ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Wijler, Etienne; Lucas, André

Working Paper

An impartial look at asset correlation stability and market structure

Tinbergen Institute Discussion Paper, No. TI 2025-051/III

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Wijler, Etienne; Lucas, André (2025) : An impartial look at asset correlation stability and market structure, Tinbergen Institute Discussion Paper, No. TI 2025-051/III, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
https://hdl.handle.net/10419/331471

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

TI 2025-051/III
Tinbergen Institute Discussion Paper

An Impartial Look at Asset Correlation Stability and Market Structure

Etienne Wijler[1]
Andre Lucas[2]

1 Vrije Universiteit Amsterdam

2 Vrije Universiteit Amsterdam, Tinbergen Institute

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and Vrije Universiteit Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at https://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

An Impartial Look at Asset Correlation Stability and Market Structure

Etienne Wijler*[1] and Andre Lucas[1]

[1] *Vrije Universiteit Amsterdam and Tinbergen Institute*

September, 2025

*Corresponding author: Department of Econometrics and Data Science, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV, Amsterdam, the Netherlands. E-mail address: e.j.j.wijler@vu.nl. Tel: +31 (0)20 598 9898

Abstract

We develop a data-driven procedure to identify which correlations in high-dimensional dynamic systems should be time-varying, constant, or zero. The method integrates a vine-based multivariate partial correlation model with sequential penalized estimation. Applied to 50 US equities and systematic risk factors, results indicate that asset-level correlation dynamics are primarily induced by time-varying exposures to systematic factors. We further uncover persistent, non-zero, and occasionally time-varying partial correlations within industries, even after controlling for standard risk and industry factors. Finally, we show how the new methodology may be used to explore the relevance of systematic risk factors in an impartial way.

Keywords: conditional correlations, score-driven models, financial market structure, regularization

JEL codes: C58, C32.

1 Introduction

Asset correlations play a dominant role in many areas of finance, ranging from modern portfolio theory to asset pricing and risk management. Though there are clear indications from the econometric literature that such asset correlations may not be constant over time (Engle, 2002; Engle et al., 2019), the debate in the finance literature still questions whether such time-varying (or even constant) non-zero correlations can be exploited at all (see, for instance, DeMiguel et al., 2009). It is therefore important to decide on the issue of time-variation in asset correlations and its exploitability in an impartial, data-driven way. To do so, there are at least three main challenges. First, we need to correct for the marginal properties of asset returns, including the well-known stylized fact of time-varying volatilities of individual returns (see, e.g., Hansen and Lunde, 2005). Such time-variation in volatilities automatically spills over into time-varying covariances, but need not imply that correlations vary over time. Second, though pairwise asset correlations are of key

interest, one typically first wants to correct these pairwise correlations for any systematic risk exposures; see Fama and French (1993), Carhart (1997), Fama and French (2015, 2016), Feng et al. (2020), and many others. It is therefore key to find out whether the correlations of asset returns with such systematic risk factors vary over time or not, and whether any pairwise asset correlations that *remain after* systematic risk exposures have been accounted for, are still time-varying or not. This shifts the focus from correlations to partial correlations. Third, once more and more assets are included in the analysis and the setting becomes high-dimensional, biases in usual estimation procedures start to emerge and need to be addressed. For instance, Ledoit and Wolf (2022) and Engle et al. (2019) consider shrinkage techniques to target large-dimensional covariance matrices using random matrix theory results. Pakel et al. (2021), by contrast, mitigate biases in the familiar DCC framework of Engle (2002) by introducing composite maximum likelihood methods and splitting the high-dimensional problem into a sequence of lower-dimensional ones. See also Fan et al. (2016) for an overview of other methods. Also in our current context of investigating the time-variation in asset correlations, some form of correction or shrinkage is needed if the number of assets grows larger.

In this paper we therefore propose a new modeling framework that addresses the above challenges and provides an impartial, data-driven perspective on the question of time-variation in partial asset correlations. For this, we build on the dynamic vine copula framework of D'Innocenzo and Lucas (2024). The vine structure allows us to address the layered format of asset correlations, first accounting for correlations with systematic risk factors, and only then looking at any remaining *partial* correlation patterns. This distinguishes our approach from, for instance, Engle et al. (2019), who instead consider the Pearson correlation matrix. It also distinguishes the paper from approaches that consider time-varying precision matrices, i.e., inverse correlation matrices, as developed in Lee et al. (2021), Zhu et al. (2024) and Chen et al. (2025). Entries in the precision matrix can also be interpreted as partial correlations between two assets, but after accounting for the correlation with *all other assets* in the system. Again, this would step

over the clear ordering in such analyses and the distinction between systematic and residual risk. By considering partial correlations, our approach also has the advantage that we do not need to account for positive definiteness of the partial correlation matrix. It suffices that all individual partial correlations in the vine structure are between $+1$ and -1, which can easily be accomplished be simple parameter transformations without the need of more complex correlation (or precision) matrix parameterizations (Creal et al., 2011; Buccheri et al., 2021; Archakov and Hansen, 2021; Hafner and Wang, 2023; Abadir and Rockinger, 2025).

To investigate the issue of time-variation in partial correlations and address the challenge of possible biases if the number of assets increases, we introduce a sequence of lasso-type penalty functions during estimation. These penalties are constructed in such a way that the estimator automatically shrinks to a constant (or even a zero) correlation patterns if the time-variation is not sufficiently pronounced. We set the penalty parameters in a data-driven way. Simulations support that the new method is able to find the time-varying versus the constant correlations both in well-specified and mis-specified settings.

We apply the new method to study the market structure of asset correlations in an application to 50 blue-chip US stock returns selected from 10 different industries and a range of systematic risk factors from the financial literature. Though the number of assets is considerably smaller than in the vast dimensional application of, for instance, Engle et al. (2019), biases already arise in these settings (see Pakel et al., 2021; Lee et al., 2021). Moreover, by concentrating on the firms with the largest average trading volumes across a range of industries, we are able to concentrate more on the financial implications of our econometric analysis. We find clear signals that correlations between individual stocks and the market factor are typically time-varying, irrespective of the frequency of the data (daily, weekly, monthly). We also find that partial asset correlations with familiar risk factors (like size or value) vary over time as well, even after accounting for (time-varying) correlations with the market factor. This latter time-variation, however, is less robust and fades into constant or even zero partial correlations for many assets if the frequency lowers to weekly or

monthly.

Interestingly, we also find that after accounting for systematic risk, many of the remaining partial correlations are set to zero by the penalized estimator, particularly at lower frequencies. This result continues to hold if we alter the number and type of risk factors. One major exception is given by several within industry partial correlation patterns. These within industry effects are non-zero and sometimes even time-varying and are highly robust to the precise specification used. They fade if the model is extended, but do not fully vanish even when we include industry factors as systematic risk factors in the model. Given that we select the five firms with the largest average trading volume in each industry, the (partial) correlations between such industry leaders cannot be fully accounted for by the systematic risk factors and broad industry developments (beyond the top five).

The remainder of this paper is set up as follows. Section 2 introduces the modeling framework. Section 3 presents simulation evidence about the performance of the new approach, both under correct and incorrect model specification. Section 4 provides our empirical application to US stocks. Section 5 concludes.

2 Model

In this section, we first describe the vine-based partial correlation framework of D'Innocenzo and Lucas (2024). Next, we discuss how to regularize the estimator of the model's parameters to distinguish between time-varying, constant, and zero partial asset correlations in a data-driven way.

2.1 Modeling dependence via partial correlations

Let $\boldsymbol{r}_t = (r_{1,t}, \ldots, r_{N,t})^\top \in \mathbb{R}^{N\times 1}$ be a vector of (excess) asset returns. Following D'Innocenzo and Lucas (2024), we split the modeling stage of $\boldsymbol{r}_t$ into two steps. In a first step, we estimate

individual $t(\nu_i)$-GARCH(1,1) volatility models and use the estimated volatility patterns of each asset. Alternative volatility models could be used as well, but the GARCH(1,1) has proven difficult to beat for most financial time series; see Hansen and Lunde (2005). We then focus on modeling the devolatilized series $\boldsymbol{y}_t$, with $\boldsymbol{y}_t = (y_{1,t}, \ldots, y_{N,t})^\top$ and $y_{i,t} = r_{i,t}/\hat{\sigma}_{i,t}$; compare the approach for the well-known Dynamic Conditional Correlation (DCC) model of Engle (2002).

Let $\mathcal{F}_{t-1} = \{\boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}, \ldots\}$ denote a filtration. Then we assume that the conditional distribution of $\boldsymbol{y}_t$ for $t \in \mathbb{Z}$ is given by a Student's t distribution with ν degrees of freedom,

$$\boldsymbol{y}_t \mid \mathcal{F}_{t-1} \sim t\left(\boldsymbol{0}, (1 - 2\nu^{-1}) \cdot \boldsymbol{R}_t, \nu\right), \qquad \nu > 2, \tag{2.1}$$

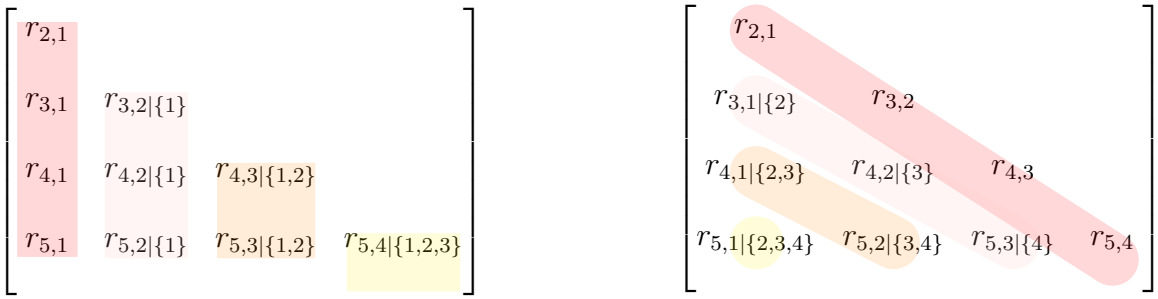
where $t(\boldsymbol{\mu}, \boldsymbol{\Omega}, \nu)$ denotes an N-dimensional Student's t-distribution with location μ, scale matrix $\boldsymbol{\Omega}$, and $\nu > 2$ degrees of freedom.

Rather than modeling the dynamics of the correlations in $\boldsymbol{R}_t$ directly using a parametrization of the Pearson correlation matrix such as the normalized covariance matrix in the DCC (Engle, 2002), the hypersphere parameterization (Jaeckel and Rebonato, 2000; Creal et al., 2011; Buccheri et al., 2021), the log-correlation matrix parameterization (Archakov and Hansen, 2021; Hafner and Wang, 2023), or the normalized spectral parameterization (Abadir and Rockinger, 2025), D'Innocenzo and Lucas (2024) use a vine structure and model the *partial* correlations instead. An advantage is that the vine-based pairwise partial correlations only need to lie between -1 and $+1$, and no further restrictions are needed to ensure that the implied Pearson partial correlation matrix $\boldsymbol{R}_t$ is positive definite.

Let $\rho_{i,j\,;t}$ denote the (i,j)th element of $\boldsymbol{R}_t$, i.e., the Pearson correlation between assets i and j at time t, conditional on the filtration $\mathcal{F}_{t-1}$. Similarly, let $r_{i,j\,;t}$ denote the corresponding *partial* correlation. Such partial correlations can be defined in different ways. For instance, from a re-normalized $\boldsymbol{R}_t^{-1}$ we easily obtain the partial correlations between assets i and j conditional on *all other* assets in the system. Here, however, we define the partial correlation sequentially

through a vine structure. This directly implies that the sequence of the vine structure determines the model, and that the model changes if the type of vine or order of the assets in the vine changes. In general, this might be deemed less desirable. In the context of our research question, however, we argue that this structure comes with two distinct advantages. First, like in D'Innocenzo and Lucas (2024), the estimation problem can easily be split into a sequence of lower-dimensional, bivariate estimation problems. This considerably simplifies the estimation process, particularly if the number of assets is larger, and also turns out to be helpful for the penalization strategy. Second, the order in the vine allows us to explicitly investigate one of our main research questions in an impartial, data-driven way. By first accounting for systematic risk factors in the vine structures, we can investigate whether there are any remaining partial correlations between individual assets *after* partialling out the correlations with the systematic risk factors, or whether time-varying Pearson correlations between individual assets are mainly due to time-varying correlations with common systematic risk factors. Such a test is easy to address in a vine-based framework, but less immediate when concentrating on the partial correlations of two assets after partialling out the correlations with *all* other assets. In addition and as mentioned before, partialling out the correlations with *all* other assets introduces complicating positive definiteness constraints back into the estimation process.

Figure 1 presents two well-known vine structures, namely a vertical and diagonal vine. For the vertical vine, we first estimate the pairwise correlations between the assets in the first column. Next, we estimate the partial correlations between the assets in the second column *after* accounting for the correlations of these assets with the first asset. Similarly, the partial correlations in the third column account for the correlations with the first two assets, and so on. A diagonal vine structure is similar, but does not account for the correlations with all *previous* assets, but rather the correlations with all *intermediate* assets; e.g., the partial correlation of assets 1 and 4 after accounting for the correlations with the intermediate assets 2 and 3. Given that we want to investigate the pairwise partial autocorrelations after accounting first for systematic risk



Vertical vine structure Diagonal vine structure

Figure 1: A comparison of two common vine structures for partial correlations

components, the vertical vine structure is most appropriate. This still leaves some sensitivity to the order of the vine for all assets after the systematic risk factors, and we investigate the robustness of our results to such variations in orderings in the empirical application in Section 4.

The mapping from Pearson correlations $\rho_{i,j\,;t}$ to partial correlations $r_{i,j\,;t}$ can be specified recursively by the equations

$$\begin{aligned} \rho_{i,j\,;t} &= \boldsymbol{R}_{i,1:i-1;t}\, \boldsymbol{R}^{-1}_{1:i-1,1:i-1;t}\, \boldsymbol{R}_{1:i-1,j;t} + r_{i,j\,;t}\, \sqrt{\boldsymbol{V}_{i,i\,;t}\, \boldsymbol{V}_{j,j\,;t}}, \\ V_{i,j} &= \rho_{i,j\,;t} - \boldsymbol{R}_{i,1:i-1;t}\, \boldsymbol{R}^{-1}_{1:i-1,1:i-1;t}\, \boldsymbol{R}_{1:i-1,j;t}, \end{aligned} \tag{2.2}$$

for $i = 1, \ldots, N-1$, and $j = i+1, \ldots, N$, where $\boldsymbol{R}_{i:j,k:\ell;t}$ is the submatrix from $\boldsymbol{R}_t$ containing rows i to j and columns k to ℓ, with $\boldsymbol{R}_{i:j,k:\ell;t}$ being empty if $j < i$ or $\ell < k$. Using this relation, we can easily convert partial correlations $r_{i,j\,;t}$ into Pearson correlations $\rho_{i,j\,;t}$, and vice versa. If all partial correlations lie strictly between -1 and $+1$, the implied Pearson correlation matrix $\boldsymbol{R}_t$ is automatically positive definite, with ones on the diagonal.

2.2 Modeling partial correlation dynamics

To model the time-variation in the partial correlations $r_{i,j\,;t}$, we define time-varying parameters $f_{i,j\,;t}$, with $r_{i,j\,;t} = \epsilon \tanh(f_{i,j\,;t}) \in (-\epsilon, +\epsilon)$ for some constant $0 < \epsilon < 1$. As a result of this parameterization, the partial correlations automatically lie in the correct interval -1 to $+1$. Time-variation in $f_{i,j\,;t}$ can now be captured via the score-driven framework of Creal et al. (2013)

and Harvey (2013). For this, we first realize from the distributional assumption in Eq. (2.1) that two elements $y_{i,t}$ and y_{jt} for $j > i$ conditional on $\mathcal{F}_{t-1}$ and on all assets $1, \ldots, i-1$ follow a bivariate Student's t distribution

$$\boldsymbol{y}_{i,j\,;t}\,|\mathcal{F}_{t-1}, \boldsymbol{y}_{1:i-1;t} \quad \sim \quad t\left(\boldsymbol{\mu}_{i,j\,;t}\;,\; \boldsymbol{D}_{i,j\,;t}^{-1/2}\,\boldsymbol{P}_{i,j\,;t}\boldsymbol{D}_{i,j\,;t}^{-1/2}\;\;,\;\nu_{i,j}\right), \tag{2.3}$$

$$\boldsymbol{D}_{i,j\,;t} = \frac{\nu - 2 + \boldsymbol{y}_{1:i-1;t}^{\top}\boldsymbol{R}_{1:i-1,1:i-1;t}^{-1}\boldsymbol{y}_{1:i-1;t}}{\nu_{i,j}} \begin{pmatrix} \boldsymbol{V}_{i,i\,;t} & 0 \\ 0 & \boldsymbol{V}_{j,j\,;t} \end{pmatrix},$$

$$\boldsymbol{\mu}_{i,j\,;t} = \begin{pmatrix} \boldsymbol{R}_{i,1:i-1;t} \\ \boldsymbol{R}_{j,1:i-1;t} \end{pmatrix} \boldsymbol{R}_{1:i-1,1:i-1;t}^{-1}\,\boldsymbol{y}_{1:i-1;t}, \qquad \nu_{i,j} = \nu + j - i - 1,$$

$$\boldsymbol{P}_{i,j\,;t} = \begin{pmatrix} 1 & r_{i,j\,;t} \\ r_{i,j\,;t} & 1 \end{pmatrix}, \quad \boldsymbol{y}_{i,j\,;t} = \begin{pmatrix} y_{i,t} \\ y_{jt} \end{pmatrix}, \quad \boldsymbol{y}_{1:i-1;t} = \begin{pmatrix} y_{1,t} \\ \vdots \\ y_{i-1,t} \end{pmatrix},$$

for $i = 1, \ldots, N-1$ and $j = i+1, \ldots, N$. The score-driven dynamics for $f_{i,j\,;t}$ are then given by

$$f_{i,j\,;t+1}(\boldsymbol{\theta}) = \omega_{i,j} + \beta_{i,j}\, f_{i,j\,;t}(\boldsymbol{\theta}) + \alpha_{i,j}\, s_{i,j\,;t}(\boldsymbol{\theta}), \qquad r_{i,j\,;t}(\boldsymbol{\theta}) = \epsilon \cdot \tanh\left(f_{i,j\,;t}(\boldsymbol{\theta})\right), \tag{2.4}$$

$$\begin{aligned} s_{i,j\,;t}(\boldsymbol{\theta}) &= s_{i,j\,;t}\left(f_{i,j\,;t}(\boldsymbol{\theta}), \boldsymbol{y}_{i,j\,;t}^{\star}; \boldsymbol{\theta}\right) \\ &= \frac{\epsilon}{1 - r_{i,j\,;t}(\boldsymbol{\theta})^2} \cdot \Big(\left(1 + r_{i,j}(\boldsymbol{\theta})^2\right)\left(w_{i,j\,;t}\, y_{i,t}^{\star}\, y_{jt}^{\star} - r_{i,j\,;t}(\boldsymbol{\theta})\right) \\ &\qquad\qquad - r_{i,j\,;t}(\boldsymbol{\theta})\,\left(w_{i,j|L_{i,j}\,;t}\boldsymbol{y}_{i,j\,;t}^{\star\,\top}\boldsymbol{y}_{i,j\,;t}^{\star} - 2\right)\Big), \end{aligned} \tag{2.5}$$

$$w_{i,j\,;t} = w_{i,j\,;t}\left(f_{i,j\,;t}(\boldsymbol{\theta}), \boldsymbol{y}_{i,j\,;t}^{\star}; \boldsymbol{\theta}\right) = \frac{\nu_{i,j} + 2}{\nu_{i,j} + \boldsymbol{y}_{i,j\,;t}^{\star\,\top}\boldsymbol{P}_{i,j\,;t}^{-1}\,\boldsymbol{y}_{i,j\,;t}^{\star}}, \tag{2.6}$$

$$\boldsymbol{y}_{i,j\,;t}^{\star} = \begin{pmatrix} y_{i,t}^{\star} \\ y_{jt}^{\star} \end{pmatrix} = \boldsymbol{y}_{i,j\,;t}^{\star}(\boldsymbol{\theta}) = \boldsymbol{D}_{i,j\,;t}^{-1/2}\,\left(\boldsymbol{y}_{i,j\,;t} - \boldsymbol{\mu}_{i,j\,;t}\right),$$

for $i = 1, \ldots, N-1$ and $j = i+1, \ldots, N$, where $\boldsymbol{\theta}$ gathers the static parameters in the model, and where we made the dependence of quantities like $f_{i,j\,;t}(\boldsymbol{\theta})$, $r_{i,j\,;t}(\boldsymbol{\theta})$, etc., on $\boldsymbol{\theta}$ more explicit in

the notation; see D'Innocenzo and Lucas (2024) for further details. As mentioned earlier, from the pairwise partial correlations $r_{i,j\,;t}(\boldsymbol{\theta})$ we can easily construct the implied pairwise Pearson correlations $\rho_{i,j\,;t}(\boldsymbol{\theta})$ via (2.2).

2.3 Sequential regularized estimation in high dimensions

D'Innocenzo and Lucas (2024) estimate their dynamic vine copula model by maximum likelihood or by a sequential estimation strategy. Full maximum likelihood estimation quickly becomes computationally impractical in high dimensions given the number of parameters and the non-linearity of the model. In our context, we therefore adopt a sequential estimation strategy. Here we first estimate the models for $j = i+1, \ldots, N$ for given i. Note that this step can be parallelized, yielding further computational gains. We then loop the estimation over $i = 1, \ldots, N-1$. We further increase the numerical speed of the algorithm by treating the degrees of freedom parameter during this stage as a tuning constant and fixing it at a reasonable value given estimation results from the literature. In our case, we fix $\nu = 7$ and investigate the sensitivity of the final results to this choice. In the empirical application in Section 4 this choice does not affect the main results. After the parameters for the pairs (i, j) for $j = i+1, \ldots, N$ have been estimated, we compute $\boldsymbol{y}^{\star}_{i+1:N;t} = \boldsymbol{y}^{\star}_{i+1:N;t}(\boldsymbol{\theta})$ and use this new standardized data vector to estimate the pairwise parameters for $i+1$ and $j = i+2, \ldots, N$.

A key difference with D'Innocenzo and Lucas (2024), however, is that we introduce a penalized objective function to estimate the pairwise parameters in $\boldsymbol{\theta}_{i,j} = (\omega_{i,j}, \alpha_{i,j}, \beta_{i,j})^{\top}$ using the penalized sequential maximum likelihood estimator (MLE)

$$\hat{\boldsymbol{\theta}}_{i,j} = \underset{\boldsymbol{\theta}_{i,j}}{\arg\min}\ \mathcal{L}_{i,j}\left(\boldsymbol{\theta}\right) + \lambda_{i,j}\sqrt{\alpha_{i,j}^2 + \beta_{i,j}^2} + \lambda_{i,j}|\omega_{i,j}|, \tag{2.7}$$

where

$$\mathcal{L}_{i,j}(\boldsymbol{\theta}_{i,j}) = \sum_{t=1}^{T} -\log(2\pi) - \tfrac{1}{2}\log|\boldsymbol{P}_{i,j\,;t}| - \frac{\nu_{i,j}+2}{2}\log\left(1+\nu_{i,j}^{-1}\,\boldsymbol{y}^{\star}_{i,j\,;t}{}^{\top}\boldsymbol{P}^{-1}_{i,j\,;t}\,\boldsymbol{y}^{\star}_{i,j\,;t}\right). \tag{2.8}$$

The penalized estimator in (2.7) introduces two types of penalties. The first penalty involving $(\alpha_{i,j}^2+\beta_{i,j}^2)^{1/2}$ introduces a joint penalty on $\alpha_{i,j}$ and $\beta_{i,j}$ and shrinks both parameters *jointly* to zero. The advantage of this approach over shrinking the parameters individually is that any sticky time-variation in the partial correlation $r_{i,j\,;t}$ is determined by both $\alpha_{i,j}$ and $\beta_{i,j}$. Absence of time-variation in the partial correlations, therefore, corresponds to both parameters being *jointly* zero. If both $\alpha_{i,j}$ and $\beta_{i,j}$ are zero, the time-varying parameter $f_{i,j\,;t}$ collapses to the intercept value $\omega_{i,j}$, yielding a *constant* partial correlation $r_{i,j\,;t} = \epsilon \cdot \tanh(\omega_{i,j})$. The joint rather than separate penalty functions also avoids an identification issue: if $\alpha_{i,j}=0$, the parameters $\omega_{i,j}$ and $\beta_{i,j}$ cannot be identified separately (see, for instance, the identification part in Blasques et al., 2022). By enforcing joint selection of $\alpha_{i,j}$ and $\beta_{i,j}$, this issue is mitigated. Finally, the second penalty in (2.7) is a standard L_1 Lasso type penalty. Whereas the first penalty shrinks toward constant partial correlations, this second penalty shrinks towards partial correlations that are (constant at or fluctuating around) *zero*.

To solve for the MLE, we use a proximal gradient descent algorithm; see Appendix C. For given i, we parallelize the optimizations over $j = i+1,\ldots,N$. Particularly for the first layers of the vine structure, this produces substantial further computational gains. Throughout the remainder of this paper, we choose $\lambda_{i,j}$ separately for each data slice using a grid of 20 different $\lambda_{i,j}$ values. This grid ranges from 1 to 1,000 and is evenly spaced on a natural logarithmic scale. After estimating all 20 solutions, the optimal solution is determined based on the BIC using the number of non-zero parameters as a proxy for the degrees of freedom. We investigate the robustness to this choice of the model selection criterion in Section 4 by considering alternative criteria like the AIC.

3 Simulations

3.1 Set-up

In this section, we explore the finite sample performance of our method through Monte Carlo simulations. First, we explore the predictive performance, estimation accuracy and selection capabilities of our estimator in a correct specification setting. Next, we evaluate the robustness of our filtering approach to model mis-specification.

We consider three variations of our estimator as defined in (2.7), which we label VCC (for $\lambda_{i,j} = 0$), standing for **V**ine **C**onditional **C**orrelation, and the penalized VCC or P-VCC (for $\lambda_{i,j} \neq 0$), respectively. For the P-VCC estimator, we consider a version that selects $\lambda_{i,j}$ using the BIC and AIC, respectively. In addition to our three VCC specifications, we also include the DCC model of Engle (2002) as a typical benchmark.

We initialize the VCC filters by first computing the sample correlation matrix $\hat{\boldsymbol{R}}^{\text{init}}$ of the observations $\boldsymbol{y}_t$ for $t = 1, \ldots, T$. Using the vine structure and $\hat{\boldsymbol{R}}^{\text{init}}$, we can calculate the implied partial correlations $\hat{r}_{i,j}^{\text{init}}$ and use them to set $\hat{f}_{i,j}^{\text{init}} = \tanh^{-1}(\hat{r}_{i,j}^{\text{init}}/\epsilon)$. Moreover, to reduce the risk of local optima for the likelihood maximization of each slice, we consider three different initializations for $\boldsymbol{\theta}_{i,j}$, namely (i) $\boldsymbol{\theta}_{i,j} = (\hat{f}_{i,j}^{\text{init}}(1 - 0.95), 0.04, 0.95)^\top$, (ii) $\boldsymbol{\theta}_{i,j} = (\hat{f}_{i,j}^{\text{init}}, 0, 0)^\top$, and (iii) zero ($\boldsymbol{\theta}_{i,j} = (0, 0, 0)^\top$. These correspond to specifications with (i) dynamic, (ii) static, and (iii) zero partial correlations, respectively. Experiments with other initializations show that the results are robust.

In each simulation s, we generate $T_{\text{train}}+T_{\text{test}}$ observations. We use the first T_{train} observations to estimate the model's parameters, and the remaining T_{test} observations to compute the performance metrics for each estimator in each simulation to avoid optimistic performance measured due to over-fitting. We consider four different performance metrics. To evaluate the predictive performance,

we use the root mean squared forecast error (RMSFE)

$$\text{RMSFE} = \left(\frac{2}{N(N-1) \cdot S \cdot T_{\text{test}}} \sum_{s=1}^{S} \sum_{t=1}^{T_{\text{test}}} \left\| \text{vech}\left(\hat{\boldsymbol{R}}_t^{(s)} - \boldsymbol{R}_t^{(s)} \right) \right\|_2^2 \right)^{1/2},$$

where S denotes the number of simulations, and $\hat{\boldsymbol{R}}_t^{(s)}$ and $\boldsymbol{R}_t^{(s)}$ denote the one-step-ahead forecast and the true value of the Pearson correlation matrix $\boldsymbol{R}_t^{(s)}$ in simulation s, respectively.

As a measure of estimation accuracy, we compute the root mean squared error (RMSE) for each estimator as

$$\text{RMSE} = \left(\frac{1}{\dim(\boldsymbol{\theta}^{(s)}) \cdot S} \sum_{s=1}^{S} \left\| \hat{\boldsymbol{\theta}}^{(s)} - \boldsymbol{\theta}^{(s)} \right\|_2^2 \right)^{1/2},$$

where $\boldsymbol{\theta}^{(s)}$ denotes the true vector of static parameters in simulation s, and $\hat{\boldsymbol{\theta}}^{(s)}$ its corresponding estimate. Since $\boldsymbol{\theta}$ is sparse, the RMSE is naturally expected to be lower for penalized versions of VCC. However, such a gain in estimation accuracy may come at the expense of substantial shrinkage bias for the non-zero parameters. To explore this possibility, we also compute the RMSE over the set of non-zero parameters as

$$\text{RMSE}_{\text{TV}} = \left(\frac{1}{\dim(\boldsymbol{\theta}^{(s)}) \cdot S} \sum_{s=1}^{S} \left\| \hat{\boldsymbol{\theta}}_{\text{TV}}^{(s)} - \boldsymbol{\theta}_{\text{TV}}^{(s)} \right\|_2^2 \right)^{1/2},$$

where $\boldsymbol{\theta}_{\text{TV}}^{(s)}$ denotes the set of non-zero parameters that governs the dynamics of the time-varying partial correlations.

Finally, to evaluate the ability of the P-VCC estimator to select the correct non-zero parameters, we also report the nonzero (NZP) and zero precision (ZP) metrics

$$\text{NZP} = \frac{1}{S} \sum_{s=1}^{S} \frac{\sum_{j=1}^{M} \mathbb{1}\{\hat{\theta}_j^{(s)} \neq 0\} \cdot \mathbb{1}\{\theta_j^{(s)} \neq 0\}}{\sum_{j=1}^{M} \mathbb{1}\{\theta^{(s)} \neq 0\}}, \quad \text{ZP} = \frac{1}{S} \sum_{s=1}^{S} \frac{\sum_{j=1}^{M} \mathbb{1}\{\hat{\theta}_j^{(s)} \neq 0\} \cdot \mathbb{1}\{\theta_j^{(s)} \neq 0\}}{\sum_{j=1}^{M} \mathbb{1}\{\theta^{(s)} \neq 0\}},$$

where M denotes the number of elements in the parameter vector $\boldsymbol{\theta}^{(s)}$, including both zero and

non-zero elements.

3.2 Correct specification

In the correctly specified setting, we draw samples directly from (2.1) with $\nu = 7$, where the time-varying correlation matrices $\boldsymbol{R}_t$ for $t = 1, \dots, T$, are generated via sequences of $\mathcal{F}_{t-1}$-measurable partial correlations. These partial correlations follow the filter process in (2.4). The corresponding Pearson correlation matrices are constructed from the partial correlations via a vertical vine structure, similar as in our empirical application in Section 4. The first column of the vine structure, containing $N - 1$ Pearson correlations, is generated through (2.4) with parameters $\boldsymbol{\theta}_{(1,j)} = (\omega_{1,j}, \alpha_{1,j}, \beta_{1,j})^\top$, where $\alpha_{1,j} = 0.1$, $\beta_{1,j} = 0.97$, and $\omega_{1,j} \sim \mathrm{N}(\mu, 0.3)$ with $\mathbb{P}(\mu = 0.7) = 0.8$ and $\mathbb{P}(\mu = -0.7) = 0.2$, for $j = i + 1, \dots, N$. The other columns in the vertical vine structure are generated as zeros, i.e., $\boldsymbol{\theta}_{i,j} = \mathbf{0}$ for all $i = 2, \dots, N$.

We draw $S = 500$ training samples of varying sample sizes $T \in \{750, 1500, 3000\}$. At the end of each training sample, we draw an additional test sample of size $T_{\text{test}} = 50$ on the basis of which we compare the out-of-sample predictive performance. In addition, we vary the dimensionality of the sample by setting $N \in \{10, 50\}$, which corresponds to a total of 135 or 3,675 parameters to be estimated, respectively. The results are reported in Table 1.

For the low-dimensional case ($N = 10$) as reported in panel A, the results show that in terms of forecasting accuracy all VCC variants obtain low RMSFEs, with a maximum discrepancy between the true and filtered out-of-sample Pearson correlation of about 0.05. The most precise forecasts are obtained by P-VCC tuned by the BIC, closely followed by P-VCC tuned by the AIC. Despite the relatively small dimension, we observe that the penalization improves the predictive performance from 11% ($T = 750$) to 20% ($T = 3,000$), respectively, vis-à-vis the unpenalized VCC. We also note that the RMSFE decreases in T for all VCC variants. The DCC model predicts substantially less accurately with a minimum deviation of around 0.136 ($T = 3,000$) from the true correlations.

Table 1: This table contains the simulation results on a correctly specified model. The reported metrics are the root mean squared forecast error (RMSFE), root mean squared error (RMSE), nonzero precision (NZP) and zero precision (ZP).

Metric	T	VCC penalty			DCC	VCC penalty			DCC
		no	BIC	AIC		no	BIC	AIC	
		Panel A: low-dimensional ($N = 10$)				Panel B: high-dimensional ($N = 50$)			
RMSFE	750	0.052	**0.046**	0.047	0.145	0.052	**0.045**	0.046	0.197
	1,500	0.035	**0.029**	0.031	0.140	0.036	**0.030**	0.031	0.187
	3,000	0.025	**0.020**	0.021	0.136	0.025	0.020	**0.021**	0.178
RMSE	750	0.469	**0.013**	0.070	____	0.513	**0.008**	0.074	____
	1,500	0.471	**0.006**	0.068	____	0.516	**0.004**	0.072	____
	3,000	0.471	**0.004**	0.068	____	0.516	**0.003**	0.069	____
RMSE_{TV}	750	**0.019**	0.026	0.020	____	**0.020**	0.029	0.021	____
	1,500	0.012	**0.012**	0.012	____	0.012	**0.012**	0.012	____
	3,000	0.008	**0.008**	0.008	____	0.008	**0.008**	0.008	____
NZP	750	1.000	0.999	1.000	____	1.000	0.995	0.997	____
	1,500	1.000	1.000	1.000	____	1.000	0.997	0.998	____
	3,000	1.000	1.000	1.000	____	1.000	0.998	0.999	____
ZP	750	0.000	0.999	0.944	____	0.000	0.999	0.946	____
	1,500	0.000	0.999	0.949	____	0.000	0.999	0.949	____
	3,000	0.000	0.999	0.950	____	0.000	0.999	0.952	____

In terms of estimation accuracy, we find that the penalized VCC estimators substantially outperform the unpenalized version. When comparing RMSES computed over the full set of parameters, the P-VCC tuned by the BIC emerges as the clear winner. This method obtains a reduction in RMSE of around 97%-99% compared to the non-regularized VCC and 81%-94% compared to P-VCC tuned by AIC, indicating clear efficiency gains by exploiting sparsity through (substantial) regularization. Perhaps even more impressively, these gains do not seem to come at the cost of substantial shrinkage bias in the estimation of the non-zero parameters. Although for $T = 750$, we actually observe a somewhat higher RMSE_{TV} for P-VCC relative to VCC, this differential completely vanishes for the two larger sample sizes.

The selection abilities of P-VCC are excellent, regardless of the choice of information criterion. As measured by the NZP, both the BIC and AIC tuned versions are able to retain virtually all non-zero parameters in the final model specification. However, as expected, the AIC results in more conservative selection and, consequently, higher ZP scores. In simpler terms, more zero parameters are estimated as nonzero in the final model when P-VCC is tuned by the AIC. This

largely explains the earlier observed differences in RMSE between the BIC and AIC variants.

The results are very similar if we increase the cross-sectional dimension to $N = 50$. All VCC-variants obtain highly similar performance metrics, demonstrating that the vine-based penalization approach scales well to high-dimensional settings. In contract, the forecast errors of the DCC model increases by about 30%-50% compared to the low-dimensional setting ($N = 10$), depending on the precise value of T.

Table A.1 in Appendix A presents another robustness check of the results in Table 1 for different degrees of freedom choices. These results show that the performance of the new methodology is robust to the precise choice of $\nu \in \{5, 10, 20\}$.

3.3 Incorrect specification

To investigate the robustness of our filtering and selection approach to mis-specification, we consider alternative dynamics for the partial correlations. We again draw samples from (2.1) with $\nu = 7$, but the time-varying partial correlations are now given by

$$\rho_{1,j;t} = \begin{cases} \mu_j + 0.2\cos\left(\left(\frac{2t}{365} - B_j\right)\pi\right) + \epsilon_t, & \text{if } j \text{ is odd,} \\ \mu_j + \frac{0.3 f_t}{\max_{1\leq t\leq T}|f_t|}, & \text{if } j \text{ is even,} \end{cases}$$

$$\rho_{i,j\,;t} = 0 \text{ for } j > i > 1,$$

where μ_j is a uniform draw from the discrete set $\{-0.5, 0.5\}$, $B_j \sim$ Bernoulli(0.5), $\epsilon_t \sim \mathrm{U}(-0.1, 0.1)$, and $f_t = 0.95 f_{t-1} + \upsilon_t$, with $\upsilon_t \sim \mathrm{N}(0, 1)$. This produces a design with slowly varying partial correlations, either with an autoregressive (j is even) or with sinusoid (j is odd) dynamics.

Note that in the mis-specified setting we can no longer report the RMSE of $\boldsymbol{\theta}_{i,j}$, as there is no corresponding true parameter value of this parameter. We can, however, still report adjusted versions of the precision metrics that count how many of the true (non)zero *correlation paths* have been identified. Letting $\boldsymbol{\theta}_{i,j} = (\omega_{i,j}, \alpha_{i,j}, \beta_{i,j})^\top$ denotes the vector of parameters governing

Table 2: This table contains the simulation results under mis-specification. The reported metrics are the root mean squared forecast error (RMSFE), nonzero precision (NZP) and zero precision (ZP).

Metric	T	VCC penalty			DCC	VCC penalty			DCC
		no	BIC	AIC		no	BIC	AIC	
		Panel A: low-dimensional ($N = 10$)				Panel B: high-dimensional ($N = 50$)			
	750	0.089	0.092	**0.087**	0.102	0.086	0.088	**0.083**	0.102
RMSFE	1,500	0.078	0.079	**0.076**	0.081	0.073	0.073	**0.071**	0.079
	3,000	0.078	0.079	0.077	**0.074**	0.074	0.075	0.073	**0.069**
	750	1.000	1.000	1.000	____	1.000	1.000	1.000	____
NZP	1,500	1.000	1.000	1.000	____	1.000	1.000	1.000	____
	3,000	1.000	1.000	1.000	____	1.000	1.000	1.000	____
	750	0.000	0.996	0.909	____	0.000	0.996	0.904	____
ZP	1,500	0.000	0.997	0.913	____	0.000	0.997	0.906	____
	3,000	0.000	0.998	0.912	____	0.000	0.998	0.910	____

the dynamics of the time-varying partial correlations between assets i and j given a vertical vine structure, we compute

$$\text{NZP} = \frac{1}{S}\sum_{s=1}^{S} \frac{\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \mathbb{1}\left\{\hat{\boldsymbol{\theta}}_{i,j}^{(s)} \neq \mathbf{0}\right\} \cdot \mathbb{1}\left\{\boldsymbol{\theta}_{i,j}^{(s)} \neq \mathbf{0}\right\}}{\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \mathbb{1}\left\{\boldsymbol{\theta}_{i,j}^{(s)} \neq \mathbf{0}\right\}},$$

$$\text{ZP} = \frac{1}{S}\sum_{s=1}^{S} \frac{\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \mathbb{1}\left\{\hat{\boldsymbol{\theta}}_{i,j}^{(s)} = \mathbf{0}\right\} \cdot \mathbb{1}\left\{\boldsymbol{\theta}_{i,j}^{(s)} = \mathbf{0}\right\}}{\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \mathbb{1}\left\{\boldsymbol{\theta}_{i,j}^{(s)} = \mathbf{0}\right\}}.$$

The results are reported in Table 2. We first note that the DCC now performs better than in Table 1 as the correlation dynamics are not score-driven and are generated directly, varying slowly over time. In Table 1, by contrast, the Pearson correlations were generated indirectly via the partial correlation dynamics. As a result, the performance metrics of the DCC and VCC estimators lie closer together. For $T = 750$ and $T = 1,500$, the penalized VCC estimators perform better in terms of forecasting power than the DCC, with the AIC-tuned variant coming out on top. Only for very long samples of $T = 3,000$, the DCC has the smallest RMSFE, but the differences are quite small for such large sample sizes anyway, as shrinkage matters less. The results are very comparable between the low-dimensional ($N = 10$) and high-dimensional ($N = 50$) setting.

The penalized estimators can still identify the zero and non-zero correlations quite well. The

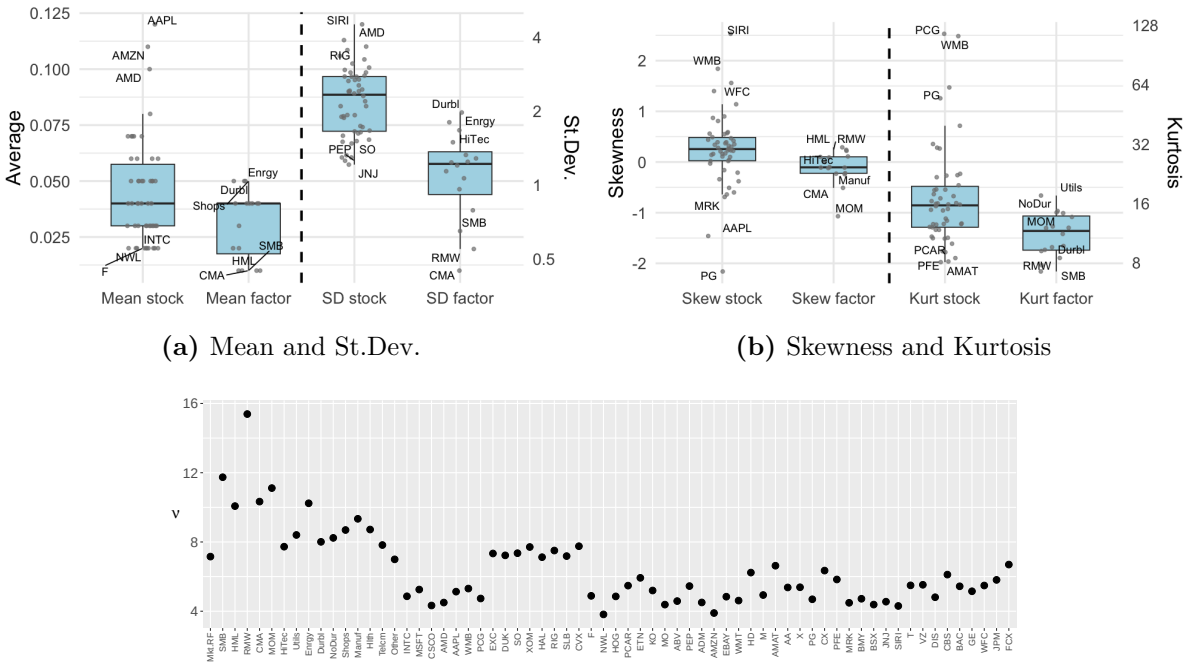
penalized VCC methods perfectly identify the non-zero partial correlations and are able to recover the zero partial correlations very well with a precision above 90%. By construction, the DCC does not allow any of the (true) zero correlation paths to be estimated as zero, which would imply trivial (non-reported) values of ZP = 0 and NZP = 1. The strong performance of our methodology in terms of selection ability even under severe model misspecification provides reassurance that our method can safely be used in real-world applications in which the interpretation of the sparsity pattern has an economic meaning, such as in the empirical application in Section 4.

4 Empirical Application

4.1 Data

In this section, we apply our approach empirically to US stock returns. We select 50 different US stocks from 10 different industries. The industries are constructed in line with the industry classification adopted for the 10 industry portfolios in Ken French's data library.[1] Using the CRSP-Compustat dataset, we select firms that have observations over the period January 3, 2000 to December 30, 2024, thus covering both the financial crisis and the Covid pandemic period. Within each period, we select the 5 stocks that have the highest average trading volume throughout the sample period. Both selections give rise to a survivorship bias in the data. Survivorship biases may be very important in various areas of finance, such as studies related to expected returns and the size of risk premia. Here, however, we argue a possible survivorship bias is less of an issue for the research question at hand, which relates to the time-variation or time-invariance of partial asset correlations after accounting for systematic risk factors and industry returns.

We construct three different data sets of excess returns, namely at daily, 5-daily (weekly), and 22-daily (monthly) frequency. For each dataset, we first estimate a t-GARCH(1,1) model to filter out the marginal volatility effects and to construct de-volatilized time series. We use these de-volatilized time series as our input $\boldsymbol{y}_t$ for the VCC model from Section 2.



(a) Mean and St.Dev. **(b)** Skewness and Kurtosis

(c) Estimated degrees of freedom

Figure 2: Descriptive statistics for the daily factor (16) and individual stock (50) returns.
Note: standard deviations and kurtosis are plotted on a logarithmically scaled axis.

We also use readily available systematic risk factors from Ken French's data library. In particular, we use the market, size (SMB) and value (HML) factors of Fama and French (1993), the momentum (MOM) factor of Carhart (1997), and the additional Robust minus Weak (RMW) and Conservative Minus Aggressive (CMA) factors of Fama and French (2015, 2016). We de-volatilize the factor return series in the same way as the individual stock returns. Finally, we download the 10 value weighted industry portfolio returns from the same site and again de-volatilize them in the same way.

Before fitting VCC, we explore some descriptive statistics of the data in Figure 2. In Figures 2a and 2b, we see the means, standard deviations, and skewness and kurtosis measures of the daily returns for all assets and factors, including the industry factors. The top and bottom three asset in each box plot are indicated by their ticker. Table B.1 in Appendix B provides the underlying data. Figure 2a reveals that the average daily returns (left axis) are slightly positive for all assets. Averages for individual stocks are typically higher than for factors. As expected, also standard deviations (right axis) of individual stocks are typically higher than those of the factors. Figure 2b

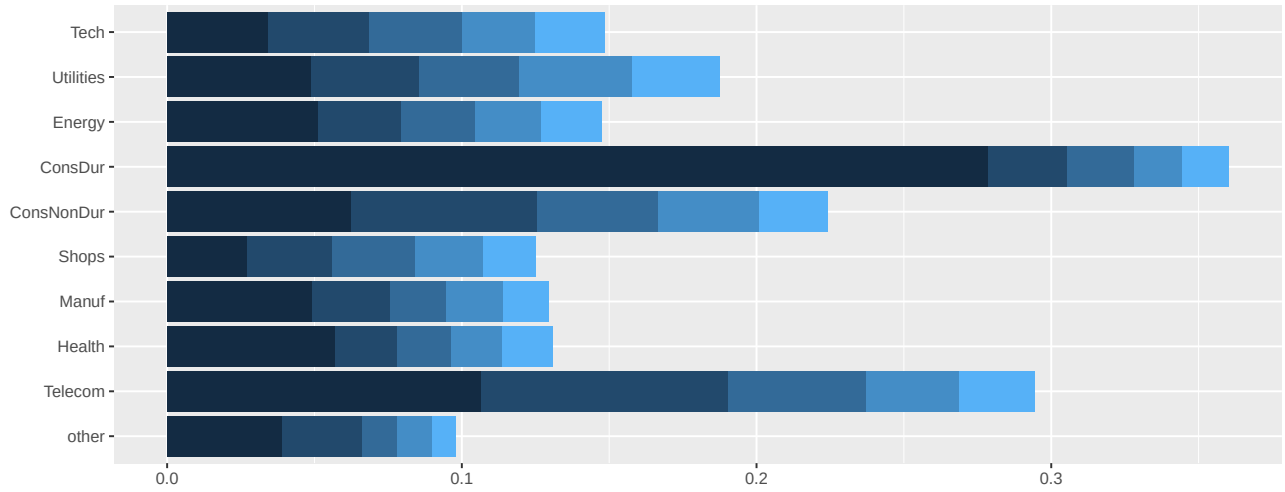


Figure 3: Average daily industry share of trade volumes

shows that some of the daily individual stock returns are right or left skewed, but that skewness (left axis) is limited for most of the stocks considered. For the factors, only the momentum factor shows some mild negative skewness. In terms of kurtosis (right axis), the picture changes. The figure shows that all stocks and factors show excess kurtosis. Kurtosis values range from 7.2 (SMB factor) to 115.6 (PCG stock). This substantiates our use of the fat-tailed Student's t distribution in the rest of our analysis. Furthermore, Figure 2c shows the estimated degrees of freedom of the univariate t-GARCH(1,1) filters and illustrates that there is also considerable *conditional* fat-tailedness in the data. Most degrees of freedom values are estimated between 4 and 10, again supporting our use of the Student's t distribution. The relative homogeneity of the different estimates also supports our use of the ν as a robustness parameter with a common value of $\nu = 7$, where we can investigate the sensitivity to this value in robustness checks. Only the SMB ($\nu \approx 12$) and the RMW ($\nu \approx 15.5$) factors have a slightly higher estimated degrees of freedom and, thus, somewhat lighter conditional tails.

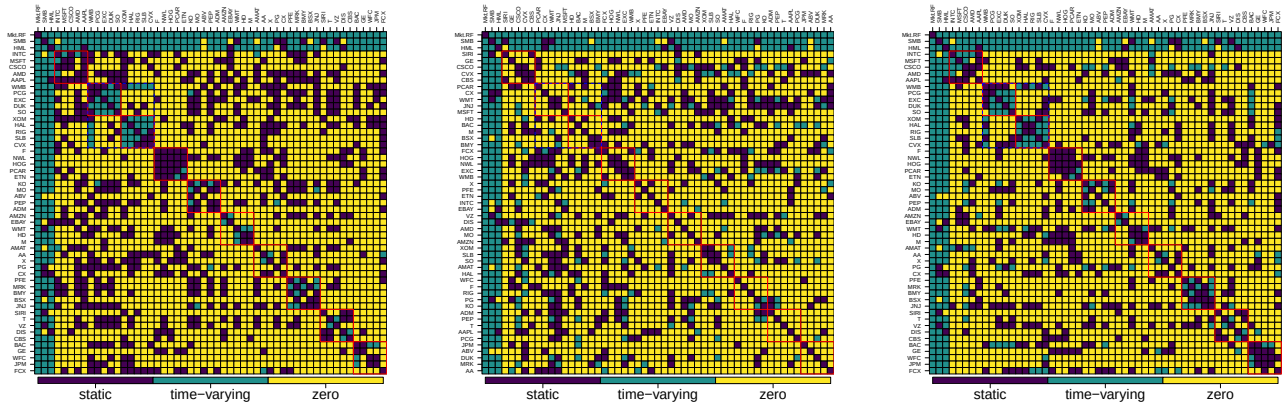
Finally, we explore the trading structure of each industry by plotting the average daily trading volume of each asset as a fraction of its industry trading volume in Figure 3. For some industries such as Tech or Shops, the shares in trading volumes for the first five firms are roughly equally sized. For other industries, the distribution of trading volumes is highly asymmetric within the

industry. For instance, for Consumer Durables we see that Ford Motor Company makes up a disproportionally large 27.9% of the industry's trading volume, and for Telecommunication the pair Sirius International Holding and AT&T Inc. represent 10.6% and 8.4% of the industry's trading volume, respectively. We also see clear differences across industries. Whereas the five firms selected from the Consumer Durables industry make up more than 35% of the industry's trading volume, the percentages for Shops, Manufacturing, and Health are each much lower, settling at levels around 12%-13%. There may thus be more additional information impounded into prices outside the top five trading volume firms for industries like Shops, Manufacturing, or Health than for an industry like Consumer Durables.

4.2 Baseline market structure result

In our first analyses, we use the three Fama and French (1993) risk factors RMRF (market), SMB (size), and HML (value) together with the 50 individual stock returns. The systematic risk factors are put into the first three positions of $\boldsymbol{y}_t$, starting with the market factor. Figure 4 shows the results. In Panel 4a, we put the industries in arbitrary order, while the firms within each industry are ordered in terms of trading volume size over the sample period. Figure 4a directly shows that the correlation between the market factor and each of the individual stocks is estimated as time-varying: all boxes in the first column are colored green. Time-varying correlations with the market factor also imply that these stocks have market betas that vary over time, above and beyond any time-variation due to their individual or the market factor's time-varying volatility; see also Engle (2016), Umlandt (2023), and Giroux et al. (2024).

After correcting for market risk, the partial correlations with SMB and HML are also non-zero and time-varying for almost all stocks (see the second and third column of the matrix in Figure 4a). Only after the effects of these three risk factors has been washed out, it turns out that most of the remaining pairwise partial correlations are set to zero, with some exceptions where the correlation is constant, but non-zero: the majority of the boxes in the 50×50 lower-right corner of

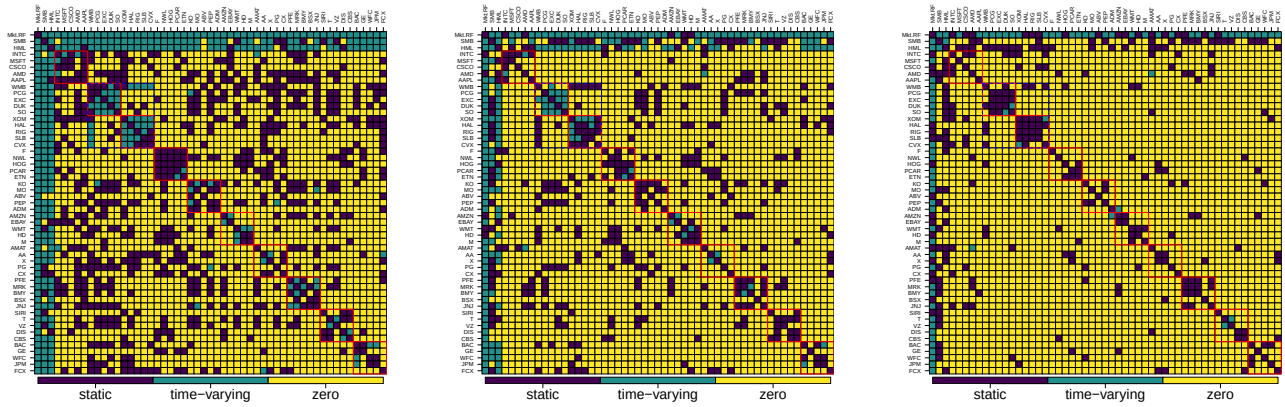


(a) BIC fixed ordering **(b)** BIC random ordering **(c)** BIC re-ordered

Figure 4: Sparsity patterns computed on daily returns for fixed industry ordering, random ordering, and a post-estimation re-ordering of the random ordering. The penalty parameter is chosen separately for each data slice based on the BIC

Figure 4a are yellow. Only in very few cases we see that the stock pairwise partial correlations are time-varying, and most of these relate to a within-industry effect. The latter can be seen by considering the 5×5 blocks on the main diagonal. These blocks are also highlighted by red boxes in Figure 5a. Our first analysis therefore results in two main takeaways. First, time-varying asset correlations may be mainly attributable to time-varying correlations with systematic risk factors. After washing out these effects, remaining partial pairwise correlations between individual assets are predominantly constant, and most often zero. Second, the main exception to the previous takeaway is the within-industry partial correlations. These appear to persist and are regularly time-varying, even after first filtering out the time-varying correlations with the systematic risk factors. We come back to this second issue in our follow-up analyses below.

A concern that one might have is that the vine structure of the VCC model induces a tightly ordered cascade of partial correlations and that this ordering might affect the previous main conclusions. To investigate this, we conduct a second analysis in Figures 4b and 4c. First, we randomly order the 50 individual stocks and redo our original analysis of Figure 4a. The result is presented in Figure 4b. We indeed see that the result changes to some extent, but also that the main first conclusion survives. In particular, we see most time-varying correlations with the main market factor. Also the partial correlations with the value factor HML and the size factor (SMB)



(a) Daily returns **(b)** Weekly returns **(c)** Monthly returns

Figure 5: Sparsity patterns computed on daily, weekly, and monthly returns. The penalty parameter is chosen separately for each data slice based on the BIC

are mostly time-varying, even after correcting for the market factor. We also see that most of the remaining pairwise correlations between individual assets after correcting for the three risk factors, are zero or constant. Again we conclude that pairwise time-variation in correlations between individual assets is likely to be due to time-varying correlations with systematic risk factors.

Figure 4b does not allow us to inspect the industry clustering effects given the random ordering of the individual stocks. Therefore, Figure 4c repeats Figure 4b, but now with the stocks put in the same order as in Figure 4a. No re-estimation is needed for this. By re-ordering the results in this way, also the second conclusion emerges clearly again: even after partialling out the time-variation with systematic risk factors, residual partial correlations remain within industries. Also this result, therefore, is a robust pattern in the data.

4.3 Robustness: data frequency, other factors, penalization choice

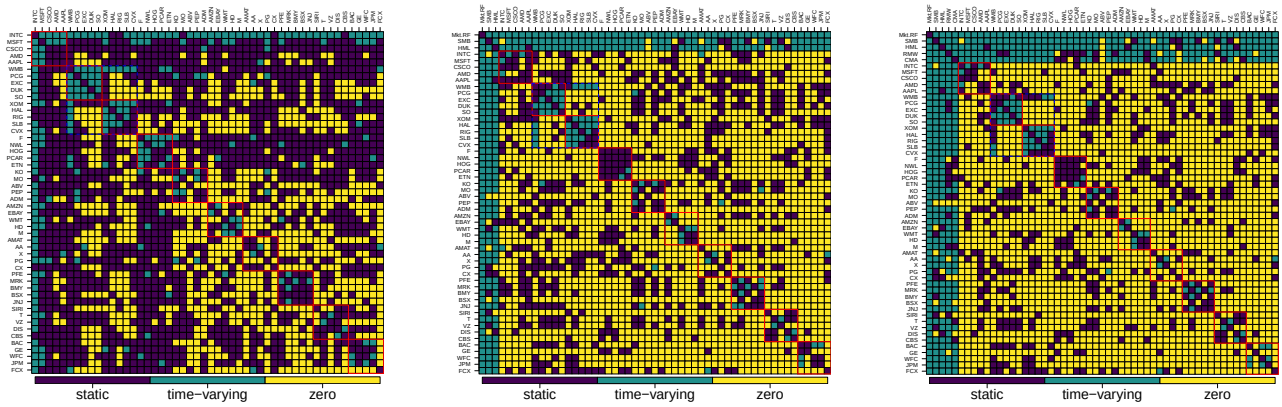
We consider several robustness analysis of our baseline results. First, in Figure 5 we investigate whether the time-variation of correlations with systematic risk factors and the remaining within-industry partial correlations are only present in the daily data, or whether these effects are also present in weekly and monthly data. Such lower data frequencies might be more practically relevant for large institutional asset managers.

Panels 5a, 5b, and 5c show the results for daily, weekly, and monthly data, respectively. Three

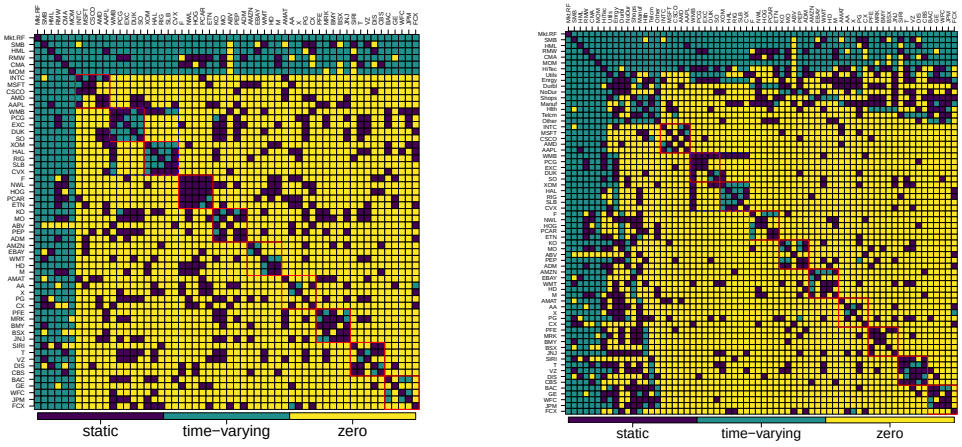
main conclusions emerge. First, the time-varying correlations with the market factor are robust, irrespective of the frequency of the data: most entries in the first column remain green in all three panels. Second, the partial correlations with the SMB and HML factors are less robust: for lower data frequencies, these partial correlations between SMB/HML and the individual stocks (after partialling out the market factor) become mostly constant and regularly even zero. This again illustrates that much of the pairwise time-variation in correlations between individual stocks might originate from time-varying correlations with the market factor. We do note, however, that the risk factors themselves remain mutually correlated with time-varying correlations irrespective of the data frequency, as can be seen from the upper-left 3×3 block in each of the panels. Third, we also see that the within-industry effects survive the variation over data frequencies, though the results are less strong than for the daily data. Within industries, residual partial correlations remain present even at monthly frequencies and after partialling out the correlations with the three systematic Fama and French (1993) factors.

An obvious follow-up question is therefore whether the within-industry effects disappear if we account for more systematic risk factors. This question is picked up in Figure 6. Panel 6b presents the baseline result for daily data, identical to Figure 5a. Panel 6a by contrast only considers the 50 individual stocks without any systematic risk factors. The difference with the baseline result is evident: many of the pairwise partial correlations are no longer zero. Still, even here it is clear that much of the time-variation is picked up by the first few series in the vine structure, which partly take over the role of the market factor and other systematic risk factors. Many of the remaining partial correlations are constant. Again, the main exceptions are formed by the within-industry blocks on the main diagonal, underlining the robustness of the presence of within industry partial correlations.

In Panels 6c and 6d we add the additional RMW and CMA factors of the 5-factor model of Fama and French (2015, 2016) and the MOM factor of Carhart (1997), respectively. Both panels confirm our earlier findings. Most of the correlations with systematic risk factors at a daily



(a) no factors **(b)** 3 factors **(c)** 5 factors



(d) 6 factors **(e)** 6 + industry factors

Figure 6: Sparsity patterns computed on daily returns for models with a different number of factors. The penalty parameter is chosen separately for each data slice based on the BIC

frequency are time-varying, but remaining pairwise partial correlations between individual assets are mostly zero, and otherwise constant, with one major exception: within industries pairwise partial correlations are often non-zero and even time-varying.

Given the persistence of the within-industry partial correlations even after including many systematic risk factors, we investigate in Figure 6e whether these effects disappear if we also include industry factors. There are three main takeaways. First, the inclusion of industry factors indeed reduces the number of non-zero pairwise partial correlations after systematic risk and 10 industry factors have been washed out. Second, each industry factor indeed shows time-varying partial correlations with the 5 individual stocks, even after partialling out the 6 systematic risk factors first. This is seen by the (typically) 5 green block-diagonal entries in columns 7 (HiTec) through 16 (Other), such as rows 17-21 (stocks INTC–AAPL) for column 7 (HiTec), through

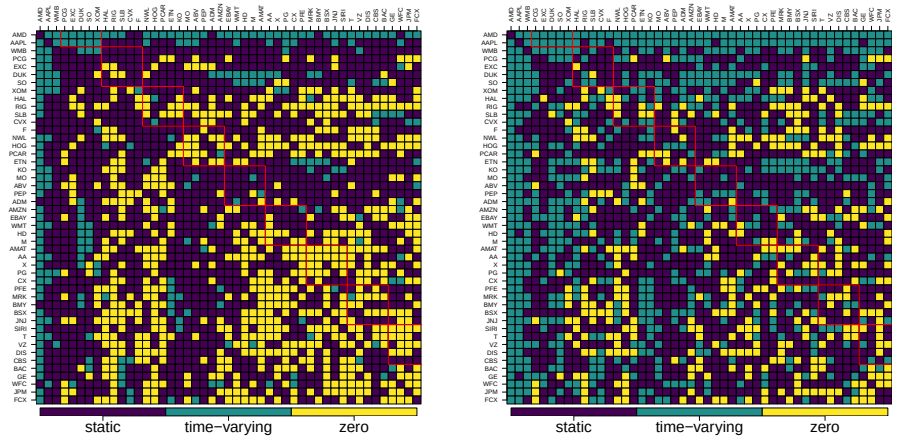
rows 112-116 (stocks BAC–FCX) for column 16 (Other), respectively. We also note that the industry factor for the Energy (Enrgy column) industry not only has strong and time-varying partial correlations with the Energy stocks, but also with the Utility (Util) stocks. This makes intuitive sense given the clear relations between these two industries.

The third main takeaway from this figure is that even after correcting for systematic risk factors *and* the industry factors, within-industry partial correlations still remain for several industries. To understand this result, we note that our five stocks chosen per industry are selected based on their trading volume over the sample period. They therefore only constitute part of the industry, which suggest that the broad industry factors may also be subject to other movements besides those of the top five firms with the highest average trading volume.

Finally, in Figure 7 we investigate the effect of choosing the penalty parameter λ. So far, we have chosen this parameter for each slice based on the BIC. The left-hand panels in Figure 7 present these baseline results for three different sets of systematic risk factors. The right-hand columns provide the corresponding results when λ is set using the AIC. The results are as expected. When using the AIC, more partial correlations are estimated to be non-zero and even time-varying. Even then, however, we still see the dominant takeaway of our data-driven analysis of the market structure. Partial correlations with systematic risk factors are typically time-varying. After accounting for this, most of the remaining pairwise partial correlations between individual stocks are mostly constant at either zero or non-zero levels. Time-variation in pairwise Pearson correlations thus appears to be mainly attributable to time-varying correlations with common systematic risk-factors. Only within industries, further partial correlations may remain, even after correcting for the systematic risk factor effects.

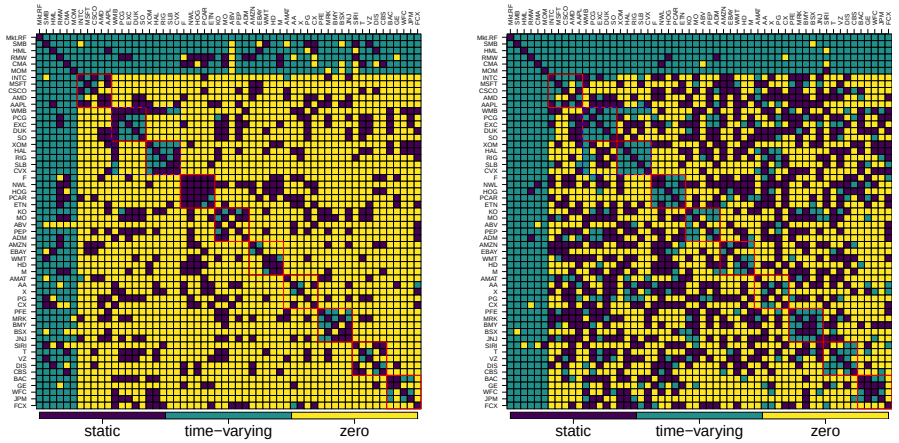
4.4 Relation to factor model comparisons

Thus far, we have concentrated on the structure of the partial correlations between factors and assets and concluded that time-varying Pearson correlations between individual stocks appear



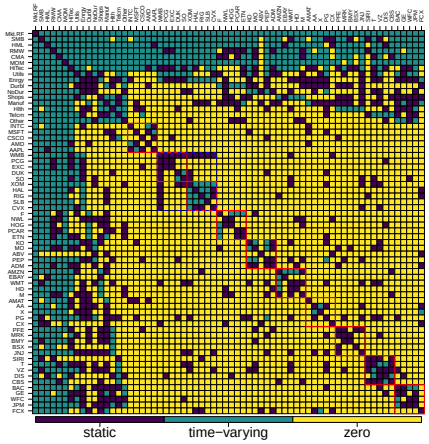
(a) BIC no factors

(b) AIC no factors

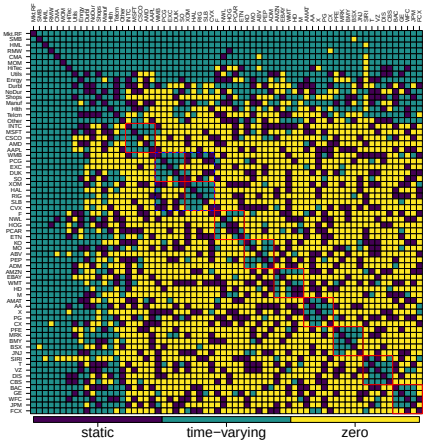


(c) BIC 6 factors

(d) AIC 6 factors



(e) BIC, 6 + industry factors



(f) AIC, 6 + industry factors

Figure 7: Sparsity patterns computed on daily returns for different models where the penalty parameter is chosen separately for each data slice based on either the BIC or AIC

to be mainly driven by time-varying partial correlations with systematic risk factors. In this section, we consider the role of our method in evaluating competing asset pricing models. As explained in Section 2, the current methodology does not allow us to perform a formal test of correct specification. However, we can use standard intuitive test statistics for various relevant

Table 3: This table contains the results of likelihood ratio tests based on varying frequencies and restrictions. For each frequency, four different sets of restrictions are considered: (i) all partial correlations between individual assets are set to zero (full); (ii) all partial correlations between assets in different industries are set to zero (ind); (iii) the zero restrictions are enforced by the VCC estimator tuned by BIC (VCC-BIC) and (iv) the zero restrictions are enforced by the VCC estimator tuned by AIC (VCC-AIC). The first column displays the model specification which may be either the single factor model (1F), three factor model (3F), six factor model (6F) or the six factor model with ten industry portfolios added (IPF). The metrics reported are the likelihood ratio test statistic (LRT), the critical value for the corresponding Chi-squared distribution (cval), the number of zero restriction imposed (df) and the corresponding p-value (pval).

		Daily				Weekly				Monthly			
		full	ind	VCC		full	ind	VCC		full	ind	VCC	
				BIC	AIC			BIC	AIC			BIC	AIC
1F	LRT	57,610	16,218	4,159	1,977	13,574	5,052	2,804	1,371	4,582	2,674	2,442	1,470
	cval	3,817	3,511	3,238	2,524	3,817	3,511	3,521	3,010	3,817	3,511	3,722	3,343
	df	3,675	3,375	3,107	2,409	3,675	3,375	3,385	2,884	3,675	3,375	3,582	3,210
	pval	0.00	0.00	0.00	**1.00**	0.00	0.00	**1.00**	**1.00**	0.00	**1.00**	**1.00**	**1.00**
3F	LRT	45,141	11,915	3,609	1,569	10,856	4,100	2,968	1,340	4,150	2,535	2,680	1,615
	cval	3,817	3,511	3,424	2,776	3,817	3,511	3,772	3,184	3,817	3,511	4,012	3,596
	df	3,675	3,375	3,289	2,655	3,675	3,375	3,631	3,054	3,675	3,375	3,866	3,458
	pval	0.00	0.00	0.00	**1.00**	0.00	0.00	**1.00**	**1.00**	0.00	**1.00**	**1.00**	**1.00**
6F	LRT	34,671	8,036	3,175	1,028	9,208	3,489	3,444	1,399	3,709	2,300	3,136	1,846
	cval	3,817	3,511	3,636	2,958	3,817	3,511	4,138	3,436	3,817	3,511	4,468	3,964
	df	3,675	3,375	3,497	2,833	3,675	3,375	3,990	3,301	3,675	3,375	4,314	3,819
	pval	0.00	0.00	**1.00**	**1.00**	0.00	**0.08**	**1.00**	**1.00**	**0.34**	**1.00**	**1.00**	**1.00**
IPF	LRT	13,445	4,638	5,812	1,358	4,499	2,522	5,833	2,444	2,874	2,241	4,951	2,946
	cval	3,817	3,511	5,024	3,899	3,817	3,511	5,677	4,706	3,817	3,511	6,169	5,434
	df	3,675	3,375	4,861	3,755	3,675	3,375	5,503	4,548	3,675	3,375	5,988	5,264
	pval	0.00	0.00	0.00	**1.00**	0.00	**1.00**	0.00	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**

hypotheses to *compare* different model specifications as a diagnostic device. A similar approach is regularly adopted in the finance literature, where comparisons of asset pricing models are commonly performed based on the so-called GRS (Fama and French, 2017, 2018; Giglio et al., 2022) or BKRS statistic (Barillas et al., 2020; Dickerson et al., 2023).[2]

Table 3 provides the results for four model specifications, four sets of restrictions, and three data frequencies. The specifications differ between the number of risk factors incorporated, where we distinguish between: (1F) the factor model that only includes the market factor; (3F) a three factor model with the market, SMB, and HML factors of Fama and French (1993); (6F) a six factor model that adds the RMW and CMA factors of Fama and French (2015, 2016) and the MOM factor of Carhart (1997) to 3F; and (IPF) a sixteen factor model that adds ten value-weighted industry portfolios to 6F. Within each specification, we further distinguish the following restrictions: (i)

'full' denotes a fully restricted model where all pairwise partial correlations between individual stocks are set to zero (after correcting for the effects of the systematic risk factors); (ii) 'ind' denotes a model where all pairwise partial correlations between individual stocks in *different industries* are zero, but possibly non-zero and time-varying between individual stocks in the same industry; (iii) 'VCC-BIC' denotes our penalized estimator for all pairwise partial correlations tuned by BIC per data slice; and finally, (iv) 'VCC-AIC' denotes our penalized estimator tuned by AIC instead. Each time, a likelihood ratio test statistic (LRT) is computed vis-à-vis the fully unrestricted model. Table 3 reports the value of this LRT statistic, a 'critical value' of the Chi-squared distribution (cval) based on the corresponding number of zero-restrictions (df), and the 'p-value' based on a chi-squared distribution with the same degrees of freedom parameter as the number of zero restrictions. We stress once more that we do not consider this to be a formal testing procedure, but instead use the LRT as a model comparison device in a similar spirit as in, for instance, Fama and French (2018). The analysis is repeated for weekly and monthly data frequencies to investigate the robustness of the results.

Focusing first on the daily frequency, the results in Table 3 make clear that systematic risk factors only do not fully account for all partial correlations between individual stocks: setting all these pairwise partial correlations between individual stocks to zero is clearly incongruent with the data. Although LRTs are monotonically decreasing if we include more systematic risk factors in the model, the statistic remains very high for all models considered, ranging from 57,610 (1F) to 13,445 (IPF). All of these are well above the informal reference level of the 5% chi-squared critical value. The models that only allow for systematic risk factor and intra-industry partial correlations ('ind' column) fare considerably better in terms of their LRTs than models that only allow for systematic risk partial correlations. The LRTs in the 'ind' column again monotonically decrease in the number of included risk factors, and are considerably smaller than before at levels ranging from 16,218 to 4,638.

When considering the models with penalization, we see a further substantial decrease in LRTs.

These models therefore appear to be much more in line with the data than their more restricted counterparts. In particular, when conservatively tuned by the AIC, the LRTs are consistently far below the corresponding informal critical values at any reasonable significance level, despite the number of zero restrictions ranging from 2,409 (1F) to 3,755 (IPF). This signals that even after accounting for intra-industry correlations, pairwise partial correlations for some stocks in different industries remain. Such remaining partial correlations may be attributable to industries themselves being in the same production line (such as the Energy and Utils sectors), or hint at further omitted risk factors.

When lowering the data frequency the results become even more favorable to the VCC specifications. An immediate finding is that the LRT statistics are consistently lower compared to the daily frequency. Nevertheless, for the weekly frequency the fully restricted model remains too restrictive as indicated by the LRTs exceeding the informal critical values by far. However, when allowing for intra-industry partial correlations, the LRTs for the six factor model and its extension with industry portfolios fall below the 'critical values', suggesting that the six factors may adequately capture the systemic risk among the assets in our sample. The VCC methods remain even more congruent with the data, with the exception of the VCC-BIC variant applied to the IPF model, which seems to introduce too much sparsity. At the monthly frequency, we observe that the fully restricted model may provide an adequate representation of the data, as long as the six risk factors and the industry portfolios are included. If we include fewer factors but allow for remaining intra-industry partial correlations ('ind' column), even the single factor model no longer exceeds the informal chi-squared critical value. For monthly data, systematic risk augmented possibly with industry factors thus appear to mop up all the (constant and time-varying) pairwise partial correlations between individual stocks, suggesting that at this frequency long-term changes in underlying fundamental values are probably the dominant driving forces in price dynamics.

5 Conclusion

In this paper we introduced a penalized vine partial correlation model to uncover the market structure and investigate the time-variation (or absence thereof) of partial pairwise correlations between individual stock returns. As opposed to direct Pearson matrix correlation modeling, the partial correlation approach combined with a vertical vine structure allowed us to first filter out any correlations with systematic risk factors. For an empirical sample of 50 US stock returns over the period 2000–2024 we found that time-variation in pairwise Pearson correlations between stocks appears largely attributable to time-varying correlations with common systematic risk factors. Once these have been corrected for, remaining time-variation in partial correlations between individual stocks largely disappears and the partial pairwise correlations are mainly constant or zero. The main exception is formed by within industry partial correlations, which remain persistently present even after correcting for several well-known systematic risk factors.

The results are robust to a variety of alternative specifications, including variations in the systematic factors, the data frequency, the way in which the penalization parameter is chosen, and the ordering of the assets in the vine structure. The approach also lent itself as a diagnostic device to compare models, similar to Fama and French (2017, 2018); Giglio et al. (2022); Barillas et al. (2020); Dickerson et al. (2023). Such comparisons corroborated the earlier results and suggest that multivariate time-varying correlation models may developed into more parsimonious structures by exploiting the relation of time-variation in pairwise Pearson correlations between individual assets to the time-variation in correlations of these assets with common systematic risk factors. We leave such developments to a next paper.

Notes

[1] See https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_10_ind_port.html.

[2] Another alternative would be to investigate the pricing errors as induced by the different models. Our approach in this paper, however, is not so much geared toward minimizing pricing errors as it is toward uncovering the market structure and the origins of the time-variation in pairwise Pearson correlations between individual stocks after correcting for systematic risk factor effects. Pricing error tests have a different focus, which is why we do not pursue this approach in the present paper.

References

Abadir, K. M. and G. M. Rockinger (2025, December). Dynamic spectral conditional correlations. Working Paper 5103673, SSRN Electronic Journal. Available at SSRN: https://ssrn.com/abstract=5103673 or http://dx.doi.org/10.2139/ssrn.5103673.

Archakov, I. and P. R. Hansen (2021). A new parametrization of correlation matrices. *Econometrica 89*(4), 1699–1715.

Barillas, F., R. Kan, C. Robotti, and J. Shanken (2020). Model comparison with sharpe ratios. *Journal of Financial and Quantitative Analysis 55*(6), 1840–1874.

Blasques, F., J. van Brummelen, S. J. Koopman, and A. Lucas (2022, April). Maximum likelihood estimation for score-driven models. *Journal of Econometrics 227*(2), 325–346.

Buccheri, G., G. Bormetti, F. Corsi, and F. Lillo (2021). A score-driven conditional correlation model for noisy and asynchronous data: An application to high-frequency covariance dynamics. *Journal of Business & Economic Statistics 39*(4), 920–936.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance 52*(1), 57–82.

Chen, J., D. Li, Y.-N. Li, and O. Linton (2025). Estimating time-varying networks for high-dimensional time series. *Journal of Econometrics*, 105941.

Creal, D., S. J. Koopman, and A. Lucas (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business & Economic Statistics 29*(4), 552–563.

Creal, D., S. J. Koopman, and A. Lucas (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics 28*(5), 777–795.

DeMiguel, V., L. Garlappi, and R. Uppal (2009, May). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The Review of Financial Studies 22*(5), 1915–1953.

Dickerson, A., P. Mueller, and C. Robotti (2023). Priced risk in corporate bonds. *Journal of Financial Economics 150*(2), 103707.

D'Innocenzo, E. and A. Lucas (2024). Dynamic partial correlation models. *Journal of Econometrics 241*(2), 105747.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of business & economic statistics 20*(3), 339–350.

Engle, R., O. Ledoit, and M. Wolf (2019). Large dynamic covariance matrices. *Journal of Business & Economic Statistics 37*, 363–375.

Engle, R. F. (2016, Fall). Dynamic conditional beta. *Journal of Financial Econometrics 14*(4), 643–667.

Fama, E. F. and K. R. French (1993, February). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Fama, E. F. and K. R. French (2015, April). A five-factor asset pricing model. *Journal of Financial Economics 116*(1), 1–22.

Fama, E. F. and K. R. French (2016). A five-factor asset pricing model. *The Review of Financial Studies 29*(1), 1–60.

Fama, E. F. and K. R. French (2017). International tests of a five-factor asset pricing model. *Journal of financial Economics 123*(3), 441–463.

Fama, E. F. and K. R. French (2018). Choosing factors. *Journal of financial economics 128*(2), 234–252.

Fan, J., Y. Liao, and H. Liu (2016). An overview of the estimation of large covariance and precision matrices. *The Econometrics Journal 19*(1), C1–C32.

Feng, G., S. Giglio, and D. Xiu (2020). Taming the factor zoo: A test of new factors. *The Journal of Finance 75*(3), 1327–1370.

Giglio, S., B. Kelly, and D. Xiu (2022). Factor models, machine learning, and asset pricing. *Annual Review of Financial Economics 14*(1), 337–368.

Giroux, T., J. Royer, and O. D. Zerbib (2024, Autumn). Empirical asset pricing with score-driven conditional betas. *Journal of Financial Econometrics 22*(5), 1310–1344.

Hafner, C. M. and L. Wang (2023). A dynamic conditional score model for the log correlation matrix. *Journal of Econometrics 237*(2), 105176.

Hansen, P. R. and A. Lunde (2005). A forecast comparison of volatility models: Does anything beat a garch(1,1)? *Journal of Applied Econometrics 20*(7), 873–889.

Harvey, A. C. (2013). *Dynamic Models for Volatility and Heavy Tails: with Applications to Financial and Economic Time Series*, Volume 52. Cambridge University Press.

Jaeckel, P. and R. Rebonato (2000). The most general methodology for creating a valid correlation matrix for risk management and option pricing purposes. *Journal of Risk 2*(2), 17–28.

Ledoit, O. and M. Wolf (2022). The power of (non-) linear shrinking: A review and guide to covariance matrix estimation. *Journal of Financial Econometrics 20*, 187–218.

Lee, T.-H., M. Y. Mao, and A. Ullah (2021). Estimation of high-dimensional dynamic conditional precision matrices with an application to forecast combination. *Econometric Reviews 40*(10), 905–918.

Pakel, C., N. Shephard, K. Sheppard, and R. F. Engle (2021). Fitting vast dimensional time-varying covariance models. *Journal of Business & Economic Statistics 39*(3), 742–759.

Umlandt, D. (2023, December). Score-driven asset pricing: Predicting time-varying risk premia based on cross-sectional model performance. *Journal of Econometrics 237*(2, Part C), 105470.

Zhu, X., Y. Chen, and J. Hu (2024). Estimation of banded time-varying precision matrix based on scad and group lasso. *Computational Statistics & Data Analysis 189*, 107849.

Online Appendix to:

An Impartial Look at Asset Correlation Stability and Market Structure

Etienne Wijler[1], Enzo D'Innocenzo[1], Andre Lucas[1]

[1]: *Vrije Universiteit Amsterdam and Tinbergen Institute*

[2]: *University of Bologna*

A Additional robustness checks

Table A.1: This table contains the robustness results for the same Monte Carlo experiments conducted in Section 3.2 across alternative degrees of freedom parameters ($\nu = 5, 10, 20$). See Table 1 for additional details.

Metric	T	$\nu = 5$			$\nu = 10$			$\nu = 20$		
		DCC	VCC		DCC	VCC		DCC	VCC	
			BIC	AIC		BIC	AIC		BIC	AIC
Panel A: $N = 10$										
	750	0.055	0.050	0.051	0.056	0.049	0.051	0.066	0.060	0.062
RMSFE	1500	0.040	0.036	0.037	0.039	0.034	0.035	0.052	0.048	0.049
	3000	0.032	0.029	0.030	0.030	0.026	0.027	0.045	0.042	0.043
	750	0.469	0.012	0.067	0.469	0.013	0.074	0.468	0.014	0.083
RMSE	1500	0.471	0.006	0.065	0.471	0.006	0.072	0.470	0.007	0.083
	3000	0.471	0.004	0.063	0.471	0.004	0.072	0.471	0.005	0.081
	750	0.019	0.026	0.020	0.020	0.027	0.020	0.021	0.028	0.022
RMSE_{TV}	1500	0.012	0.012	0.012	0.012	0.012	0.012	0.014	0.014	0.014
	3000	0.008	0.008	0.008	0.008	0.008	0.008	0.011	0.011	0.011
	750	1.000	0.999	1.000	1.000	0.999	1.000	1.000	0.999	1.000
NZP	1500	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	3000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	750	0.000	0.999	0.947	0.000	0.998	0.939	0.000	0.997	0.929
ZP	1500	0.000	0.999	0.952	0.000	0.999	0.946	0.000	0.999	0.933
	3000	0.000	0.999	0.954	0.000	0.999	0.946	0.000	0.999	0.934
Panel B: $N = 50$										
	750	0.055	0.049	0.050	0.056	0.048	0.050	0.066	0.059	0.061
RMSFE	1500	0.041	0.036	0.037	0.040	0.034	0.035	0.052	0.047	0.048
	3000	0.032	0.029	0.030	0.030	0.025	0.026	0.044	0.040	0.041
	750	0.514	0.008	0.073	0.513	0.008	0.077	0.513	0.008	0.083
RMSE	1500	0.516	0.004	0.070	0.516	0.004	0.075	0.515	0.005	0.081
	3000	0.516	0.003	0.068	0.516	0.003	0.071	0.516	0.004	0.078
	750	0.019	0.030	0.021	0.020	0.028	0.021	0.021	0.029	0.022
RMSE_{TV}	1500	0.012	0.012	0.012	0.013	0.012	0.012	0.014	0.014	0.014
	3000	0.009	0.008	0.008	0.009	0.008	0.008	0.011	0.011	0.011
	750	1.000	0.995	0.997	1.000	0.995	0.997	1.000	0.995	0.997
NZP	1500	1.000	0.997	0.998	1.000	0.997	0.998	1.000	0.997	0.998
	3000	1.000	0.998	0.999	1.000	0.998	0.999	1.000	0.998	0.999
	750	0.000	0.999	0.947	0.000	0.998	0.944	0.000	0.998	0.936
ZP	1500	0.000	0.999	0.950	0.000	0.999	0.947	0.000	0.999	0.940
	3000	0.000	0.999	0.954	0.000	0.999	0.950	0.000	0.999	0.944

B Full data descriptives

Table B.1: Descriptives

Abr.	Name	mean	variance	skewness	kurtosis
Systematic factors					
Mkt.RF	Market	0.03	1.54	-0.22	11.97
SMB	Small-minus-big	0.01	0.42	0.11	7.20
HML	High-minus-low	0.01	0.62	0.29	9.70
RMW	Robust-minus-weak	0.02	0.30	0.24	7.75
CMA	Conservative-minus-aggressive	0.01	0.20	-0.51	12.00
MOM	Momentum	0.01	1.13	-1.07	14.35
Industry factors					
HiTec	Technology	0.04	2.78	0.22	9.56
Utils	Utilities	0.04	1.44	0.09	17.49
Enrgy	Energy	0.05	3.23	-0.12	14.24
Durbl	Consumer Durables	0.05	3.88	0.00	8.43
NoDur	Consumer Non-durables	0.04	0.92	-0.23	14.61
Shops	Shops	0.05	1.52	-0.10	9.14
Manuf	Manufacturing	0.04	1.64	-0.24	11.17
Hlth	Healthcare	0.04	1.29	-0.11	9.27
Telcm	Telecommunications	0.02	1.75	0.10	12.19
Other	Other	0.04	2.22	-0.11	13.62
Technology stocks (HiTec)					
INTC	INTEL CORP	0.02	5.79	-0.34	12.62
MSFT	MICROSOFT CORP	0.05	3.62	0.16	12.53
CSCO	CISCO SYSTEMS INC	0.03	5.29	0.47	14.69

Continued on next page

Table B.1 – continued from previous page

Abr.	Name	mean	variance	skewness	kurtosis
AMD	ADVANCED MICRO DEVICES INC	0.10	15.07	0.49	13.04
AAPL	APPLE INC	0.12	5.97	-1.46	39.54
Utilities stocks (Util)					
WMB	WILLIAMS COS	0.08	11.21	1.84	112.76
PCG	P G & E CORP	0.05	8.62	1.40	115.85
EXC	EXELON CORP	0.04	2.68	0.23	15.46
DUK	DUKE ENERGY CORP NEW	0.04	2.27	0.10	15.77
SO	SOUTHERN CO	0.05	1.67	0.57	19.72
Energy stocks (Enrgy)					
XOM	EXXON MOBIL CORP	0.03	2.76	0.23	12.12
HAL	HALLIBURTON CO	0.05	8.22	-0.60	22.55
RIG	TRANSOCEAN LTD	0.03	13.36	0.90	22.08
SLB	SCHLUMBERGER LTD	0.03	5.84	-0.16	10.82
CVX	CHEVRON CORP NEW	0.04	3.01	0.11	22.19
Consumer Durables stocks (Durbl)					
F	FORD MOTOR CO DEL	0.02	7.01	0.55	16.93
NWL	NEWELL RUBBERMAID INC	0.02	5.76	0.44	30.30
HOG	HARLEY DAVIDSON INC	0.03	6.45	0.56	13.36
PCAR	PACCAR INC	0.07	4.37	0.26	8.42
ETN	EATON CORP	0.07	3.51	0.39	12.06
Consumer Non-durables stocks (NoDur)					
KO	COCA COLA CO	0.02	1.67	0.02	12.55
MO	ALTRIA GROUP INC	0.06	2.33	-0.03	14.95
ABV	COMPANHIA DE BEBIDAS DAS AMERS	0.07	7.59	0.39	15.96

Continued on next page

Table B.1 – continued from previous page

Abr.	Name	mean	variance	skewness	kurtosis
PEP	PEPSICO INC	0.03	1.57	0.17	17.78
ADM	ARCHER DANIELS MIDLAND CO	0.04	3.70	-0.38	15.73
Shops stocks (Shops)					
AMZN	AMAZON COM INC	0.11	9.71	1.14	18.79
EBAY	EBAY INC	0.07	7.21	0.59	14.81
WMT	WAL MART STORES INC	0.03	2.17	0.30	11.71
HD	HOME DEPOT INC	0.05	3.58	-0.51	18.62
M	MACYS INC	0.04	9.01	0.37	9.96
Manufacturing stocks (Manuf)					
AMAT	APPLIED MATERIALS INC	0.06	7.62	0.34	8.09
AA	ALCOA INC	0.04	7.14	0.22	12.45
X	UNITED STATES STEEL CORP NEW	0.06	12.51	0.35	9.12
PG	PROCTER & GAMBLE CO	0.03	1.74	-2.16	61.89
CX	CEMEX S A B DE C V	0.03	8.21	0.42	11.72
Healthcare stocks (Hlth)					
PFE	PFIZER INC	0.02	2.51	0.02	8.03
MRK	MERCK & CO INC NEW	0.03	2.72	-0.69	21.67
BMY	BRISTOL MYERS SQUIBB CO	0.02	2.95	-0.64	16.39
BSX	BOSTON SCIENTIFIC CORP	0.05	4.79	0.04	16.02
JNJ	JOHNSON & JOHNSON	0.03	1.46	-0.21	15.88
Telecommunications stocks (Telcm)					
SIRI	SIRIUS X M HOLDINGS INC	0.05	20.26	2.53	54.52
T	A T & T INC	0.02	2.62	0.16	10.62
VZ	VERIZON COMMUNICATIONS INC	0.02	2.24	0.32	10.57

Continued on next page

Table B.1 – continued from previous page

Abr.	Name	mean	variance	skewness	kurtosis
DIS	DISNEY WALT CO	0.04	3.69	0.28	12.52
CBS	C B S CORP NEW	0.02	7.74	0.14	17.37
Other stocks (Other)					
BAC	BANK OF AMERICA CORP	0.05	7.60	0.87	30.61
GE	GENERAL ELECTRIC CO	0.02	4.37	0.25	10.68
WFC	WELLS FARGO & CO NEW	0.05	5.72	1.56	31.95
JPM	JPMORGAN CHASE & CO	0.06	5.54	0.81	18.74
FCX	FREEPORT MCMORAN COPPER & GOLD	0.07	10.50	0.21	9.02

C Proximal Gradient Descent Algorithm

Algorithm 1 Proximal gradient descent

1: At $k = 0$, set initial values $\hat{\boldsymbol{\theta}}_{i,j}^{(0)}$ and compute $f^{(0)} = L_{i,j}\left(\hat{\boldsymbol{\theta}}_{i,j}^{(0)}\right)$.
2: Update $k = k + 1$.
3: Compute $\nabla L_{i,j}\left(\hat{\boldsymbol{\theta}}_{i,j}^{(k-1)}\right)$.
4: Update[1,2] $\hat{\boldsymbol{\theta}}_{i,j}^{(k)} = h_{t\lambda}\left(\hat{\boldsymbol{\theta}}_{i,j}^{(k-1)} - t \cdot \nabla L_{i,j}\left(\hat{\boldsymbol{\theta}}_{i,j}^{(k-1)}\right)\right)$.
5: Compute $f^{(k)} = L_{i,j}\left(\hat{\boldsymbol{\theta}}_{i,j}^{(k)}\right)$
6: While $\left|\frac{f^{(k)} - f^{(k-1)}}{f^{(k-1)}}\right| > c$, repeat steps 2-5.

1. The function h is the proximal operator. Letting $\boldsymbol{\gamma}_{i,j} = (\alpha_{i,j}, \beta_{i,j})^T$,

$$h_\lambda(\boldsymbol{\theta}_{i,j}) = \left[\frac{\omega_{i,j}}{|\omega_{i,j}|}\left(|\omega_{i,j}| - \lambda\right) \quad \frac{\boldsymbol{\gamma}_{i,j}}{\|\boldsymbol{\gamma}_{i,j}\|_2}\left(\|\boldsymbol{\gamma}_{i,j}\|_2 - \lambda\right)\right]^T.$$

2. The step size t is computed via backtracking. First set $t = t_{\text{init}}$. For step size t, compute the candidate solution

$$\hat{\boldsymbol{\theta}}_{i,j;t}^{(k)} = h_{t\lambda}\left(\hat{\boldsymbol{\theta}}_{i,j}^{(k-1)} - t \cdot \nabla L_{i,j}\left(\hat{\boldsymbol{\theta}}_{i,j}^{(k-1)}\right)\right).$$

Then, while

$$L_{i,j}\left(\hat{\boldsymbol{\theta}}_{i,j;t}^{(k)}\right) > L_{i,j}\left(\hat{\boldsymbol{\theta}}_{i,j}^{(k-1)}\right) + \nabla L_{i,j}\left(\hat{\boldsymbol{\theta}}_{i,j}^{(k-1)}\right)^\top\left(\hat{\boldsymbol{\theta}}_{i,j;t}^{(k)} - \hat{\boldsymbol{\theta}}_{i,j}^{(k-1)}\right) + \frac{1}{2t}\left\|\hat{\boldsymbol{\theta}}_{i,j;t}^{(k)} - \hat{\boldsymbol{\theta}}_{i,j}^{(k-1)}\right\|_2^2,$$

shrink $t = mt$, for some $0 < m < 1$. In our implementation, $t_{\text{init}} = 1$ and $m = 0.5$.