outliers are captured by the time-varying loadings.

3.2 Observed factor

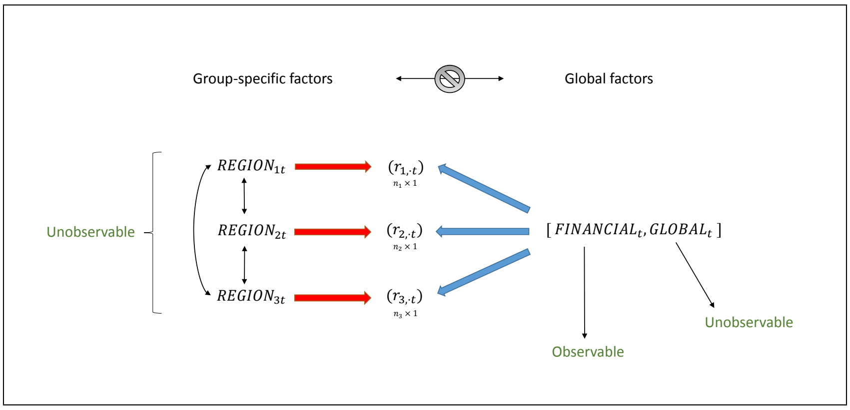
We define O_t as the returns on the S&P500 Financials Index, which is a known source of risk that impacts all stocks in all regions. Other authors show that this factor plays an important role in explaining the cross-section of returns [e.g. Bekaert et al. (2014)]. After orthogonalising all series against O_t, we are then able to identify the remaining unobservable global and regional risk factor that affect the stocks in our portfolio.

3.3 Region classification

There is evidence that shocks to share prices are more region specific than sector specific. Common currency, geographic proximity, similar stage of economic development or distribution of wealth are more important than sector membership. Heston and Rouwenhorst (1995), Griffin (2002), Bekaert et al. (2009) provide evidence for the superiority of country factors. Recently, Ando and Bai (2017) estimated the group membership of a stock as an unknown parameter in the model, and reached the same conclusion. We classify our universe of 1815 stocks into six geographical regions, in line with Bekaert et al. (2014) and Breitung and Eickmeier (2015). The regions are: North America, Latin America, Asia-Pacific, Western Europe, Emerging Europe, and Middle-East & Africa (MEA). This classification will be used to identify the region-specific factors in our model. The region composition is reported in Table 1. We use the following six sectors classified by Bloomberg: Basic Materials, Communications, Consumer Cyclical, Consumer Non-cyclical, Diversified, Energy, Financial, Industrial, Technology and Utilities.

Below we report a graphical representation of our model.

Visual representation of our model



3.4 Summary statistics

Table 2 reports the summary statistics of simple returns (in %), market capitalisation, total assets and total debt of the 1815 companies contained in the final dataset. Simple returns are non-Gaussian for all regions and sectors. In our empirical application, we use log-returns such that the empirical distribution appears more Gaussian. We report the average pair-wise Pearson correlation coefficient among the stocks in each group (region or sector), which gives a snapshot of the dependence between firms in each region and sector. The region Middle East & Africa has the lowest value, with a coefficient of 0.159, which can be expected given the economic diversity of this area. North America and Western Europe have a correlation of 0.378 and 0.424, respectively. The sector with the highest level of linear dependence is Energy, at 0.362. The balance sheet data are in line with expectations. For instance, North-American and Western European stocks have the highest average market capitalisation. The biggest companies in the final dataset are Apple (US); Vodafone (WestEur); GazProm (EmEur); China Petroleum (Asia). We exclude financial stocks from our analysis, given the peculiar nature of their balance sheet. Consequently, energy stocks are the highest capitalised, with a mean value of $25 billion and a median of $8 billion. Utilities and energy stocks are the ones with larger assets and debt, in line with the infrastructures needed for the business.

[Table 2 about here.]

4 Estimation results

In this section, we present the estimated factors (one global and six regional), their corresponding time-varying loadings and we explore the benefits of allowing the loadings to vary over time in terms of model fit and misspecification.

4.1 Mapping the estimated factors to exogenous variables

Figure 1 plots the estimated global factor and the regional factors for Asia-Pacific, Emerging Europe, Latin America, Middle-East Africa, North America and Western Europe. The factors are estimated by PCA from the model with static loadings shown in Eq. (6) and they are rotated such that they are positively correlated with the stock market index of the biggest country in the region.

The global factor in Panel 1a captures a source of global risk uncorrelated with US financial stocks, and its fluctuations are related to the overall trend of equity markets during our sample period. Together with the global factor, we plot a double-sided two-month moving average. Even after taking into account the shocks coming from US financial institutions, in 2008 the stocks of the major stock markets around the world experienced a substantial drop. This can be attributed to financial contagion [Bekaert et al. (2014)]. Other periods of consecutive negative returns can be seen in 2011 (around the sovereign-debt crisis) and between the end of 2014 and the beginning of 2015 (the period corresponding to a large drop in oil price).

Our model also allows to disentangle global and region-specific factors. The regional factors capture shocks that affect only the firms in that region. We find that the impact of the global financial crisis in 2007-2008 was so pervasive that in 2008 there were very large region-specific negative shocks in all six regions (though the regional factors are orthogonal to the global factor and the observable US financial risk). However, their magnitude is smaller than the shocks coming from the global factor, and other shocks become more evident. All the regional factors display heteroscedasticity, in particular around periods of market turmoil. For instance, the estimated factors for Latin America and Emerging Europe display large negative consecutive shocks and excess volatility during the oil shocks of 2014-2015. Specifically, countries of Latin America with large oil reserves (e.g. Mexico) may have been particularly hit by the decrease in oil price that started at the end of 2014. Lower oil prices have a negative impact on oil producers and on countries whose GDP highly depend on oil exports, while they have a positive effect on companies whose costs depend on oil price (e.g. airlines) and for net importer countries. This could also explain the divergence of performance between Western Europe and Emerging Europe (which includes Russia) in the last part of the sample.

[Figure 1 about here.]

Table 3 reports, in Panel A, the correlation among the six regional factors. The sparsity of the loading matrix allows the estimated regional factors to be correlated because they do not interact in the model. The North America factor is most highly correlated with Latin America (which includes Mexico) and Western Europe, with correlation coefficients of 0.429 and 0.483, respectively. Emerging Europe has the highest correlation with Latin America with a correlation coefficient of 0.441. Since the former region includes Russia and the latter Brazil, this connection may be due to the presence of large oil companies in the stock market indexes of these countries. The Middle-East Africa factor is uncorrelated with North America, Latin America and Western Europe, while it is mildly correlated with Asia-Pacific and Emerging-Europe, with a correlation coefficient of 0.210 and 0.131, respectively.

As a robustness check, in Panel B of Table 3 we report the correlation coefficients between the first principal component of the six regional portfolios and the global PC, the S&P500 index, and S&P500 Financials index, respectively, where the factors are extracted separately from portfolios of stock returns in the relevant region. As expected, the first PC for North America correlates almost perfectly with the S&P500 index and the S&P500 Financials, with a correlation coefficient of 0.962 and 0.847, respectively. The firms based in Western Europe also closely mimic the pattern of the US stock markets, with correlations as high as 0.821 for the S&P500 index. The Middle-East&Africa (MEA) factor is the least correlated with the fluctuations of the US stock market. All factors, except the MEA, are highly correlated with a global factor, with coefficients ranging from 0.889 to 0.968. This pattern motivates the distinction of global and regional variation in our model.

Finally, in Panel C of Table 3 we report the correlation coefficients between the estimated global and regional factors, the S&P500 index, and the S&P500 Financials index, without orthogonalising the dependent variables against the S&P500 Financials index. We notice that in this case there is a high correlation between Global and North America factors and both S&P500 and S&P500 Financials, which implies the presence of contagion effects of shocks coming from the US stock markets. The orthogonalisation with respect to the S&P500 Financials index therefore is a natural approach to separate the global risk from the US financial sector risk.

[Table 3 about here.]

4.1.1 The connection with Fama-French factors

To understand what type of risk our factors are capturing, we use canonical correlation analysis (CCA) to map the estimated factors to the three Fama-French (FF) factors (market, value, and size).

Table 4 reports the (squared) maximum canonical correlation between the linear combinations of estimated factors and FF market, value and size factors. Panel A reports the canonical correlation between all three factors, i.e. $[O_t, G_t, F_{j,t}]$, for $j = 1, \ldots, R$, while Panel B reports the results using only the regional factors. We use five sets of FF region-specific factors: Asia-Pacific (excluding Japan), Europe, Global, Global excluding US, and North America. Overall, we find that the spaces spanned by the FF factors and by the factor estimated in this paper are very similar. This leads us to conclude that our model to a large degree captures the risk to which firms are exposed. In particular, our three factors for Asia-Pacific, Western Europe and North America agree with the FF factors constructed with stocks in the respective regions, with canonical correlations up to 0.958 in North America's case. Only for Europe's case, the FF factors have a comparable canonical correlation with both our Western Europe and North America estimated factors, as reported in Panel B. This is possibly due to the integration between Europe and the US.

[Table 4 about here.]

4.2 Factor loadings

Our model allows to estimate, for each stock, the parameters (mean, variance and persistence) driving the dynamics of the stock's exposure to each factor. In this section, we present the results for two selected stocks, and for the aggregation by either region or sector.

4.2.1 The factor exposure of two large firms

Figure 2 displays the estimated time series of the three factor loadings for two large firms: IBM and Tenaris. We choose to analyse these two firms individually because they are likely to be exposed to global determinants that are difficult to quantify, and our model can provide guidance in their identification.

Figures 2a - 2c plot the financial, global and regional factor loadings for IBM, respectively. On average, this firm is most exposed to financial risk, with a factor loading fluctuating around a level of 0.6Ṡince the dependent variable is standardised, the factor loadings' economic magnitude corresponds to β standard deviations for one standard deviation increase in the factor. The global and regional factor loadings fluctuate around a similar long-run mean of about 0.3 but with different AR(1) parameters (0.5 for the global and -0.22 for the regional factors, respectively). A large AR(1) implies that the process spends a long time away from the long-run average.

Figures 2d - 2f plot the financial, global and regional factor loadings for Tenaris, respectively. Tenaris is a global company, headquartered in Luxembourg and with business in over 20 countries. It deals with the construction, distribution and service of steel pipes. Our model identifies this company as highly exposed to both observed and unobserved global factors. However, these exposures have very different dynamics. On the one hand, the exposure to the financial factor is constant around a value of 0.537 with a negligible variance that makes the AR(1) not identifiable. On the other hand, the exposure to the global factor follows an AR(1) coefficient of 0.94 close to a unit root, suggesting a persistent exposure of Tenaris to global shocks.

Our factor loadings closely follow the idiosyncratic variations of stock returns and identify some firm-specific events that the factors are not able to capture. For instance, the financial factor in Figure 2a shows some negative spikes, which appear regularly in October from 2011 to 2015. These can be caused by either third quarter earning announcements or the expiration of stock options. We can rule out dividend payments as a cause, as these happened quarterly in that period of time. Furthermore, the regional factor in Figure 2c exhibits a large drop in the third week of October 2014, which corresponds to the announcement by IBM of a large fall in sales,[4] while the stock market was rallying upwards. In that moment, the covariance between the market and IBM return switched from positive to negative. These events caused the loading to become temporarily negative, before reverting to its long run mean.

[Figure 2 about here.]

4.2.2 Aggregate results

Table 5 reports the average magnitude of the factor loadings, their persistence (AR(1) parameter) and their volatilities, aggregated by either region or sector. The loading magnitude is estimated via OLS from a static loading model, while the AR(1) parameter and the variance are estimated via maximum likelihood estimation from Eq. (1). The table also reports the percentage of stocks whose loadings vary so little that they are indistinguishable from the OLS estimates. We set a loading volatility threshold to 0.01, under which it is very difficult to identify the autoregressive parameter, and consider such loadings static. Assuming that the goal of an equity investor is to estimate the systematic risk of a set of stocks correctly, factor loadings with large AR(1) coefficients indicate stocks with very persistent shocks to their exposure to the factors.

First, we discuss the results by regions. The firms listed in North America are the most exposed to the financial factor. The second most exposed group is Western Europe, with an average factor loading of 0.461. These firms, together with firms listed in Latin America and Emerging Europe are relatively more exposed to financial shocks than to global or regional shocks. Conversely, firms listed in MEA are more exposed to regional than global shocks. Firms listed in the Asia-Pacific region are least exposed to the US financial risk. We report the percentage of firms, within each group, with an AR(1) parameter larger than 0.5, defining them as firms with very persistent factor loadings. Table 5 shows that the regional factor loadings are the most persistent, implying higher predictability of regional systematic risk compared to financial and global risk, in particular for Asia-Pacific and MEA.

Second, we analyse the results by sectors. As to be expected, financial firms are the most exposed to the financial factor. However, we do not find much variation across

[4]Source: "IBM shares tumble as profits and sales fall", Financial Times (20 October 2014).

sectors. Utilities stocks are the least exposed to the financial factor. Energy stocks are most highly exposed to the estimated global factor and Consumer Non-cyclical stocks are the least exposed. This is in line with expectations, since Consumer Non-Cyclical stocks (e.g. food, beverages and tobacco) should be less likely to follow global trends.

[Table 5 about here.]

5 The benefits of using time-varying factor loadings

In this section, we compare the residuals obtained from a multi-level factor model with static loadings with those obtained from our model. We provide evidence that our model has an unambiguously higher goodness of fit. We discuss the significant deviations between our factor loadings and the static ones estimated by OLS. Finally, we show that allowing the loadings to be time-varying has an impact on standard residual-based misspecification tests and on the estimation of the number of factors following Bai and Ng (2002).

5.1 Model fit

The model specification in Eq. (5) nests the two-level factor model with static loadings as special case, where it is assumed that $\Lambda_t \equiv \Lambda$. The static specification for the N stocks takes the form:

$$r_t = \Lambda F_t^* + e_t, \tag{12}$$

where Λ is an $N \times m$ matrix of loadings parameters and F_t^* contains m factors, some global and some region-specific. The model in Eq. (12) has been used by Breitung and Eickmeier (2015) to study the comovements of real economic variables and by Goyal et al. (2008) to estimate NYSE- and NASDAQ-specific factors for stock returns.

Panel A of Table 6 reports the goodness of fit of specifications (5) and (12), measured by the R^2 coefficient of the regressions averaged within regions and sectors. We choose this method over a joint likelihood-ratio test of the overall fit because this shows in which regions/sectors the time-varying loading model provides the largest improvements. We find that there is an improvement of around 25% by using time-varying parameters. The biggest improvement of 28% is for the firms listed in Middle East & Africa, which are the ones where the idiosyncratic component plays a prominent role. As showed in Table 2, this group of stocks has the lowest average pair-wise correlation, which is caused by the economic and political differences of these countries that might have limited their economic integration. Regions where the OLS estimator provided the highest average R^2 are the ones that are more integrated (North America and Western Europe). Thus, the time-varying specification provides a connection between the common factors and the idiosyncratic characteristics of each stock.

Furthermore, we can compare the distribution of $\lambda_{i,t}$ with the theoretical distribution of the OLS estimator of λ_i in Eq. (12). Since the least-squares estimator is normally distributed, we expect 5% of the observations ($\lfloor T(0.05) \rfloor = 26$) to lie outside the 95% confidence interval (CI). We define the number of times the factor loadings a_i, b_i and c_i are outside the 95% confidence interval of the OLS estimator as:

$$n_i^F = \sum_{t=1}^{T} \mathbb{1}_{\left(a_{i,t} \notin \{\hat{a}_i \pm 1.96\sqrt{Var(\hat{a}_i)}\}\right)}, \tag{13}$$

$$n_i^G = \sum_{t=1}^{T} \mathbb{1}_{\left(b_{i,t} \notin \{\hat{b}_i \pm 1.96\sqrt{Var(\hat{b}_i)}\}\right)}, \tag{14}$$

$$n_i^R = \sum_{t=1}^{T} \mathbb{1}_{\left(c_{i,t} \notin \{\hat{c}_i \pm 1.96\sqrt{Var(\hat{c}_i)}\}\right)}, \tag{15}$$

for $i = 1, \ldots, N$ stocks. We then average n_i^F, n_i^G and n_i^R by region and sector and report the results in Panel B of Table 6. We find that in all groups of stocks there is a larger than expected number of significant deviations from the OLS loadings estimates.

[Table 6 about here.]

Since we estimate the loadings conditionally at each time t, we can also identify in which year and for which factors there are more deviations from the OLS estimates, i.e. at what time a static specification will tend to underestimate or overestimate the model factor exposure. Fig. 3 reports, for every year, the cross-sectional average of the number of significant deviations from OLS in the six world regions.

We find that in North America, in 2008, the exposure of the average firm to financial risk would have been underestimated or overestimated 25 out of 52 weeks, if it was estimated assuming constant loadings. In 2009, the number is very similar and the same holds true for stocks listed in Western Europe, Emerging Europe, and Asia-Pacific. Even though the great financial crisis is affecting the estimation of the factor loadings of all stocks, in the Middle-East&Africa and Latin America, the estimation of the regional factor loadings is the most affected.

[Figure 3 about here.]

5.2 Misspecification tests

We further explore the effect of time-varying factor loadings on the Bai and Ng (2002) estimator of the number of factors, on residual heteroscedasticity, and on serial correlation.

First, we estimate the number of factors using the Bai and Ng (2002) criteria on the returns matrix, the residuals from a static model and the residuals from a time-varying loading model. Bai and Ng (2002) propose various modifications of information criteria for model selection with an additional penalty that is a function of both N and T. We use

$$IC_{p1}(k) = \ln(V(k, \widetilde{F}^k)) + k\left(\frac{N+T}{NT}\right)\ln\left(\frac{N+T}{NT}\right), \tag{16}$$

where $V(k, \widetilde{F}^k)$ is the average residual variance of a factor model with k factors. $IC_{p1}(k)$ can be used for estimating the number of factors as a standard information criterion:

$$\hat{k} = \underset{0\leq k\leq k_{max}}{\arg\min}\ PC_1(k). \tag{17}$$

Second, we implement a White-type test under the null of homoscedasticity, estimating the following auxiliary regression:

$$\hat{e}^2_{i,t} = \alpha_i + \gamma_i\hat{F}^{*2}_{i,t} + u_{i,t}, \quad u_{i,t} \sim N(0, \sigma^2_{u,i}), \tag{18}$$

for $i = 1\ldots, N$ stocks, where $\hat{F}^*_{i,t}$ contains the three factors specific to stock i, estimated by PCA. The test statistic is equal to the R^2, times the sample size T, and it is distributed as a χ^2 with degrees of freedom equal to the number of factors. The properties of the test are studied by Mikkelsen (2017), who proves that this test is equivalent to testing for constant loadings. Finally, we test the null hypothesis of no serial correlation of the error term $\hat{e}^2_{i,t}$ up to the p-th lag using the Breusch-Godfrey test (up to lags two and five). The test statistic is equal to R^2 times the sample size and it is distributed as a χ^2 with degrees of freedom equal to $T - p$.

Table 7 reports the results. The Bai and Ng (2002) IC_{p1} criterion finds that ten factors are needed to describe the variation in our panel of 1815 firms from 50 countries. Given that we include eight factors, we expect the residuals to have two omitted factors. However, when we apply the IC_{p1} criterion to the residuals, it suggests to use five factors, and only when allowing the factors to vary over time this number decreases to three. Thus, the structural instability of the loadings has an influence on the Bai and Ng (2002) number of factors estimator.

For the heteroscedasticity and serial correlation tests, in Table 7 we report the percentage of stocks for which we reject the null at 99% confidence level. We find that 51% of firms have time-varying loadings, which implies that the volatility of returns is not entirely captured by the static-loadings model. When loadings are allowed to change over time, only 5% of stock returns have time-varying volatility. The results on the serial correlation tests are less strong. The percentage of firms that exhibit residual serial correlation up to lag five is reduced from 31% to 21%.

[Table 7 about here.]

6 Global comovements of stock returns: new evidence

In this section, we use variance decomposition methods to analyse the comovements of the large panel of stock returns. Variance decomposition has been used extensively to interpret the estimates of factor models. Among others, Breitung and Eickmeier (2015) assess the degree of comovement of groups of variables. In particular, the higher the (average) share of variance explained by common factors compared to the idiosyncratic variance, the higher the comovements. However, this method provides one number for the whole sample and often two or more sample periods are compared to assess whether there was an increase in commonality. Our model allows to overcome this limitation, estimating a conditional variance decomposition at each time t. In what follows, we compare the two decompositions.

6.1 Static variance decomposition

Table 8 reports the static variance decomposition. The firms listed in North America and Western Europe have the highest commonality, while the ones listed in the MEA region have the largest idiosyncratic component. This is in line with the results reported in Table 5.

[Table 8 about here.]

6.2 Time-varying variance decomposition

Our model allows calculation of the share of variance explained by the factors at each point in time. Hence, we can capture possible shifts in the importance of some factors and connect them to macro events. The variance of the returns on stock i at time t, conditional on the estimated factor loadings, can be written as:

$$\mathrm{var}_t(r_{i,t} \,|\, \hat{\lambda}_{i,t}) = \hat{a}^2_{i,t}\ \mathrm{var}(O_t) + \hat{b}^2_{i,t}\ \mathrm{var}(G_t) + \hat{c}^2_{i,t}\ \mathrm{var}(F_{k,t})\mathbb{1}_{\{i\in R_k\}}, \tag{19}$$

for $i = 1, \dots, N$, $t = 1, \dots, T$, and assuming that the factors and the errors are conditionally orthogonal. O_t is an observable global factor, G_t is an unobservable (latent) global factor and $F_{k,t}$ is an unobservable factor specific to stocks in region R_k, for $k = 1, \dots, K$. All these factors have unconditional variance equal to one. Thus, if we are interested in the share of variance explained by each of the factors at each time t, we can define the following quantities:

$$FV_{i,t} = \frac{\hat{a}_{i,t}^2}{\text{var}_t(r_{i,t} \,|\, \hat{\lambda}_{i,t})} \qquad \text{(Financial)},$$

$$GV_{i,t} = \frac{\hat{b}_{i,t}^2}{\text{var}_t(r_{i,t} \,|\, \hat{\lambda}_{i,t})} \qquad \text{(Global)},$$

$$RV_{i,t} = \frac{\hat{c}_{i,t}^2}{\text{var}_t(r_{i,t} \,|\, \hat{\lambda}_{i,t})} \qquad \text{(Regional)},$$

where $FV_{i,t}$ is the share of variance explained by the financial factor at time t, and $GV_{i,t}$ and $RV_{i,t}$ are defined accordingly. Calculating the cross-sectional average of the quantities above provides a measure of the importance of different drivers for the comovements of groups of stocks (e.g. within one region, one sector, one country).

Fig. 4 shows the share of variance explained by each factor, averaged across all N stocks in the portfolio. The graphs shows that, on average, the financial factor is the most pervasive, with a share of variance explained around 15%. However, during the GFC, there was a considerable increase in the exposure of stocks to financial shocks, which corroborates the evidence of contagion from the financial sector to other areas of the economy. Regional and sectoral figures shed more light on the heterogeneity of this effect. Furthermore, Fig. 4b shows that the share of variance explained by all the common factors increased by 10% by the end of 2008.

[Figure 4 about here.]

Figs. 5 - 6 plot the share of variance explained by the three factors for each region. For the stocks in all regions, the contribution of all factors increases during the Financial Crisis, hence the comovements increase. The increase in comovement at the outset of the crisis varies across regions and, for instance in Asia-Pacific and Western Europe, some stocks start to become more sensitive to global shocks already at the end of 2007 or the beginning of 2008. We find evidence of increased comovements during the European sovereign-debt crisis, recording positive spikes in the financial factor contribution in 2011 for Western Europe and for Emerging Europe. In Middle East & Africa, the regional factor is the most important throughout the sample.

[Figure 5 about here.]

[Figure 6 about here.]

Figs. 7 - 9 plot the share of variance explained by the three factors in each sector. The most interesting cases are the ones where the variance shares increase during the sample.

For example, the variance of the stocks in the Energy sector start to be highly explained by the global factor from the beginning of 2015. This is the direct effect of shocks from the oil market: Energy companies were hit by a very low price of oil, due to overproduction, shale-gas as a substitute product, and a reduced demand from China. In addition, the Utilities and Basic Materials sector experiences a large increase in the importance of the global factor from 2015. As expected, these firms, providing for instance gas and electricity, are subject to demand shocks that are specific to their area, and they are also very sensitive to interest rate changes due a high debt/equity ratio.

[Figure 7 about here.]

[Figure 8 about here.]

[Figure 9 about here.]

7 Dynamics of the loadings and the profile of the firm

In this section, we investigate if the firm-specific estimates that we obtain from our model - in particular loading persistence and variance - are associated with different types of firms. This research question is motivated by Ang et al. (2017), who show that creating portfolios for asset pricing tests destroys information and leads to larger standard errors than using individual stocks. Our goal is to test whether the dynamics of the beta of the factors are related to the size and the leverage of the firm and to its expected returns. This information, gathered only from market prices, could be used in asset allocation models.

7.1 Size effect

Figure 10 reports the median market capitalisation (at the end of the sample) by loading variance, persistence, and magnitude quantiles. The figure is composed of three panels and each one reports three sets of bars, one for each factor. Since the data set of securities comprises the constituents of large stock market indexes, we do not expect a strong size effect for every quantile.

We find that bigger firms have larger exposure to financial and regional common factors, while there is no clear difference across global factor loading quantiles [Figure 10c]. For instance, the firms in the bottom quantile of financial factor loadings have a median market capitalisation of $1 billion, while the ones in the top quantile have 15$ billion. This evidence is in line with the finding of Fama and French (2017) that a global version of their factor model is not able to price the cross-section of stock returns. Furthermore, we find that stocks with high loading uncertainty, approximated by loading variance, tend to be larger (from two to three times bigger) than firms with little variation of factor loadings. In Fig. 10a we can see that the effect exists for the financial and global factors only. This is an extension of Armstrong et al. (2013), who also analyse the cross-section of firms with loading uncertainty, but using a single-factor CAPM with US stocks. There is no difference in size between firms with difference persistence parameters [Figure 10b].

In conclusion, we find that large firms tend to have large exposures to US financial and regional factors, and these exposures are more volatile than those of small firms.

[Figure 10 about here.]

7.2 Leverage effect

Figure 11 reports the expected change in the leverage ratio from January 2010 and January 2016, by loading variance, persistence and magnitude quantiles. Each panel reports three sets of bars, one for each of the factors. Financial stocks are excluded. We use the average of the change in leverage ratio (at quarterly frequency) instead of the leverage ratio because various authors have shown that firms adjust their leverage towards a target ratio [see, e.g. Halling et al. (2016)], and we find that the average leverage does not vary substantially across stocks.

Figure 11a shows that firms with higher variance of the financial factor loading have an average change in leverage ratio closer to zero than firms with low variance of the financial factor loading. This implies that large changes in factor loadings are connected to positive and negative changes to leverage ratio that cancel each other out, which result in large changes in the systematic risk of the firm.

One would expect the leverage of a firm to be connected to regional factors such as interest rate shocks. However, Figure 11c and 11a show that firms with high leverage have a low average exposure to the regional factor and a high variance of its loading.

[Figure 11 about here.]

7.3 Expected returns

Figure 12 reports the expected weekly returns, expressed in basis points, as a function of loading variance, persistence and magnitude. Expected returns are the average log-return for each stock in the data set, excluding financial stocks, from January 2010 until January 2016.

We find that expected returns are increasing in the variance of financial and global factor loadings, while they are decreasing in the variance of the regional factor loading [Figure 12a]. The decreasing relationship is in line with the finding of Armstrong et al.

(2013) for US stocks. However, our model suggests that there is a premium for holding stocks with large variance in the exposure to the global factor. This pattern cannot be explained by cross-sectional differences in returns volatility. Furthermore, expected returns are decreasing in the persistence of financial and global factor loadings, implying that there is no premium for holding firms with highly persistent factor exposures.

[Figure 12 about here.]

8 Conclusions

In this paper, we studied the dynamics of the systematic risk in a large portfolio of 1815 firms from 54 countries with weekly return observations covering the period 13 January 2006 to 8 January 2016. We proposed a two-level factor model with time-varying loadings that captures financial, global, and regional risk to estimate common components in stock returns. The global and regional factors are latent and estimated via principal components. The loadings evolve as autoregressive processes and are estimated via maximum likelihood.

Our analysis yields three main findings. First, we find that the estimated factors are linear combinations of Fama and French's market, value, and size factors. Thus, we are able to capture the same source of risk. Second, we find that the relative importance of unobserved regional and global factors is time-varying: when unexpected events happen globally, loadings on global factors increase. For instance, Energy stocks were more highly exposed to global shocks, both during the Great Financial Crisis and from the beginning of 2015. Finally, the dynamics of the factor loadings are related to the profile of a company. Expected returns are higher when the variance of financial and global factor loadings is large, while not so pronounced when the variance of the regional factor loading is large, in line with Armstrong et al. (2013). We find that expected returns are lower when the variance of the regional factor loadings is large and higher when the variance of financial and global factor loadings is large. Furthermore, our model suggests that there is a premium for holding stocks whose global systematic risk is more volatile. Finally, expected returns are decreasing in the persistence of financial and global factor loadings, implying that there is no premium for holding firms with highly persistent factor exposures.

References

Andersen, T. G., T. Bollerslev, F. X. Diebold, and G. Wu (2006). Realized beta: Persistence and predictability. *Advances in Econometrics 20*(1), 1–39.

Ando, T. and J. Bai (2017). Clustering huge number of financial time series: A panel data approach with high-dimensional predictors and factor structures. *Journal of the American Statistical Association 112*(519), 1182–1198.

Ang, A. and D. Kristensen (2012). Testing conditional factor models. *Journal of Financial Economics 106*(1), 132–156.

Ang, A., J. Liu, and K. Schwarz (2017). Using stocks or portfolios in tests of factor models. *Available at SSRN: https://ssrn.com/abstract=1106463*.

Armstrong, C. S., S. Banerjee, and C. Corona (2013). Factor-loading uncertainty and expected returns. *The Review of Financial Studies 26*(1), 158–207.

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

Barigozzi, M., M. Hallin, and S. Soccorsi (2018). Identification of global and local shocks in international financial markets via general dynamic factor models. *Journal of Financial Econometrics 16*(1), 1–33.

Bates, B. J., M. Plagborg-Moller, J. H. Stock, and M. W. Watson (2013). Consistent factor estimation in dynamic factor models with structural instability. *Journal of Econometrics 177*(2), 289–304.

Bekaert, G., M. Ehrmann, M. Fratzscher, and A. Mehl (2014). The global crisis and equity market contagion. *Journal of Finance 69*(6), 2597–2649.

Bekaert, G., R. J. Hodrick, and X. Zhang (2009). International stock return comovements. *Journal of Finance 64*(6), 2591–2626.

Belvisi, M., R. Pianeti, and G. Urga (2016). Modelling financial markets comovements during crises: A dynamic multi-factor model approach. *Advances in Econometrics Vol. 35, Dynamic Factor Models 35*, 317–360.

Berk, J. B., R. C. Green, and V. Naik (1999). Optimal investment, growth options, and security returns. *The Journal of Finance 54*(5), 1553–1607.

Boivin, J. and S. Ng (2006). Are more data always better for factor analysis? *Journal of Econometrics 132*(1), 169–194.

Bollerslev, T. and B. Y. B. Zhang (2003). Measuring and modeling systematic risk in factor pricing models using high-frequency data. *Journal of Empirical Finance 10*(5), 533–558.

Breitung, J. and S. Eickmeier (2015). Analyzing business cycle asymmetries in a multi-level factor model. *Economics Letters 127*, 31–34.

Carlson, M., A. Fisher, and R. Giammarino (2006). Corporate investment and asset price dynamics: Implications for SEO event studies and long-run performance. *The Journal of Finance 61*(3), 1009–1034.

Carlson, M., A. Fisher, and R. Giammarino (2010). SEO risk dynamics. *The Review of Financial Studies 23*(11), 4026–4077.

Cooper, I. and R. Priestley (2011). Real investment and risk dynamics. *Journal of Financial Economics 101*(1), 182–205.

Del Negro, M. and C. Otrok (2008). Dynamic factor models with time-varying parameters: measuring changes in international business cycles. *Staff report 326, Federal Reserve Bank of New York, New York*.

Dungey, M. and L. Martin, Vance (2007). Unravelling financial market linkages during crises. *Journal of Applied Econometrics 22*(1), 89–119.

Durbin, J. and S. J. Koopman (2012). *Time Series Analysis by State Space Methods* (2nd ed.), Volume 38. Oxford: Oxford University Press.

Eickmeier, S., W. Lemke, and M. Marcellino (2015). Classical time varying factor-augmented vector auto-regressive models: estimation, forecasting and structural analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 178*(3), 493–533.

Fama, E. F. and K. R. French (1997). Industry costs of equity. *Journal of Financial Economics 43*(2), 153–193.

Fama, E. F. and K. R. French (2017). International tests of a five-factor asset pricing model. *Journal of Financial Economics 123*(3), 441.

Ferson, W. E., S. Kandel, and R. F. Stambaugh (2002). Where is beta going? the riskiness of value and small stocks. *Unpublished PhD dissertation, MIT*.

Flood, R. P. and A. K. Rose (2005). A new approach to asset integration: Methodology and mystery. *Pacific Economic Review 10*(1), 5–28.

Gagliardini, P., E. Ossola, and O. Scaillet (2016). Time-varying risk premium in large cross-sectional equity data sets. *Econometrica 84*(3), 985–1046.

Ghysels, E. (1998). On stable factor structures in the pricing of risk: Do time-varying betas help or hurt? *The Journal of Finance 53*(2), 549–573.

Goyal, A., C. Pérignon, and C. Villa (2008). How common are common return factors across the NYSE and nasdaq? *Journal of Financial Economics 90*(3), 252–271.

Griffin, J. M. (2002). Are the Fama and French factors global or country specific? *The Review of Financial Studies 15*(3), 783–803.

Hackbarth, D. and E. Morellec (2008). Stock returns in mergers and acquisitions. *The Journal of Finance 63*(3), 1213–1252.

Halling, M., J. Yu, and J. Zechner (2016). Leverage dynamics over the business cycle. *Journal of Financial Economics 122*(1), 21.

Harvey, A. C. (1990). *The Econometric Analysis of Time Series* (2nd ed.). Philip Allan.

Heston, S. L. and K. G. Rouwenhorst (1995). Industry and country effects in international stock returns. *Journal of Portfolio Management 21*(3), 53–58.

Lewellen, J. and S. Nagel (2006). The conditional CAPM does not explain asset-pricing anomalies. *Journal of Financial Economics 82*(2), 289–314.

Ludvigson, S. C. and S. Ng (2007). The empirical riskreturn relation: A factor analysis approach. *Journal of Financial Economics 83*(1), 171–222.

Mikkelsen, G., J. (2017). Testing for time-varying loadings in dynamic factor models. *CREATES Research Paper 2017-22*.

Mikkelsen, G., J., E. Hillebrand, and G. Urga (2018). Maximum likelihood estimation of time-varying loadings in high-dimensional factor models. *Journal of Econometrics*, Forthcoming.

Patton, A. J. and M. Verardo (2012). Does beta move with news? Firm-specific information flows and learning about profitability. *Review of Financial Studies 25*(9), 2789–2839.

Robichek, A. A. and R. A. Cohn (1974). The economic determinants of systematic risk. *The Journal of Finance 29*(2), 439–447.

Rosenberg, B. and J. Guy (1995). Prediction of beta from investment fundamentals. *Financial Analysts Journal 51*(1), 101.

Santos, T. and P. Veronesi (2004). Conditional betas. *Available at SSRN: http://ssrn.com/abstract=649462*.

Savor, P. and M. Wilson (2016). Earnings announcements and systematic risk. *The Journal of Finance 71*(1), 83–138.

Shanken, J. (1990). Intertemporal asset pricing - an empirical investigation. *Journal of Econometrics 45*(1), 99–120.

Stock, J. H. and M. Watson (2009). Forecasting in dynamic factor models subject to structural instability. *Hendry, D.F., Castle, J., Shephard, N. (Eds.), The Methodology and Practice of Econometrics. A Festschrift in Honour of David F. Hendry, Oxford University Press, Oxford.*, 173–205.

Wang, P. (2010). Large dimensional factor models with a multi-level factor structure. *Mimeo, Hong Kong University of Science and Technology*.

Table 1: **Universe of securities**

The table reports the countries that make up each region, and for each country it reports the following variables: *Ticker* is the Bloomberg ticker that identifies the stock market index; *#Stocks* is the number of companies that became members of the index during the period from 10 January 2003 to 19 May 2017; *Avg.Active* is the average number of index members at the beginning of every month in the sample period; *Full* is the number of stocks with complete price time series - missing values are filled with the previous value as long as there are no more than four consecutive missing; *Jan06-Jan16* is the number of complete price series, when restricting the sample from 13 January 2006 to 8 January 2016.

					#Complete	
Region	Country	Ticker	#Stocks	Avg.Active	Full	Jan06-Jan16
North America	Canada	SPTSX60 Index	101	60	61	65
North America	US	OEX Index	167	100	120	123
Latin America	Mexico	MEXBOL Index	72	35	23	33
Latin America	Argentina	MERVAL Index	53	16	32	34
Latin America	Brazil	IBOV Index	129	66	43	56
Latin America	Chile	IPSA Index	71	40	41	55
Latin America	Peru	SPBLPGPT Index	89	33	20	27
Latin America	Venezuela	IBVC Index	18	15	0	12
Asia-Pacific	Japan	TPXL70 Index	125	70	96	101
Asia-Pacific	China	SSE50 Index	151	50	24	33
Asia-Pacific	HongKong	HSCEI Index	83	40	28	48
Asia-Pacific	India	SENSEX Index	48	30	34	39
Asia-Pacific	Indonesia	LQ45 Index	122	45	46	59
Asia-Pacific	Korea	KOSPI50 Index	86	50	49	57
Asia-Pacific	Taiwan	TW50 Index	83	50	66	73
Asia-Pacific	Thailand	SET50 Index	110	50	61	75
Asia-Pacific	NewZealand	NZSX15G Index	33	15	16	20
Asia-Pacific	Australia	AS31 Index	95	50	49	57
Western Europe	Austria	ATX Index	39	20	23	24
Western Europe	Belgium	BEL20 Index	37	20	20	26
Western Europe	Denmark	KFX Index	34	20	24	27
Western Europe	Finland	HEX25 Index	37	25	23	26
Western Europe	France	CAC Index	63	40	43	47
Western Europe	Germany	DAX Index	46	30	36	38
Western Europe	Ireland	ISEQ Index	91	53	15	22
Western Europe	Luxembourgh	LUXXX Index	21	10	5	5
Western Europe	Netherlands	AEX Index	47	20	20	23
Western Europe	Norway	OBX Index	67	26	22	34
Western Europe	Portugal	PSI20 Index	38	20	20	22
Western Europe	Spain	IBEX Index	61	35	30	33
Western Europe	Sweden	OMX Index	41	30	35	35
Western Europe	Switzerland	SMI Index	34	21	26	26
Western Europe	UK	UKX Index	203	102	110	133
Emerging Europe	Croatia	CRO Index	61	26	7	21
Emerging Europe	CzechRepublic	CCTX Index	14	8	6	7
Emerging Europe	Estonia	TALSE Index	25	16	5	7
Emerging Europe	Hungary	BUX Index	28	13	11	14
Emerging Europe	Latvia	RIGSE Index	44	29	7	9
Emerging Europe	Malta	MALTEX Index	26	18	4	4
Emerging Europe	Lithuania	VILSE Index	46	29	9	13
Emerging Europe	Poland	WIG20 Index	46	20	20	32
Emerging Europe	Romania	ROTXEUR Index	24	12	3	6
Emerging Europe	Russia	CRTX Index	49	16	5	7
Emerging Europe	Serbia	BELEX15 Index	26	15	0	15
Emerging Europe	Turkey	XU030 Index	77	30	44	51
Emerging Europe	Ukraine	PFTS Index	27	19	0	7
MEA	Egypt	HERMES Index	80	39	42	55
MEA	Qatar	DSM Index	38	20	2	25
MEA	UAE	ADSMI Index	69	61	11	19
MEA	Morocco	MOSEMDX Index	81	51	27	35
Total			3256	1709	1464	1815

Table 2: **Descriptive statistics**

The table reports the summary statistics for the 1815 companies that remained in the final data set. The original number of firms in each country is reported in Table 1. Panel A reports the cross-sectional average of the summary statistics of simple returns. Panel B reports average market capitalisation, total assets and debt. *Min* and *Max* are the minimum and maximum over time and across all stocks in a group (i.e. the absolute min and max). The remaining statistics are N-averages of the relevant coefficient: *Mean*, *Med* and are the cross-sectional averages of mean and median; *StDev, Skw* and *Krt* are the average standard deviation, skewness and kurtosis; $\rho(1)$ is the OLS estimate of the first autocorrelation coefficient; *ADF* is the statistics for the Augmented Dickey-Fuller test, which is run with a constant, time trend and one lag. The critical value at 95% significance is -3.41 and the null hypothesis is the presence of a unit root. Finally, *Pearson* is the average pair-wise correlation of the stocks in the relevant group.

Panel A: Stock returns

	Mean	Med	Min	Max	StDev	Skw	Krt	$\rho(1)$	ADF	Pearson	#
Returns(%)											
North America	0.202	0.223	-35.903	45.019	4.298	-0.007	4.440	-0.057	-16.535	0.378	188
Latin America	0.275	0.082	-28.376	48.919	5.353	0.311	4.671	-0.023	-15.923	0.265	217
Asia-Pacific	0.265	0.123	-35.674	42.507	5.248	0.181	4.131	-0.027	-16.037	0.237	562
Western Europe	0.186	0.213	-58.848	55.163	5.087	-0.020	4.369	-0.046	-16.590	0.424	521
Emerging Europe	0.099	0.029	-31.425	39.595	5.863	0.169	4.697	0.004	-15.360	0.269	193
MEA	0.143	-0.065	-25.218	38.627	4.997	0.345	5.247	-0.004	-16.158	0.159	134
Basic Materials	0.190	0.055	-35.674	48.919	5.954	0.208	4.490	0	-15.823	0.268	208
Communications	0.173	0.119	-30.810	47.807	4.886	0.119	4.335	-0.047	-16.260	0.246	147
Energy	0.158	0.096	-35.714	48.309	5.527	0.035	4.275	-0.035	-16.382	0.362	122
Consumer, Cyclical	0.261	0.170	-28.809	38.201	5.348	0.153	4.358	-0.026	-16.015	0.245	212
Financial	0.205	0.112	-40.133	55.163	5.294	0.156	4.924	-0.032	-16.159	0.272	366
Technology	0.159	0.165	-25.818	34.320	5.067	0.050	3.960	-0.036	-16.162	0.267	79
Industrial	0.196	0.123	-58.848	48.148	5.357	0.126	4.364	-0.018	-16.038	0.256	293
Consumer, Non-cyclical	0.283	0.211	-31.425	41.759	4.343	0.129	4.288	-0.050	-16.465	0.206	262
Utilities	0.176	0.131	-26.030	28.125	4.276	0.029	4.075	-0.048	-16.551	0.225	103
Diversified	0.210	0.048	-23.529	30.195	5.211	0.181	4.436	-0.013	-15.876	0.256	23

Panel B: Balance sheet

	Market Cap ($bil.)	**Tot Assets ($bil.)**	**Tot Debt ($bil.)**
North America	48.914	123.886	32.957
Latin America	7.167	25.520	8.288
Asia-Pacific	8.613	30.583	7.715
Western Europe	16.858	97.682	29.290
Emerging Europe	2.844	10.154	2.135
MEA	1.892	4.852	1.056
Basic Materials	8.542	12.783	3.379
Communications	17.890	24.288	7.650
Energy	24.144	39.474	7.284
Consumer, Cyclical	10.100	17.669	6.051
Financial	14.880	198.829	56.687
Technology	24.436	16.209	2.529
Industrial	8.835	14.648	4.461
Consumer, Non-cyclical	17.451	13.819	3.540
Utilities	10.148	28.427	10.190
Diversified	4.394	17.717	3.129

Table 3: **Correlation between factors and exogenous variables**

Panel A reports the correlation matrix of the six estimated factors. Panel B reports the correlation between the regional factors (each estimated by the first PC of a portfolio of the relevant stocks) and a global factor, the S&P500 index and S&P500 Financials index. Panel C reports the correlation between the estimated global and regional factors, the S&P500 index and S&P500 Financials index, i.e. before orthogonalising against the S&P500 Financials index.

Panel A: Correlation between regional factors

	(1)	(2)	(3)	(4)	(5)	(6)
North America (1)	1					
Latin America (2)	0.429	1				
Asia-Pacific (3)	0.294	0.460	1			
Western Europe (4)	0.483	0.374	0.361	1		
Emerging Europe (5)	0.215	0.441	0.406	0.422	1	
MEA (6)	-0.044	0.097	0.210	-0.025	0.131	1

Panel B: PC1

	Glob PC	S&P500	S&P Fin
North America	0.902	0.962	0.847
Latin America	0.909	0.766	0.638
Asia-Pacific	0.904	0.685	0.557
Western Europe	0.968	0.821	0.708
Emerging Europe	0.889	0.688	0.597
MEA	0.348	0.225	0.177

Panel C: Global and regional factors

	S&P500	S&P Fin
Global	0.718	0.570
North America	0.663	0.689
Latin America	0.291	0.277
Asia-Pacific	0.183	0.158
Western Europe	0.402	0.445
Emerging Europe	0.212	0.250
MEA	0.009	0.007

Table 4: **Mapping estimated factors and Fama-French 3 factors**

The table reports the maximum squared canonical correlations between the market, size and value factors constructed by Fama and French, and the three orthogonal factors (S&P500 Financials, a global factor and a regional factor). Panel B reports the same but using only the regional factors.

Panel A: Three factors and Fama-French

	North America	Latin America	Asia-Pacific	Western Europe	Emerging Europe	MEA
Asia_Pacific_ex_Japan_3_Factors	0.765	0.780	0.876	0.760	0.763	0.746
Europe_3_Factors	0.830	0.814	0.817	0.961	0.834	0.790
Global_3_Factors	0.922	0.875	0.894	0.927	0.874	0.856
Global_ex_US_3_Factors	0.840	0.816	0.857	0.927	0.817	0.777
North_America_3_Factors	0.958	0.891	0.888	0.895	0.886	0.885

Panel B: Regional factor and Fama-French

	North America	Latin America	Asia-Pacific	Western Europe	Emerging Europe	MEA
Asia_Pacific_ex_Japan_3_Factors	0.056	0.064	0.265	0.055	0.078	0.088
Europe_3_Factors	0.214	0.078	0.051	0.204	0.047	0.015
Global_3_Factors	0.308	0.090	0.115	0.133	0.030	0.014
Global_ex_US_3_Factors	0.129	0.060	0.133	0.179	0.044	0.034
North_America_3_Factors	0.371	0.098	0.072	0.122	0.024	0.004

Table 5: **Model estimates**

The table reports the average magnitude of the factor loadings, their persistence (AR(1) parameter) and their volatilities, aggregated by either region or sector. The magnitude of the loadings is estimated via OLS from a static loadings model, while the AR(1) parameter and variance are estimated via maximum likelihood estimation from Eq. (1). The table also reports the percentage of stocks whose loadings vary so little that we consider them constant.

$$r_{i,t} = a_{i,t}O_t + b_{i,t}G_t + \sum_{j=1}^{R} c_{i,t}F_{j,t}\mathbb{1}_{\{i\in\mathcal{J}_j\}} + e_{i,t}$$

$$\begin{aligned} a_{i,t} &= (1-\phi_i^O)\bar{a}_i + \phi_n^O a_{i,t-1} + \eta_{i,t}^O, \quad \eta_{i,t}^O \sim iid\, N(0, q_i^O) \\ b_{i,t} &= (1-\phi_i^G)\bar{b}_i + \phi_n^G b_{i,t-1} + \eta_{i,t}^G, \quad \eta_{i,t}^G \sim iid\, N(0, q_i^G) \\ c_{i,t} &= (1-\phi_i^j)\bar{c}_i + \phi_n^j c_{i,t-1} + \eta_{i,t}^j, \quad \eta_{i,t}^j \sim iid\, N(0, q_i^j) \end{aligned}$$

	Financial				Global				Regional				
	Avg a_i^{OLS}	AR(1)>0.5	Std $a_{i,t}$	#static(%)	Avg b_i^{OLS}	AR(1)>0.5	Std $b_{i,t}$	#static(%)	Avg c_i^{OLS}	AR(1)>0.5	Std $c_{i,t}$	#static(%)	Tot
North America	0.519	16	0.222	5	0.225	20	0.193	5	0.248	29	0.179	10	188
Latin America	0.330	19	0.156	20	0.285	29	0.126	29	0.282	33	0.140	15	217
Asia-Pacific	0.266	17	0.234	5	0.267	28	0.187	11	0.317	36	0.156	9	562
Western Europe	0.461	18	0.244	4	0.354	30	0.148	12	0.302	28	0.154	14	521
Emerging Europe	0.312	13	0.241	8	0.307	24	0.154	15	0.276	28	0.168	16	193
MEA	0.088	11	0.203	13	0.158	29	0.157	28	0.353	40	0.223	9	134
Basic Materials	0.322	11	0.227	11	0.370	30	0.167	13	0.260	27	0.171	11	208
Communications	0.357	20	0.202	8	0.255	26	0.149	20	0.299	33	0.149	17	147
Energy	0.369	22	0.190	7	0.432	45	0.145	7	0.239	29	0.173	11	122
Consumer, Cyclical	0.358	14	0.243	5	0.249	21	0.176	15	0.325	35	0.150	15	212
Financial	0.387	11	0.263	6	0.269	27	0.175	13	0.313	36	0.165	8	366
Technology	0.352	23	0.219	5	0.250	24	0.154	18	0.304	34	0.157	18	79
Industrial	0.350	15	0.213	10	0.301	27	0.164	15	0.297	32	0.160	12	293
Consumer, Non-cyclical	0.306	23	0.221	5	0.234	23	0.171	16	0.300	30	0.165	10	262
Utilities	0.295	22	0.194	13	0.249	34	0.140	17	0.339	29	0.177	15	103
Diversified	0.346	17	0.188	13	0.264	30	0.146	13	0.365	39	0.117	17	23

Table 6: **Goodness of fit**

In Panel A, the table reports the goodness of fit of our model compared with a static model where the loadings are estimated using OLS; in Panel B, the number of times the time-varying λ_{it} is outside the 95% confidence interval of the static λ_i, estimated with OLS. The numbers reported are averages of the total number in each group. Note that $T = 521$ and $T \times 0.05 = 26$.

Panel A: R^2 comparison

	R^2	R^2-OLS	Δ
North America	0.707	0.459	0.248
Latin America	0.493	0.325	0.168
Asia-Pacific	0.550	0.279	0.271
Western Europe	0.686	0.464	0.222
Emerging Europe	0.586	0.329	0.256
MEA	0.506	0.226	0.280
Basic Materials	0.623	0.381	0.242
Communications	0.538	0.327	0.211
Energy	0.642	0.427	0.215
Consumer, Cyclical	0.607	0.354	0.253
Financial	0.668	0.401	0.268
Technology	0.537	0.310	0.228
Industrial	0.588	0.361	0.227
Consumer, Non-cyclical	0.547	0.289	0.258
Utilities	0.542	0.322	0.219
Diversified	0.574	0.389	0.185

Panel B: Significant deviations from OLS

	Fin	Glob	Reg
North America	168	177	146
Latin America	89	81	138
Asia-Pacific	130	134	124
Western Europe	187	141	136
Emerging Europe	124	108	118
MEA	77	82	193
Basic Materials	133	153	133
Communications	131	99	112
Energy	128	172	134
Consumer, Cyclical	149	131	139
Financial	176	141	178
Technology	133	110	102
Industrial	121	111	121
Consumer, Non-cyclical	139	113	119
Utilities	121	108	136
Diversified	122	100	133

Table 7: **Misspecification tests**

The table reports in the first column the number of factors implied by the Bai and Ng (2002) IC_{p1} criterion for the returns matrix, the residual matrix derived from a static loadings factor model and the residual matrix implied by the time-varying factor loading model; in the second column the percentage of stocks for which we reject the null at 99% confidence level using the White's test; in the last two columns, the Breusch and Godfrey with 2 and 5 lags, respectively

	Bai-Ng02 (#)	White (%)	BG 1-2 (%)	BG 1-5 (%)
Returns	10			
Static loadings	5	51	12	31
TV loadings	3	5	9	21

Table 8: **Static variance decomposition**

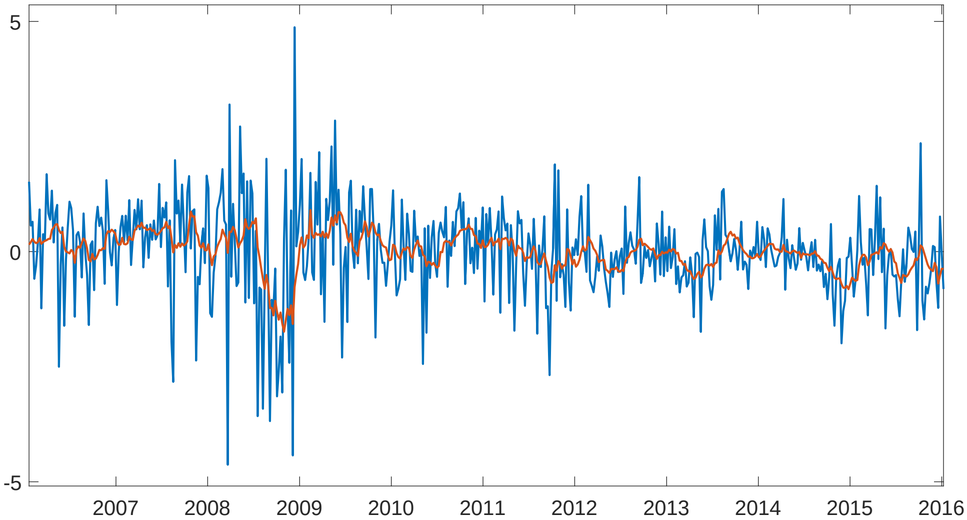
The table reports the average share of variance explained by the common factors in the relevant region or sector.

	Fin	Glob	Reg	Idio
North America	29.419	8.905	7.590	54.086
Latin America	12.938	9.354	10.159	67.549
Asia-Pacific	8.513	8.723	10.600	72.163
Western Europe	22.573	13.707	10.080	53.641
Emerging Europe	10.976	10.250	11.657	67.118
MEA	1.052	3.114	18.421	77.412
Basic Materials	13.261	15.854	8.929	61.956
Communications	15.152	7.534	10.012	67.302
Energy	15.263	20.461	6.888	57.387
Consumer, Cyclical	15.513	7.603	12.264	64.620
Financial	19.084	8.531	12.405	59.980
Technology	14.354	6.824	9.759	69.063
Industrial	15.246	10.538	10.272	63.943
Consumer, Non-cyclical	11.400	6.992	10.454	71.155
Utilities	10.704	8.423	13.081	67.792
Diversified	14.379	7.898	16.624	61.099

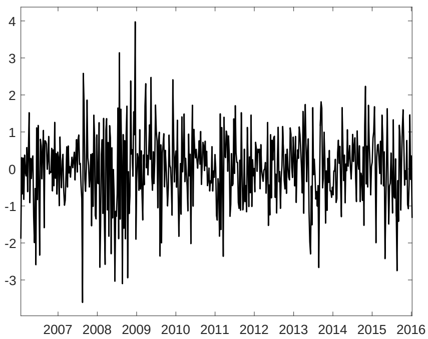
Figure 1: **Estimated global and regional factors**

The figure plots the estimated global factor and the regional factors for Asia-Pacific, Emerging Europe, Latin America, Middle-East Africa, North America and Western Europe. Together with the estimated factor, we plot a double-sided two-month moving average for the global factor only. The factors are estimated by PCA from the model with static loadings in Eq. (6). The factors are rotated to ensure that they are positively correlated with the stock market index of the biggest country in the region.

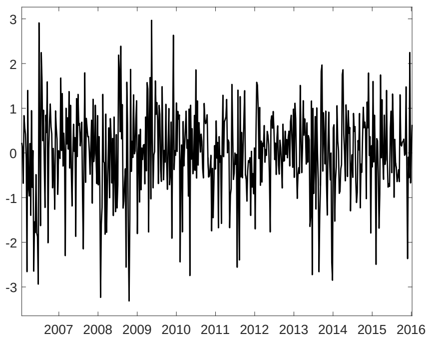
(a) Global



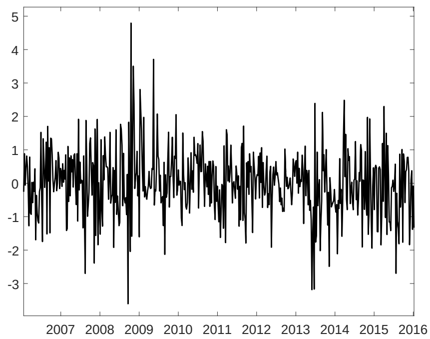
(b) Asia-Pacific



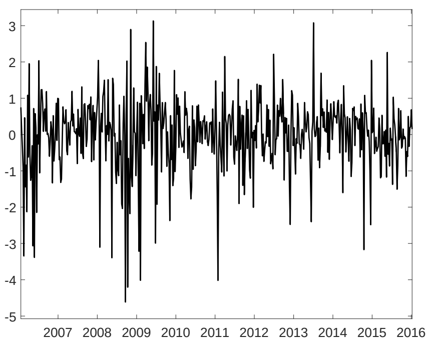
(c) Emerging Europe



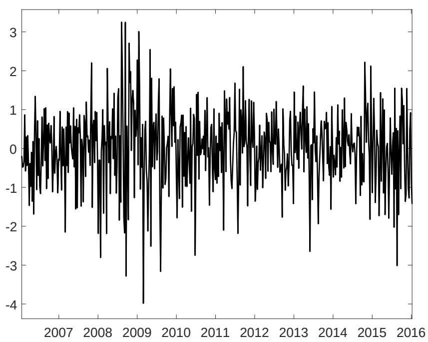
(d) Latin America



(e) Middle-East and Africa



(f) North America



(g) Western Europe

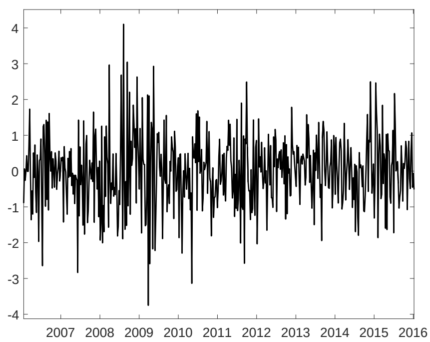
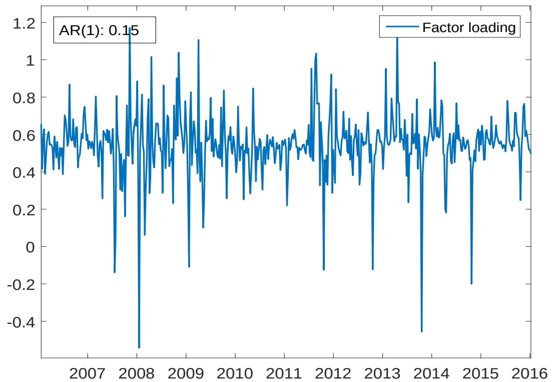


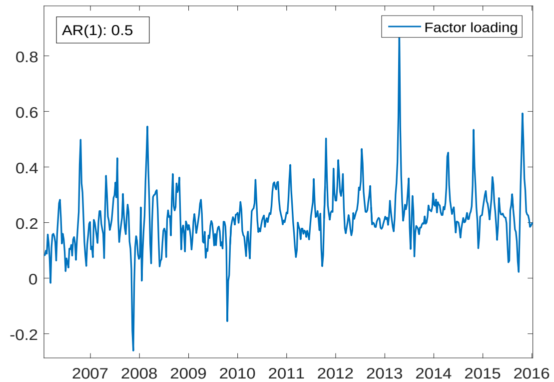
Figure 2: **The conditional factor exposure of two large firms**

The figure plots time-varying loadings estimated for IBM and Tenaris, respectively. The loadings are exposures of each stock's returns to financial, global and regional factors.

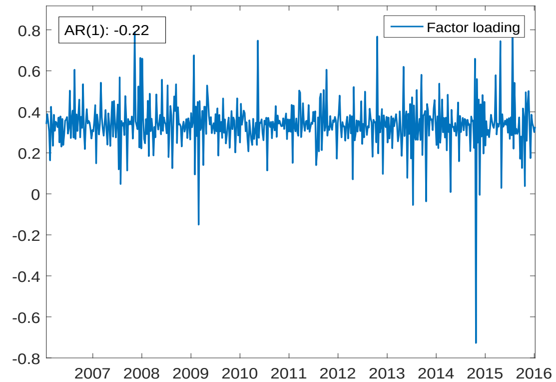
(a) IBM, financial factor



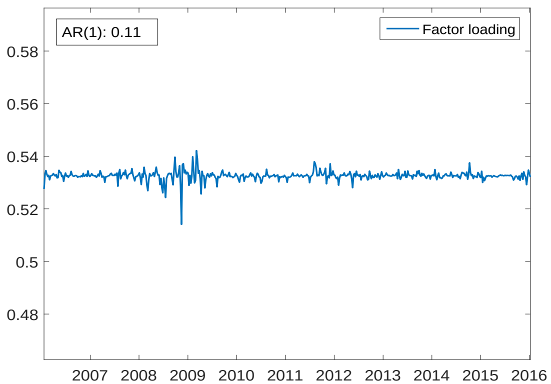
(b) IBM, global factor



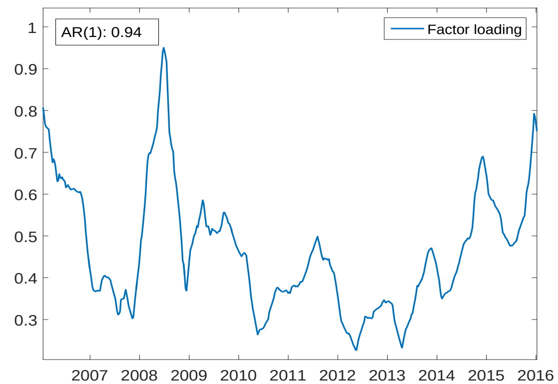
(c) IBM, regional factor



(d) Tenaris, financial factor



(e) Tenaris, global factor



(f) Tenaris, regional factor

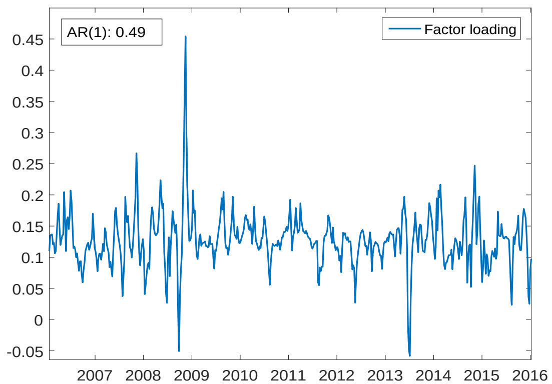
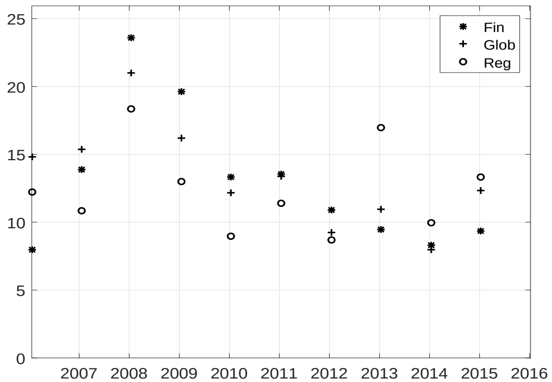


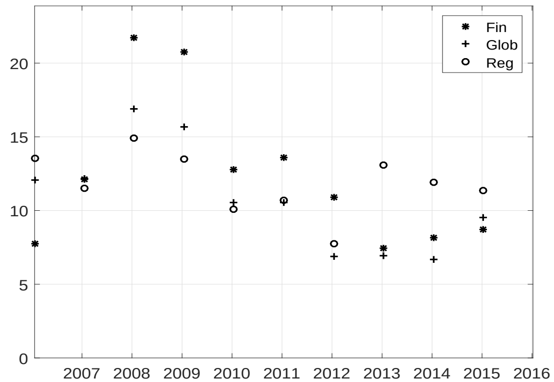
Figure 3: **Average deviations from OLS per year**

The figure plots, for every year, the cross-sectional average of the number of significant deviations from OLS in the six world regions (Asia-Pacific, Emergin Europe, Latin America, Middle-East and Africa, North America, Western Europe).

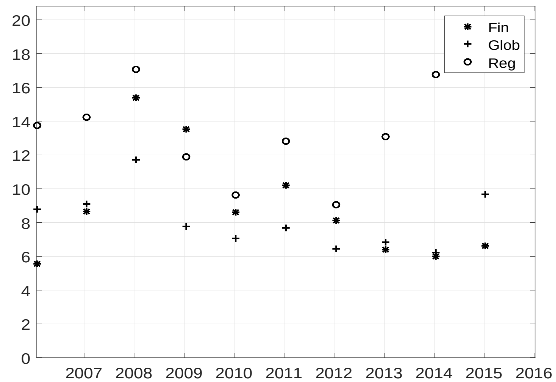
(a) Asia-Pacific



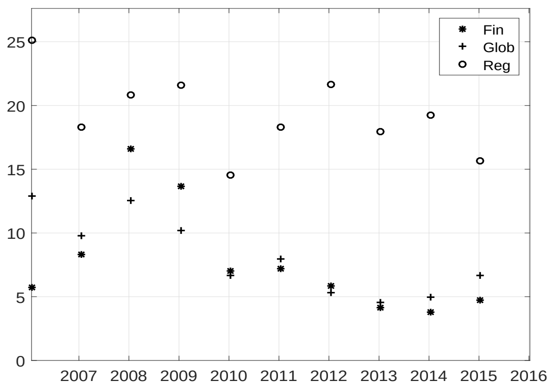
(b) Emerging Europe



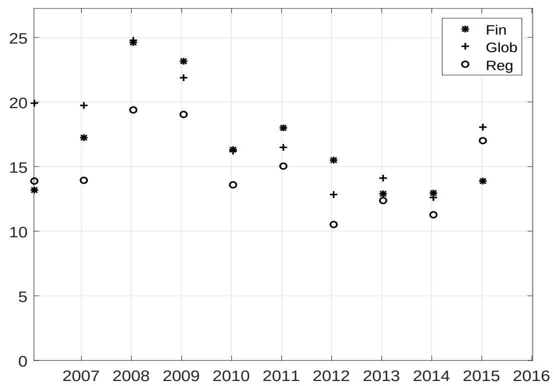
(c) LatinAmerica



(d) Middle-East and Africa



(e) North America



(f) Western Europe

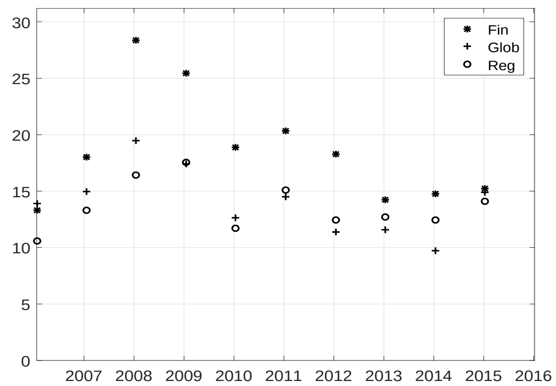
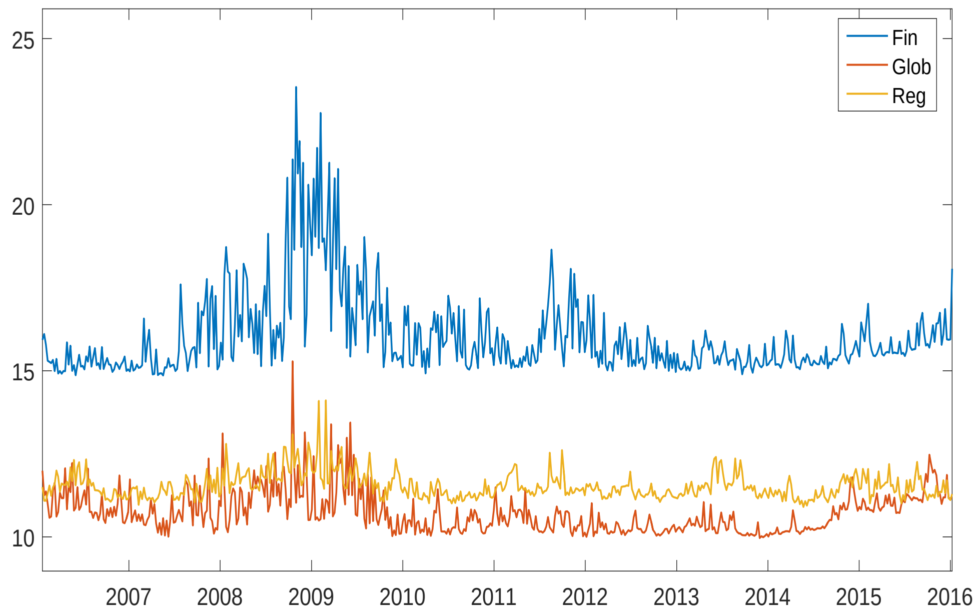


Figure 4: **Time-varying variance decomposition**

The figure reports the average estimated conditional variance decompositions. Panel (a) shows the cross-sectional average of the share of variance explained by each factor at each point in time. Panel (b) reports the total share of variance explained by the factors as the sum of the three series in Panel (a).

(a) Variance decomposition



(b) Variance decomposition

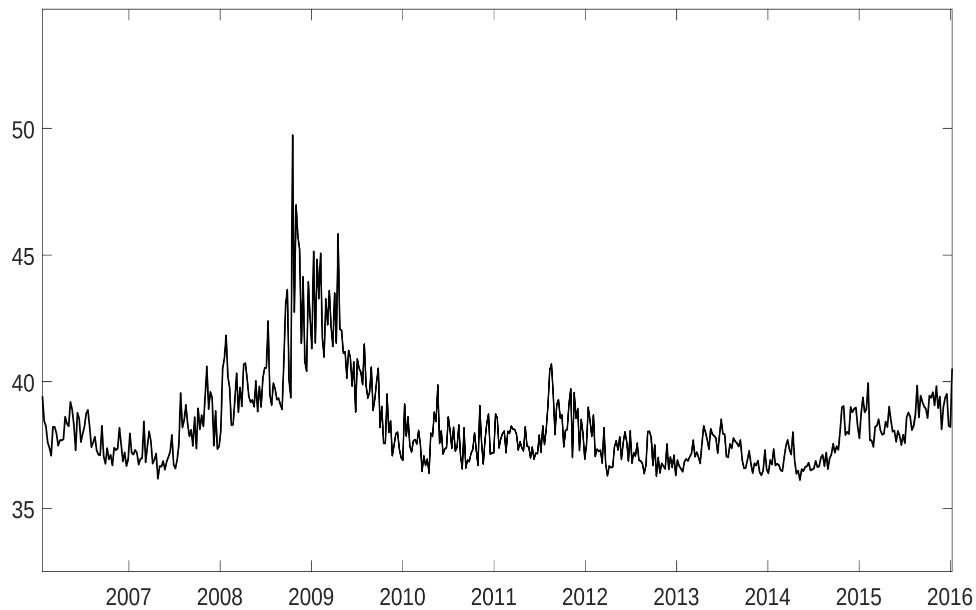
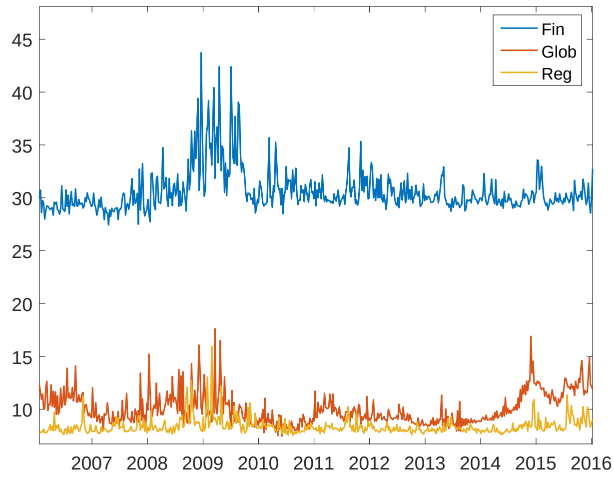


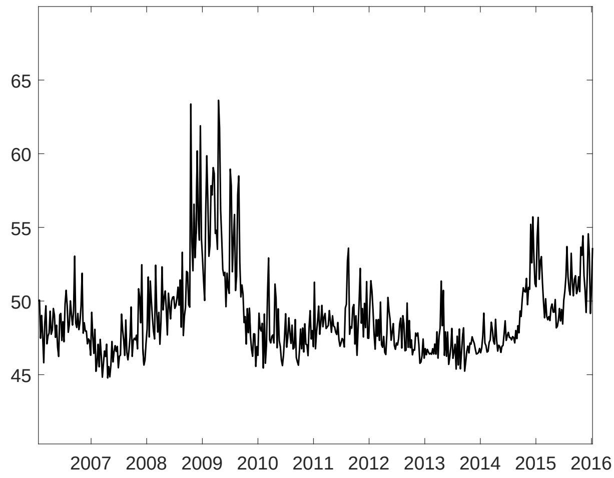
Figure 5: **Time-varying variance decomposition (by region)**

The figure reports the average estimated conditional variance decompositions, aggregated by region. In the left column of panels, the blue line represents the percentage of variance explained by the financial factor, yellow and orange lines represent regional and global factors, respectively. The right column of panels shows the total share of variance explained by the factors, the sum of the left panel.

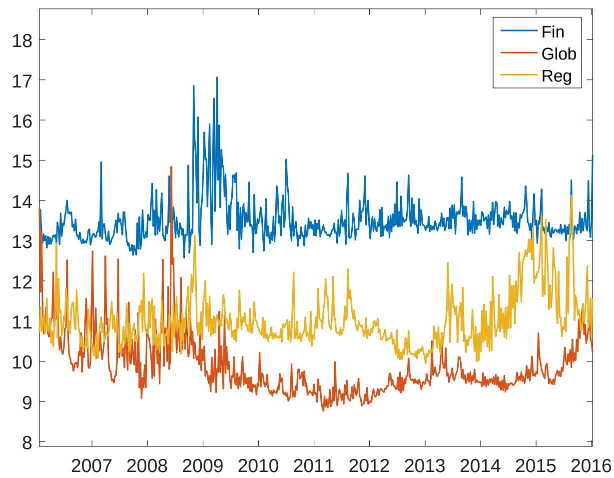
(a) North America



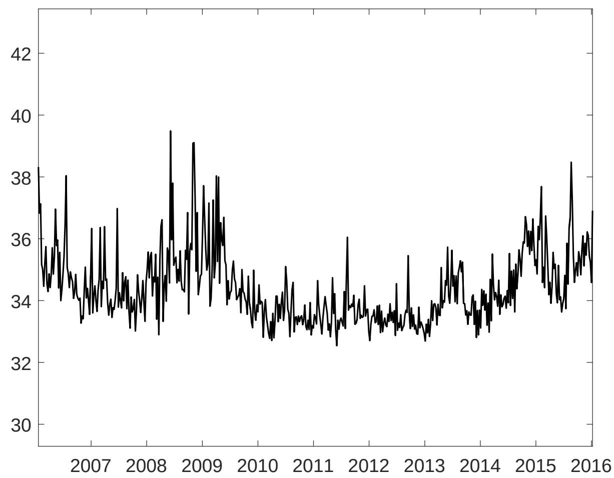
(b) North America total



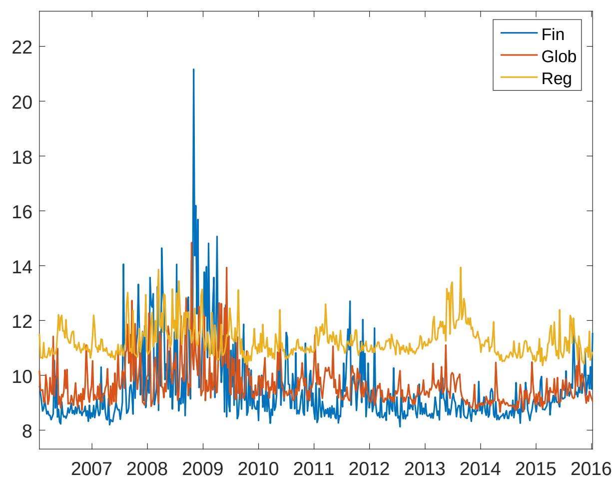
(c) Latin America



(d) Latin America total



(e) Asia-Pacific



(f) Asia-Pacific total

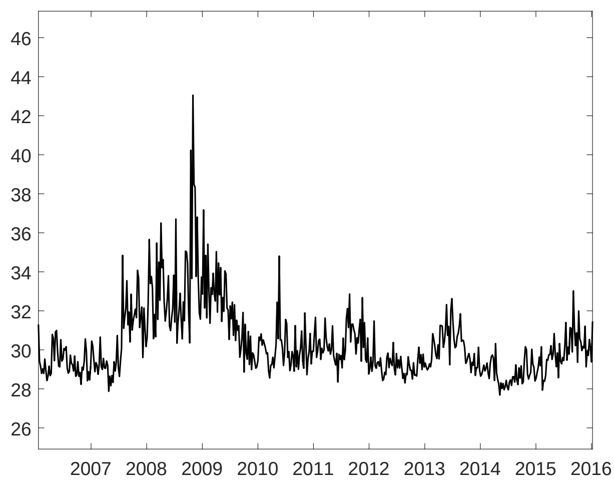
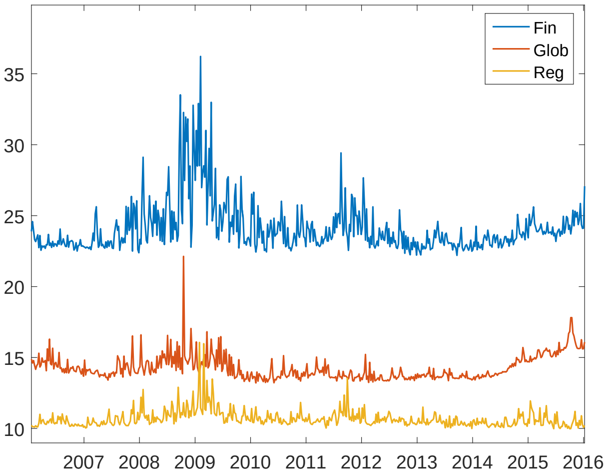


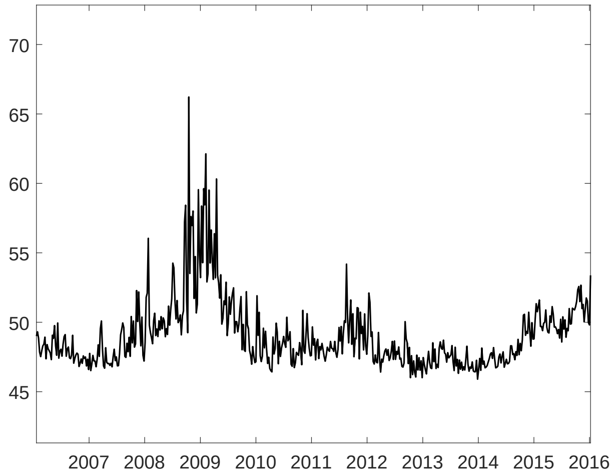
Figure 6: **Time-varying variance decomposition (by region) - continued**

The figure reports the average estimated conditional variance decompositions, aggregated by region. In the left column of panels, the blue line represents the percentage of variance explained by the financial factor, yellow and orange lines represent regional and global factors, respectively.The right column of panels shows the total share of variance explained by the factors, the sum of the left panel.

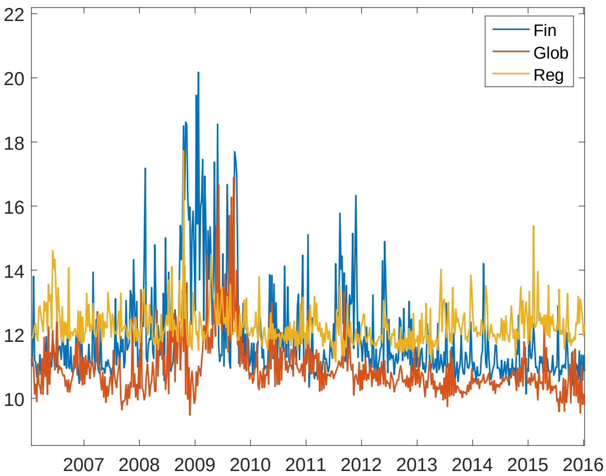
(a) Western Europe



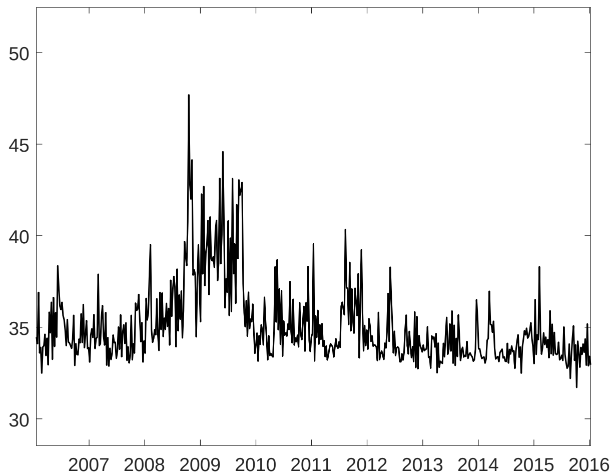
(b) Western Europe total



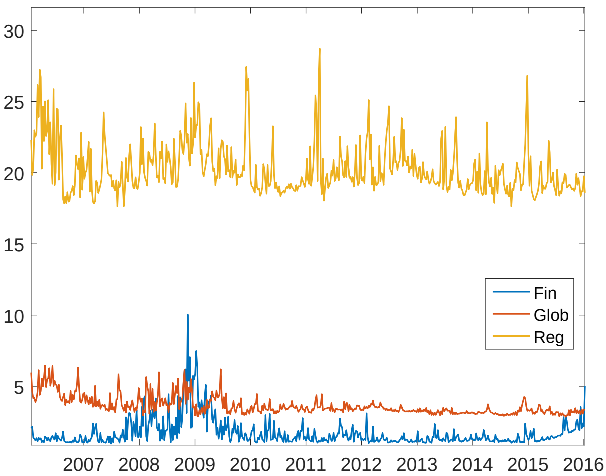
(c) Emerging Europe



(d) Emerging Europe total



(e) Middle-East Africa



(f) Middle-East Africa total

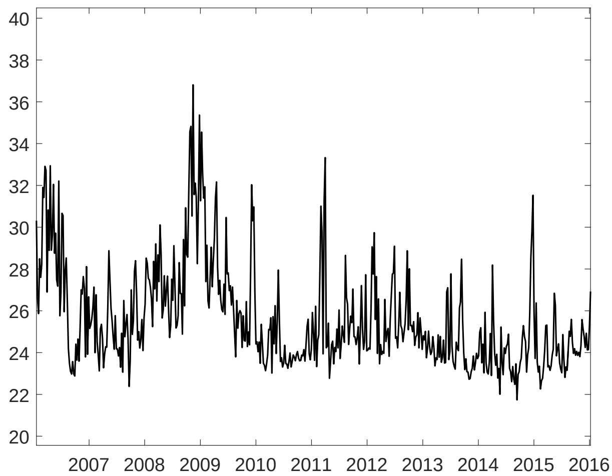
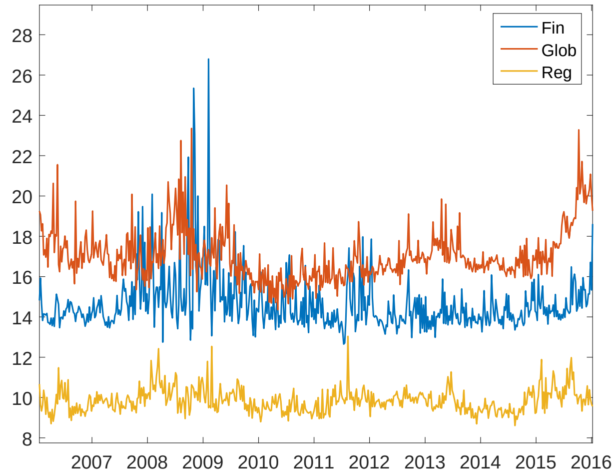


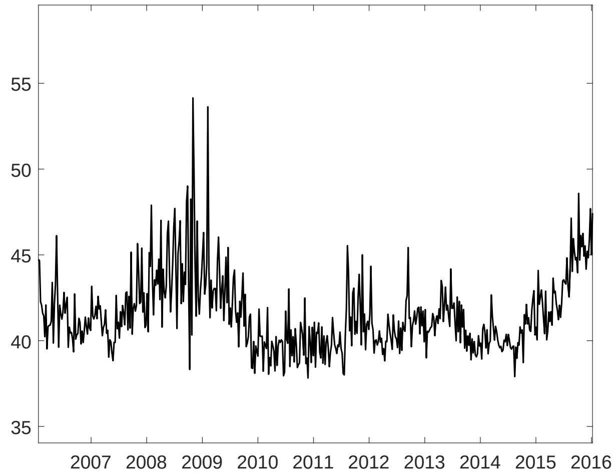
Figure 7: **Time-varying variance decomposition (by sector)**

The figure reports the average estimated conditional variance decompositions, aggregated by sector. In the left column of panels, the blue line represents the percentage of variance explained by the financial factor, yellow and orange lines represent regional and global factors, respectively. The right column of panels shows the total share of variance explained by the factors, the sum of the left panel.

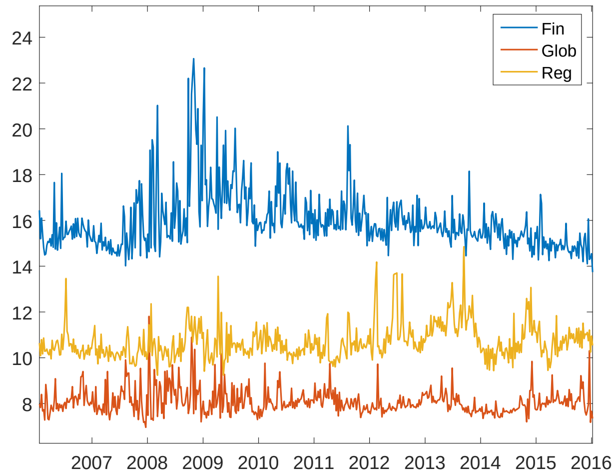
(a) Basic Materials



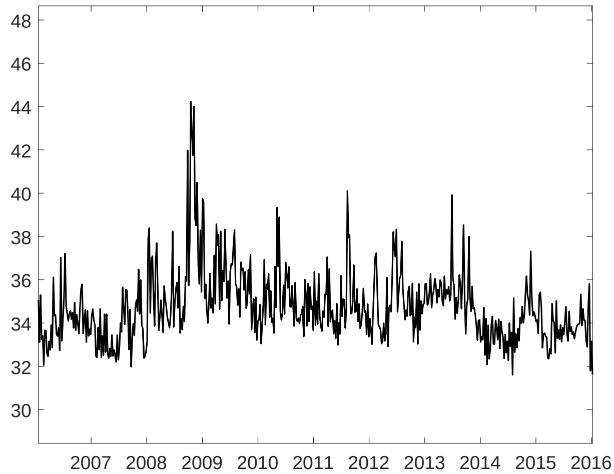
(b) Basic Materials total



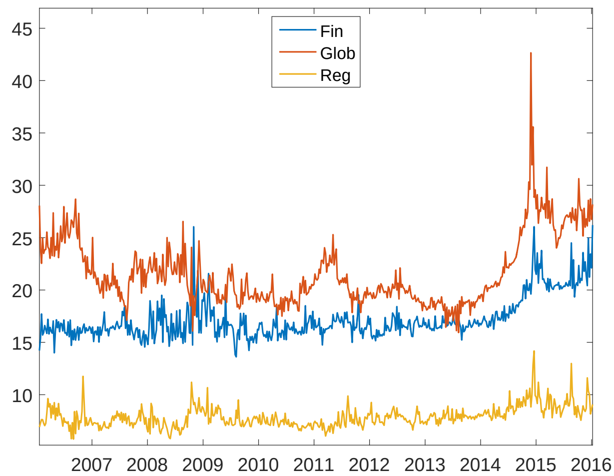
(c) Communications



(d) Communications total



(e) Energy



(f) Energy total

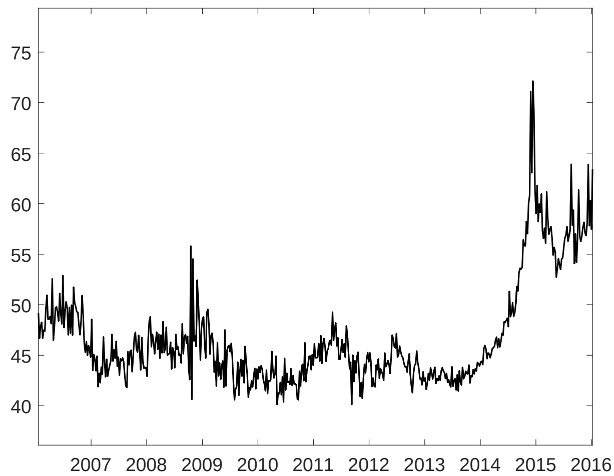
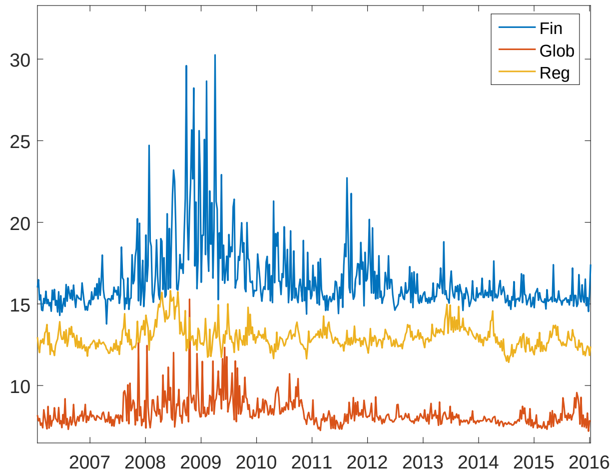


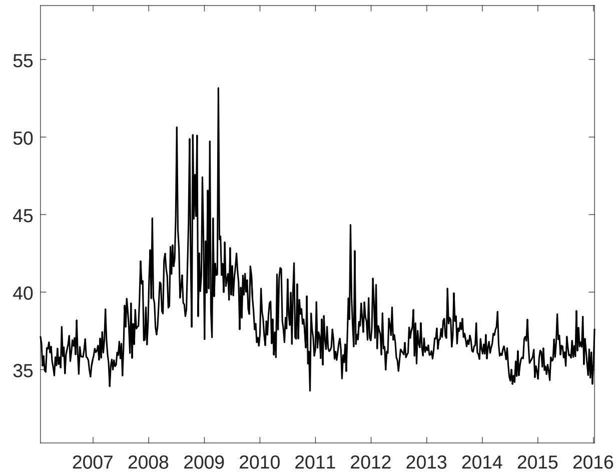
Figure 8: **Time-varying variance decomposition (by sector) - continued**

The figure reports the average estimated conditional variance decompositions, aggregated by sector. In the left column of panels, the blue line represents the percentage of variance explained by the financial factor, yellow and orange lines represent regional and global factors, respectively. The right column of panels shows the total share of variance explained by the factors, the sum of the left panel.

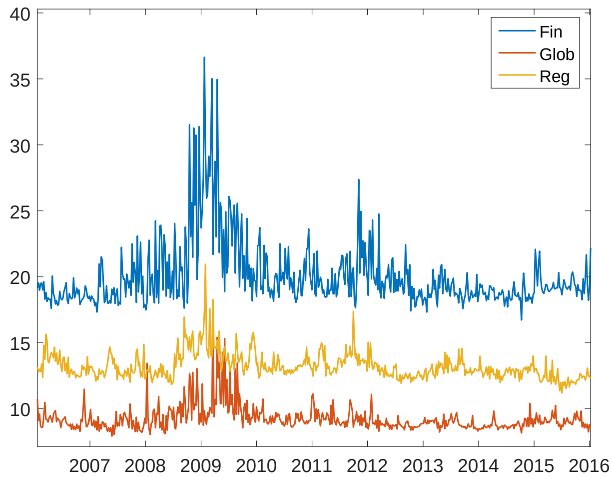
(a) Consumer Cyclical



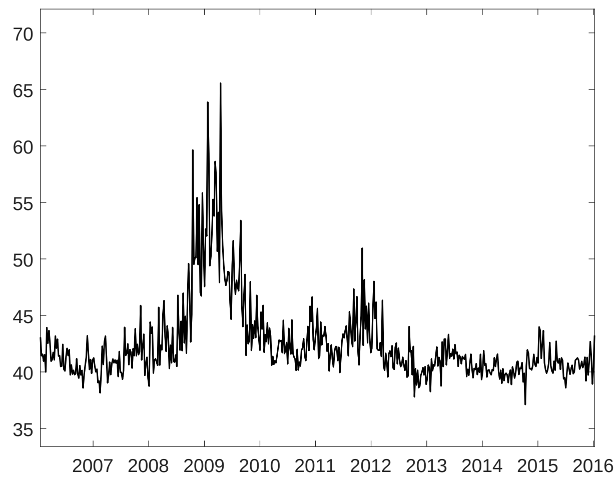
(b) Consumer Cyclical total



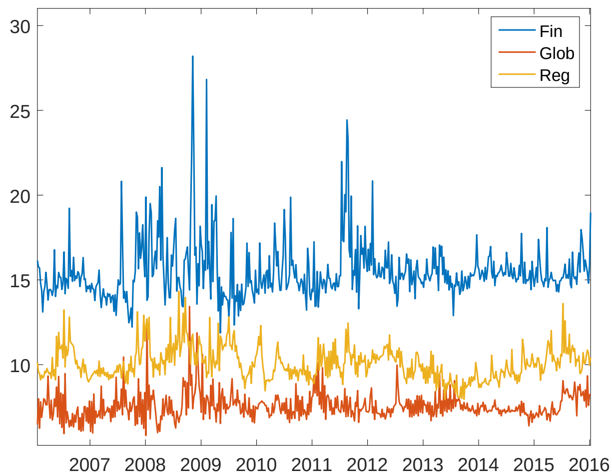
(c) Financial



(d) Financial total



(e) Technology



(f) Technology total

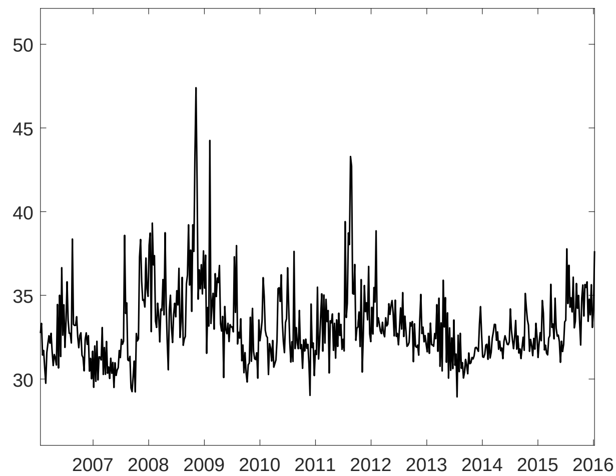
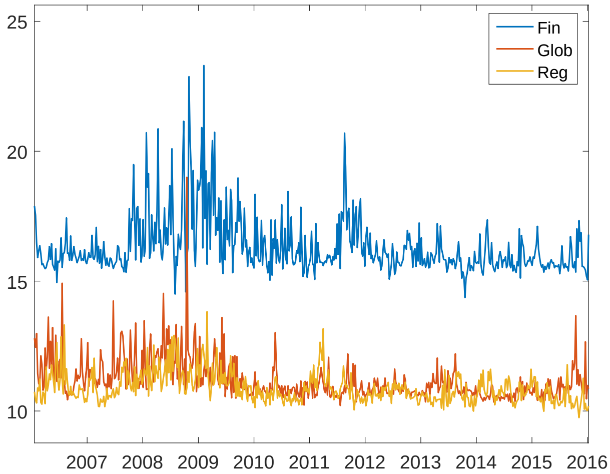


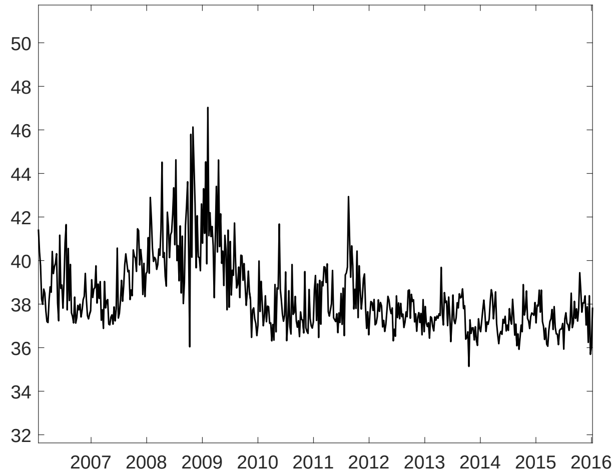
Figure 9: **Time-varying variance decomposition (by sector) - continued**

The figure reports the average estimated conditional variance decompositions, aggregated by sector. In the left column of panels, the blue line represents the percentage of variance explained by the financial factor, yellow and orange lines represent regional and global factors, respectively. The right column of panels shows the total share of variance explained by the factors, the sum of the left panel.

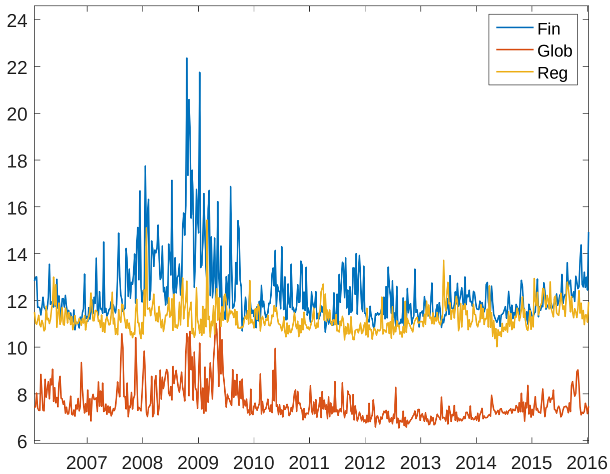
(a) Industrial



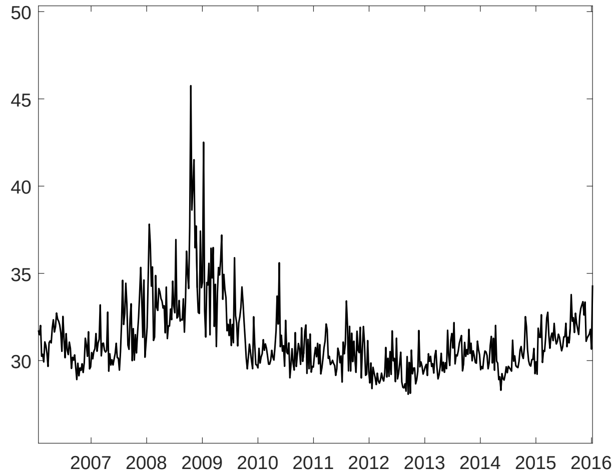
(b) Industrial total



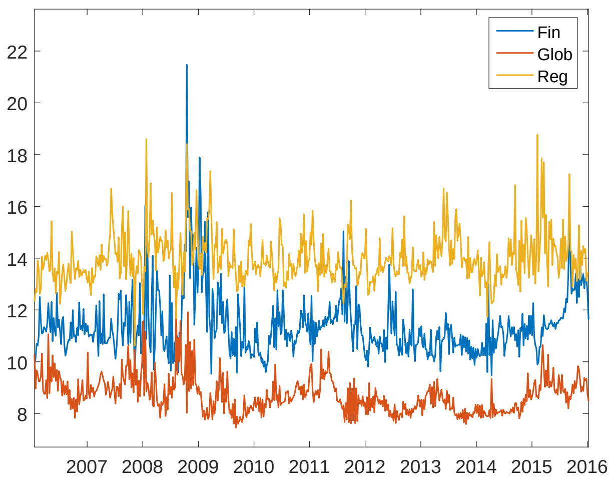
(c) Consumer Non-cyclical



(d) Consumer Non-cyclical total



(e) Utilities



(f) Utilities total

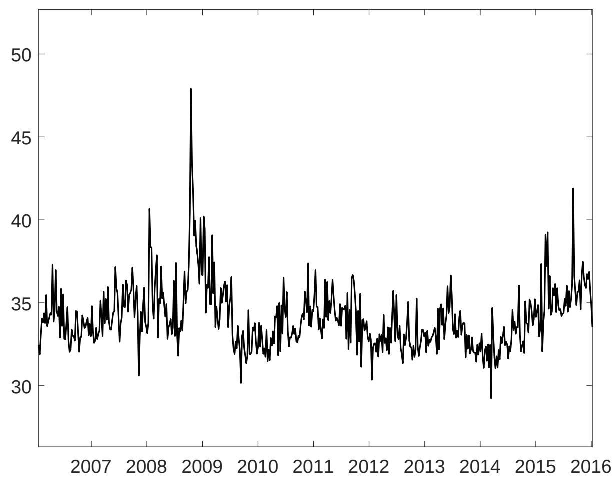
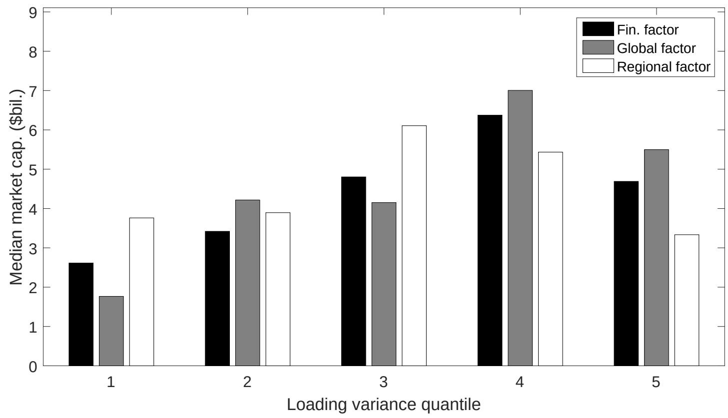


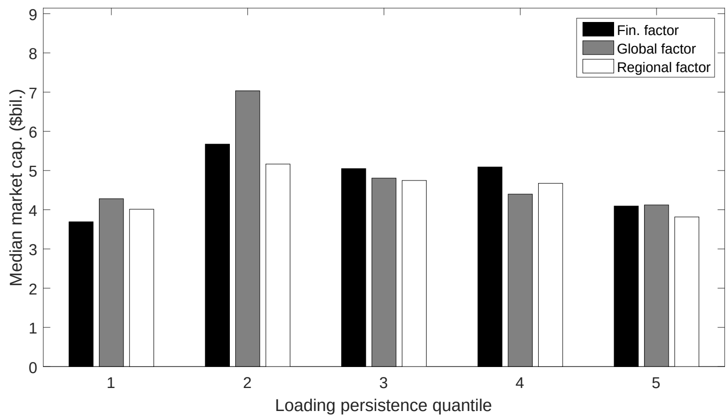
Figure 10: **Relationship with firm size**

The figure reports the median market capitalisation (at the end of the sample) ordered by factor loadings variance (Panel (a)), persistence (Panel (b)) and magnitude (Panel (c)). At the end of the sample, stocks are sorted in quantiles of either loading variance, or persistence, or magnitude. Quantile five contains the largest value. Then, for each quantile we calculate the median market market capitalisation, and we plot it against the quantile number.

(a) Variance of loading



(b) Persistence of loading



(c) Loading magnitude

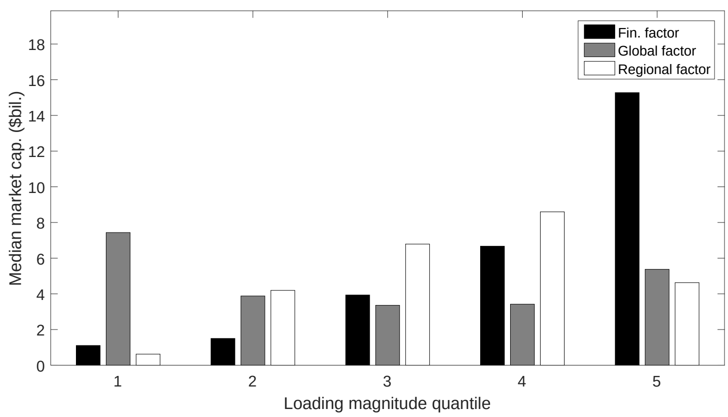
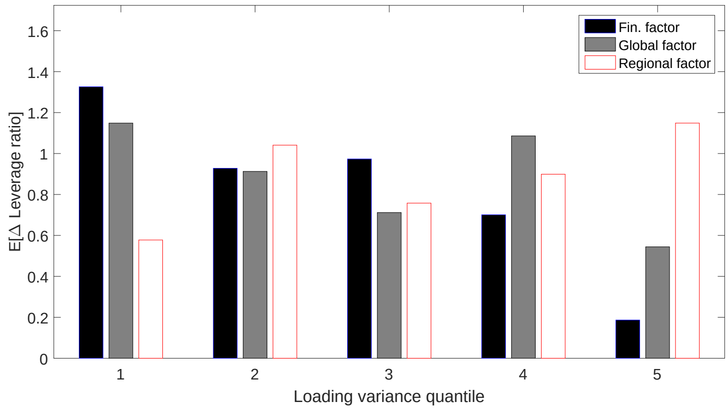


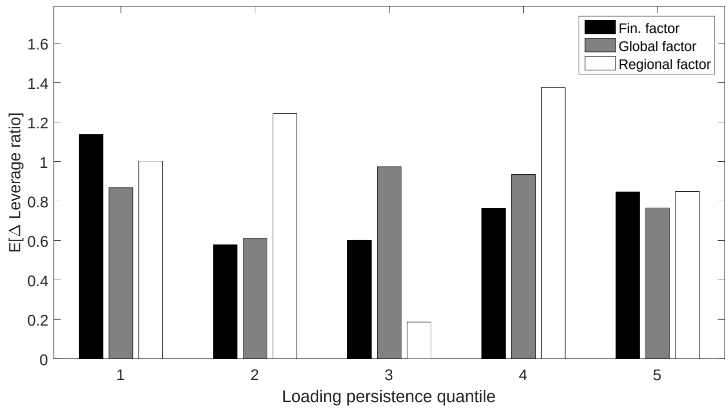
Figure 11: **Relationship with firm leverage**

The figure reports the average quarterly change in leverage ratio (from January 2010 and January 2016) ordered by factor loadings variance (Panel (a)), persistence (Panel (b)) and magnitude (Panel (c)). At the end of the sample, stocks are sorted in quantiles of either loading variance, or persistence, or magnitude. Quantile five contains the largest value. Then, for each quantile we calculate the average quarterly change in leverage ratio (debt over assets), and we plot it against the quantile number. Financials are excluded.

(a) Variance of loading



(b) Persistence of loading



(c) Loading magnitude

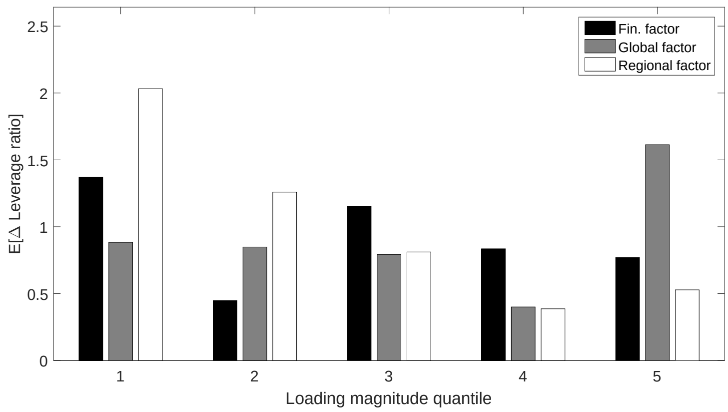
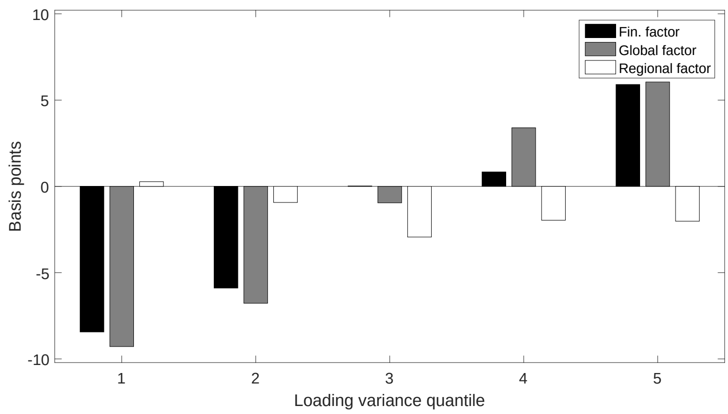


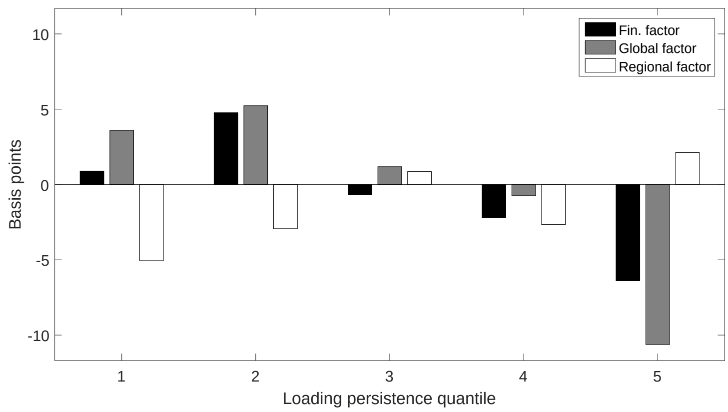
Figure 12: **Relationship with expected returns**

The figure reports the average weekly stock returns from January 2010 until January 2016, expressed in basis points (one basis point = 0.01%), ordered by factor loadings variance (Panel (a)), persistence (Panel (b)) and magnitude (Panel (c)). Financial stocks are excluded. At the end of the sample, stocks are sorted in quantiles of either loading variance, or persistence, or magnitude. Quantile five contains the largest value. Then, for each quantile we calculate the average log-returns, and we plot it against the quantile number.

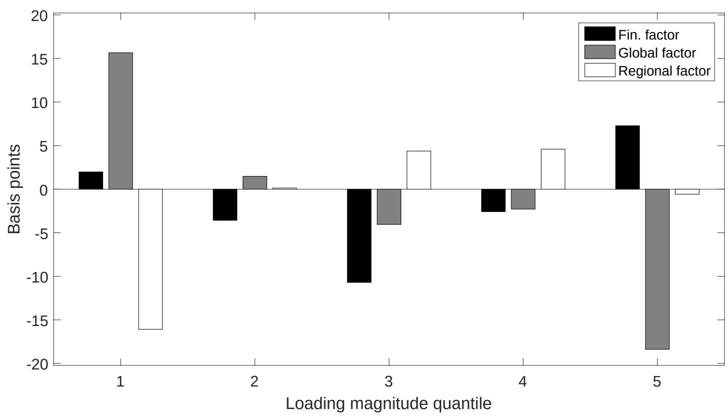
(a) Variance of loading



(b) Persistence of loading



(c) Loading magnitude

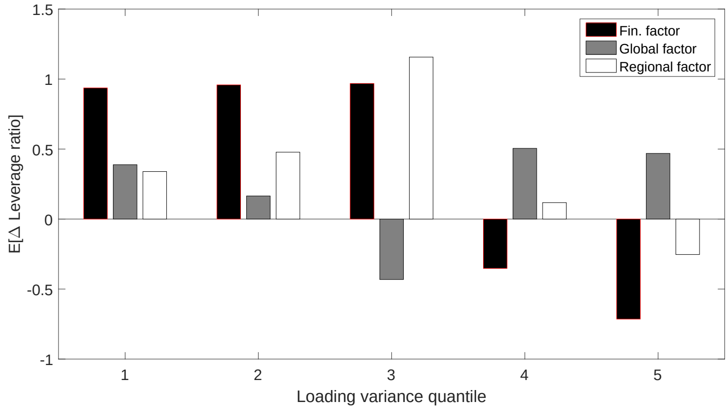


Appendix A Leverage for Financials

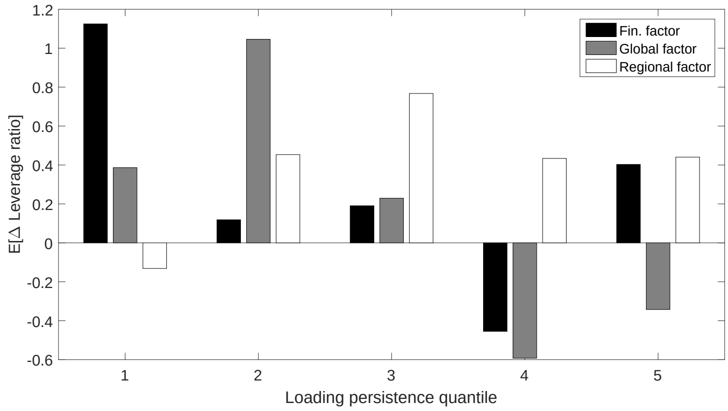
Figure A.1: **Relationship with firm leverage for Financials**

The figure reports the average quarterly change in leverage ratio (from January 2010 and January 2016) for financial firms ordered by factor loadings variance (Panel (a)), persistence (Panel (b)) and magnitude (Panel (c)). At the end of the sample, stocks are sorted in quantiles of either loadings variance, or persistence, or magnitude. Quantile five contains the largest value. Then, for each quantile we calculate the average quarterly change in leverage ratio (debt over assets), and we plot it against the quantile number.

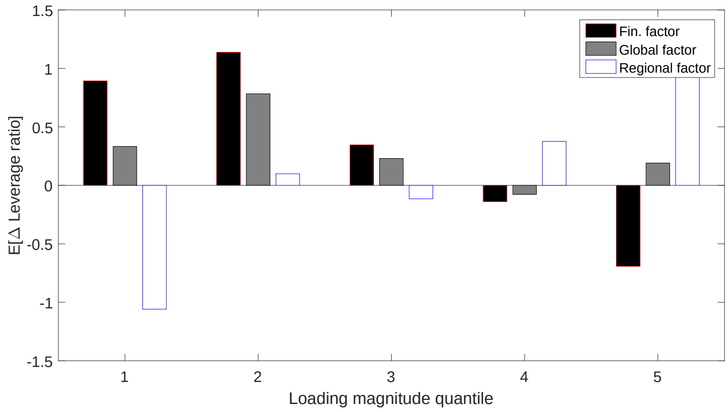
(a) Variance of loading



(b) Persistence of loading



(c) Loading magnitude



Appendix B Data cleaning

The number of time series observations for most stock markets is $T = 835$.

Egypt. Following the "Egyptian Revolution of 2011" that started on the 25^{th} January 2011, the stock exchange closed from the 27^{th} January until the 23^{rd} March, which results in 7 consecutive missing data cells for all stocks. Thus, we repeat the last value which results in 7 zero-return observations.

Russia, Ukraine and India. Table 1 shows that both countries' time series start on 23^{rd} January 2000 and ends on the 10^{th} January 2016. This is unexpected because both days are Sundays, and Bloomberg should have assigned the last working day of the week to the weekly observation. However, in these countries the stock exchange operates normally, from Monday to Friday. Thus, prices refer to the Friday close (or the last available data point of the week) but Bloomberg reports the Sunday date. We checked with other data providers (Datastream) that this is the case. Thus, since we are downloading Friday to Friday data, the last observation is missing and $T = 834$, where the last observation refers to the penultimate week of the data set. Note that we do not need to shift the time series because all the other data points match across countries.

Egypt, Israel, Qatar, UAE. Weekday reference is Thursday, for religious reasons. That is weekends are on Fridays and Saturdays. $T = 835$.

Korea and Taiwan. The time series start on the 22^{nd} January 2000, and they end on 9^{th} January 2016, which are both Saturdays. Hence, an observation is missing at the end of the sample. $T = 834$.

In both countries the stock exchange operates on a traditional trading calendar. Thus, the weekly data refers to the last working-day traded price.

Serbia. For all companies, the time series starts in December 2005. Thus, if we want to include this country in the study we have to trim all time series. Note that, according to Bekaert et al. (2014), Table IV, pag. 2618, the Serbian equities experienced the most negative return over the crisis period, so it is a country worth including.

We eliminate the last observation for countries with 835 observations, and the final time series length is $T = 834$.

Research Papers 2018

CREATES
Center for Research in Econometric Analysis of Time Series

2018-20: Ruijun Bu, Kaddour Hadri and Dennis Kristensen: Diffusion Copulas: Identification and Estimation

2018-21: Kim Christensen, Roel Oomen and Roberto Renò: The drift burst hypothesis

2018-22: Russell Davidson and Niels S. Grønborg: Time-varying parameters: New test tailored to applications in finance and macroeconomics

2018-23: Emilio Zanetti Chini: Forecasters' utility and forecast coherence

2018-24: Tom Engsted and Thomas Q. Pedersen: Disappearing money illusion

2018-25: Erik Christian Montes Schütte: In Search of a Job: Forecasting Employment Growth in the US using Google Trends

2018-26: Maxime Morariu-Patrichi and Mikko Pakkanen: State-dependent Hawkes processes and their application to limit order book modelling

2018-27: Tue Gørgens and Allan H. Würtz: Threshold regression with endogeneity for short panels

2018-28: Mark Podolskij, Bezirgen Veliyev and Nakahiro Yoshida: Edgeworth expansion for Euler approximation of continuous diffusion processes

2018-29: Isabel Casas, Jiti Gao and Shangyu Xie: Modelling Time-Varying Income Elasticities of Health Care Expenditure for the OECD

2018-30: Yukai Yang and Luc Bauwens: State-Space Models on the Stiefel Manifold with A New Approach to Nonlinear Filtering

2018-31: Stan Hurn, Nicholas Johnson, Annastiina Silvennoinen and Timo Teräsvirta: Transition from the Taylor rule to the zero lower bound

2018-32: Sebastian Ankargren, Måns Unosson and Yukai Yang: A mixed-frequency Bayesian vector autoregression with a steady-state prior

2018-33: Carlos Vladimir Rodríguez-Caballero and Massimiliano Caporin: A multilevel factor approach for the analysis of CDS commonality and risk contribution

2018-34: James G. MacKinnon, Morten Ørregaard Nielsen, David Roodman and Matthew D. Webb: Fast and Wild: Bootstrap Inference in Stata Using boottest

2018-35: Sepideh Dolatabadim, Paresh Kumar Narayan, Morten Ørregaard Nielsen and Ke Xu: Economic significance of commodity return forecasts from the fractionally cointegrated VAR model

2018-36: Charlotte Christiansen, Niels S. Grønborg and Ole L. Nielsen: Mutual Fund Selection for Realistically Short Samples

2018-37: Niels S. Grønborg, Asger Lunde, Kasper V. Olesen and Harry Vander Elst: Realizing Correlations Across Asset Classes

2018-38: Riccardo Borghi, Eric Hillebrand, Jakob Mikkelsen and Giovanni Urga: The dynamics of factor loadings in the cross-section of returns